BOOK OF ABSTRACTS

53rd Annual Meeting of the Societas Linguistica Europaea
26 August – 1 September 2020

Place: SLE 2020 Platform
SOCIETAS LINGUISTICA EUROPAEA: **President:** Teresa Fanego (Santiago de Compostela), **Vice-President:** Mira Ariel (Tel Aviv), President-Elect: Johannes Kabatek (Zurich), **Secretary:** Bert Cornillie (Leuven), **Treasurer:** Lachlan Mackenzie (Amsterdam), **Editor FL:** Olga Fischer (Amsterdam) **Editor FLH:** Muriel Norde (Humboldt-Berlin), **Conference Manager:** Olga Spevak (Toulouse).

**EXECUTIVE COMMITTEE (the above plus):**
Marcin Kilarski (Poznań), Rik van Gijn (Leiden), Annemarie Verkerk (Saarlandes), Francesco Gardani (Zurich), Drinka (Texas, San Antonio), Eva Schultze-Berndt (Manchester)

**PROGRAMME COMMITTEE: Chair of the General Session:** Eitan Grossman (Jerusalem), **Chair of Workshops:** Jacek Witkoś (Poznań).

**Members:** Werner Abraham (University of Vienna), Karin Aijmer (University of Gothenburg), Shanley Allen (University of Kaiserslautern), Cormac Anderson (MPI-SHH Leipzig), Reli Argus (Tallinn University), Mira Ariel (Tel Aviv University), Jóhanna Barðdal (Ghent University), Edith L. Bavin (La Trobe University), Dorothee Beermann (NTNU Trondheim), Valeria Belloro (Autonomous University of Querétaro), Theresa Biberauer (Cambridge University), Walter Bisang (University of Mainz), Anna Bondaruk (John Paul II Catholic University of Lublin), Kasper Boye (University of Copenhagen), Wayles Browne (Cornell University), Concepción Cabrillana (University of Santiago de Compostela), Anne Carlier (Sorbonne University), Michelarena Cennamo (University of Naples Federico II), Bernard Comrie (University of California (Santa Barbara), Denis Creissels (University of Lyon), Sonia Cristofaro (University of Pavia), Hubert Cuyckens (KU Leuven), Pierluigi Cuzzolin (University of Bergamo), Andriy Danylenko (Pace University), Jesús de la Villa (Autonomous University of Madrid), Walter De Mulder (University of Antwerp), Nicole Delbecque (KU Leuven), Adina Dragomirescu (Institute of Linguistics Iorgu Iordan - Al. Rosetti (Bucharest), Michael Dunn (Uppsala University), Katarzyna Dziubalska-Kolaczky (Adam Mickiewicz University in Poznań), Martin Ehala (University of Helsinki), Victoria Escandell-Vidal (University Complutense of Madrid), Urtzi Etxeberria (CNRS-IKER), Caleb Everett (University of Miami), Thorhallur Eythoraðsson (University of Iceland), Benjamin Fagard (CNRS), Teresa Fanego (University of Santiago de Compostela), Ad Foonen (Radboud University Nijmegen), Livio Gaeta (University of Turin), María Del Pilar García Mayo (University of the Basque Country UPV/EHU), Antonio García-Gómez (University of Alcalá), Michalis Georgiafentis (National and Kapodistrian University of Athens), Klaus Geyer (University of Southern Denmark), Anna Giacalone (University of Pavia), Chiara Gianollo (University of Bologna), Spike Gildea (University of Oregon), Alessandra Giorigi (University of Venice), Francisco González García (University of Almería), Stefan Th. Gries (UC Santa Barbara), Eitan Grossman (Hebrew University of Jerusalem), Eva Hajicová (Charles University in Prague), Camiel Hamans (Adam Mickiewicz University in Poznań), Bjorn Hansen (University of Regensburg), Alexander Haselow (University of Rostock), Martin Haspelmath (MPI-SHH Leipzig), Katharina Haude (CNRS), Dag Haug (University of Oslo), Lars Hellan (NTNU Trondheim), Eugen Hill (University of Cologne), Hans Henrich Hock (University of Illinois at Urbana-Champaign), Richard Huyghe (University of Fribourg), Aritz Iruztun (CNRS-IKER), Andra Kalnača (University of Latvia (Riga), Katalin É. Kiss (Research Institute for Linguistics of the Hungarian Academy of Sciences (Budapest), Seppo Kittila (University of Helsinki), Ekkehard Koenig (FU Berlin), Bernd Kortmann (University of Freiburg), Gitte Kristiansen (Complutense University of Madrid), Lívia Körtvélyessy (P. J. Šafárik University in Košice), Martin Kimmel (Friedrich Schiller University Jena), Meri Larjuvaara (Åbo Akademi (Helsinki), Pierre Larivée (University of Caen), Nikolaos Lavidas (National and Kapodistrian University of Athens), Elisabeth Leiss (Ludwig Maximilian University of Munich), Diana Lewis (Aix Marseille University), Suliko Liv (Tallinn University), María-Rosa Lloret (University of Barcelona), Lucía Loureiro-Porto (University of the Balearic Islands), Silvia Luraghi (University of Pavia), María José López-Couso (University of Santiago de Compostela), Ian Maddieson (University of New Mexico), Martin Maiden (University of Oxford), Andrej Malchukov (University of Mainz), Francesca Masini (University of Bologna), Dejan Matic (University of Münster), Helle Metslang (University of Tartu), Amina Mettouchi (EPHE and CNRS LLACAN), Susanne Maria Michaelis (Leipzig University), Matti Miestamo (University of...
Helsinki), Marianne Mithun (University of California (Santa Barbara), Liljana Mitkovska (FON University), Edith Moravcsik (University of Wisconsin-Milwaukee), Nicola Munaro (Ca' Foscari University of Venice), Pieter Muysken (Radboud University), Belen Méndez Naya (University of Santiago de Compostela), Maria Napoli (University of Eastern Piedmont), Kiki Nikiforidou (University of Athens), Tatiana Nikitina (CNRS), Jan Nuyts (University of Antwerp), Karl Pajusalu (University of Tartu), Marco Passarotti (Catholic University of the Sacred Heart), Theodossia-Soula Pavlidou (Aristotle University of Thessaloniki), Thomas Payne (University of Oregon), Vladimir Plungian (Institute of Linguistics RAS (Moscow), Lola Pons-Rodríguez (University of Seville), Rodrigo Pérez-Lorido (University of Oviedo), Angela Ralli (University of Patras), Eric Reuland (Utrecht University), Nikolaus Ritt (University of Vienna), Jorge Emilio Rosés Labrada (University of Alberta), Anna Roussou (University of Patras), Francisco Ruiz de Mendoza Ibáñez (University of La Rioja), Cinzia Russi (The University of Texas at Austin), Andrea Sansò (University of Insubria), Barbara Schlücker (Leipzig University), Stephan Schmid (University of Zurich), Ilja Seržant (Leipzig University), Petra Sleeman (University of Amsterdam), Elena Smirnova (University of Neuchâtel), John Charles Smith (University of Oxford), Augusto Soares Da Silva (Catholic University of Portugal), Andrey N. Sobolev (Institute for Linguistic Studies (RAS (St. Petersburg), Andrew Spencer (University of Essex), Olga Spevak (University of Toulouse), Elisabeth Stark (University of Zurich), Sabine Stoll (University of Zurich), Cristina Suarez-Gomez (University of the Balearic Islands), Inesa Šeškauskienė (Vilnius University), Magda Ševčíková (Charles University in Prague), Catherine Travis (Australian National University), Beata Trawinski (IDS Mannheim), Jarmila Tárnýiková (Palacký University in Olomouc), Johan van der Auwera (University of Antwerp), Guido Vanden Wyngaerd (KU Leuven), Arie Verhagen (Leiden University Centre for Linguistics), Anna Verschik (Tallinn University), Jean-Christophe Verstraete (KU Leuven), Letizia Vezzosi (University of Florence), Nigel Vincent (University of Manchester), Jacqueline Visconti (University of Genoa), Ferdinand von Mengden (FU Berlin), Søren Wichmann (Leiden University), Björn Wiemer (Johannes Gutenberg University Mainz), Jacek Witkoś (Adam Mickiewicz University in Poznań), Alena Witzlack-Makarevich (University of Kiel), Magdalena Wrembel (Adam Mickiewicz University in Poznań), Il-Il Yatziv-Malibert (INALCO (Paris)).

SLE 2020 BOOK OF ABSTRACTS EDITOR Olga Spevak
Editor’s note: Abstracts that had not been updated in due time (especially, abstracts in pdf) have not been included as well as abstracts that would damage the whole document.
# TABLE OF CONTENTS

**GENERAL SESSION** ........................................................................................................ 14

- Aroldo de Andrade ........................................................................................................ 15
- Ekaterina Aplonova ...................................................................................................... 16
- Peter Arkadiev .............................................................................................................. 17
- Jóhanna Barðdal, Peter Alexander Kerkhof & Roland Pooth ........................................ 18
- Andreas Baumann & Nikolaus Ritt ............................................................................. 18
- Karin Beijering & Katarzyna Janic ............................................................................. 21
- Delia Bentley & Michela Cennamo ............................................................................. 22
- Henrik Bergqvist & Karolina Grzech ......................................................................... 24
- Matea Birtić & Ivana Brač ....................................................................................... 28
- Eva-Marie Bloom Ström & Matti Miestamo ............................................................... 29
- Anna Bondaruk & Ewa Willim .................................................................................... 31
- Joseph Brooks & Russell Barlow ................................................................................ 34
- Jorina Brysbaert & Karen Lahousse ......................................................................... 36
- Svetlana Burkova ........................................................................................................ 37
- Matthew Carter ............................................................................................................ 39
- Yi-Yang Cheng ............................................................................................................. 41
- Maria Cheremisinova ................................................................................................. 42
- Eleanor Coghill ............................................................................................................ 43
- Greville G. Corbett & Sebastian Fedden ..................................................................... 44
- Sonia Cristofaro ........................................................................................................... 46
- Carlo Dalle Ceste ........................................................................................................ 47
- Saverio Dalpedri .......................................................................................................... 47
- Don Daniels ................................................................................................................ 49
- Anna De Marco & Emanuela Paone ......................................................................... 51
- Stefan Dedio & Paul Widmer ..................................................................................... 52
- Pedro Díaz Lammertyn & María Sol Sansiñena ......................................................... 53
- Quy ngoc thi doan ....................................................................................................... 55
- Carmen Dobrovie Sorin ............................................................................................. 56
- Nina Dobrushina ........................................................................................................ 57
- Bridget Drinka ............................................................................................................. 58
- Katarzyna Dziubalska-Kołaczyk & William A. Kretzschmar, Jr. ................................ 59
Elad Eisen, Dmitry Nikolaev, Eitan Grossman & Steven Moran .................................................. 60
Thorhallur Eythorsson & Sigridur S Sigurdardottir ................................................................. 62
Benjamin Fagard & Alexandru Mardale ................................................................................. 62
Sebastian Fedden .................................................................................................................... 64
Maksim L. Fedotov, Sofia A. Oskolskaya & Natalia M. Zaika ............................................. 65
Alon Fishman, Daniel Asherov, Nicole Katzir, Inbal Arnon & Mira Ariel .......................... 67
Christian Forche .................................................................................................................... 68
Michael Franjieh, Greville G. Corbett & Alexandra Grandison .......................................... 70
Minoru Fukuda, Takeshi Furukawa & Koichiro Nakamura .................................................. 71
Livio Gaeta .............................................................................................................................. 71
Jerzy Gaszewski .................................................................................................................... 73
Ion Tudor Giurgea ................................................................................................................ 74
Michał Głuszkowski ............................................................................................................. 74
Matías Guzmán Naranjo & Olivier Bonami ....................................................................... 76
Camiel Hamans ..................................................................................................................... 78
Maj-Britt Mosegaard Hansen ............................................................................................. 79
Stefan Hartmann, Nikolas Koch & Antje Endesfelder Quick ............................................. 80
Martin Haspelmath ............................................................................................................... 81
Rachel Hatchard .................................................................................................................... 82
Lars Hellan ............................................................................................................................ 82
Yoselin Henriques ................................................................................................................ 84
Helen Hint & Elsi Kaiser ....................................................................................................... 85
Daven Hobbs ......................................................................................................................... 87
Ioana Jieanu .......................................................................................................................... 89
Nina Julich-Warpakowski .................................................................................................... 90
Johannes Kabatek .................................................................................................................. 92
Csilla Kász ............................................................................................................................ 93
Liudmila Khokhlova & Boris Zakharin ............................................................................... 94
Sergei Klimenko .................................................................................................................... 96
Ezequiel Koile & George Moroz .......................................................................................... 98
Viktorija Kostadinova .......................................................................................................... 100
Olga Krasnoukhova & Marc Tang ....................................................................................... 101
Joanna Wrzesień-Kwiatkowska & Björn Wiemer ............................................................. 102
Cristina Lastres-López ........................................................................................................ 104
Nikolaos Lavidas & Leonid Kulikov ................................................................................... 107
Mary C. Lavissière ............................................................................................................... 108
<table>
<thead>
<tr>
<th>Name</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oksana Lebedivna</td>
<td>110</td>
</tr>
<tr>
<td>Paulina Łęska &amp; Jacek Witkoś</td>
<td>111</td>
</tr>
<tr>
<td>Diana Lewis</td>
<td>112</td>
</tr>
<tr>
<td>Liina Lindström, Maarja-Liisa Pilvik &amp; Helen Plado</td>
<td>113</td>
</tr>
<tr>
<td>Andrea Listanti, Jacopo Torregrossa &amp; Liana Tronci</td>
<td>114</td>
</tr>
<tr>
<td>Christian Locatell</td>
<td>117</td>
</tr>
<tr>
<td>David Lorenz</td>
<td>118</td>
</tr>
<tr>
<td>Dana Louagie</td>
<td>118</td>
</tr>
<tr>
<td>Marcia dos Santos Machado Vieira</td>
<td>118</td>
</tr>
<tr>
<td>Anastasia Makarova &amp; Olivier Winistörfer</td>
<td>120</td>
</tr>
<tr>
<td>Maxim Makartsev &amp; Max Wahlström</td>
<td>121</td>
</tr>
<tr>
<td>Michalis Marinis</td>
<td>123</td>
</tr>
<tr>
<td>Lidia Federica Mazzitelli</td>
<td>125</td>
</tr>
<tr>
<td>Giulia Mazzola, Bert Cornillie &amp; Malte Rosemeyer</td>
<td>127</td>
</tr>
<tr>
<td>Susanne Maria Michaelis</td>
<td>127</td>
</tr>
<tr>
<td>Vítor Míguez</td>
<td>129</td>
</tr>
<tr>
<td>Steven Moran &amp; Eitan Grossman</td>
<td>130</td>
</tr>
<tr>
<td>George Moroz and Nina Dobrushina</td>
<td>130</td>
</tr>
<tr>
<td>Garbiñe Bereziartua &amp; Beñat Muguruza</td>
<td>132</td>
</tr>
<tr>
<td>Lars Naborn, Eline Zenner &amp; Dorien van de Mieroop</td>
<td>133</td>
</tr>
<tr>
<td>Oana Niculescu</td>
<td>134</td>
</tr>
<tr>
<td>Peter Juul Nielsen</td>
<td>135</td>
</tr>
<tr>
<td>Mads Nielsen &amp; Maria Messerschmidt</td>
<td>137</td>
</tr>
<tr>
<td>Dmitry Nikolaev</td>
<td>138</td>
</tr>
<tr>
<td>Elissavet Nouchoutidou</td>
<td>139</td>
</tr>
<tr>
<td>Verónica Orqueda &amp; Francisca Toro</td>
<td>140</td>
</tr>
<tr>
<td>Mojgan Osmani</td>
<td>142</td>
</tr>
<tr>
<td>Tania Paciaroni</td>
<td>143</td>
</tr>
<tr>
<td>Rebecca Paterson</td>
<td>145</td>
</tr>
<tr>
<td>Nantke Pecht</td>
<td>147</td>
</tr>
<tr>
<td>Erika Petrocchi</td>
<td>149</td>
</tr>
<tr>
<td>Dirk Pijpops</td>
<td>150</td>
</tr>
<tr>
<td>Margarita Popova</td>
<td>152</td>
</tr>
<tr>
<td>Tihomir Rangelov &amp; Julie Barbour</td>
<td>153</td>
</tr>
<tr>
<td>Monica-Mihaela Rizea</td>
<td>155</td>
</tr>
<tr>
<td>Oleksandra Romaniuk &amp; Inna Stupak</td>
<td>156</td>
</tr>
</tbody>
</table>
Malte Rosemeyer & María Sol Sansiñena ................................................................. 158
Jorge Emilio Rosés Labrada .................................................................................. 160
Paulette Roulon-Doko ......................................................................................... 161
Alba E. Ruz ............................................................................................................. 163
Vieri Samek-Lodovici ......................................................................................... 164
Stefan Savić ........................................................................................................... 166
Sergey Say ............................................................................................................. 168
Marlieke Shaw ..................................................................................................... 170
Leon Shor & Michal Marmorstein ....................................................................... 171
Kaius Sinnemäki & Francesca Di Garbo ............................................................... 173
Piotr Sobotka & Magdalena Żabowska ................................................................ 175
Barbara Sonnenhauser & Anastasia Makarova .................................................. 177
Przemysław Staniewski & Adam Golębiowski .................................................... 179
Piia Taremaa, Helen Hint, Maria Reile & Renate Pajusalu ................................... 180
Miriam Thegel ...................................................................................................... 182
Marili Tomingas ................................................................................................... 183
Rossella Varvara ................................................................................................. 185
Albert Ventayol-Boada ....................................................................................... 186
Paola Vernillo & Rossella Varvara ...................................................................... 188
Albert Wall, Philipp Obrist, Senta Zeugin & Johannes Kabatek ........................... 189
Nikolaus Wildner .................................................................................................. 191
Olivier Winistörfer .............................................................................................. 192
Jacek Witkoś ......................................................................................................... 194
Sally Wong, Eric Reuland & Martin Everaert ........................................................ 195
Rodica Zafiu ........................................................................................................ 196
Alexander Zahrer ................................................................................................. 198
Eva Zehentner ...................................................................................................... 199
Paulina Zydorowicz ............................................................................................. 200

POSTER PRESENTATIONS .................................................................................... 204
Omri Amiraz ......................................................................................................... 205
Roberto Batisti ..................................................................................................... 206
Anastasia Bauer & Svetlana Burkova ................................................................... 208
Đorđe Božović ..................................................................................................... 210
Sara Cerutti, Valentina Scarda & Francesca Volpato ............................................ 211
Man-ni Chu .......................................................................................................... 213
Riku Erkkilä, Iida Lankoski, Felix Mäkelä & Tuomas Koukkari ............................ 214
WORKSHOP PRESENTATIONS ................................................................. 252

WORKSHOP 1 .................................................................................. 253

The grammaticalization of manner expressions into complementizers or quotatives .............................. 253

Dmitry Gerasimov ............................................................................. 215
Tiit Hennoste, Külli Habicht, Helle Metslang & Külli Prillop .................. 217
Mihaela Iloiaia .................................................................................. 219
Asma Izadi Bidani, Nafiseh Taghva & Vahideh Abolhasanizadeh .............. 220
Irantzu Epelde & Orotz Jauregi ........................................................ 221
Sandra Jiménez-Pareja & Salvador Valera ........................................... 222
Andra Kalnača & Ilze Lokmane .......................................................... 225
Anastasia Kharlamova ....................................................................... 228
Hideki Kishimoto ............................................................................... 229
John J. Lowe, Adriana Molina-Muñoz & Antonia Ruppel ......................... 230
Ellison Luk & Jean-Christophe Verstraete ........................................... 232
Ilya Makarchuk ............................................................................... 233
Stepan Mikhailov ................................................................................ 235
Carmen Mirzea Vasile ........................................................................ 236
James Myers ..................................................................................... 239
Ricardo Napoleão de Souza & Kaius Sinnenäki ..................................... 241
Bastian Persohn ................................................................................ 242
Elisa Piccoli, Anna Cardinaletti & Francesca Volpato ............................ 244
Shu-Ing Shyu .................................................................................... 245
Dmitri Stichinava ............................................................................. 246
Vasileios Symeonidis ........................................................................ 247
Jen Ting ............................................................................................. 247
Annemarie Verkerk, Luigi Talamo & Shahar Shirtz ................................. 249
Niina Ning Zhang ............................................................................. 250

The grammaticalization of manner expressions into complementizers or quotatives .............................. 253

Borja Ariztimuño-Lopez .................................................................. 255
Wojciech Guz ................................................................................... 256
Rodrigo Hernández ......................................................................... 258
Dorota Krajewska ............................................................................ 261
Eda-Riin Leego & Denys Teptiuk ....................................................... 262
Julian Rentzsch ................................................................................. 264
Natalia Serdobolskaya & Irina Kobozeva ........................................... 265
Stef Spronck ...................................................................................... 267
Yvonne Treis ....................................................................................... 268
Nigel Vincent .......................................................... 269
Björn Wiemer .......................................................... 271

**WORKSHOP 2** .......................................................... 274
Towards a diachronic typology of middle voice systems .......................................................... 274
Uldis Balodis .................................................................. 277
Carlota de Benito Moreno .................................................. 277
Maria Cristina Lo Baido .................................................. 279
Yiwei Fang, Xianhui Ye, Song Lu & Foongha Yap .................................................. 280
Silvia Luraghi & Dionysios Mertyris .................................................. 282
Yankee Modi & Mark Post .................................................. 284
Maarten Mous .................................................................. 284
Naonori Nagaya .................................................................. 285
Lengson Ngwasi .................................................................. 287
Sara Pacchiarotti & Leonid Kulikov .................................................. 289
Cristina Sánchez López & Margot Vivanco .................................................. 290
Andrea Sansò & Dawid Gajewski .................................................. 292
Emily Smith .................................................................. 293
Kenneth Van Bik .................................................................. 294

**WORKSHOP 3** .......................................................... 298
‘The grammar of thinking: Comparing reported thought and reported speech across languages’ .................................................................................. 298
Bert Cornillie .................................................................. 301
Sophia Fiedler .................................................................. 302
Alessandra Giorgi .................................................................. 304
Karolina Grzech .................................................................. 305
Anja Hennemann .................................................................. 307
Prapatsorn Tiratanti .................................................................. 308
Sofia Yaroshevich .................................................................. 309

**WORKSHOP 4** .......................................................... 312
Derivational zero affixes .................................................................................. 312
Diana Anițescu .................................................................. 314
Pavel Caha, Karen De Clercq & Guido Vanden Wyngaerd .................................................. 316
Bożena Cetnarowska .................................................................. 318
Serena Dal Maso & Sabrina Piccinin .................................................. 319
Anna Maria Di Sciullo .................................................................. 321
Matthias Gerner .................................................................. 322
Heidi Harley .......................................................... 324
Nikos Koutsokos & Angela Ralli .................................................. 324
Costanza Marini & Elisabetta Jeżek .............................................. 326
Chiara Melloni & Gianina Iordăchioia ....................................... 327
Verginica Barbu Mititelu, Svetlozara Leseva & Ivelina Stoyanova ........................................................................... 329
Yoshiki Ogawa ........................................................................... 330
Matteo Pellegrini & Fabio Montermini ..................................... 332
Marc Richards .......................................................................... 332
Magda Ševčíková ..................................................................... 334
Petra Sleeman .......................................................................... 336
Salvador Valera ....................................................................... 336

WORKSHOP 5 ........................................................................... 337

Multifunctionality and syncretism in non-finite forms .................. 337
  Bernat Bardagil ...................................................................... 339
  Daria Bikina, Denis Rakhman, Aleksey Starchenko & Svetlana Toldova ......................................................... 339
  Natalia Cáceres Arandia .......................................................... 342
  Ivan Gamov ............................................................................. 342
  Magdalena Lemus Serrano .......................................................... 345
  Nikita Muravyev ...................................................................... 346
  Eric Reuland ............................................................................ 346
  Krzysztof Stroński and Saartje Verbeke .................................. 348
  Jonathan North Washington, Francis M. Tyers, & Ilnar Salimzianov .............................................................. 348

WORKSHOP 6 ........................................................................... 351

Neglected syntactic functions and non-syntactic functions of applicative morphology .............................................. 351
  Valeria Generalova .................................................................. 360
  Hannah Gibson, Lutz Marten, Maarten Mous, & Kristina Riedel .......................................................... 361
  Lilíán Guerrero ...................................................................... 363
  Hilde Gunnink ...................................................................... 365
  Katarzyna Janic ...................................................................... 366
  Shuan Karim .......................................................................... 368
  Dominique Knuchel ................................................................ 369
  Martin Kohlberger .................................................................. 371
  Yury Lander & Irina Bagirokova .................................................. 372
  Marianne Mithun ..................................................................... 373
  Manoah-Joël Misago, Koen Bostoen & Sara Pacchiarotti .............................................................. 374
  Sergio Ibáñez, Alejandra Ortiz & Armando Mora ........................................................................... 376
Doris Payne ................................................................. 378
Thomas E. Payne & Voltaire Q. Oyzon ......................................................... 378
Mark Post & Yankee Modi ............................................................................. 379
Ali Salehi ...................................................................................................... 379
Camille Simon ............................................................................................... 380
Christina L. Truong & Bradley McDonnell .................................................. 382
An Van linden .................................................................................................. 383
Jozina Vander Klok .......................................................................................... 385
Fernando Zúñiga ........................................................................................... 386

WORKSHOP 7 ......................................................................................... 388

Discourse phenomena in typological perspective ......................................... 388
Doriana Cimmino ............................................................................................ 390
Annette Herkenrath & Birsel Karakoç ............................................................ 392
Katsunobu Izutsu & Mitsuko Narita Izutsu .................................................... 394
Łukasz Jędrejowski & Wojciech Guz .............................................................. 395
Atsuko Kanda Utsumi .................................................................................... 398
Timur Maisak .................................................................................................. 398
Elena Martínez Caro ....................................................................................... 400
Tatiana Nikitina & Ekaterina Aplonova ......................................................... 401
Jesús Olguín Martínez .................................................................................... 401
Vladimir Panov ............................................................................................... 403
Ekaterina Rakhilina & Polina Bychkova ....................................................... 404
Seongha Rhee ................................................................................................. 405
Stefan Schnell, Geoffrey Haig, Nils-Norman Schiborr & Maria Vollmer ......... 406
Martine Vanhove ............................................................................................. 408

WORKSHOP 8 ......................................................................................... 410

Contact and the architecture of the language faculty .................................... 410
Theresa Biberauer ........................................................................................... 412
Adnana Boioc Apintei & Ştefania Costea ....................................................... 413
Ricardo Etxepare, Georg Kaiser & Simon Dold ............................................ 414
Cristina Guardiano & Melita Stavrou ............................................................ 415
Alexandru Nicolae & Adina Dragomirescu .................................................... 417
Diego Pescarini ............................................................................................... 417
Cecilia Poletto & Alessandra Tomaselli .......................................................... 417
Anna Roussou ................................................................................................. 418
Leonardo M. Savoia & Benedetta Baldi ......................................................... 419
Alessandra Tomaselli, Andrea Padovan Ermenegildo Bidesi ........................................ 421

**WORKSHOP 9** ............................................................................................................. 422

**New perspectives on word order flexibility** ............................................................... 422
Rodrigo Becerra ................................................................................................................ 424
Laura Becker & Matías Guzmán Naranjo ....................................................................... 425
Christian Ebert, Balthasar Bickel & Paul Widmer .......................................................... 427
Dina El Zarka .................................................................................................................... 428
Alex Kramer ...................................................................................................................... 430
Natalia Levshina .............................................................................................................. 432
Zoey Liu ............................................................................................................................ 432
Chiara Naccarato, Anastasia Panova & Natalia Stoynova ............................................... 433
Savithry Namboodiripad ................................................................................................. 435
Rachel Nordlinger, Gabriela Garrido Rodriguez, Sasha Wilmoth & Evan Kidd .......... 435
Eva Schultze-Berndt ........................................................................................................ 437
Luigi Talamo & Annemarie Verkerk ............................................................................. 438
Marc Allassonnière-Tang ............................................................................................... 440

**WORKSHOP 10** ......................................................................................................... 442

**Theoretical, Applied and Experimental Perspectives** .................................................. 442
**on the Influence of English Today** .............................................................................. 442

Anabella-Gloria Niculescu-Gorpin .................................................................................. 442
Ana-Maria Barbu .............................................................................................................. 444
Valentina Cojocaru ......................................................................................................... 445
Henrik Gottlieb ................................................................................................................ 445
Aaditya Kulkarni ............................................................................................................. 447
Cristian Moroianu, Anabella Gloria Niculescu-Gorpin & Monica Vasileanu .................. 448
Rania Papadopoulou & George J. Xydopoulos ............................................................... 449
Gillian Roberts, Eline Zenner & Laura Rossee ............................................................... 450
Cristina Andreea Stan ....................................................................................................... 452
Tobias Ungerer & Alex Lorson ....................................................................................... 453
Monica Vasileanu & Anabella-Gloria Niculescu-Gorpin ................................................ 454

**WORKSHOP 11** .......................................................................................................... 455

**Discourse Marker use: from production to comprehension** ....................................... 455
Mira Ariel .......................................................................................................................... 456
Joanna Blochowiak, Cristina Grisot & Liesbeth Degand .................................................. 457
Zoé Broisson & Liesbeth Degand .................................................................................... 459
Maria-Josep Cuenca ....................................................................................................... 460
Eva Hajičová, Jiří Mírovský & Barbora Štěpánková ................................................................. 462
Maria Cristina Lo Baido ............................................................................................................. 464
Óscar Loureda, Inés Recio Fernández & Adriana Cruz .......................................................... 465
Hilla Polak-Yitzhaki & Yael Maschler ...................................................................................... 467
Pekka Posio & Malte Rosemeyer ............................................................................................. 469
Ariadna Ştefănescu & Sorina Postolea .................................................................................... 470
Valentina Schiattarella & Valentina Serreli ........................................................................... 472
Jennifer Schumann & Sandrine Zufferey ............................................................................... 473
Ekaterina Tskhovrebova & Sandrine Zufferey ...................................................................... 474
Frances Yung, Jana Jungbluth & Vera Demberg .................................................................... 476
GENERAL SESSION
The discourse role of NP Pseudocleft Types in Portuguese

Aroldo de Andrade
(Federal University of Minas Gerais)

Keywords: Pseudoclefts, Discourse functions, Rhetorical relations, SDRT, Portuguese

This work examines the discourse functions of three NP Pseudocleft types in Portuguese, traditionally dubbed canonical, reverse and extraposed, following the terminology for Brazilian Portuguese cleft constructions in Braga, Kato & Mioto (2010):

1. Canonical pseudocleft:
   \[ \text{CLEFT CLAUSE} \rightarrow \text{COPULA} \rightarrow \text{CLEFT CONSTITUENT} \]
2. Reverse pseudocleft:
   \[ \text{CLEFT CONSTITUENT} \rightarrow \text{COPULA} \rightarrow \text{CLEFT CLAUSE} \]
3. Extraposed pseudocleft:
   \[ \text{COPULA} \rightarrow \text{CLEFT CONSTITUENT} \rightarrow \text{CLEFT CLAUSE} \]

This classification assumes the type of relativizer as crucial. In de Cesare (2017), on the other hand, (1c) would be classified together with cleft sentences.

The guiding theoretical framework is Segmented Discourse Representation Theory or SDRT (Asher & Lascarides 2003). The main assumption is that discourse structure connects to information structure in a regular way. Therefore, the hypothesis is that the types of rhetorical relations (RRs) linking an NP pseudocleft to its context constrain the informational packaging to be found, following de Andrade (2019).

The corpus consisted of 34 interviews broadcast on the Brazilian TV show Roda Viva. The exclusion criteria were sentences with neutral demonstratives as cleft constituents, which may have propositional reference (Gast & Wiechmann 2012), and sentences in turn-taking points. After collecting the sentences, which amounted to 80, they were classified according to the RR linking them to the context.

The results have confirmed the hypothesis. First, canonical pseudoclefts occur only in subordinating RRs, or in the coordinating RR Correction, respectively serving to mark new information focus or contrastive focus; the example below illustrates the former case, and is linked by Comment:

(2) Carlos Velloso
   [Context: In 1995 there was a thematic subcommittee on Party Reform.]
   \[ e \ o \ que \ se \ preconizou \ foi \ mesmo \ a \ fidelidade \ partidária \]
   ‘And what was advocated was indeed partisan allegiance.’

Reverse pseudoclefts occur only in coordinating RRs, serving to mark a contrastive topic; below, the RR is Result:

(3) Carlos Lyra
   [Context: I do not believe that Elizeth Cardoso’s disc marked the beginning of Bossa Nova, but João Gilberto’s]
   \[ \text{então aquele LP do João Gilberto foi o que marcou o início da bossa nova} \]
   ‘So what marked the beginning of bossa nova was that album by João Gilberto.’
Finally, extraposed pseudoclefts occur either in environments combining a subordinating and a coordinating relation, or two coordinating RRs, one of them being Correction, serving to mark focus and contrastive topic on different clausal constituents, as below:

(4) Ferreira Gullar
   [Context: There is nobody else able to make the changes for a human being.]
   **é ele quem tem que fazer**
   ‘He (the human being) is the one who must do it.’

This work confirms that RRs are a relevant tool for carrying out detailed studies on the discourse functions of cleft constructions than those based on information structure alone (e.g. Prince 1978). More specifically, different NP pseudocleft constructions are distinguishable in terms of signaling a focus, a contrastive topic or a combination thereof.

Acknowledgments
I thank CNPq for the financial support (grant nr. 437100/2018-9) and the Collaborative Research Center Prominence in Language at the University of Cologne for support during a visiting period in December 2019.

References
Asher, Nicholas & Alex Lascarides (2003), Logics of Conversation, Cambridge: Cambridge University Press.

An unusual reported speech construction in Bashkir and Chuvash
Ekaterina Aplonova

<not updated>
Grammaticalization by semantic enrichment: from progressive to proximative to avertive in Lithuanian

Peter Arkadiev

(Institute of Slavic Studies of the Russian Academy of Sciences and Russian State University for the Humanities, Moscow)

Keywords: Lithuanian, grammaticalization, avertive, proximative, implicature

Avertive is a cross-linguistic gram type expressing an event that was “potentially imminent but did not get realized” (Kuteva 1998). According to Kuteva (2009), it is a “semantically elaborate” gram combining meanings from the temporal (pastness), aspectual (imminence) and modal (counterfactuality) domains. Kuteva (1998) also proposes a recurrent grammaticalization path from past volition to avertive and further to the proximative, which is a semantically simpler aspectual gram denoting pure imminence (Heine 1994). This development results from the loss of the semantic components of counterfactuality and pastness, i.e. semantic bleaching commonly assumed to be a hallmark of grammaticalization.

In this paper I argue for a different grammaticalization path of the avertive involving semantic enrichment rather than impoverishment, probably more compatible with a “loss-and-gain” model (e.g. Brems 2011). The evidence comes from the historical development and synchronic distribution of the Lithuanian (Indo-European > Baltic) construction involving a past tense auxiliary and a present active participle with the continuative prefix be-. In Old Lithuanian (Ambrazas 1990: 180–181), the construction (with the prefix be- on the participle being optional and tense of the auxiliary unrestricted) expressed the progressive (1). In modern Lithuanian, the construction has two main functions: avertive (2) and past proximative (3) (examples from LtTenTen14 corpus).

1) Ir buw-a piem-is t-o ye s chal-y ant lauk-a
   and AUX-PST.3 shepherd-NOM.PLDEM-LOC.SG.F place-LOC.SG.on field-GEN.SG
   be-gan-a
   CNT-pasture-PRS.PA.NOM.PL.M
   ‘And there were in the same country shepherds abiding in the field.’ (1579, Ambrazas 1990: 180)

2) Jau buv-o be-lip-qas ant žem-ės, bet už-kliuv-o
   already AUX-PST.3 CNT-climb-PRS.PA.NOM.SG.M on ground-GEN.SG but
   PVB-stumble-PST.3
   ‘He was already climbing ashore, but stumbled...’

3) J-is j-si-my lėj-o mergin-q, kur-i buv-o
   3-NOM.SG.M PVB-RFL-love-PST.3 girl-ACC.SG which-NOM.SG.F AUX-PST.3
   be-iš-vyk-sta nt-i į Amerik-q, ved-ė ir iš važiav-o.
   CNT-PVB-go-PRS.PA.NOM.SG.F in America-ACC.SG marry-PST.3 and PVB-drive-PST.3
   ‘He fell in love with a girl who was about to leave for America, married her and left.’

The avertive function, accounting for about 75% of the tokens in the corpus and almost invariably occurring in elicitation, appears to be the most prominent use of the construction. However, my corpus study (Arkadiev 2019) has revealed that in most cases the avertive interpretation arises when the context explicitly indicates non-occurrence of the event, while in the absence of such indications the construction is ambiguous between the avertive and the past proximative readings. This suggests that its semantics involves a combination of pastness and imminence, whereas counterfactuality is due to an implicature naturally arising when past proximative is used (Ziegeler 2000): if the speaker utters (i)
I was about to fall instead of (ii) I fell, the hearer can infer that (ii) is not true, because if it were true the speaker would have used the stronger statement (ii) instead.

I suggest that the Lithuanian construction started as a (focalized) progressive which, probably due to competition with other tense-aspect forms, got restricted to the past contexts and turned into a past proximative (cf. Johanson 2000: 153–154), and is now further developing into an avertive by the familiar process of the conventionalization of implicatedness (Traugott & König 1991).

References

Ambrazas, Vytautas (1990), Sravnitel’nyj sintaksis pričastij baltijskiх jazykov [Comparative Syntax of Participles in Baltic Languages], Vilnius: Mokslas.

Oblique Anticausatives: Dependent-Marked Anticausativization
Jóhanna Barðdal, Peter Alexander Kerkhof & Roland Pooth

<not updated>

Abundance and other correlates of linguistic stability: formal predictions and evidence from sparse lexical items

Andreas Baumann & Nikolaus Ritt
(University of Vienna)

Keywords: linguistic stability, English lexicon, language change, acquisition, abundance

Linguistic stability, i.e. the extent to which linguistic constituents are stably established and resist change, has been operationalized in diverse ways, such as frequency in synchronic use, diachronic growth, diachronic fluctuation (drift) or ease of acquisition (Pagel et al. 2007, Monaghan 2014,
Newberry et al. 2017). In this study, we focus on the stability of words, i.e. lexical constituents, and analyze in which way various measures of linguistic stability hang together. Most prominently, we assess the validity of ‘lexical abundance’, i.e. the fraction of speakers using a lexical item, for measuring stability. We approach the problem from two methodologically different angles. First, we analyze correlations between the chosen measures empirically based on English lexical data. Second, we derive relationships between these measures analytically by inspecting a population-dynamic model of linguistic spread.

The empirical analysis is based on diachronic English data (Corpus of Historical American English, Davies 2010; 18 decades) and token frequencies (using base-form frequency as proxy for lemma frequency), crowd-sourced age-of-acquisition (AoA) ratings and abundance data, i.e. estimates of the fraction of speakers knowing an item (Kuperman et al. 2012). Since we focus on the relevance of abundance to lexical stability, we restricted our analysis to relatively sparse items (words with abundance less than 0.9, resulting in exactly 9728 items). To cater for the diachronic aspects of linguistic stability, a linear model was fit to the normalized frequency trajectory of each item. From this model, the slope and residual standard deviation were extracted to measure diachronic growth and drift, respectively. Pairwise relationships between all derived variables were assessed (Figure 1).

Formal predictions are based on a population-dynamical ODE model of learner-user dynamics (Nowak 2000, Baumann & Ritt 2018). In short, it models the dynamics of the fraction of users of a given linguistic item on the population level under the assumption of spread through learner-user interactions and homogeneous mixing. We derive abundance and AoA by means of the population-dynamic equilibrium, while growth is derived as the model’s intrinsic growth rate. We show that abundance is expected to depend positively on frequency and growth, and negatively on drift and AoA.

Our results show, first, that acquisition and abundance are closely related to each other. In fact, frequency in use is less relevant to acquisition than is the number of people using a lexical item. This hints at the role that generalization plays during learning as well as at the communicative function of language. Second, we find that diachronic growth is only weakly related to abundance, AoA and frequency, which contradicts our formal prediction. This clearly has methodological reasons, since rate of growth—as a measure of linguistic success—can only be reasonably estimated for innovations. Third, we see that abundance behaves as expected with respect to frequency, AoA and drift. We conclude that abundance qualifies as a reasonable proxy for linguistic stability, in spite of the methodological difficulties involved in estimating it (Eisenstein et al. 2014, Gries 2016) and we stress the relevance of sparse items for understanding language change.
Figure A1. Pair-plots together with linear model (orange) and corresponding beta coefficients (all non-trivial at p<.05).

References
Nowak, M. A. (2000), The basic reproductive ratio of a word, the maximum size of a lexicon. Journal of Theoretical Biology, 204(2), 179–189.

A cross-linguistic study of benefactive double object constructions: Constraints, variation and grammaticalization
Keywords: double object construction, grammaticalization, plain beneficiary, recipient-beneficiary, typology

Double object constructions (DOCs) contain two objects as in *Tom bought Mary a cat* (Larson 1988). Semantically, they can code benefactive events (as shown above), where the beneficiary assumes a plain beneficiary, recipient-beneficiary or deputative-beneficiary role (Van Valin & LaPolla 1997). Benefactive DOCs display peculiar lexico-semantic and morphosyntactic constraints on the verb (Colleman 2010). In Santome, the verb ‘steal’ can occur in a DOC to express a maleficiary argument, (1a). However, to express a plain beneficiary with the same verb, the language must use a serial verb construction, (1b).

(1) Santome (Hagemeijer 2000:106)
   a. Ladlon futa mu djelu.  
      thief steal 1SG money  
      ‘The thief stole money from me.’
   b. Ladlon futa djelu da mu.  
      thief steal money give 1SG  
      ‘The thief stole money for me.’

Yet, in Tukang Besi, some verbs are not licenced in (benefactive) DOCs unless accompanied by special morphology. For instance, the verb ‘fetch’ introduces only one theme argument in the basic construction, (2a). However, to introduce an additional beneficiary argument, the same verb must occur in a DOC and make use of applicative morphology, (2b).

(2) Tukang Besi (Austronesian; Donohue 1999a:256)
   a. no-ala te kau 3.REALIS-fetch the wood  
      ‘She fetched the wood.’
   b. no-ala-ako te ina-su te kau 3.REALIS-fetch-APPL the mother-my the wood  
      ‘She fetched the wood (as a favor) for my mother.’

Languages also display variation with regard to how to code a beneficiary in benefactive DOCs. In Thai, the recipient-beneficiary is preferably coded without extra dative marking, as in (3a).

(3) Thai (Kra-Dai; Jenny 2010:382)
      student TAKE flower give teacher  
      ‘The student gives a flower to her teacher.’
   b. Nák.rion háy dɔɔk.máay (kɔ) khruu.  
      student give flower (DAT) teacher  
      ‘The student gives a flower to her teacher.’

This study aims to document (cross-)linguistic variation and constraints on benefactive DOCs in the world’s languages, and to investigate their diachronic changes with particular reference to Scandinavian. Three research questions guide our investigation: (i) What are the key-defining
properties of benefactive DOCs in typologically different languages? (ii-iii) How do these properties change and constrain benefactive DOCs over time, and why?

We will conduct the investigation within Basic Linguistic Theory (Dixon 2012), using two data collection methods: gathering (secondary) data from grammars, and conducting a diachronic case study based on (digital) original texts from prominent (Danish-) Norwegian authors to trace changes in benefactive DOCs from the 17th century to now.

We expect benefactive DOCs to be grammatically unstable and prone to change. The preliminary results from Norwegian, where (benefactive)DOCs are possible with ever-fewer verbs, being gradually replaced by alternative prepositional (tolfor-) objects (i.e. dative alternation) point in this direction. We also expect languages to pass by lexico-semantic and syntactic constraints on (benefactive) DOCs by using other linguistic strategies such as serial verb construction, applicative morphology or prepositional objects.

References
Van Valin, Robert D. and Randy J. LaPolla (1997), Syntax: structure, meaning, and function, Cambridge: Cambridge University Press.

Thematic and lexico-aspectual constraints on V-S agreement:
Evidence from Northern Italo-Romance

Delia Bentley & Michela Cennamo
(University of Manchester/University of Naples Federico II)

Keywords: agreement, subjecthood, event structure, definiteness, Italo-Romance

It has long been known that if, in a given language, the verb agrees with a following argument in a given syntactic function, it will also agree with a preceding argument in the same syntactic function (Moravcsik 1978). While traditionally the relatedness of agreement to word order has been attributed to its anaphoric properties (Lehmann 1982), the study of the dialects of Italy has brought to light the sensitivity of V(erb)-S(ubject) agreement to verb class and to the definiteness/topicality of the postverbal S (Parry 2000, 2013; Cennamo/Sorace 2007; Bentley 2018, a.o.). In particular, in many dialects of the North, agreement of the verb with a following argument in subject function varies in a

(1) a. Chel film lì, l’ an vist tìti i me amis / tanti person. [TRANSITIVE]
   that film DEM OCL have.3PL seen all the my friends many people
   ‘That film, all my friends / many people have seen it.’

   b. An ciamà i tò gent / tanti malà. [UNERGATIVE]
   have.3PL called the.PL your people many patients
   ‘Your parents / Many patients have called.’

   c. In / gh’ è rivà i to surèi / / di pac. [UNACCUSATIVE]
   be.3PL.CL be.3SG left the your sisters of parcels
   ‘Your sisters / some parcels have arrived.’

(San Giovanni di Persiceto, Bologna)

(3) a. I an / (à) ciamè i tò genitûr. [UNERGATIVE]
   SCL have.3PL have.3SG called the.PL your parents
   ‘Your parents have called.’

   b. (Ala fèsta) i an / (à) balè dimónndi invidê. [UNERGATIVE]
   at.the party SCL have.3PL have.3SG danced many guests
   ‘Many guests danced at the party.’

   c. Ai è arivê al tò surèl/ socuànt pâc. [UNACCUSATIVE]
   CL.CL be.3SG arrived the your sisters some parcels
   ‘Your sisters / some parcels have arrived.’

We test the hypothesis that V-S agreement is sensitive to the thematic scale of subjecthood in (2), attempting to shed light on the role played by the interaction between the elements of meaning
lexicalized in the verbal root (e.g., the type of change, whether target or result, Parsons 1990) and the event structure of the predicate in determining the agreement implicational scale proposed in Bentley (2018) and the variation displayed by the data investigated.

References
Bentley, Delia (2018), Grammaticalization of subject agreement on evidence from Italo-Romance, *Linguistics* 56(6), 1245-1301.
Tortora, Christina (1997), The Syntax and Semantics of the Weak Locative, University of Delaware, PhD thesis.

What’s theoretical about evidentiality?

Henrik Bergqvist & Karolina Grzech  
(Stockholm University)

Keywords: evidentiality, epistemic authority, grammaticalization, language use, language documentation
A narrow view of evidentiality states that the semantic core of (grammaticalized) evidentials can be characterized as the speaker’s “source of information” for a talked-about event (Aikhenvald 2004, 2014). This definition has become a reference point for cross-linguistic explorations of evidentiality, but there are some noted problems with a narrow approach and the definition it advocates (e.g. Cornillie 2010, Guentchéva 2018). In this talk, we focus our discussion on two problems: i) a sufficient level of abstraction is lacking for making adequate theoretical generalizations, ii) adhering to “information source” as the core meaning of evidentiality confines analyses of evidentials based on their use, to the margins as exceptions, or “effects” (Aikhenvald 2004: 179). We address these problems using a comparative, corpus-based approach.

With respect to the first point, we focus on the grammaticalization of evidentials. We argue that theoretical claims concerning evidentiality has failed to abstract away from the meaning of individual markers, thus leading to inadequate generalisations. It is a well-known fact that verbs of perception can be polysemous and that ‘seeing’ may be equated to ‘knowing’, and ‘hearing’ to ‘understanding’ (Viberg 1983, Sweetser 1990). Verbs of “saying” can also be used to attribute mental states to others (e.g. ‘feelings’, ‘wants’, ‘thoughts’, McGregor 1994, Loughnane 2005) in addition to utterances. It has not been demonstrated how these polysemous meanings and uses relate to the encoded meaning of corresponding evidentials that result from a process of grammaticalization. Lexico-semantic content gives way to functional meaning as part of the process of grammaticalization (e.g. Bybee et al. 1994), so why would a mode of access, such as ‘seeing’ be preserved semantically in grammaticalized evidentials? We argue that the field of evidential research has oversimplified the categories it regards as theoretically grounded by permitting folk definitions and direct translation/paraphrase to be used in lieu of abstracted analysis.

To address the issue of what role evidentials-in-use should play in the definition of such forms, we discuss data from distinct varieties of Quechua. In the typological literature, languages belonging to this language family are viewed as having a set of grammaticalized evidentials (e.g. Aikhenvald 2004, 2018, de Haan 2013). We show that the synchronic picture emerging from descriptions of individual varieties is much more complex than the literature suggests. Using examples taken from corpus data, we demonstrate that cognates of purported evidential markers encode meanings related to ownership of and authority over knowledge rather than information source (Grzech 2016, Hintz & Hintz 2017). The narrow approach to evidentiality rejects such findings instead of allowing this semantic complexity to be incorporated in the analysis of evidential systems.

We would be well served to revisit some of the long-standing assumptions and conceptions that have been cemented in the literature, and we need to make room for new data resulting from the ever-growing documentation of evidentials systems. This would allow us to shape a coherent theory of evidentiality; one that makes adequate predictions about the systems we actually encounter, rather than confining us to making neatly-structured taxonomies.

References


Sweetser, Eve (1990), From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure, Cambridge: Cambridge University Press.


**A semantic constraint on optional that in English: the role of factivity**

Gabriela Bîlbîie & Israel de la Fuente  
(University of Bucharest & University of Lille)

Keywords: complementizer drop, embedding, factivity, acceptability judgment task, English

The present study constitutes an experimental investigation into the semantic factor of factivity in relation to the complementizer omission in English.

The complementizer omission in English has been extensively studied from different perspectives and by taking into account various linguistic and non-linguistic factors, as summarized by Jaeger (2006, 2010): dependency length and position of Complement Clause (CC), overt production difficulty at CC onset, lexical retrieval at CC onset, lexical retrieval before CC onset, ambiguity avoidance at CC onset, grammaticalization, etc. In addition to the previously identified factors, Jaeger (2010) observes a significant effect of information density, i.e. the lower the information density of the embedded clause onset, the more likely the omission of the complementizer. However, to our knowledge, no previous work considered the semantic class of the embedding predicate, with the exception of Dor (2005), who takes into account the lexical semantics of the embedding predicates but not factivity.

In recent experimental work, we tested the role of verb semantics in embedded gapping and we observed that factivity has indeed an effect on the behaviour of embedded gapping. Despite the observed cross-linguistic variation (English and French vs. Spanish and Romanian vs. Persian), a general semantic constraint is at work: non-factive verbs (e.g. think, believe, suspect, imagine) embed more easily than factive ones, and semi-factives (knowledge verbs, e.g. realize, observe, see, know) embed more easily than true factives (emotion verbs, e.g. regret, be surprised, be bothered, like),
independently of ellipsis (Kiparsky & Kiparsky 1970, Karttunen 1971). On the basis of our previous experimental results, we test at present the interaction between factivity and complementizer omission in English, outside coordination and ellipsis contexts. Our hypothesis is that true (emotion) factive predicates are less likely to cooccur with complementizer omission than non-factive ones, and that semi-factive (knowledge) predicates are less resistant to complementizer drop than true factives.

We ran an acceptability judgment task (20 experimental items, 24 distractors and 16 ungrammatical controls) by using a 2x2 design with two factors (FACTIVITY and COMPLEMENTIZER), giving rise to 4 conditions, cf. (1). Within the [+FACT] conditions, half of the predicates used were true factive and half were semi-factive. The rating (Likert) scale was 1-7. We tested 51 participants from the USA (on Amazon’s Mechanical Turk).

\begin{enumerate}
\item a. [+COMP, +FACT] I regret that Albert visited very strange places during his trip.
\item b. [+COMP, –FACT] I suspect that Albert visited very strange places during his trip.
\item c. [–COMP, +FACT] I regret Albert visited very strange places during his trip.
\item d. [–COMP, –FACT] I suspect Albert visited very strange places during his trip.
\end{enumerate}

Our results show no differences in judgments with non-factive verbs with or without a complementizer, unlike with factive verbs where the absence of the complementizer results in significantly lower acceptability judgments than when the complementizer is present (cf. Figures 1 and 2). Linear mixed models reveal significant main effects of Complementizer and Factivity (both at p<.001) and a significant Complementizer*Factivity interaction (p<.001). Therefore, true factives and semi-factives do not give rise to exactly the same acceptability judgments: bare clauses (without that) embedded under a semi-factive verb are more acceptable than those embedded under a true factive verb.

In order to fully account for the complementizer omission phenomenon in English, we need thus a multifactorial analysis (cf. Jaeger 2006, 2020) that also takes into consideration a semantic distinction in terms of factivity. Unlike previous studies which consider semi-factives either as closer to true factives (Hooper & Thompson 1973) or to non-factives (Dor 2005), our approach is based on a three-fold semantic analysis of factivity, which seems crucial as evidenced by the gradience in acceptability judgments.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{factive vs. non-factive}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{semi-factives vs. true factives}
\end{figure}

References
Dor, Daniel (2005), Toward a semantic account of that-deletion in English, *Linguistics* 43(2), 345-382.

**Psychological object experiencer verbs in Croatian**

Matea Birtić & Ivana Brač  
(Institute of Croatian Language and Linguistics, Zagreb)

Keywords: psychological verbs, object experiencer verbs, Croatian language, causative alternation, reflexive verbs

Much attention has been paid to the psychological predicates in the last few decades (Beletti & Rizzi 1988, Grimshaw 1990, Croft 1933, Arad 1988, Landau 2010, Grafmiller 2013, Alexiadou & Iordăchioaia 2014). The main topic regarding research of psychological predicates was a possibility of experiencers to be lexicalized either as subjects or objects, a pattern observed in many languages. The same phenomenon has been noted for Croatian. In Croatian experiencer are expressed as a nominative subject (1), accusative object (2), and dative object (3):

(1)  Ivan se boji pasa. ‘Ivan fears dogs’  
Ivan.nom. refl. fear.3.sg.pres. dogs.gen.pl.

(2)  Psi plaše Ivana. ‘Dogs frightens Ivan.’  
dogs.nom.pl. frighten.3.sg. pres. Ivan.acc.

(3)  Ivanu se sviđa taj pas. ‘Ivan likes this dog.’  
Ivan.dat. refl. likes.3.sg.pres. this.nom.sg. dog.nom.sg.

Our talk will focus on the Croatian Object Experiencer verbs (ObjExp) whose experiencers are marked with accusative case (2). We have noted that almost all Croatian psychological ObjExp verbs have a corresponding reflexive verb with similar meaning. In the reflexive variant, the experiencer is a nominative subject, and the second argument (non-experiencer) argument is realized either as a prepositional phrase or a noun phrase marked with the genitive, dative, and instrumental case. The transitive and reflexive variant of the verb *ljutiti* ‘enrage / anger / be angry at’ is illustrated in (4) and (5).

(4)  Političare ljute najavljeni prosvjedi sindikata.  
politicians.acc.pl. anger.3.pl.pres. announced union’s protests.nom.pl. ‘The announced union’s protests anger politicians.’

(5)  Političari se ljute na najavljenje prosvjede sindikata.

28
politicians.nom.pl. refl. anger.3.pl.pres. at/of announced union’s protests.acc.pl.
‘The politicians were angry at the announced union’s protests.’

Our research will be based on the corpus of 95 psychological ObjExp verbs which we have collected so far. It aims to thoroughly investigate the grammatical and semantic properties of ObjExp verbs to see if the ObjExp psychological verbs resemble prototypical transitive verbs or whether they form a special class with some particular behavior which tells them apart from the other prototypical transitive verbs. As has been noted already for other languages, a transitive variant and a reflexive variant of the same psychological verb could be considered as a kind of psych-causative alternation (Alexiadou & Iordâchioaia 2014). In the Croatian language, as well as in other Slavic languages, the anticausative member of causative alternation is marked with the reflexive marker se (‘self’), e.g. Ivo je otvorio prozor. ‘John opened the window’, Prozor se otvorio. ‘The window opened.’, the same morphology which is visible with the pairs of psych verbs. Hence, our research will focus on a few points: 1. establishing whether all ObjExp verbs have a reflexive counterpart or not, 2. how many ObjExp verbs can be passivized, 3. whether passivization is related to perfective aspect and prefixation, 4. whether the morphological form of the non-experiencer argument with a reflexive variant is predictable of the behavior of passivization of the transitive variant, etc.

Eventually, we hope to give a better description of the Croatian ObjExp verbs and perhaps to get some insight into the nature of psychological verbs in general.

This work has been fully supported by Croatian Science Foundation under the project IP 2019-04-7896 (Syntactic and Semantic Analysis of Arguments and Adjuncts in Croatian, SARGADA).

References
Alexiadou, Artemis & Gianina Iordâchioaia (2014), The psych causative alternation. https://www.academia.edu/11844275/The_Psych_Causative_Altention (15. 2. 2019.)

The use of the augment in Nguni languages with special reference to the referentiality of the noun
Eva-Marie Bloom Ström & Matti Miestamo
(University of Gothenburg & University of Helsinki)

Keywords: augment, Bantu languages, negation, Nguni languages, referentiality

In a sub-set of Bantu languages the noun class prefix (NCP) is preceded by a prefix referred to as the augment (AUG). This paper explores the connection between (non-)referentiality and the use of the augment in a number of language varieties in the Nguni sub-group of Bantu languages. The structure
of a typical Bantu noun is: AUG-NCP-STEM, e.g. a-ma-doda ‘men’ in Xhosa. There has been considerable interest in this morpheme in the Bantu literature. Its occurrence vs. non-occurrence “often involves an intricate interplay between phonological, morphological, syntactic and semantic/pragmatic factors” (Hyman and Katamba 1993: 209) as well as perhaps sociolinguistic factors (Petzell and Kühl 2017). The contexts in which the augment is used vary across the language family and it has proved difficult to give a unified syntactic or semantic/pragmatic account of its function (van de Velde 2018; Halpert, in press).

The use of the augment has been connected to referentiality, the presence of the augment marking the noun as referential and its absence expressing non-referentiality. Negatives (and other non-fact contexts) induce non-referential readings: indefinite NPs in the scope of negation are typically non-referential. In the Xhosa examples in (1), the affirmative has the augment and can be interpreted either referentially or non-referentially. In the negative, the augment is typically absent and the reading is non-referential (referential readings under negation tend to be definite and they involve object cross-reference on the verb).

(1) Xhosa (Taraldsen 2010: 1526–1527)

   a. ndi-bon-a a-ba-fundi
      1SG-see-FV AUG2-NCP2-student
      ‘I see the/some students’

   b. a-ndi-bon-i ba-fundi
      NEG-1SG-see-FV NCP2-student
      ‘I don’t see any students’

The connection between negation and non-referentiality has been observed, e.g., by Givón (1978), and surveyed in a typological perspective by Miestamo (2014) in the larger context of investigating structural differences (asymmetries) between affirmatives and negatives. The above-described behaviour of the augment in Nguni and some other Bantu languages is in line with the typological tendencies observed.

In this study, we set out to investigate to what extent the referential function of the augment can be identified in contemporary Nguni languages. The study of these closely related varieties gives new insights into developmental tendencies of the augment. Our data come from our own fieldwork as well as from various textual and descriptive sources. All contexts in which the augment can be omitted in our data are either negatives or other non-fact contexts. Contrary to what has previously been argued for some varieties, however, the presence or absence of the augment is not exploited to make a referentiality distinction. Referentiality provides a diachronic semantic/pragmatic explanation for the absence and presence of the augment in different contexts, but this function is eroded in present-day Nguni. What remains is a limited referentiality distinction for some speakers in some varieties. The loss of function explains why the augment is included in the noun in nearly all contexts in some varieties, and omitted everywhere in others. Due to its loss of function, the augment has become free to participate in sociolinguistic and stylistic variation in some Nguni languages.
Abbreviations
AUG augment, FV final vowel, NCP noun class prefix, NEG negative/negation, SG singular. Numbers preceding an abbreviation refer to person (e.g. 1SG). Numbers following an abbreviation refer to noun class (e.g. AUG2).

References
Taraldsen, Knut Tarald, (2010), The nanosyntax of Nguni noun class prefixes and concords, Lingua 120, 1522–1548.

Causation in Adversity Impersonals in Polish

Anna Bondaruk & Ewa Willim
(John Paul II Catholic University of Lublin & Jagiellonian University)

Keywords: Adversity Impersonals, adjunct control, causer, location verbs, the Polish language

Adversity Impersonals (AIs) found in Slavic languages, including Russian and Polish, are causative structures referring to events of external causation out of human control, which may be caused by natural forces and may have adverse effects (Schlund 2020). Since only externally caused eventualities can form AIs, obligatorily agentive verbs (zamordować ‘to murder’) or internally caused unaccusatives (zwiędnąć ‘to wilt’) are absent from AIs. The questions whether the causer argument is overtly present in AIs and how it is syntactically represented have been extensively discussed in the literature (Kibort 2008, Junghanns et al. 2017, Lavine and Babby 2019, a.o.).

The aim of the paper is to check whether Polish AIs exhibit a syntactically represented overt or covert causer argument. Another question addressed is whether the presence of a causer argument is responsible for the structural accusative, as in (1):

(1) Las zasnuło mgłą.
‘The forest was enveiled in fog.’
Polish AIs are incompatible with *by itself* modification, as in (2), which shows that they denote a caused event.

(2) Zniosło samochód na pobocze (*samo z siebie).
   pulled car-ACC on roadside by itself
   ‘The car was pulled onto the road side (*by itself).’

There are two approaches to the presence of a causer in AIs: (i) the causer is implicit, and is projected in the syntax as an external argument, and (ii) the causer is overtly realized as an instrumental case-marked nominal, (1). The Polish data analysed in this paper seem to argue against both of these stances. First, it is shown that AIs in Polish do not contain a syntactically projected implicit causer, either as a null indefinite pronoun coś ‘something’ (Kibort 2008), or as a referentially reduced weather pronoun (for Russian, Szucsich 2007). The evidence that a causer argument is syntactically projected in AIs based on adjunct control (Szucsich 2007, Junghanns et al. 2017) is problematic, since the data evoked either exhibit eventive control or the controller is determined pragmatically/semantically, thus representing non-obligatory, rather than obligatory control (Landau 2013). Instances of obligatory control, like (3), are unacceptable, which indicates that there is no syntactically projected implicit causer in AIs:

(3) *Zniosło samochód na pobocze [PRO włączając alarm].
   pulled car-ACC on roadside turning-on alarm
   ‘The car was pulled onto the roadside turning on the alarm.’

An instrumental nominal in Polish AIs, (1), found with location verbs (Rappaport Hovav 2018), represents a Theme argument co-occurring with a Locative argument – an accusative NP. It is not a causer, as it cannot refer to some natural forces such as the wind, the sun, etc. The instrumental is a complement of the verb (1), just like the locative PP in (2).

Consequently, Polish AIs lack a syntactically projected causer, but have it represented in the semantics. The accusative case assignment to the Location (1) and the Theme (2) is divorced from either an external argument in Spec, VoiceP or an instrumental nominal in VP activating the v_{CAUS}, as in Lavine and Babby (2019), in violation of Burzio’s Generalisation (1986) (e.g., Lavine 2014).

References

(Non)finiteness in concessive conditionals

Tom Bossuyt
(Ghent University)

Keywords: concessive conditionals; finiteness; subordination; typology

This talk is concerned with adverbial subordinate clauses called concessive conditionals. Three subtypes are distinguished, exemplified for English in (1)–(3) below. Like conditionals, concessive conditionals express a relationship between a protasis and an apodosis (‘if \( p \), then \( q \)’), but instead of one \( p \)-value, their protasis contains a set of antecedents: ‘if \{ \( p_1, p_2, \ldots \} \), then \( q \)’ (Haspelmath & König 1998: 564–566).

In the present paper I will present preliminary results from an on-going typological investigation into concessive conditionals. Its main goals are to investigate

(a) which coding strategies are used to express concessive conditionality in the languages of the world;
(b) which grammatical factors determine the coding strategy used in any given language.

In their study on European languages, Haspelmath & König (1998) show that many languages, including English, have formally heterogeneous concessive conditionals.

(1) scalar concessive conditional
   Even if it rains, we will go outside.
(2) alternative concessive conditional
   Whether it rains or not, we will go outside.
(3) universal concessive conditional
   No matter what / Whatever the weather is like, we will go outside.

Other languages, however, mark them uniformly as conditionals, thus reflecting their basically conditional meaning. An example of such a language is Godoberi (Nakh-Daghestanian), in which all concessive conditionals consist of a conditional converb plus an additive focus particle (Haspelmath & König 1998: 628):

(4) scalar concessive conditional
   \( \text{čai} \text{-}\text{a}\-\text{ala-la}, \text{iLe} \text{išqa-ru} \text{ma-n-iLibu-da}. \)
   [rain PL:NT-come-COND-also] we:ABS home-ELAT PL:H-go-FUT.PART-COP
   ‘Even if it rains, we will go outside.’
(5) alternative concessive conditional
   \( \text{čai} \text{-}\text{a}\-\text{ala-la}, \text{mili} \text{b-ax-ala-la}, \)
   [rain PL:NT-come-COND-also] sun N-fall-COND-also
   iLe išqa-ru ma-n-iLibu-da.
   we:ABS home-ELAT PL:H-go-FUT.PART-COP
   ‘Whether it rains or the sun shines, we will go outside.’
(6) universal concessive conditional
Whatever the weather will be, we will go outside.’

Haspelmath & König (1998: 625) hypothesize that the differences in the marking of concessive conditionals are associated with different subordinating strategies: finite-subordinating languages like English mark all three subtypes differently, whereas nonfinite-subordinating languages like Godoberi have uniformly-marked triplets. This suggests a bi-directional universal ‘finite subordination ↔ differential marking’ (or its contrapositive: ‘nonfinite subordination ↔ uniform marking’).

I will investigate whether Haspelmath & König’s (1998) hypothesis holds on a global scale. Recently several articles have been published, examining concessive conditionals in individual languages (e.g. AnderBois 2014; Menz 2016; several papers in Xrakovskij, ed., 2012). Initial results suggest that there are basically two groups of languages:

(i) a homogeneous group consisting of languages in which all concessive conditionals are uniformly marked as conditionals (like in Godoberi)

(ii) a heterogeneous group of languages with differentially-marked concessive conditionals, using a wide variety of coding strategies (as in English)

Whereas the former group shows strong correlations with nonfinite subordination and OV word order, the determining factors for the latter group are much harder to pin down.

References

Plural and non-plural nominal number in northeast New Guinea

Joseph Brooks & Russell Barlow
(University of Virginia & the Max Planck Institute for the Science of Human History)

Keywords: nominal number, non-plural category, relative values, Papuan languages, typology

This paper concerns a little-known nominal number opposition between PLURAL and NON-PLURAL in several Papuan languages. Nominal number in these languages is remarkable in that, as in the Pama-Nyungan language Kala Lagaw Ya (Comrie 1981), a SINGULAR nominal category (i.e., indexing one-and-only-one referent) is absent. The evidence from these languages suggests not only that SINGULAR is not obligatory, but also that number distinctions need not include an absolute value, but may instead be based on a relative distinction (cf. Durie 1986, Corbett 2001, Dixon 2012).

Our approach is based on fieldwork in two parts of northeastern Papua New Guinea, where four languages possessing this type of number distinction are known to exist: Pondi, Ambakich, Yigaves,
and Chini. Our evidence comes from elicitation of noun phrases and some textual evidence. For Chini there is a documentary corpus of narrative and conversation.

After discussing the genealogical relationships among these languages, we provide an overview of the formal marking of the categories PLURAL and NON-PLURAL. We propose that non-plurality encompasses but extends beyond the enumeration of singular entities in a way that challenges universalist assumptions about SINGULAR as an obligatory category.

In Pondi, although pronouns and determiners exhibit a crosslinguistically common singular/dual/plural distinction, nouns distinguish between (the default) non-plural (one or two) and plural (three or more):

(1)  
\begin{align*}
\text{\textit{kula-m mi}} & \quad \text{\textit{boy-NPL}} & \quad \text{3SG} \\
\text{‘the boy’}
\end{align*}

(2)  
\begin{align*}
\text{\textit{kula-m min}} & \quad \text{\textit{boy-NPL}} & \quad \text{3DU} \\
\text{‘the two boys’}
\end{align*}

(3)  
\begin{align*}
\text{\textit{kula-wi ndin}} & \quad \text{\textit{boy-PL}} & \quad \text{3PL} \\
\text{‘the (more than two) boys’}
\end{align*}

In Pondi, PLURAL is an absolute category: it applies invariably to sets of more than two referents. In Chini, however, absolute values alone are not predictive of usage; instead, the values are relative. Parallel to the Pondi examples, one pattern in Chini is for highly differentiated animates to be expressed as non-plurals in one’s and two’s, but as plurals in three’s or more. This is seen in (4), about a man’s two (non-plural) younger brothers and their three (plural) wives:

(4)  
\begin{align*}
\text{\textit{ŋi=} & \quad \text{\textit{anŋhi}} & \quad \text{\textit{ŋuñi}} & \quad \text{\textit{twaví …}} & \quad \text{\textit{ñji=} & \quad \text{\textit{amur-iyi}}} \\
\text{POSS.REFL=} & \quad \text{\textit{younger.brother.NPL}} & \quad \text{two} & \quad \text{with} & \quad \text{POSS.PL=} & \quad \text{\textit{wife-PL}} \\
\text{‘with his two younger brothers (and …) all their (three) wives’}
\end{align*}

The use of these categories is relative in that it is constrained by lexical semantics (e.g. high differentiation and animacy), and by semantic oppositions in specific contexts of use: collective/distributive, minimal/maximal, and subset/full set, among others (and, crucially, not by the presence of explicit numerals). In (4), the wives and the younger brothers are both total sets, but only the wives are expressed as such (through the plural). The relativity of values for subsets versus full sets is seen in (5), where the same value and referent type as in (4) (i.e., three people) is expressed through the non-plural, conveying their subset status (in a description of a scene involving four children in total).

(5)  
\begin{align*}
\text{\textit{twamingañi}} & \quad \text{\textit{ŋuñi noyi}} \\
\text{\textit{child.NPL}} & \quad \text{three} \\
\text{‘three children’}
\end{align*}

We conclude with crosslinguistic observations about the non-monolithic nature of non plurality as well as the possibility for number categories to encode absolute values serving enumerative functions or to express relative values serving expressive functions (cf. Mithun 1988).

References


**Contrastive subjects modified by contrastive adverbs in different registers of French**

Jorina Brysbaert & Karen Lahousse
(KU Leuven and Research Foundation – Flanders; & KU Leuven)

Keywords: contrastive subjects, contrastive adverbs, French, register, corpus analysis

I. The goal of this talk is to shed new light on contrastive subjects in French, by an analysis of the understudied case in which a contrastive adverb such as *par contre*, *en revanche* ‘on the other hand’ shows up between the subject and the verb (1):

(1) *En ce qui concerne les compétitions 2010-2011, si les dates sont d’ores et déjà fixées, les lieux, par contre*, ne seront pas connus avant la mi-septembre.

‘Regarding the 2010-2011 competitions, if the dates have already been fixed, the places, on the other hand, will not be known before mid-September.’ (*Est Républicain*)

Previous research mainly deals with the different types of contrastive discourse relations expressed by contrastive adverbs and are mostly based on a limited analysis of formal (written) data (Danjou-Flaux 1980). Studies which do consider their syntactic position (Csűry 2001, Hamma & Haillet 2002, and Dupont 2015) give hardly any attention to cases as (1). In addition, nothing is known about their distribution in different registers.


III. We first show that the frequency of contrastive subjects modified by a contrastive adverb varies depending on the corpus and the register. Contrastive adverbs occur more often between the subject and the verb in formal than in informal French. We hypothesize that this is due to a different trade-off between contrastive adverbs and other markers of contrastivity across registers.

IV. We then show that subjects modified by a contrastive adverb and the entities they are contrasted with are mostly subtopics of a more general discourse topic (Erteschik-Shir 2007) (*the dates* and *the places* in (1) both concern ‘practical information of the competitions’). This is further supported by the fact that the modified subject is most frequently a lexical NP. In addition, the two verbs are often
oppositional, which is marked lexically or grammatically (negation in (1)). Furthermore, our data show that contrastive adverbs are only rarely combined with other markers indicating contrastivity (e.g. *mais*, adversative *si* (1)), which confirms our hypothesis in III above. We also show that the contrastive subject is hardly ever emphasized by means of another construction (e.g. dislocation) on top of the contrastive adverb.

**References**


The expression of modality in Russian Sign Language

Svetlana Burkova

<note updated>

**Come here vs. will you come here:**
A corpus investigation of the social value of requests

Anouk Buyle
(KU Leuven; FWO)

Keywords: requests; address terms; power and solidarity; 20th-century; British English

In politeness research, requests are considered face-threatening acts (Brown & Levinson 1987; Blum-Kulka, House & Kasper 1989). A distinction can be made between more directly realized requests (e.g. imperatives) and more indirectly realized requests (e.g. references to ability or willingness). Furthermore, requests can be modified by strategies that are characterized as mitigating (e.g. hedges) or as aggravating (e.g. intensifiers). The different request types are often described as polite or impolite, and tentative distinctions are made between negatively and positively polite request types. Still, little research has focused on the interaction between request types and social dimensions such as power and solidarity (as noted by Culpeper & Archer (2008)), which are essential components in Brown & Levinson’s politeness theory (1987). The aim of this paper is to investigate how request types distribute over social space, answering the question who uses what to whom.

To that end, imperatives, ability requests and willingness requests as well as internal and external modification strategies (following Blum-Kulka, House & Kasper 1989 and Flöck 2016) have been
identified in a corpus of British English plays published between 1899 and 1912. The analysis of the different request types combines a top-down approach with a bottom-up approach. The top-down approach relies on previously annotated power and solidarity parameters (following Buyle forthc.), as well as additional parameters for ranking the degree of face threat (Brown & Gilman 1989) and affect (Kopytko 1995). The bottom-up approach makes use of the address system as a proxy to describing social relations between speakers, which can then be exploited to assess the social value of the different request types.

Based on a number of experimental studies (House & Kasper 1981; Faerch & Kasper 1989; Trosborg 1995), it is hypothesized for Present-Day English that (1) requests are realized more directly when the speaker is more powerful than the addressee; and that (2) mitigating modification strategies are more frequent in distant interactions, while aggravating modifiers are more frequent in close or conflicted interactions. For the corpus inquiry, then, it is expected that direct or aggravating request types are more frequent than indirect or mitigating request types in relations where speakers address social inferiors or intimates (top-down analysis). The same is expected in (1) relations where the speaker uses FIRST NAME for the addressee than in relations where the speaker uses TITLE + LAST NAME, and (2) relations where the speaker use FIRST NAME for the addressee, while the addressee does not, than in relations where FIRST NAME is reciprocal (bottom-up analysis).

A first study has revealed that results for the bottom-up analysis confirm results for the top-down analysis. Imperatives are most frequent when speakers address social inferiors with non-reciprocal FIRST NAME. Ability and willingness requests are not more frequent than mitigated imperatives in distant interactions, but occur more often when speakers address socially superior intimates.

References
Buyle, Anouk. Forthc. Dear, my dear, my lady, your ladyship: Meaning and use of address term modulation by my. Pragmatics.
What Licenses Polyfunctionality?: The Case of /b\textsuperscript{3}/ in Ket

Matthew Carter
(University of California, San Diego)

Keywords: Ket, polyfunctionality, morphology, Word-and-Paradigm, Siberian

This study investigates polyfunctionality in Ket (Yeniseian, central Siberia), a polysynthetic language with extensive polyfunctionality in its verbal system. Variability in the function of individual markers raises questions about how Ket learners reliably determine the meaning of verbs without deriving it compositionally. Based on analysis of data drawn from Vajda & Zimm (2004) and Kotorva & Nefedov (2015), I demonstrate that systematic, though non-compositional, patterns of co-occurrence between markers, together with generalization across paradigmatically-related forms, provide sufficient information to understand novel forms.

Polyfunctionality refers to recurrent associations of the same marker with different functions. Recent work in morphologic theory (Blevins 2016, \textit{inter alia}) argues that the internal structure of words primarily serves to discriminate between paradigmatically-related forms, rather than to build up words compositionally. Under this view, polyfunctionality emerges as simply one strategy for discriminating between related forms. The same pieces of words can be used in different combinations to mark morphological distinctions, without assigning a consistent function to a given marker in all combinations.

Work on Ket over the past two decades (Vajda 2001, 2003, 2004 \textit{inter alia}, Nefedov & Vajda 2015, Georg 2007) analyzes its verbal system according to a templatic model of 10 position classes (P\textsuperscript{8}-P\textsuperscript{1}). The order and formal identity of the markers in the verb is fixed, but formally and morphotactically identical markers may differ radically in their function.

I focus particularly on determining the argument structure of verbs containing the marker /b/ in P\textsuperscript{3} (hereafter \textit{“/b\textsuperscript{3}/”}). /b\textsuperscript{3}/ shows a consistent form and morphotactic distribution, always occurring in P\textsuperscript{3}, and always marks 3\textsuperscript{rd} person inanimate when it marks a verbal argument. However, /b\textsuperscript{3}/ can reference either the subject (1) or the object (2).

(1) /b\textsuperscript{3}/ as subject

\begin{verbatim}
anjipta /an\textsuperscript{7}-b\textsuperscript{3}-ta\textsuperscript{0}/
\hspace{1em} hang-b\textsuperscript{3}-ROOT
\hspace{1em} ‘it is hanging’

danista /du\textsuperscript{8}-an\textsuperscript{7}-ta\textsuperscript{0}/
\hspace{1em} 3.SG.M-hang-ROOT
\hspace{1em} ‘he is hanging’
\end{verbatim}

(2) /b\textsuperscript{3}/ as object

\begin{verbatim}
dusqavra /di\textsuperscript{8}-us\textsuperscript{7}-q\textsuperscript{5}-a\textsuperscript{4}-b\textsuperscript{3}-da\textsuperscript{0}/
\hspace{1em} 1.SG.M-warm-CAUS-LEX-B\textsuperscript{3}-ROOT
\hspace{1em} ‘I warm it up’

dusqavyra /di\textsuperscript{8}-us\textsuperscript{7}-q\textsuperscript{5}-a\textsuperscript{4}-ku\textsuperscript{1}-da\textsuperscript{0}/
\hspace{1em} 1.SG.M-warm-CAUS-LEX-2.SG-ROOT
\end{verbatim}
‘I warm you up’

In addition, /b³/ can be thematic, with no apparent meaning (3).

(3) Thematic /b³/

inbayavra
in⁷-ba⁶-k⁵-a⁴-b³-da⁰
guess-1.SG.-LEX-LEX-B³-ROOT
‘I guess’

inayavra
in⁷-a⁶-k⁵-a⁴-b³-da⁰
guess-3.SG.M-LEX-LEX-B³-ROOT
‘he guesses’

In some cases, the syntagmatics are sufficient to determine the argument structure. Alone, /b³/ always marks the sole argument of an intransitive verb (1). Together with a marker in P⁸, /b³/ always marks the object (2). With a marker in P⁶, /b³/ is always thematic (3). In other cases, knowledge of paradigmatically-related forms is necessary. As some verbs mark their subjects twice, and as /b³/ can mark an argument or be thematic, forms combining P⁸, P⁶ and P³ are syntagmatically ambiguous (4). However, knowledge of another form resolves the ambiguity.

(4)

dabugbit /da⁸-bu⁶-k⁵-b³-t⁰/ 3.SG.F-SG.F-LEX-B³-carry
‘she carries it’
*‘she carries her’
*‘she carries’

dabukkut /da⁸-bu⁶-k⁵-ku¹-t⁰/ 3.SG.F-SG.F-LEX-2.SG-carry
‘she carries you’

This study contributes to the literature on polyfunctionality and the organization of complex morphological systems, bringing together for the first time insights from Word-and-Paradigm morphology and the enigmatic and typologically anomalous Ket verb.

References
Keywords: differential agent marking; argument structure; symmetrical voice; Austronesian languages; Formosan languages

Western Austronesian languages are typically characterized with symmetrical voice systems, in which two or more morphologically basic voice constructions can be found (Himmelmann 2005). In addition to the number/types of voice constructions, one point of crosslinguistic variation regarding Western Austronesian voice lies in core argument alignment, with two major patterns observed. First, a language may show a binary case distinction, with either the Agent/Actor argument or the Patient/Undergoer argument marked as nominative, and the other core argument consistently marked in a different case, e.g. genitive ng in Tagalog (the Philippines) and oblique taltolnocalne in Tsou (Formosan), depending on the voice construction. A second commonly observed pattern involves a three-way case distinction, in which the non-nominative argument is marked differently across different voice constructions, e.g. the Patient/Undergoer as oblique in the Agent-/Actor-voice construction and the Agent/Actor as Genitive in the Non-actor-/Undergoer-voice construction in Atayal (Formosan).

The present paper examines the symmetrical voice system of a critically endangered, but highly understudied, Formosan language—Kanakanavu. Although there has been a thread of research on the implications of its voice system for Austronesian higher-order subgrouping (e.g. Zeitount & Teng 2016), one aspect that has been glossed over concerns how the argument alignment pattern does not correspond perfectly to either of the two patterns above. Specifically, there are three case alignment patterns that can be observed. Similar to the Atayal pattern, the Agent-voice construction (1) involves the Agent marked as nominative, and the Patient as oblique: in the Patient-voice construction (2), the Patient/Undergoer is marked as nominative, and the Agent as genitive:

(1) m-áca=ku
V-NEG-laugh=1SG.NOM 2SG.OBL.
(A) (P)
‘I laugh at you.’

(2) ni-pácaca= máku
PV-NEG-laugh=1SG.GEN 3SG.NOM
(A) (P)
‘I laughed at her/him.’

Intriguingly, Kanakanavu also exhibits a third construction, where although the Patient/Undergoer is in the Nominative case, the Agent is obligatorily marked as oblique; note, however, that the oblique agent is interpreted as especially high in agentivity:
The pattern observed in Kanakanavu presents a typological anomaly under the Western Austronesian context, as its Non-Actor-/Undergoer-voice constructions can be analyzed as showing *differential agent marking* (Fauconnier 2011), as seen in the coding of the Agent arguments across (2) and (3). Motivated by language-internal and comparative evidence (regarding especially the conservative status of the -el-ai voice suffix in (3)), it is hypothesized that Kanakanavu had a simpler, core-oblique, case distinction, which is preserved in its synchronically canonical transitive construction (instantiated as in (3) above), but that the overall alignment pattern became complexified (into a three-way, nominative-genitive-oblique, system) as nominalization strategies were reanalyzed as verbal voice (cf. Starosta et al. 1981, Ross 2009), introducing the genitive agent. As the reanalysis is still in progress, a case split ensues in the Agent argument of Non-actor-/Undergoer-voice constructions. The Kanakanavu alignment pattern, therefore, has implications not only for the evolution of voice in Austronesian languages, but also for how differential agent marking may arise (Seržant & Witzlack-Makarevich 2018) in a language with symmetrical voice.

**References**


**Ges What: Comparative-Attenuative Polysemy in Beserman Udmurt**

Maria Cheremisinova

<not updated>
Affinal kinship terms in North-Eastern Neo-Aramaic dialects

Eleanor Coghill
(University of Uppsala)

Keywords: kinship, Aramaic, Semitic, lexicon, dialectology

Affinal kinship covers relationships which involve marriage (English ‘in-laws’). Compared to consanguineal kinship it is less well-researched cross-linguistically. As with other kinship, its typology can be complex. Categories can involve the same factors as consanguineal kinship (as presented in Kroeber 1909 and Greenberg 1990), e.g. difference between generations, difference between lineal and collateral relationships, difference of age and difference of sex of the relative, ego or person through whom the relationship exists. Languages can vary widely in what distinctions they make among affinal relations, what categories they merge (what terms are colexified) and what kinds of forms they use for given concepts (e.g. monolexic or dilexic).

In the case of North-Eastern Neo-Aramaic (NENA), native to Iraq, Syria, Turkey and Iran and spoken by Christian and Jewish communities, we have the advantage of a large dataset from dozens of varieties (from grammars, fieldwork and the Cambridge North-Eastern Neo-Aramaic database), as well as from earlier Aramaic varieties. Thus, one can see variation and change in affinal terminology within a language family over thousands of years. Several affinal terms are deep-rooted, going back to Proto-Semitic. Some of the inherited Aramaic vocabulary is universal or near-universal across the dialects, but loanwords from Arabic and Kurdish have increased the amount of variation.

Conceptually, affinal kinship terms in NENA can be divided into four types: (1) spouses of ego’s consanguineal relations, (2) consanguineal relationships of ego’s spouse, (3) self-reciprocal relations of the third degree and (4) spouse of ego’s spouse.

Among terms for type (1), there is a tendency towards category merger. The words for women or men who marry one’s close blood relations are usually colexified (synchronically or historically) with the words for ‘bride’ and ‘bridegroom’ respectively, e.g. šəθma ‘bridegroom, daughter’s husband, sister’s husband etc.’

Among terms for type (2), there are some very specific monolexic terms, e.g. ʔəðma ‘husband’s brother’. This is a reflex of Late Aramaic yəbmā with the same meaning (Tezel 2003: 222–223) and is cognate with Hebrew yāḥām which is associated with the custom of Levirate marriage. ʔəθma has no such associations, at least not among Christian speakers. The historical importance of one’s father-in-law (xəmyāna ~ xmā) is reflected in the etymology of the term for ‘wife’s brother’, baraxmāya < *bar xmā [son.of father_in_law], an example of a dilexic term which has become monolexic.

The terms for type (3) include yəṭisa ‘wife’s sister’s husband’, an inherited lexeme. This has been replaced by an (adapted) Arabic loanword, ʃādīla, in some dialects.

For type (4), only one term, ʔarṭba ‘rival wife’ is attested, and only for a few dialects. Polygamy is forbidden among Christians and was rare among Iraqi Jews.

These and other affinal kinship terms will be discussed and the typology of NENA affinal kinship related to that of other Semitic languages and neighbouring languages, as well as cross-linguistically.

References
German gender: A statistical approach

Greville G. Corbett & Sebastian Fedden
(University of Surrey & Université Sorbonne Nouvelle and University of Surrey)

Keywords: gender, German, assignment rules, morphosyntax, typology

German has one of the most complex gender systems, one of great typological interest. Considerable progress has been made by researchers such as Klaus-Michael Köpcke and David Zubin (1982, 1983, 1984, and 1996). They have demonstrated clear regularities in the assignment of gender to German nouns, involving semantic, morphological and phonological assignment principles. But while we know many of the parts, there is no comprehensive analysis of the entire system, as has been done for other languages which rely on a combination of semantic and formal assignment rules, such as Russian (Corbett 1991, Corbett & Fraser 2000).

The basic semantic assignment rules are straightforward. Sex-differentiable nouns are assigned gender based on biological sex: der Mann ‘man’, die Frau ‘woman’, der Bulle ‘bull’, die Kuh ‘cow’. For most German nouns, however, gender assignment is formal, that is, it is associated with their phonology (Köpcke 1982, Köpcke & Zubin 1983) or their morphology (Augst 1975, Pavlov 1995, Bittner 1999, Kürschner & Nübling 2011).

To establish the predictive power of any given assignment rule, an important step is to put numbers to predictions. Thus when evaluating a proposed morphological or phonological assignment rule, we compare this rule to the overall distribution of nouns over the gender values. We can then ascertain whether prediction by the rule is better than chance. Table 1 gives percentages according to noun frequency, based on the CELEX database (Baayen et al. 1995).

<table>
<thead>
<tr>
<th>Most frequent …</th>
<th>MASCULINE</th>
<th>FEMININE</th>
<th>NEUTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>… 100 nouns</td>
<td>34%</td>
<td>35%</td>
<td>31%</td>
</tr>
<tr>
<td>… 1,000 nouns</td>
<td>37%</td>
<td>41%</td>
<td>22%</td>
</tr>
<tr>
<td>… 10,000 nouns</td>
<td>36%</td>
<td>45%</td>
<td>19%</td>
</tr>
</tbody>
</table>

We discuss both morphological and phonological assignment of German nouns, focusing on prediction rates of individual rules and how well they compare to the overall percentages in the database. For at least four inflection classes, gender can be predicted unambiguously (or nearly unambiguously) from the inflection class. These are powerful rules with an efficiency of up to 100%. Then there are several inflection classes which allow partial prediction of gender. Thus der Knauf ‘knob’ cannot be feminine based on its paradigm. To distinguish masculine nouns from neuter nouns
is less straightforward, but often further phonological cues help: for example, of the 10 monomorphemic, monosyllabic nouns with initial /kn/ in our database 9 are masculine and 1 is neuter.

For investigating the interaction between semantic and formal criteria, we examine an interesting suggestion, namely that German has semantic clusters, smaller than those of the semantic assignment rules, which predict a certain gender (Enger’s ‘crazy’ rules, 2009). For example, most nouns denoting alcoholic drinks are masculine. We evaluate ‘crazy’ rules by comparing their predictive power to the formal rules established from the corpus.

The German gender system involves many interacting regularities, of different types: semantic, phonological, morphological and ‘crazy’ rules. We report progress from treating the system as a whole and by putting numbers to assignment rules so that we can evaluate their effectiveness.

Acknowledgments
The support of the ESRC (grant: ES/R00837X/1: Optimal categorisation: the origin and nature of gender from a psycholinguistic perspective) is gratefully acknowledged.

References
Enger, Hans-Olav (2009), The role of core and non-core semantic rules in gender assignment, Lingua 119, 1281–1299.
Köpcke, Klaus-Michael (1982), Untersuchungen zum Genussystem der deutschen Gegenwartssprache, Tübingen: Niemeyer.
Zubin, David & Klaus-Michael Köpcke (1984), Affect classification in the German gender system, Lingua 63, 41–96.
On the diachronic emergence of alienability splits cross-linguistically: Two types of counterexamples to frequency-based explanations

Sonia Cristofaro
(University of Pavia)

Keywords: possession, typology, diachrony, overt marking, frequency

Cross-linguistically, if overt possessive markers are used for inalienable possession, they are usually also used for alienable possession, but not the other way round. An influential explanation for this pattern is that alienable nouns are not usually possessed, so the possession relationship is more difficult to identify, and hence more in need to be disambiguated through overt marking (Nichols 1988, Koptjevskaja-Tamm 1996, Haspelmath 2017; see Haiman 1985 for an alternative explanation in terms of iconicity).

The paper discusses the diachronic origins of overt possessive marking (case markers, adpositions, linkers, possessive classifiers) for alienable and inalienable possession cross-linguistically, based on combined evidence about the etymology of possessive markers in some fifty genetically and areally diverse languages and the distribution of individual markers across different types of alienable and inalienable possession. This evidence, it will be argued, provides two types of counterexamples to the idea that the use of overt marking reflects the relative need to disambiguate possession.

First, in many cases the development of individual markers cannot be assumed to be related to the need to encode the possession relationship. The markers are originally attributive demonstratives (‘that house (of) John > ‘John’s house’) or appositive pronouns or nouns (‘house, that (of) John’ > ‘John’s house’; ‘fish, John(‘s) one/thing/food’ > ‘John’s fish’. These elements are used to modify or denote the possessee, not to encode a relationship between possessor and possessee. While they evolve into possessive markers over time, this is plausibly because they take on the possessive meaning associated with the construction as a whole as their original meaning is bleached. The fact that the resulting possessive markers are restricted to alienable possession reflects the fact that the source construction is not usually used with inalienable nouns, for several reasons.

Second, several restrictions in the distribution of individual markers contrast with the relative need to disambiguate alienable or inalienable possession, in that the marker is sometimes used for inalienable, but not alienable possession, or may or may not be used for the same possession type (alienable or inalienable) in different contexts. This, however, can be related to the etymology of the marker. For example, markers derived from elements expressing origin from a (usually) nonhuman source are only used with nonhuman possessors, irrespective of inalienability (e.g. ‘the castle’s wall’, but not ‘John’s house’). Markers derived from locative and benefactive/recipient markers are used, respectively, when the possessee can be conceptualized in terms of its location and in terms of its function vis-a-vis the possessor, but they are not used when this is not the case (e.g. ‘the book’s cover’ vs. ‘John’s mother’; ‘John’s mother’ vs. ‘John’s book’), again independently of alienability.

In line with previous work by diachronically oriented typologists (Aristar 1991, Bybee 2008, Claudi and Heine 1989, Creissels 2001), these facts suggest that explanations for alienability splits and recurrent cross-linguistic patterns in general should refer to recurrent source constructions and developmental processes that give rise to these patterns in individual languages, rather than synchronic properties of these patterns in themselves.
References


Debonding beyond degrammaticalization: a view from the Pacific

Carlo Dalle Ceste

<not updated>

On modal verbal adjectives in Ancient Greek

Saverio Dalpedri

(Georg-August-Universität Göttingen)

Keywords: verbal adjectives; modality; Ancient Greek; semantic change; pragmatic inference.

Modality has long been a favourite topic in the scholarship on Ancient Greek language. However, most studies have mainly concentrated on the expression of modality and the speaker’s attitude (subjectivity) through finite verb forms, with particular reference to the modal particles, e.g. áν/κε(n), ára, dé, pou. In the realm of complex predicates, relatively well-studied are the so-called gerundives, i.e. verbal adjectives built with the suffix -téo- and expressing necessity or obligation.

The other class of Ancient Greek verbal adjectives, those carrying the suffix -tó-, is commonly taken to be equivalent to past passive participles (though they also can be described as resultative participles in the sense of Haskelmath (1994), for they are compatible with some unaccusative verb
roots. Cf. ex. 1), but it can express the modal notions of possibility (cf. ex. 2) and, sporadically and only in Homer, necessity (cf. ex. 3), too. When the adjective is used predicatively, only the modal usage is possible (cf. ex 2), while when it modifies an NP both usages are attested (cf. ex. 1 and 3). The explanation for this syntactic behaviour is straightforward: a passive (or resultative) verbal adjective used along with the copula would be functionally equivalent to the perfect construction.

While the function of the verbal adjectives in -tó- as resultative participles is inherited from PIE, their modal values seem to be an innovation. Benveniste (1948) already proposed to recognize the source for this innovation in the negated forms, both with syntactic negation and with affixation (a(n)-). In my presentation, I will show that the Homeric data seem to confirm Benveniste’s hypothesis. In particular, I will propose a pragmatic account for the development of modal predicates out of resultative verbal adjectives triggered by negation. The notion of possibility seems to be primary: the non-veridicality of an event can lead to a pragmatic inference over the possibility of the realization of the event—or, in other words, the hearer infers that something undone is such since it cannot be done, is undoable (epistemic modality). This process would thus be an instance of subjectification as described by Traugott (2003, 2010) and Narrog (2012), i.e. of grammaticalisation of the speaker’s attitudes towards the content of an uttered sentence. Although this way of interpreting semantic change as the conventionalisation of pragmatic meaning is promising, I reject Traugott’s emphasis on the speaker’s intentionality, as posited by her Invited Inferencing Theory of Semantic Change (e.g. Traugott & Dasher 2002).

In the scenario I will sketch, the semantic change at stake is not necessarily speaker-driven: instead, it is the hearer who infers an implicature (the non-possibility) over the circumstances of the non-veridicality of the negated event and, through reanalysis, can subsequently overextend this modal interpretation to non-negated, factual events. This is in line with Eckardt’s theory of Avoiding Pragmatic Overload (cf. Eckardt 2009, 2012). The non-perspicuous contexts which would favour a new, less costly interpretation by the hearer would be those where the verbal adjective is used predicatively with the copula: in order to avoid a tensed, periphrastic interpretation, which would otherwise be equivalent to the finite middle perfect, the listener tries to accommodate the quirky context providing their inference with salience.

The notion of necessity surfacing in a few instances of verbal adjectives with -tó- is found similarly (and exclusively) with negation, but it is explained in a different way. The starting point is the potential meaning: from this, a shift from possibility to necessity (as attested, e.g., for Eng. must = Germ. müssen – cf. van der Auwera & Plungian 1998, 99 and Narrog 2012, 187–190) could take place thanks to a change in the scope of the negation, which was made possible by the logical equivalence between ‘not possible’ (negation of modality) and ‘necessary not’ (negation of proposition), as per Palmer (2001, 90–91).

(1) peiśma d’ élusan apò trētoĩo
hawser:ACC.SG PTCL unleash:AOR.3PL off perforate:VA.GEN.SG
líthoio
stone:GEN.SG
‘they unleashed the ship’s cable off the perforated stone’ (Od. 13, 77)

(2) éntha málista ambatós esti polis
where mostly access:VA.NOM.SG is city:NOM.SG
‘where the city is most easily accessible’ (Il. 6, 433f.)

(3) ōikhet epopsómenos Kaka-ión ouk
leave:IMPF.3SG see:PTCP.FUT.NOM.SG.M evil-Ilión:ACC NEG
onomastên

48
mention: VA.ACC.SG
‘he (scil. Odysseus) left to see evil, not-to-be mentioned Ilion’ (Od. 6, 260)

References
Narrog, Heiko (2012), Modality, subjectivity, and semantic change: A cross-linguistic perspective, Oxford: Oxford University Press.

Aisi semi-embedded clauses resolve an insubordination paradox

Don Daniels
(University of Oregon)

Keywords: insubordination, subordination, syntactic change, discourse, Trans New Guinea

There is a growing literature on insubordination, the phenomenon by which formally subordinate clauses are used as stand-alone main clauses (Evans 2007; Evans & Watanabe 2016). Subordinate clauses usually convey backgrounded, non-asserted information (Cristofaro 2003), and subordinate constructions often retain this backgrounded flavor (Mithun 2008). But a particularly common kind of insubordination in Papuan languages is non-embedded nominalization, in which subordination is achieved via nominal morphology, and non-embedded nominalizations often convey vividness and foregrounding (Schapper & San Roque 2011). This creates a paradox. If insubordinate constructions originate in subordinate, backgrounded structures (which they do, by definition), how do they acquire a meaning that is precisely the opposite of that?

Aisi, a Trans New Guinea language of Papua New Guinea (Daniels 2015), offers a clue. Aisi subordinate clauses take the form of a finite clause followed by a demonstrative. This construction serves as a noun phrase in the matrix clause, and the demonstrative bears a case marker that signals its syntactic role. In (1), the middle accusative demonstrative gakug subordinates the clause abey ‘I’m talking’, and the subordinate clause is the object of iro ‘perceive’.
[Ab-eng] ga-kug ir-o!
talk-1SG.IPST MID-ACC perceive-2SG.IMP
‘Listen to what I’m saying!’

These nominalized clauses convey given, presupposed information (Reesink 2014). In special cases, though, they can convey new, asserted information. In (2), from a folk tale about giants, the subordinate clause introduces a new giant and asserts that it got up.

[Mo ga-ku kip-is-i] ga-ku, kibi niku yo-s-i.
another MID-NOM get.up-PST-3SG MID-NOM in.law 3SG.POSS hit-PST-3SG
‘Another one got up and killed his in-law.’

This use of subordinate clauses is subject to certain restrictions: the demonstrative must be the middle form, not proximal or distal, and it must be in nominative or topic case. Because these clauses are asserted and occupy high positions in the matrix clause structure (they are less “hierarchically downgraded,” per Lehmann 1988), I call them “semi-embedded.” This structure can also become fully insubordinate, subject to similar restrictions.

The key observation is that Aisi semi-embedded clauses are only seen at climactic points in narratives. As in (2), they contain a rise in narrative tension that sets the stage for a climax in the matrix clause. We can thus see that semi-embedded clauses were created by a process of extension (Mithun 2008, 2019) based on the following analogy. Just as the presupposed information in a subordinate clause sets the scene and prepares the hearer for the new information in the matrix clause, so the rise in a semi-embedded clause sets the scene and prepares the hearer for the matrix clause climax.

In the next step, the salience and vividness inherent in narrative climaxes becomes associated with the first clause, which eventually can be uttered on its own. This is a kind of elision (Evans 2007; Cristofaro 2016). This paper thus illustrates that the mechanisms of extension and elision can work in tandem to create insubordination, and in doing so resolves an apparent paradox in the meaning of non-embedded nominalizations.

References


“Mi scusi!”. Prosodic cues and perceived politeness in L2 Italian learners’ apologies

Anna De Marco & Emanuela Paone

The present study examines the relationship between prosodic cues and perceived politeness in L2 Italian learners’ apologies. Cross-linguistic research has shown that one of the most robust correlates of perceived politeness is pitch range (Nadeu, Prieto 2011). Although few studies have addressed this issue so far, some results seem to suggest that language and culture may play also a significant role in the overall perception of (im)politeness and more generally in the interpretation of pitch range variations (Chen et al. 2004; Brown et al. 2014).

In the light of the above, the study will first focus on the prosodic cues learners of Italian rely on to convey the speech act. For a comparative analysis, the study involved also native speakers of Italian (NSs). Secondly, the research will verify to what extent the variation of prosodic cues in NSs and NNSs’ apologies could affect the perception of politeness. We hypothesized that: a) learners would experience difficulties in modulating their intonation in accordance with different social contexts; b) NSs’ pitch range would be more variable and modulated than NNSs’; c) at a perceptual level, learners’ apologies might sound less emphatic, less sincere and less polite than NSs’. The implications of these results on SLA should be investigated in order to shed lights on learners’ prosodic and socio-pragmatic competence.

20 Spanish learners of Italian as L2 (with B1 level of proficiency) and 5 NSs of Italian participated in the study. Participants were required to enact six role plays and to perform the apologies in accordance with the given context.

Dialogues have been recorded by means of a professional audio recorder. Sequences containing the expression of apology (*mi scusi/scusami/mi dispiace, “I’m sorry”*) have been isolated and analyzed. Acoustic analysis has been carried out by using PRAAT, considering fundamental frequency mean values and pitch range.

A perception test was administered to a sample of 40 NSs of Italian. These were supposed to listen to a portion of a dialogue conveying the expression of apology uttered by native and non-native speakers and to rate the “tone of voice” employed by the speakers on different scales from 0 to 10: not sincere/sincere; not emphatic/emphatic; not deferent/deferent, not polite/polite; not appropriate/appropriate to the context. Participants were also asked to leave a personal comment to explain their evaluation.

Preliminary results have revealed that learners’ productions were less variable and were characterized by a narrower pitch range than NSs’ performances. At a perceptual level, learners’ apologies have been perceived as less emphatic, sincere and polite than NSs’.
References

Object multirepresentation in the history of the Indo-European language family

Stefan Dedio & Paul Widmer
(University of Zurich)

Keywords: object multirepresentation, language change, phylogenetic comparative methods, Indo-European

Some properties, such as subject–verb agreement, are pervasive to an extent that they can safely be considered to have played an important role at almost any point in time in the long attested and unattested history of Indo-European. For other properties, the evidence is less clear: for example, quite a few Indo-European languages, ancient and modern ones, allow for multirepresentation of direct and/or indirect objects at clausal level, cf. ex. (1)–(2).

(1) n=an=šan QEMANNAM išhûwâi
CONN=3SG.P=FTCL flour.P scatters
‘She scatters the flour.’ (Hittite, Middle script (CTH 433.2) KBo 17.105+ rev. iii 2–31; Sideltsev 2011: 85)

(2) Je-le-lui-ai donné son livre.
1SG.A-3SG.T-3SG.G-have.PST.3SG give.PTCP POSS.3SG book
‘I gave him/her his/her book.’ (Québécois French; Auger 1993: 178)

For obvious reasons, object multirepresentation (OMR) doesn’t figure prominently among the typical features of IE: its morphological matter is not obviously cognate (cf. -an- vs. -le- above), the morphosyntactic behaviour doesn’t match (Wackernagel vs. verbal affix), and the frequency as well as the restrictions on appearance differ substantially (rare in Hittite, common in Québécois with pronouns and full NPs). However, the very presence of OMR in two distant clades is intriguing and may very well bear a phylogenetic signal. In order to get a better understanding of the role of OMR in the history of the IE family, we collected information about its presence and frequency (ABSENT, RARE, COMMON) in a sample of 47 IE languages, modern and ancient, across all clades, cf. Fig. 1.

We then performed a phylogenetic analysis in a Bayesian framework and modelled the evolution of OMR as a Continuous Time Markov Chain (CTMC) using BayesTraits (Pagel and Meade 2017) in order to estimate the transition rates between three ordered states ABSENT ↔ RARE ↔ COMMON. The rate estimates were then exploited to perform stochastic character mapping (Huelsenbeck, Nielsen, and Bollback 2003) as implemented in the r package phytools (Revell 2012) and we simulated possible histories along the phylogeny of IE.

The results obtained suggest that given our data, the phylogeny, and the model we used, IE
languages spent an equal amount of time in the state of not having any OMR (ca 50%) and having OMR (RARE 25%, COMMON 25%; cf. the aggregated histories in figure [fig2]), changing states (ABSENT ↔ RARE, RARE ↔ COMMON) on average ± every 200 years. This is in line with the observation that the morphological matter and the morphosyntax have been subject to constant renewal (i.e. lost and developed anew) as mirrored by the fact that the morphological matter doesn’t reconstruct well to PIE across languages and clades.

We conclude that despite the lack of cognate matter in its morphological makeup, OMR is firmly rooted in the history of the IE language family. This paves the way for exploring in much more detail the possible processes that initiate, foster, and inhibit language change, in our case possibly cognitive benefits (cf. the prevalence of OMR in the signed modality?), contact events (e.g. on the Balkans), and genealogical biases (no OMR in Germanic, but present in all other clades).

References


The Spanish *<conditional + progressive>* at the interplay between lexical and grammatical aspect

Pedro Díaz Lammertyn & María Sol Sansiñena
(KU Leuven)

Keywords: conditional, lexical aspect, progressive marking, Spanish

Besides the well-described uses of the Spanish conditional (see Azzopardi 2013, Vatrican 2013, a.o.) in conditional constructions and with the temporal meaning of ‘future in past’, other meanings pointed out in the literature are various evidential uses. A unitary account of all the uses is proposed by Haverkate (2002:25): they concern “assertions on counterfactual states of affairs”. However, this explanation does not account for recent uses such as (1)-(2), which are on the increase in South American varieties, particularly in Rioplatense Spanish.

(1) *Estaría necesitando mucha plata*. (Twitter)
   Lit. I would be needing much money ‘I need a lot of money’

(2) A: *Alguien que frene esta intensidad que manejo.*
B: Estarías necesitando una bofetada de realidad. (Twitter)
A: Someone that slows down this intensity that I handle.
B: Lit. You would be needing a reality slap ‘You need a reality slap’

The evidential-attitudinal uses of <V-ría+ V-ndo> in (1)-(2) do not receive plain conjectural, prospective or attenuated interpretations. In these undescribed uses, the conditional, often in combination with progressive marking and stative verbs, gives rise to an ironic effect by pragmatic inferences. The use in (1) pragmatically conveys the opposite of the attenuated meaning associated with the progressive, viz. an urgent need or longing of the speaker. The construction in (2) counts as a dispreferred response: a previous statement or behavior is criticized as unrealistic. In a sense, these uses parody the hearsay or ‘journalistic’ conditional (Bermúdez 2016, a.o.) in order to express the speaker’s subjective stance. Gerund periphrases like <estar+V-ndo>, usually favor dynamic interpretations (Schwenter & Torres Cacoullos 2010) and have been traditionally considered incompatible with state predicates (Garachana 2016) except for a restricted set of contexts where the periphrasis coerces dynamic interpretations of state predicates, e.g. as a sequence of temporally linked states (García Fernández 2009), or progressive or ingressive meanings (Yllera 1999). Van Rompaey (2014) argues that the English <be+V-ing> structure with state predicates overrides the lexical aspect of the predicate and imposes a meaning of temporariness. Progressive marking a state predicate coerces a contingent reading less indicative of a structural state and involving more agency, and its markedness reflects the speaker’s intention to construe the state as atypical (De Wit et al. 2018:11). Accordingly, the state predicates in<V-ría+ V-ndo> may be coerced into different dynamic meanings such as ingressive, progressive or habitual.

This study aims to contribute to the discussion on the grammatical and lexical aspect of state and non-state predicates in progressive constructions by analyzing a recent set of constructional innovations in Rioplatense Spanish. The objectives are to describe their aspectual interpretations, and to explain the mechanisms of aspectual coercion and the role of markedness in their development. In particular, the study seeks to establish the inherent lexical aspect of the predicates, how it is affected by the aspectual configuration of the construction, and what role the conditional plays in the construal of aspect. The hypothesis is that these changes provide further evidence for the coercion of dynamic readings and that changes in the perception and construal of aspect have allowed the structure to extend to a wider group of predicates.

A series of case studies were performed on an ad-hoc corpus of 5,000 Tweets with <AUX-ría+V-ndo> (2010-2019). 500 random tokens were annotated for morphosyntactic, semantic-pragmatic and discourse-interactional features, e.g. the inherent lexical aspect of the predicates vs. the overall aspect of the construction, type of evidentiality and resources recruited for markedness. The data confirms that the constructional innovations co-occur more frequently with state predicates such as necesitar ‘to need’ and extrañar ‘to miss (someone)’. This study further contributes to the description of Rioplatense Spanish, language use in online communication and constructionalization.

References
Blocking effect in Vietnamese: The interaction of mình-binding with the 1st person value

Quý Ngọc Thị Đoán
(Utrecht University)

Key words: Blocking effect, non-local binding, the 1st person value, the performative frame, mình

The blocking effect is a phenomenon that has been observed in languages with non-local binding such as Mandarin Chinese (Huang 1984, Cole, Hermon and Sung 1990, Cole and Sung 1994, Giblin 2016 among others). Like Mandarin, Vietnamese has an anaphoric element mình ‘body’ that needs not be bound in the local domain and there are cases in which the non-local binding of mình is blocked, as illustrated in (1), (2) and (3):

(1) \[}\text{Nam}_i \text{ nghi} \text{ Hùng/mày}_j \text{ biết Mai thích mình}_i^{\text{giap}}\]
   Nam think Hung/you know Mai like body
   ‘Nam thinks Hung/you know(s) Mai likes him/you/me.’

(2) \[}\text{Nam}_i \text{ nghi} \text{ tôi}_j \text{ biết Mai thích mình}_i^{\text{siği}}\]
   Nam think I know Mai like body
   ‘Nam thinks I know Mai likes me.’

(3) \[}\text{Tôi}_i \text{ nghi} \text{ Nam}_j \text{ biết Mai thích mình}_i^{\text{rışj}}\]
   I think Nam know Mai like body
   ‘I think Nam knows Mai like me.’
In (1), the antecedent of mình can be the intermediate subject Hung/you or the matrix subject Nam, or the silent speaker from the discourse, marked as sp. In contrast, mình in (2-3) can only be coreferential with the first person antecedent, namely the first person pronoun tôi ‘I’, and the binding of mình by Nam is blocked. Hence, a blocking pattern is indeed observed. Generally, what is important to notice is that Vietnamese mình can always refer to the speaker even when the first person antecedent is non-overt.

In the literature, the blocking pattern in Vietnamese has not been discussed and this contribution will be the first step towards a solution.

Our general mechanism of syntactic binding is based on Multiple Agree (Hiraiwa 2005 and Giblin 2016), see (4) and (5). C\(^0\) and mình are initially unvalued, as in (4). After valuation by the local NP, C\(^0\) shares its value with mình by the agree operation, yielding (5). To account for the availability of a first person interpretation, I rely on the proposal of Ross (1970) that a sentence contains a performative frame with a silent, first person pronoun as the subject of a silent verb of saying or thinking. I assume that this performative frame is optionally realized in the syntax. If it is, the silent first person pronoun qualifies as the NP\(_{\text{val}}\) in (4), and mình is valued as the speaker.

\[
\begin{align*}
(4) \quad [ & C^0_{\text{up}} \left[ \begin{array}{c}
NP_{\text{valp}} \\
T^0_{\text{up}} \\
\end{array} \right] \left[ T^0_{\text{up}} \ldots \left[ T^0_{\text{up}} \ldots \text{minh}_{\text{up}} \ldots \right] \right]\right] \rightarrow \\
(5) \quad [ & C^0_{\text{valp}} \left[ \begin{array}{c}
NP_{\text{valp}} \\
T^0_{\text{valp}} \\
\end{array} \right] \left[ T^0_{\text{valp}} \ldots \left[ T^0_{\text{valp}} \ldots \text{minh}_{\text{valp}} \ldots \right] \right]\right]
\end{align*}
\]

To account for the other options in (1), I assume that a C\(^0\) domain can be introduced for any verb of saying or thinking. If so, binding by the subject of that verb obtains through the same mechanism. To account for the facts in (2-3), I assume that the presence of the first person pronoun tôi in the intermediate clause and the matrix clause respectively, obligatorily triggers the selection of the performative frame. Technically, then, it is the silent first person pronoun that is the binder, not the lower blocking first person pronoun.

References
Giblin, I (2016), Agreement Restrictions in Mandarin Long-distance Binding, Doctoral Dissertation, MIT.
Huang, Yun-Hua (1984), Reflexives in Chinese, Studies in English Literature and Linguistics.
Hiraiwa, K (2005), Dimensions of symmetry in syntax: Agreement and clausal architecture, Doctoral Dissertation, MIT.

(In)definiteness in quality and quantity superlatives in Romance and beyond

Carmen Dobrovie Sorin

<not updated>
Optatives in the Caucasus. Evidence for a linguistic or discourse area?

Nina Dobrushina
(HSE University, Moscow)

Keywords: optative, Caucasus, East Caucasian family, language contact, discourse area

Inflectional optative is the morphological form dedicated to the expression the wish of the speaker, as Mehweb (<East-Caucasian) form in –b:

qu b-alq-a-b
field N-grow.IPFV-IRR-OPT
‘May the field grow!’

As shown in Dobrushina, van der Auwera & Gussev (2006), inflectional optatives are a typologically uncommon phenomenon. They are found in 48 languages out of 319 languages of the sample, and are distributed unevenly, with two areas rich in optatives, the Caucasus and northern India and Nepal.

Caucasian optatives are present in languages that belong to different families (East Caucasian, West Caucasian and South Caucasian). They are also found in Turkic languages spoken in the Caucasus (Kumyk and Nogai), while many other Turkic languages lack them (Dobrushina 2011).

A closer look at the East Caucasian family shows that optatives in related languages may have different origins. The optative in Bagvalal is marked with the suffix –la which is attached to the imperative. In the negative polarity (negative optative), it is attached to the prohibitive, a form which is morphologically unrelated to other negative forms (Kibrik 2001). In Tabassaran, on the other hand, the optative has suffix -ri, is not related to the imperative, and is negated in a regular way (Babalieva 2013). This seems to exclude the possibility of a common origin.

Therefore, inflectional optatives in the Caucasus show every hallmark of a contact origin: they are present in every language of the area (while being typologically infrequent), they occur in genetically unrelated languages, and they often have a different origin in genetically related languages.

I suppose that such an even spread of the optatives may be partly due to the fact that optatives are not a purely linguistic phenomenon. Blessings and curses are an ethnocultural discourse feature in the Caucasus (cf. the notion of discourse area, introduced by Beier et al. (2002) for the Amazon). Blessings are a common type of ritualistic speech behaviour, frequent in everyday speech practices of the village life. The spread of the optative can be due to the spread of speech rituals resulting in an emergence of a new grammatical category.

The input of language contact can be observed in Dagestan where optatives share many structural features, such as persons covered by the optatives, or coexpression of several different functions in one form. The spread of the category itself is also likely to be attributed to cultural contact mediated by languages. Linguistic phenomena resulting from language contact are rarely distributed so evenly.

I intend to test the hypothesis that linguistic phenomena linked to culture may spread more evenly than purely linguistic phenomena by closely considering contact phenomena that seem to be ‘highly contagious’, covering areas with dozens of languages. I will compare them with the Caucasian optatives according to two parameters: whether they occur in all languages, or only in a part of them, and whether they can have cultural motivation.

References


Reassessing the Indo-European data: An integrated approach

Bridget Drinka
(University of Texas at San Antonio)

Keywords: Indo-European, DNA, genome, Steppe Hypothesis, morphology

Paul Heggarty (2018) examines the most recent evidence from DNA analysis concerning the origin and spread of the Indo-European (IE) languages, and claims that neither the arguments for the “steppe” hypothesis (Gimbutas 1970, Mallory 1989) nor the “Anatolian” hypothesis (Renfrew 1987) are fully supported. While recent studies (Haak et al. 2015; Allentoft et al. 2015) suggest that massive population movement from the steppe into northern Europe did occur c. 2,500 BCE, Heggarty assumes the date was too late to be identifiable as the earliest movement of the Indo-Europeans into Europe. He goes on to point out that the Yamnaya steppe ancestry detected in many modern and ancient northern Europeans is found in the Uralic-speaking population as well as in the IE-speaking population, implying that the genetic evidence does not correspond well with the linguistic evidence. In addition, DNA evidence suggests that southern IE varieties like Greek and Hittite were less affected by steppe incursions than northern IE varieties. Heggarty proposes that the DNA evidence calls for a sort of hybrid version of the two hypotheses, with agriculture developing in the Fertile Crescent and specializing as pastoralism when it spread to steppe lands north of the Caucasus.

While endorsing the syncretizing and stratified nature of this proposal, this paper presents an alternative interpretation of the data: a later date for the arrival of the Indo-Europeans in Europe is congruent with the concept of a dynamic “proto-conglomerate” from which the Indo-European varieties would have separated at different points in space and time. The presence of Yamnaya ancestry in the Uralic-speaking population may reflect a later influx, c. 500 BCE (Mittnik et al. 2018). With regard to both Greek and Indo-Iranian, evidence for small but perhaps significant late incursions from the north exist in both groups (Lazaridis et al 2017; Damgaard et al. 2018; Narasimhan et al. 2018). Besides such arguments based on genomic and archeological data, the paper presents extensive lexical and morphological evidence that corroborates the steppe hypothesis, but that points to a need for a more dynamic model for PIE.

References
Complex Natural Systems for Language

Katarzyna Dziubalska-Kołaczyk & William A. Kretzschmar, Jr.
(Adam Mickiewicz University, University of Georgia)

Keywords: natural linguistics, complex systems, linguistic theory, phonology, phonetics

In this paper we propose a theory of modeling linguistic phenomena deriving from complex systems and natural linguistics. It is a transdisciplinary eclectic approach in which dynamic systems theory is applied to enhance and support the epistemology of natural linguistics to arrive at complex natural systems for language.

We will focus on major concepts and assumptions of complex systems and natural linguistics and examine their compatibility. In complex systems structures emerge without any specific cause (Ellis and Larsen-Freeman 2009, Kretzschmar 2009), while in natural linguistics they result from functional behavior and teleology of change (Dressler and Dziubalska-Kołaczyk, 1994). In complex systems frequency of usage drives a structure towards stability (Mitchell 2009, Kretzschmar 2015, Burkette 2016) while in natural linguistics a degree of preferability is a major force behind survival of structures. Universals are understood as post-factum generalizations in complex systems (where the only true universal is human interaction) while they derive from general semiotic and functional principles in natural linguistics. Typological, sociological and psychological factors are interpretable and matter in both theories.
Despite the apparent divergences, we will demonstrate the applicability of the complex systems apparatus to the natural linguistics framework without depriving the latter of its basic ideology (cf. Dressler 2011). The key to the convergence of these methods is to understand that the semiotic and functional ideas of natural linguistics are perceptual (not essential) phenomena that arise from the underlying frequency profiles that emerge from the complex system of human speech. For instance, in the emergent environment of complex systems the traditional idea of a “vowel system” cannot apply as essential phenomena; what people actually say does not come down to simple targets. The actual picture is much more complicated, with substantial overlaps between vowels in the “system”, and no central points that could be thought of as a “target.” Thus, the notion of language learning corresponding to nonlinear patterns of how each vowel may be realized cannot be said to be “correct,” but rather more “in an understandable range of what speakers do.” This is how perceptual ideas from natural linguistics can be associated with complex systems; what people perceive about the speech around them is subject to preferences, semiotics, and function. The principles of natural linguistics are thus not generative of structure, but are nonetheless useful characterizations for maintenance of structure. To illustrate the workings of complex natural systems for language we will exercise modeling language acquisition and phonotactics, alongside large-scale phonetics. We can show that an understanding of the distributional properties of a complex systems analysis of speech can inform natural linguistics.

References

Defining and operationalizing “borrowability” in phonology

Elad Eisen, Dmitry Nikolaev, Eitan Grossman & Steven Moran
(The Hebrew University of Jerusalem, Stockholm University, The Hebrew University of Jerusalem & University of Zurich)

Keywords: typology, language contact, language change, phonology, linguistic database

Borrowability is a central notion in language contact research. It reflects the idea that some linguistic items or categories (e.g., sounds, word classes) are more easily borrowed than others. Generalizations about borrowability have typically been formulated as borrowing hierarchies or scales (Haugen 1950, Weinreich 1953, Thomason & Kaufman 1988, Muysken 1999, and Matras 2009), e.g., NOUNS >
VERBS, suggesting that nouns are more borrowable than verbs. However, such hierarchies have multiple interpretations (Haspelmath 2009).

Crucially, the notion of borrowability has never been properly defined. A typical assumption is that categories are “more likely” to be borrowed if they are borrowed more frequently in cross-linguistic comparison (Matras 2009: 154). In phonology, for example, it is observed that consonants are borrowed more frequently than vowels, but it has been suggested that this is a result of languages typically having more consonants than vowels (Matras 2007: 37).

Here, we show that the empirical frequency of borrowing depends, to a large extent, on the probability of a contact situation that permits borrowing to occur. Following Eisen (2019), we propose a definition of borrowability that takes this bias into account: borrowability is the probability of a given item/category to be borrowed, given a contact situation that permits borrowing.

We illustrate this by examining phonological segment borrowing. When a language is said to borrow a phonological segment from another language, it is assumed to have lacked this segment beforehand, while the donor language is assumed to have had it. That is, a contact situation that permits the borrowing of a phonological segment is one in which a language that has a certain segment is in contact with a language lacking it.

In mathematical terms, if a segment’s typological frequency is $f_S$, the probability of a language to have it is $f_S$, and the probability of a language to lack it is $1 - f_S$. The probability of a contact situation that permits borrowing is then the product of both probabilities, $f_S \cdot (1 - f_S)$. Marking the borrowability of a segment $S$ as $b_S$, the probability of a segment to be borrowed is expressed as:

$$P = b_S \cdot f_S - b_S \cdot f_S^2$$

This results in a series of quadratic functions of $f_S$, illustrated in Figure 1. From their form we predict that the most frequently borrowed segments are segments that occur in around 50% of the world’s languages, since they are the most probable to occur in an appropriate contact situation to begin with. Data from a new typological database of borrowed segments, called SegBo (Grossman et al. 2020), confirm this prediction.

We use the SegBo data to assess borrowability scores of various segments, showing that the most borrowable segments are /ʕ/ and /ʃ/. We calculate average borrowability scores of several phonological classes, showing that consonants are more borrowable than vowels, even when taking into account the differences in their typological distributions. Results also show that fricatives are more borrowable than affricates, which are more borrowable than stops.

This study provides the first explicit operationalization of borrowability in segmental phonology as distinct from the empirical frequency of borrowing. This allows us to quantitatively assess the borrowability of phonological segments and classes and to make predictions about what will likely be borrowed in language contact situations.
Figure 1: Predicted probability of borrowing vs. typological frequency

References
Haugen, Einar (1950), The Analysis of Linguistic Borrowing, Language 26(2), 210-231.
Matras, Yaron (2009), Language Contact, Cambridge University Press.
Weinreich, Uriel (1979 [1953]), Languages in Contact: Findings and Problems, Mouton.

Distinguishing between pronouns and expletives in Icelandic

Thorhallur Eythorsson & Sigridur S Sigurdardottir

<not updated>

Motion event description in Romanian: Romance core and language contact

Benjamin Fagard & Alexandru Mardale
Lattice (CNRS, ENS, Paris Sorbonne Nouvelle ; PSL & USPC) / Inalco (SeDyL, UMR 8202)

Keywords: motion events, space, typology, language contact, Romanian

Our aim is to establish the place of Romanian in Talmy’s (1991) typology, which distinguishes two types of languages, Satellite-framed (SF) and Verb-framed (VF). In SF languages, e.g. Germanic and Slavic, the Path of motion is typically expressed by a verb prefix or verb particle (a Satellite), leaving the verb root free to express other semantic features, for instance Manner of motion, as in English he ran out. In VF languages, e.g. Romance, the Path of motion is typically expressed by the verb root:
since that locus is taken, Manner of motion is therefore typically omitted, or expressed elsewhere, as in French *il est sorti (en courant)* ‘he exited (at a run)’. A known consequence of this systemic difference is that of a rich paradigm of Manner verbs in SF languages, compared with VF languages (Slobin 2004). This typology has been widely discussed, subsequent studies showing that, despite serious limitations, it accurately predicts motion-event descriptions in many languages (Zlatev and Yangklang 2004, Beavers et al. 2010, Croft et al. 2010). Romanian has been described as a potentially VF language (Dragomirescu 2013, Dragomirescu & Geană 2011): our aim here is to see, on the basis of elicited data, to what extent Romanian aligns with VF languages, esp. other Romance languages: what types of motion verbs are used in Romanian, and do they stem from Latin or were they borrowed as a result of language contact?

To do this, we gathered data with an elicitation tool in French (17 speakers), Italian (10) and Romanian (11). It is a series of video-clips (the *Trajectoire* videos, Ishibashi et al. 2006), containing 55 target videos with scenes of human motion events – various characters entering, passing and exiting; coming and going; walking, jumping and running. The resulting dataset totals 2,867 descriptions and 33,429 words.

Our results confirm that all three are typical VF languages, expressing Path and avoiding Manner except in specific contexts, e.g. marked Manner (Slobin et al. 2014) and in the absence of boundary-crossing (Aske 1989). Romanian appears to be the most markedly VF of the three, with a strong avoidance of Manner in scenes with boundary-crossing (1) and, conversely, frequent Manner expression in scenes with no boundary-crossing (2).

(1) copilul   în apă
child.DET.M.SG in water
“the child enters the water”

(2) un  băiat  pe nisip
a boy run.IND.PRS.3SG on sand
“a boy runs on the sand”

In terms of lexical renewal, the motion verbs used in our dataset are overwhelmingly of Latin origin. There is some renewal, however, especially for Manner verbs, which are much more often borrowed from surrounding SF languages, especially in French (with borrowings from Germanic in the early Middle-Ages, e.g. *marcher* “to walk”) and more strongly yet in Romanian, which borrowed from the Slavic superstrate in the late Middle-Ages (e.g. *a ocoli* “to bypass”). An interesting feature of this renewal is that it does not seem to have changed the type of Romanian, which remains structurally very close to other Romance languages, at least concerning Motion event descriptions (cf. also Reinheimer 1965).

References


Agreement and argument realization in Mian discourse

Sebastian Fedden
(Université Sorbonne Nouvelle and University of Surrey)

Keywords: agreement, Complementarity Principle, Mian, overt/null arguments, referential density

Mian is a Mountain Ok language spoken in Papua New Guinea. Its curious features include the fact that agreement with the object is restricted to a small subset of transitive verbs. Such “sporadic” agreement (Corbett 2006) is a rare and non-canonical type of agreement. Transitive verbs that agree with their object fall into two lexical classes. Both these classes agree in person and number with the object, and – depending on class – according to one of two different nominal classification systems. They are identified as systems 1 and 2 below. These two systems are based on different semantic distinctions and use different means of formal marking (Corbett, Fedden & Finkel 2017), see (1) and (2). However the majority of transitive verbs do not index their object (3).

(1) ̧ ō máam=e a-nā’-n-o=be
3SG.F mosquito=ART 3SG.M.OBJ-hit-REALIS-3SG.F.SBJ=DECL
‘She hit the mosquito.’

(2) nē fūt=e tob-ō-n-i=a
1SG tobacco=ART 3SG.LONG.OBJ-take-SS.SEQ-1SG.SBJ=and
‘I take the tobacco leaf and (then I) …’

(3) ̧ ō máam=e bou-n-o=be
3SG.F mosquito=ART swat-REALIS-3SG.F.SBJ=DECL
‘She swatted the mosquito.’
We should ask how such a system works in discourse; in particular, how the presence or absence of agreement relates to the overt vs. null realization of arguments. A hypothesis that has been put forward concerning this relation is the *Complementarity Principle* (CP) (see Kibrik 2011, Haig & Schnell 2016). This is a principle of economy, and it claims that null arguments are favoured by overt agreements and vice versa.

To verify this hypothesis, I conducted a discourse study of Mian, similar to the study by Nichols (2018) on the Nakh-Daghestanian language Ingush. I measured the referential density (Bickel 2003), i.e. the ratio of actually overt arguments to grammatically permitted arguments. It is very low in Mian: only about one quarter of grammatically possible arguments is expressed. We do find expected discourse effects, such as subjects tending to be null and objects among all argument types being most likely realized overtly; this is helpful, given that the study looks at objects. The table below summarizes the comparison of null vs. overt objects to agreement vs. non-agreement. I include the expected percentages under full complementarity (i.e. 0% of objects overt with agreeing verbs, 100% of objects overt with non-agreeing verbs) to help contextualize the actual figures.

<table>
<thead>
<tr>
<th></th>
<th>Verb agrees with object</th>
<th>Verb does not agree with object</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>system 1</td>
<td>system 2</td>
</tr>
<tr>
<td>null object</td>
<td>50</td>
<td>134</td>
</tr>
<tr>
<td>overt object</td>
<td>21</td>
<td>97</td>
</tr>
<tr>
<td>% overt object</td>
<td>0.30</td>
<td>0.42</td>
</tr>
<tr>
<td>% prediction from CP</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The results do not support the CP. Despite the difference between the two systems the overall percentages of overtly realized objects are very similar when looking at agreeing verbs (systems 1 and 2 combined) and non-agreeing verbs; under complementarity we would expect significant differences, specifically more overt objects with non-agreeing verbs. Furthermore, the majority of transitive verbs do not agree with their object. According to the CP, referential density should be much higher.

**References**


**Caritive as a negative marker**

Maksim L. Fedotov, Sofia A. Oskolskaya & Natalia M. Zaika (Institute for Linguistic Studies, RAS)

**Keywords**: caritive, negation, privative, typology, morphosyntax.
The notion of caritive (= abessive) usually refers to grammatical markers such as the one in *John came home without Mary*. Scholars have given different definitions of caritive. It has often been described as expressing the absence of a participant (Stolz et al. 2007: 66), as a negative comitative (Haspelmath 2009: 514), or as negative possession (Ivanov 1995: 5). In the talk, we will attempt at characterizing the semantic nature of caritive markers. In particular, we will discuss the connection of caritives with negation, comitative relations, and possession. Our observations are based on data from a sample of 50 languages from different parts of the world. The data comes from grammars, as well as from corpora and elicitation, where possible.

First, it may not be obvious that caritives are negative markers at all. Dedicated caritive markers are usually not connected morphologically with sentential-negation markers — unlike their close relatives, privative markers, cf. Russian *On ne prišël* ‘He didn’t come’, *nekrasivyj* ‘not beautiful’, but *bez knigi* ‘without a book’.

However, caritives can be shown to convey a negative semantic component. They introduce to the semantic structure of the clause a low-level negative operator (= narrow-scope negation, or “constituent negation” in the semantic sense of Jackendoff 1969). This is similar to the low-level negation introduced by privative (“antonymic”) derivates (Zimmer 1964), such as *unequal, unhappily*, etc. Caritives are paraphrased with sentences containing overt negation (cf. *I came without money* → *I came not having money on me*), and in some languages are always expressed/translated by periphrastic constructions with constituent or sentential negation (as in Gban, a South Mande language). Moreover, caritives allow negative polarity items, cf. *He came without any money*.

The next question we address is what types of relations are negated by caritives. Caritives often mark the absence or non-involvement of a companion, a possessee (cf. *John left without money*), or an instrument. Moreover, there is a number of other meanings (e.g. absence of an accompanying action, cf. *He bought a car without thinking*) which form a periphery around the core ones. Our sample at the moment include few languages with a dedicated caritive marker and sufficient data (such as full grammatical descriptions, corpora) on it. But, in all these languages (several Indo-European languages, Hill Mari, Nanai, Chuvash, Mongolian, Basque), caritive markers can express absence of a companion, a possessee, an instrument, and at least some of the peripheral meanings. Therefore, caritives in general seem to have broad meanings, not confined to one particular type of relation.

We can analyze caritive markers as expressing low-level negation together with an operator that takes a term (a non-predicate) and converts it into a (negated) predication, by inserting this term into a (negated) predication. Therefore, caritives are paraphrased with sentences containing overt negation (cf. *He opened the can without a knife* ≈ ‘He opened the can, not using a knife for it’) or an NP (*beardless men* ≈ ‘men that do not have beards’).

Acknowledgments: The study is supported by the Russian Science Foundation, grant 18-78-10058, project “Grammatical periphery in the languages of the world: a typological study of caritives”. We would also like to thank all other participants of the project.

References

Ivanov, Vjacheslav V. (1995), Tipologija lishitel’nosti (karitivnosti) [Typology of absence (caritivity)], in V. V. Ivanov, T. N. Moloshnaja (eds.), *Etudy po tipologii grammaticheskix kategorij v slavjanskix i balkanskix jazykax* [Sketches on the typology of grammatical categories in Slavic and Balkan languages], Moscow: Indrik. 5–59.

What do or constructions ‘say’?

Alon Fishman, Daniel Asherov, Nicole Katzir, Inbal Arnon & Mira Ariel

Linguists commonly assume an ‘inclusive disjunction’ core meaning for or constructions (e.g. Aaron came with Betty or Cathy), deriving ‘exclusive disjunction’ readings via scalar implicatures (Horn 1972, Carston 1990, Chierchia 2004, inter alia). Two types of experimental tasks have been used to identify participants’ interpretations of scalar expressions. In an INTERPRETATION experiment, participants match a state of affairs to a given target proposition (e.g. Author5, Degen and Goodman 2014). In the more common, VERIFICATION task, participants judge the sentence as T/F given some state of affairs (e.g. Paris 1973, Noveck 2001). Current theories don’t predict a difference between these experimental tasks, since the speaker-intended ‘what is said’ (‘w.i.s.’) fully specifies the truth conditions of the proposition. We propose that truth judgments don’t necessarily reflect ‘w.i.s.’, since these may also factor in Truth-Compatible Inferences (TCIs) (Author5). TCIs, if mobilized, bridge the potential gap between ‘w.i.s.’ (e.g. ‘One of Betty and Cathy’) and actual states of affairs (e.g. ‘Both Betty and Cathy’). We therefore predict a potential difference between INTERPRETATION and VERIFICATION tasks.

We adopt Author3 and Mauri’s (2018, 2019) Alternativist theory for or, whereby or only imposes an ‘alternativity’ relation between the disjuncts. When the context-relevant ‘alternativity’ points to a ‘One of Betty and Cathy’ ‘w.i.s’, participants performing INTERPRETATION should uniformly opt for a state of affairs in which only one of the women arrived. At the same time, in performing VERIFICATION, the same participants may judge the sentence as true given a state of affairs in which both women arrived. The variability in truth judgments depends on whether the participant takes ‘Both Betty and Cathy’ as compatible with ‘One of Betty and Cathy’ in the specific discourse.

We present an experiment testing how speakers map Hebrew or sentences to states of affairs (INTERPRETATION) and how they judge their truth (VERIFICATION). Participants (N=64) read short stories, where the difference between ‘One of’ and ‘Both’ was irrelevant (e.g. ‘Aaron will get a discount if he came with at least one woman’). Each story was followed by a dialogue including the target sentence. We manipulated the type of task (INTERPRETATION/VERIFICATION) and the type of target sentence (with/without or) within-subjects in a 2X2 design. The resulting four conditions were paired with four stories randomly for each participant. Two filler stories were added. Each story was followed by five multiple-choice questions about it, one of which was the task.

Results show that participants predominantly chose ‘One of’ (83%) in the INTERPRETATION task. In VERIFICATION, however, only 55% of the participants judged ‘Both’ as false, which would correspond to a ‘One of’ interpretation (see Fig. 1).
We take the variability of responses in VERIFICATION, as well as the discrepancy between INTERPRETATION and VERIFICATION, as support for the reality of TCIs, which can only play a role in T/F judgments. The overwhelming convergence between the participants in INTERPRETATION is evidence that INTERPRETATION, rather than VERIFICATION should be used for probing or’s ‘w.i.s.’.

References

A Noun in Verb’s Clothing? When a verb won’t take finite forms

Christian Forche
(University of Leipzig)

Keywords: finiteness, verb second, complex verbs, verbal pseudo-compounds, NonV2 verbs

This contribution provides further evidence for the important role that finiteness, here understood as the use of a verb in a synthetically conjugated form, plays in syntax – even where this is not expected. In German, certain complex verbs cannot appear in V2 position (e.g. German *bausparen* ‘to save with a building society’, *ersterscheinen* ‘to be first published’, *schutzimpfen* ‘to inoculate’, *uraufführen* ‘to première’). Sentences like (1a) or (1b) are generally dismissed as ungrammatical. However, in verb-final position (VL), i.e. in subordinate clauses (1c), the occurrence of these verbs is unproblematic. The same applies to infinite components of periphrastic verb forms (1d).

(1) a. *Der Regisseur ur-auf-führt das Stück.*
   the director fore-on-put-PRS.3SGV2 the play
b. *f* Der Regisseur *führt das Stück ur-auf.*
   the director put-PRS.3SGV2 the play fore-on
   ‘The director premières the play.’
c. ... *ob der Regisseur das Stück ur-auf-führt-t.*
   whether the director the play fore-on-put-PRS.3SGVL
   ‘... whether the director premières the play.’
d. *Der Regisseur will das Stück ur-auf-führen.*
"The director wants to première the play."

Previous accounts for this puzzling positional restriction usually resort to violations of well-formedness conditions or structural uncertainties that lead to a syntactic conflict (cf. Haider 1993; 2010; Fortmann 2007; 2015; Vikner 2005; Booij 2010). There was never any doubt cast though as to whether these complex verbs are at all verbs and whether they could form verbal finite forms.

This empirically informed paper challenges this view. German data from the Corpora from the Web (COW, Schäfer & Bildhauer 2012) show that the relevant complex verbs rarely occur finitely at all (neither in V2 nor in VL position) but rather in passive (2) or nominal constructions (3).

(2) Am 30. September 1935 wird “Porgy and Bess" am Colonial Theatre in Boston uraufgeführt. ‘On 30 September 1935 “Porgy and Bess” is (being) premiered at the Colonial Theatre in Boston.’

(3) a. Mit der 1797 in Bremen uraufgeführten Oper liegt […] die erste umfangreiche Vertonung des Faust-Stoffes vor. ‘The opera, premièred in Bremen in 1797, is the first comprehensive setting of the Faust material.’

b. So ist für jeden ein passendes Angebot für das Bausparen aufzuspüren. ‘In this way a suitable offer for building society savings can be found for everyone.’

I argue that examples such as (3) can be analyzed as nominal or adjective compounds whose head is a nominalized infinitive or adjectivized participle. What factors, then, do have an influence on (complex) nouns being (re)converted to verbs, or, as I propose, nouns in verb’s clothing?

A series of logistic regression analyses that use predictors usually discussed in the literature as V2-inhibiting factors reveals that all significant position-restricting factors actually more generally inhibit the formation of (or the transition from infinite to) finite verb forms. Hence, contrary to what has been advocated, the positional restriction with complex verbs is not a syntactic problem but one that can be traced back to finiteness.

References
Haider, Hubert (2010), The Syntax of German (Cambridge Syntax Guides), Cambridge: Cambridge University Press.
How classifiers become gender in Oceania: An experimental approach

Michael Franjieh, Greville G. Corbett & Alexandra Grandison
(University of Surrey)

Keywords: categorisation, gender, classifiers, Oceanic, psycholinguistics

The possessive classifier systems of some Oceanic languages show signs of shifting towards fixed gender systems. Typically, a noun can occur with different classifiers, depending on how the possessed item is used by the possessor (Lichtenberk, 1983). For example, nompui ‘pig’ in Lewo (Vanuatu) occurs with either the edible or the general classifier.

1a. ka-na nompui
   CL.ED-3SG pig
   ‘her pig (to eat)’

b. ke-na nompui
   CL.GEN-3SG pig
   ‘her pig (being raised)’ (fieldnotes)

In marked contrast, North Ambrym’s (Vanuatu) cognate for pig – bu – occurs only with the edible classifier (2a), not the general classifier (2b):

2a. a-n bu
   CL.ED-3SG pig
   ‘her pig (for any purpose)’

b. *mwena-n bu
   CL.GEN-3SG pig
   intended: ‘her pig’ (fieldnotes)

North Ambrym’s innovative system resembles a gender system: a noun occurs with a particular classifier regardless of context.

We designed a suite of experiments to compare possessive classifier systems in six representative Oceanic languages: Merei, Lewo, Vatlongos, North Ambrym (Vanuatu), Nêlêmwa and Iaai (New Caledonia). Each has a different inventory of classifiers, from a two-way distinction (Merei) up to 23 (Iaai). The classifier systems in these languages also represent varying degrees of informativeness: from transparent semantic motivation to opaque assignment. Effective categorisation needs to be informative to maximise communicative efficiency, but also needs to be simple, to minimise cognitive load (Hawkins, 2004). Our sample languages allow us to investigate the trade-off between these two principles of informativeness and simplicity. Our experiments were designed to uncover how and why languages would relinquish a useful, meaningful classificatory system, and adopt a rigid, apparently unmotivated gender system.

We discuss results from two recently run experiments, involving data from 122 speakers across our sample languages:

(i) free listing, where participants are asked to list the members of each classifier. This establishes central members of a classifier’s semantic domains.
(ii) video vignettes, which depict typical and atypical interactions with different items. These determine whether speakers are free to use different classifiers (as generally believed) or whether assignment of nouns to classifiers is rigid.

Data from the free listing experiment were analysed using a number of salience measures including Smith’s Salience (Smith & Borgatti, 1997), which integrates noun frequency and mean list position to provide a meaningful measure of salience. Initial results indicate that maximum salience is greater for languages with smaller classifier inventories, and that mean salience differs significantly across classifiers within languages. Additionally, a noun is less likely to be a member of two or more different classifier categories in North Ambrym and Merei than in Vatlongos or Lewo. But, interestingly, the results of the video vignettes indicate that assignment of nouns to classifiers is far more rigid in North Ambrym than in Merei, Lewo and Vatlongos. These results point to North Ambrym’s system being the most gender-like, followed by Merei, with Lewo and Vatlongos functioning more like typical classifier systems. Together our results suggest that the classifier systems tested do indeed represent different stages of grammaticalization, from classifier to gender marker.

References


A Cartographic Analysis of “Clausal” Nominal Expressions in Colloquial Japanese

Minoru Fukuda, Takeshi Furukawa & Koichiro Nakamura

<not updated>

The challenge of complexity: Morphological change and language contact in Walser German

Livio Gaeta
(University of Turin)

Keywords: complexity, language change, morphology, language contact, linguistic island

Recently, a big debate has started focusing on the role of complexity (and of its conceptual counterpart simplicity) in language and specifically in morphology (cf. Miestamo et al. 2008, Sampson et al. 2009, Baerman et al. 2015, Baechler et al. 2016). In particular, the hypothesis has been discussed according to which language contact is taken to favor processes of simplification, and accordingly to militate against morphological complexity: “Contact-induced grammatical change is likely to produce outcomes simpler (in some sense) than the original ones, affecting thus the overall typology of a language” (Karlsson et al. 2008: viii; see critically Meakins et al. 2019).
In this paper, the relevance of the notion of complexity for morphological change will be discussed with the help of data coming from Walser German varieties spoken in linguistic islands which stand in close contact with the surrounding Romance varieties and in the last decades have undergone a dramatic process of language decay (cf. Dal Negro 2004). Besides clear phenomena of simplification like case-reduction (cf. Baechler 2016), a number of changes are observed representing interesting innovations whose status in terms of increase or decrease of complexity is not easy to define. In particular, in Guryner Titsch in which verb classes normally display syncretism of the 1st ps. sg./pl. (e.g. fri:ba ‘I/we write’, fri:ht / fri:bat ‘thou / you write’, etc.) while modals diverge insofar as they distinguish 1st ps.sg. and pl. by means of a suffix -u: myas / myassu ‘I / we must’, etc., the so-called short-formed verbs which go back to the model of fri:ba but largely consist of auxiliary and other semi-grammaticalized verbs developed a number differentiation in the 1st ps.pl. similar to modals recruiting a suffix -v: ln: / la:v ‘I / we let’, etc. In Greschönnytitsch the traditional opposition between strong and weak verbs has been remodeled according to the constructional context. The strong verbs, traditionally displaying a nasal suffix in the past participle, retain the nasal suffix when the past participle is found in a construction displaying agreement – as in the passive construction: d’grösso lougò èsch gwäschn-e kanget ‘the big laundry[F.SG] has been washed-F.SG’ – while the dental suffix typical of the weak verbs is found in a construction in which no agreement is found – as in the (active) perfect construction: heintsch ... d’husgspòntto wollschrangna gwäschet ‘they have washed the home-spun wool strand’.

What these (and other) cases have in common is that the morphological system as a whole is improved in terms of its overall regularity insofar as overt coding and new constraints are introduced. In this sense, the system has become more manageable – and therefore simple – for the speakers because the inflectional behavior is predictable on the basis of contextual properties. On the other hand, it cannot be denied that these changes have rendered the system more complex in accordance with Nichols’ (2009: 112) suggestion that also “information required for describing the system should enter into the definition of complexity, and therefore constraints should be regarded as increasing complexity”.

References
Meakins, Felicity, Xia Hua, Cassandra Algy and Lindell Bromham (2019), Birth of a contact language did not favor simplification, Language 95(2), 294–332.

Affinity analysis reveals Central European areal traits in verb argument marking

Jerzy Gaszewski
(University of Wroclaw)

Keywords: valency, cases, adpositions, Central European languages, areal linguistics

Parallels in argument marking are often claimed to link Central European languages (Newerkla 2000, Pusztay 1996, Bláha 2015), though not without dissenting voices (Pilarský 2001). Most of the authors list few examples representing similar argument marking. This proves little as there are hundreds of verbs to choose (or cherrypick) from. Moreover, patterns observed among languages of any region can be its areal trait only if they differentiate it from adjacent areas (Haspelmath 2001, Campbell 2017).

The paper shows that certain patterns of argument marking are an areal feature of Central Europe. Applying affinity analysis to an extensive valency database (about 150 verbal meanings selected independently of language comparison) I identify groups of cases and adpositions that are equivalent across a number of argument positions of individual verbal meanings across the languages of the region. The tables below illustrate one such set.

The comparison involves valency markers in “obvious” Central European languages (Czech, German, Hungarian), marginal members of the area/close neighbours (e.g. Croatian, Polish, Romanian) and languages clearly outside the region (e.g. Spanish).

Apriori algorithm applied to the data returns frequent itemsets (i.e. sets of markers that are often enough equivalent in the data), many of which seem to reflect broader patterns (e.g. the equivalence of dative cases). However, in some groups of markers those of the Central European languages are found in a greater number of itemsets than those of other languages.

For example, West Slavic na + Accusative, Hungarian Sublative case and German auf + Accusative are strongly linked, cf. Table 1. The Croatian and Romanian markers fit less well, and the Spanish counterpart suggested by the algorithm is only loosely connected. Table 2. shows how many retrieved itemsets each of the languages is involved with. The gradation of numbers makes a strong case for Central European areal patterning.

Table 1. Example structures in seven languages

<table>
<thead>
<tr>
<th>Croatian</th>
<th>Czech</th>
<th>German</th>
<th>Hungarian</th>
<th>Polish</th>
<th>Romanian</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘to answer sth’</td>
<td>odgovoriti na+Acc.</td>
<td>odpovědět na+Acc.</td>
<td>antworten auf+Acc.</td>
<td>válaszol +Sublative na+Acc.</td>
<td>odpowiadać na+Acc.</td>
<td>rašpuniť na+Acc.</td>
</tr>
<tr>
<td>‘to look at sth’</td>
<td>*</td>
<td>čítat se na+Acc.</td>
<td>schauen auf+Acc.</td>
<td>néz +Sublative na+Acc.</td>
<td>patrzieć na+Acc.</td>
<td>a se uita la+Acc.</td>
</tr>
<tr>
<td>‘to wait for sth’</td>
<td>*</td>
<td>čekat na+Acc.</td>
<td>warten auf+Acc.</td>
<td>vár +Sublative na+Acc.</td>
<td>czekać na+Acc.</td>
<td>*</td>
</tr>
<tr>
<td>‘to watch’</td>
<td>*</td>
<td>dát pozor na+Acc.</td>
<td>acht en auf+Acc.</td>
<td>figyel na+Acc.</td>
<td>uważać na+Acc.</td>
<td>a avea grijă</td>
</tr>
</tbody>
</table>
Table 2. Rankings of the seven languages on the basis of itemsets they participate in (note: these are shorter itemsets, subset of the full group of equivalent markers)

<table>
<thead>
<tr>
<th>language</th>
<th>no. of itemsets length 4</th>
<th>no. of itemsets length 3</th>
<th>no. of itemsets length 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungarian</td>
<td>13</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td>Polish</td>
<td>13</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>Czech</td>
<td>9</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>German</td>
<td>6</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Romanian</td>
<td>5</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Croatian</td>
<td>3</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Spanish</td>
<td>1</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

References

Newerkla, Stefan Michael (2000), Language Affinity in Central Europe – Some Thoughts on the Interrelations of German, Czech, Slovak and Magyar, Opera slavica – Slavistické rozhledy X(4), 1-16.

Temporal modifiers and the Romance imperfective

Ion Tudor Giurgea

<not updated>

Voicing and devoicing in non-standardized insular language in the conditions of Slavic-Slavic language contact

Michał Głuszkowski
(Nicolaus Copernicus University, Toruń, Poland)
Voicing and devoicing assimilation in word-final and word-internal positions have been widely described in literature (see e.g. Wetzels & Mascaró 2001). Despite existing exceptions and acceptable variants of normative pronunciation, it is possible to characterize the patterns of the given phenomenon in the orthoepic norms of standard languages. Due to the lack of norms or their unstability, non-standard languages are more internally differentiated. Although phonetics of Slavic dialects has been already a subject of in-depth phonetic studies, also in the field of voicing and devoicing assimilation (Wojtkowiak & Schwartz 2018), the given phenomenon in Polish insular vernaculars in the East has not been analyzed so far.

Our attention is focused on sandhi and word-internal consonant clusters of the Lesser Polish dialect in Eastern Siberia: of the village of Vershina in Irkutsk Oblast’, a 110-years-old language and cultural island founded by voluntary settlers from Little Poland before the World War I. The Lesser Polish dialect belongs to the voicing zone of Polish regional varieties (Dejna 1993), in which the obstruents in the word-final are pronounced as voiced before the following types of sounds in the initial segment of the next word: vowels – powiś i już ‘you say it and’ [ʒi]; sonorants - szed napszód ‘he was moving on’ [d n]; voiced obstruents – dziadek był Polak ‘grandfather was Polish’ [g b]. In word-internal positions even non-voiced obstruents may undergo voicing: jestem ‘I am’ [zd]; wielki ‘great’ [lg].

The insular non-standard language in Vershina is in contact with standard Russian language, in which only an initial voiced obstruent in the next word can result in pronouncing voiced obstruents in the word-final (cf. Avanesov 1984), and voicing in word-internal positions occur only before voiced obstruents except of [v] (regressive assimilation). To a lesser extent Vershinian vernacular contacts with local Russian dialects (of the surrounding villages) with the same pattern of voicing/devoicing assimilation (Kasatkin 1999), and standard Polish language (acquired from vestigial teaching of minority language in the local school and from Polish tourists visiting Vershina every summer) with voicing assimilation in word-final and word-internal positions only before voiced obstruents (Gussmann 1992; Rubach 1996). Thus, language contact situation stimulates the processes of devoicing on word boundaries and inhibits word-internal voicing and its stability is endangered.

My aim is to analyze the stability of voicing and devoicing in word-final and word-internal positions and present their sociolinguistic background, since the language situation of the village is changing quickly and the generational differences strongly affect their bilingualism (Głuszkowski 2012). I am going to answer the question which of the following variables co-occur with the level of maintenance of dialectal sandhi: age, sex, level of education, ethnic homogeneity/heterogeneity of the family. The language and sociolinguistic material was gathered during subsequent field expeditions to Vershina in the years 2008-2019.

References
Distributional assessment of derivational semantics

Matías Guzmán Naranjo & Olivier Bonami
(Université de Paris, LLF, CNRS)

Introduction A main challenge for the study of word formation is the elusive nature of the semantic relations between derivationally related words. That situation is especially problematic for paradigmatic approaches to derivation, which typically rely on the identification of morphosemantic relations to organize paradigms (Štekauer, 2014).

A basic tenet of distributional semantics is that the distribution of a word is informative on its meaning (Lenci, 2008). In that context, the semantics of a morphological process can be seen as a function that maps input vectors to output vectors (Marelli and Baroni, 2015). Our research hypothesis is the following: if functions mapping input vectors to output vectors capture the semantics of derivational processes, then semantically similar processes should lead to similar functions. We test this on French data.

Data collection Our dataset combines information from various sources: Hathout and Namer (2014) for relations between verbs and nouns; Tribout (2010) for nouns and verbs related by conversion; Koehl (2012) for fordejectival nouns; Strnadová (2014) for derived adjectives; and Bonami and Thuilier (2019) for derived verbs in -iser and -ifier.

We use distributional vectors to represent the meaning of the lexemes under study. For this study we make use of the vectors provided by Fares et al. (2017). The final dataset contained 22,356 pairs of lexemes classified into 33 distinct processes. While that dataset does not give a full picture of the French derivation system, it is diverse enough that different degrees of similarity are instantiated.

Experimental setup For each derivation process we fit a model from the semantics of the base (B) to the semantics of the derived form (D). Our models assume that (i) each dimension of D is linearly related to the same dimension of B, and, in addition, (ii) there is a non-linear relation between each dimension in D and all the other dimensions in B. In other words, the model for a derivational process predicts each dimension of D from each corresponding dimension of B plus all the other dimensions of B. For the models we used a perceptron with a single hidden layer.

We computed the within-process performance of the model and compared it to the performance of the model when tested on other derivation processes. We measured model performance as the mean cosine distance between the actual meanings of the derived lexemes and the meaning predicted by the model from the meaning of the base. For the cross-model performance, we trained a model on all pairs of each derivation process and used that model to predict the derived meaning of bases found in other derivation processes. The idea is that if two derivation processes, P1 and P2, are semantically similar, then the model trained on P1 should perform well on base-derivation pairs of P2. After repeating the method just described for derivation processes, we end up with two estimations of the similarity between P1 and P2: one from the model trained on P1 predicting P2.
and one from the model trained on P2 and predicting P1. We define the semantic similarity between P1 and P2 as the mean of this two values. The result is a similarity matrix on which we perform hierarchical clustering.

**Results** The results shown in Figure 1 are remarkably close to expectations. The models end up grouping together broad classes of processes, such as deadjectival nouns, denominal adjectives, denominal verbs, and deverbal adjectives. The overall conclusion then is that vector-to-vector mappings do an excellent job of capturing similarities and differences between derivational processes.

![Figure 1: Hierarchical clustering](image-url)

**References**


From Matter to Pattern Borrowing

Camiel Hamans

(University of Amsterdam/Adam Mickiewicz University Poznań)

Keywords: non-morphemic word formation, morphological borrowing; embellished clipping, libfixing and blending

This paper wants to show that the distinction between morphological MATTER borrowing and PATTERN borrowing (cf. Sakel 2007, Gardani, Arkadiev & Amiridze 2015 and Gardani 2018) is not absolute. In addition, the results of the research presented here question the precise borrowing hierarchies (as proposed by a.o. Thomason and Kaufman 1988 and Thomason 2001). Three processes of morphological borrowing from English into Dutch will be discussed: embellished clipping (journo < journalist), libfixing (extracting segments from opaque wordforms like -topia from utopia) and blends (stagflation < stagnation + inflation). All three processes are instances of non-morphemic word formation.

The Dutch data that will be discussed come from a huge newspaper corpus ‘Delpher’ that includes more than 12 million newspapers from the period 1618-1995, from the digital ‘ANW’, a corpus of more than of more than 102,5 million tokens from 1970 till now, compiled by the Institute for the Dutch Language, INT and the Dutch ‘Etymologiebank’, in which all Dutch etymological and major general dictionaries are included or are found through targeted internet searches.

Embellished clippings (Bauer & Huddleston 2002) appeared in American English for the first time recently (Hamans 2004). The data found in the Dutch corpus show that first a few English words were borrowed into Dutch by speakers with a certain, albeit limited, bilingual Anglo-Dutch competence. However, these first loans appeared in Dutch newspapers some years after their first attestation in English. Subsequently, speakers of Dutch recognised a certain similarity between these loanwords within a short period, discovered a pattern based on this similarity and used this pattern in their own language. The first attestations of new non-English forms are found only a few years after the first loans. Finally, the pattern became completely integrated into the recipient language by applying it to native bases as well.

A similar process occurred in the cases of libfixing (Zwicky 2010) and blending, although there are found a few older instances of both processes in Dutch. However, the popularity of both non-morphemic processes of word formation is very recent in Dutch. In addition, Renner (2018), claims that blends rarely appeared in languages other than English until recently. However, nowadays one finds blends not only in Dutch but also in Bulgarian, Latvian, Polish, Serbian, Slovak and Ukrainian.

The method of borrowing turns out to be the same in these three different processes: it starts with MATTER borrowing followed by PATTERN recognition and borrowing and finally resulting in integration into the recipient language. This makes it clear that the distinction between MATTER and PATTERN borrowing is not justified in the case of non-morphemic word formation processes. The same must be said about Johanson’s distinction between global and selective copying. Consequently, the data under discussed demonstrates that these word formation processes can blur the hierarchical
distance between lexical and morphological borrowing. Instead of a clear cut distinction between Matter and Pattern or lexical and morphological borrowing, a borrowing cline will be proposed.

References

Cyclic changes to the French negative coordinating conjunction

Maj-Britt Mosegaard Hansen
(University of Manchester)

The evolution of the negative coordinating conjunction (‘(neither’/’nor’) from Latin to Modern French appears to have taken an almost perfectly circular path.

Thus, in Classical Latin, as is consonant with the typological status of that language as a Double Negation language, neque/nec was exclusively used in negative contexts, sometimes with an added scalar meaning (Gianollo 2017). Medieval French being a Negative Concord language, on the other hand, its negative coordinating conjunction, ne, a direct descendant of neque/nec, was able to develop a full range of weak negative polarity uses, but appears to have lost the scalar use (Antoine 1962, Foulet 1970, Martin & Wilmet 1980, Buridant 2000). In a range of contexts, ne was thus in competition with both the additive conjunction et (‘and’) and disjunctive ou (‘or’).
This development from a fully negative conjunction in Latin to one that has a stable range of weak negative polarity uses in Medieval French adds ne/ni to a growing list of grammatical items in French whose evolution can be shown to run counter to the strong unidirectionality hypothesis put forward by Haspelmath (1997), for whom the only possible direction of evolution within the domain of polarity is from positive/neutral polarity > weak negative polarity > strong negative polarity.

During the Renaissance, ne was gradually replaced by ny/ni (< ne+icelui ‘this one’ or ne+il ‘he/it’, cf. Badiou-Monferran fc a). It is unclear that there was initially any functional differentiation between the two forms ne vs ny/ni (pace Badiou-Monferran fc b). What we do observe, however, is that by the end of the Classical French period (17th-18th c.), the conjunction has lost all of its weak negative polarity uses, and is used only in strong negatively polar environments in Modern Standard French (Antoine 1962, Badiou-Monferran 2004, fc a/b). Interestingly, Modern French ni seems to have gained a (relatively marginal) scalar reading in the construction et ni (‘not even’, lit. ‘and nor’, cf. Antoine 1962).

Based on data from the electronic corpora Frantext and Base de Français Médiéval, I will analyze these changes with a view to uncovering what motivates them, and by what mechanism(s) they took place. I will argue that pragmatics played a crucial role, both in the change from Latin to Medieval French and from Medieval to Modern French.

I will argue, moreover, that the morphosyntactic status of Modern French ni is ambiguous: On the one hand, the conjunction functions like an n-word and thus combines with other n-words such as rien (‘nothing’) or personne (‘no-one’), but not with the postverbal standard clause negator pas, within the same clause. On the other hand, ni is unlike other Modern French n-words (with the exception of nul ‘no[ne]’) in that (a) it formally incorporates a negative morpheme, and (b) it is unable to appear in weak negative polarity contexts. I suggest that these characteristics may explain the low rate of contemporary usage of ni, as well as the fact that it appears to be increasingly confined to fairly formal registers.

References
Martin, Robert and Marc Wilmet (1980), Syntaxe du moyen français, Bordeaux: Bière.

Emerging Networks: The Traceback method and the early constructicon
Stefan Hartmann, Nikolas Koch & Antje Endesfelder Quick

<not updated>
A first version of the Grammaticon: A reference catalogue for cross-linguistic grammatical databases

Martin Haspelmath
(MPI-SHH Jena & Leipzig University)

Keywords: comparative grammar, language typology, grammatical terminology, databases

The study of Human Language requires systematic cross-linguistic comparison at a worldwide scale, which is rather difficult in the domain of morphosyntactic patterns (more difficult than in the domains of the lexicon or segment inventories). It is thus important to preserve and curate available worldwide datasets in a way that makes reuse by other researchers not only possible, but also straightforward.

In the domain of lexical comparison, the Concepticon (List et al. 2016) has made it possible to compare languages at an unprecedented scale, especially in the domain of colexification patterns (List et al. 2018; Rzymski et al. 2020), and some of this work has found its way into high-profile studies interfacing with other disciplines (Jackson et al. 2019). The Concepticon is a standard reference catalogue that allows linguists to link their lexical resources to a general concept list that can be extended flexibly when needed. Other reference catalogues that are required for cross-linguistic comparison are a standard list of languages and other languoids (e.g. Glottolog, Hammarström et al. 2019), as well as a standard list of comparable sound segments (e.g. PHOIBLE, Moran & McCloy 2019). The usefulness of such reference catalogues for has been highlighted by Forkel et al. (2018), who make a concrete proposal for a standard data format for all kinds of cross-linguistic datasets.

Likewise, in the domain of large-scale grammatical comparison, we need a standard reference catalogue for morphosyntactic features, to help us compare data from resources like WALS (Dryer & Haspelmath 2013), DiaCL (Carling 2017), eWAVE (Kortmann & Lunkenheimer 2013), SAILS Online (Muysken et al. 2016), APiCS Online (Michaelis et al. 2013), or Syntactic Structures of the World’s Languages (http://test.terraling.com/groups/7). Moreover, datasets from more specific projects (e.g. Dryer 2018; Handschuh 2019) should eventually be published online and linked to such a reference catalogue, rather than in the form of data tables as an appendix to a paper.

The present talk reports on the first version of the Grammaticon, which is intended to facilitate the comparison of different grammatical datasets on the analogy of the Concepticon for lexical datasets. The goal of allowing comparability, also for machines (i.e. interoperability), is similar to that of the GOLD ontology (General Ontology for Linguistic Description, Farrar & Langendoen 2003), and my talk will compare the Grammaticon with GOLD. The first version is primarily based on the WALS and APiCS features, but it will also take into account the Grambank features (https://glottobank.org/#grambank) as well as a number of features from other datasets. The publication of the Grammaticon is planned for late 2020.

References


Whole-form storage and frequency effects beyond the word: An analysis of preserved grammatical constructions in aphasia

Rachel Hatchard

<not updated>

Syntax and Semantics of Light Reflexives in Norwegian

Lars Hellan
(NTNU, Trondheim)
Keywords: Light reflexive, syntactic argument, syntactic function, transitivity, Norwegian, valence lexicon

Light reflexives (LRs) are mono-morphemic reflexive pronouns occurring in positions generally available to syntactic arguments, where semantically their role can be described as indicating an orientation of the verbal action towards the subject. A question of dispute is whether the LR in such cases can has the kind of function that an argument could otherwise hold in the position. For instance, for a generally transitive verb V, should V in the configuration “V_LR” be counted as transitive, with LR as its object, rather than, e.g., as intransitive, with LR as a kind of de-transitivizing marker (see, e.g., Frajzyngier (2019) for discussion). In this paper we argue that in Norwegian (Bokmål), LRs should syntactically be assigned full functional status (also cf. Hellan and Beermann (to appear)).

In a Norwegian valence lexicon with 12,900 entries, reflecting about 250 verb construction types, 1580 entries have an LR as a distinctive element in the frame (See https://typecraft.org/tc2wiki/Norwegian_Valency_Corpus). The table below indicates some of these construction types in terms of syntactic patterns, with seg as the LR, and with the most sparing amount of glossing. Under ‘Dictionary’ is given the number of entries in the valence dictionary with a verb with the frame in question, and under ‘Corpus’ are indicated number of instances of such verb-with-frame occurrences in a corpus of 22,000 sentences.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Example</th>
<th>Dictionary</th>
<th>Corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 NP_V_LR</td>
<td>hun vasker seg (she…wash)</td>
<td>667</td>
<td>319</td>
</tr>
<tr>
<td>2 NP_V_LR_P_NP</td>
<td>hun befatter seg med dem (deal…with…them)</td>
<td>333</td>
<td>242</td>
</tr>
<tr>
<td>3 NP_Vobjcontrol_LR_P_Inf</td>
<td>hun tvinger seg til å sitte (force…to…sit)</td>
<td>79</td>
<td>40</td>
</tr>
<tr>
<td>4 NP_V_LR_LOC</td>
<td>hun oppholder seg her (stay…here)</td>
<td>15</td>
<td>47</td>
</tr>
<tr>
<td>5 NP_V_LR_DIR</td>
<td>hun snygger seg hit (slither…here)</td>
<td>76</td>
<td>5</td>
</tr>
<tr>
<td>6 NP_V_LR_Prtcl</td>
<td>de dummer seg ut (they…make-fool…out)</td>
<td>139</td>
<td>14</td>
</tr>
<tr>
<td>7 NP_Vobjcontrol_LR_Inf</td>
<td>hun tillater seg å komme (allow-come)</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>8 NP_V_LR_P_Sdecl</td>
<td>hun undrer seg over at vi kommer (wonder…over…come)</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>9 NP_V_P_LR</td>
<td>hun rører på seg (move…on)</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>10 NP_V_LR_SCPRcaus-Ap</td>
<td>hun løper seg frisk (run…healthy)</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>11 NP_V_LR_SCPRstate-AP</td>
<td>hun befinner seg vel (find…healthy)</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>12 NP_V_LR_SCPRprop-PrtclP</td>
<td>hun oppfører seg som en kandidat (behave…like…candidate)</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>13 NP_V_LR_PPpossrais</td>
<td>hun gnir seg i nakken (rub…in…neck)</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>14 NP_Vraiss-to-obj_LR_Inf</td>
<td>hun viser seg å komme (show…come)</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

We leave open how these observations may replicate in other languages (the LR being instantiated pervasively throughout the Indo-European languages, for instance), but we discuss some formal consequences. For instance, one may, for a semantically loaded concept like ‘transitive’, distinguish between a notion covering ‘prototypically transitive’ acts, and a notion covering any syntactic configuration morpho-syntactically similar to the one typically exposing a ‘prototypically transitive’ act (cf. Creissels 2016). We thus argue that “V_LR” (frame number 1) be counted as a morpho-
syntactically transitive frame, whatever the semantic status of constructions instantiating the pattern. Likewise the other patterns listed should count as instances of patterns where $LR$ has the same syntactic function as a full NP (where lexically allowed – covering about half of the frames) would have in that position. An analytic task-complex is then identifying syntactic functions for these (and remaining) cases, and their possible semantic contents.

References


On the possible emergence of a new 1st plural person marker in rural varieties of Madeiran Portuguese

Yoselin Henriques
(University of Zurich)

Keywords: Portuguese, dialectology, morphosyntax, person markers

The expression of the first person in the plural (henceforth 1PL) in Portuguese has been repeatedly addressed in synchronous and diachronic studies of linguistic variation. The form $a$ gente coexists with the canonical 1PL pronoun of European Portuguese nós, often associated with southern dialects or Brazilian Portuguese varieties. Note that because $a$ gente originates in a noun phrase meaning ‘the people’, it normally shows 3rd person singular agreement.

(1) $a$.gent e faz-ia um-a vid-a, ali onde nós viv-famos
We make-IPFV.3SG ART.INDF-ART.SG.IP var-life(F)-SG.F there wher we live-IPFV.1P L

‘We built a life, there where we used to live’

Portuguese insular varieties of Madeira present a predominant use of $a$ gente. Interestingly, they also exhibit constructions in which this form co-occurs with the impersonal clitic se. However, these structures have received little attention in dialectological studies. A pioneering study is found in Martins (2003; 2009). Using spoken samples of European Portuguese varieties, the author provides a detailed description of the syntactic properties of what she calls “double subject impersonal constructions”, which seem to be more frequent in Madeira. According to this proposal, the stressed pronoun in these constructions restricts the referential range of the impersonal marker $se$, triggering an inclusive interpretation.
Nonetheless, as it is shown in the examples below, the combination of the two elements does not necessarily have to have an indefinite interpretation, but it can also have deictic reference (to participants of the speech event).

(2) Hav-ia um-∅ arraial-∅ ali; a.gente cham-a=se “Maio
have- ART.INDF- fair(M)- there we call- “Maio
IPFV.3SG SG.F SG ; 3SG=SE
‘There used to be a fair, we call it Maio’

The goal of this presentation is twofold. First, I will offer a description of these innovative constructions. Second, I will provide an interpretation of the function of these apparently hybrid constructions, proposing that they show the emergence of a new 1PL person marker (se). The study relies on spoken data from semi-directed interviews and free-speech conversations with elderly speakers of rural Madeiran Portuguese. I will use a quantitative variationist methodology to provide a description of the wide scope of the referential spectrum of these constructions, from indefinite to deictic reference. Focusing on the latter, I will propose that the impersonal clitic se in these contexts is reanalyzed as a 1PL person marker, given the ambiguity of verbal agreement patterns in third person singular, predominant with the innovative form a gente. To support this hypothesis, I also consider a number of syntactic factors such as co-referentiality and word order.

References

Overt and zero third person reference in Estonian: When and why

Helen Hint & Elsi Kaiser
(University of Tartu and Tallinn University; & University of Southern California)

Keywords: anaphora, personal pronouns, zero reference, experimental methods, Estonian
We examined the referential properties of two Estonian near-synonymous anaphoric 3rd person referential expressions: the overt 3SG pronoun *ta* ‘s/he’ and its phonetically null ‘zero’ counterpart. Both refer to highly-salient discourse entities (e.g. Gundel et al. 1993). However, overt *ta* is regarded as the default referential form in most contexts (Pajusalu 2009); occurrence of the zero is more constrained (Hint et al. 2020). Yet, the exact referential differences between these forms are not yet understood. We report experimental work that directly tests which factors affect the choice of Estonian overt or zero 3SG reference in discourse.

Our work has two theoretical vantage points. (1) We adhere to the form-specific multiple constraints approach to reference resolution, which claims that only one factor (e.g. salience) cannot fully explain all the different usage patterns of referential expressions (Kaiser & Trueswell 2008). (2) We follow the “no synonymy” principle from Construction Grammar (Goldberg 1995): even when two (slightly) distinct forms refer to the same entity in similar contexts, this does not indicate identical meaning of these forms.

To investigate which factors affect the choice between Estonian overt vs. zero 3SG reference, we designed two experiments. First, we conducted a speech-restoration experiment (Mack et al. 2012) where we manipulated the information status (given/new) and referential form (overt pronoun/fullNP) used in the preceding utterance for the referent. 30 participants heard pre-recorded narratives with a 3SG agent. In the final utterance of each narrative, the critical form was hidden under computer-generated noise. Participants repeated verbatim the last sentence of each narrative. This allowed us to see whether participants produced a zero form or an overt pronoun for the noise-covered part.

The results did not support our hypothesis about the information status and type of previous referential expression as influential factors. However, the number of utterances that a participant repeats as a sentence affects which pronoun form is restored: when a participant repeats two utterances (e.g. *A boy was watching TV. Then (he) ate an apple.*), the zero form is preferred, but when only one utterance (e.g. *Then he ate an apple.*) is repeated, the overt pronoun dominates.

To further assess these findings, we have designed an acceptability-judgement task as a second study. The participants (at least 35) will be asked to rate the acceptability/naturalness of written informal sentences (4-pt scale). The experiment tests the influence of coordination between two utterances as a crucial factor in guiding the choice of zero/overt forms. We manipulate (i) the number of sentences (1 sentence with conjunction *ja* ‘and’, 2 sentences separated by a period), (ii) type of referential form (zero/overt form) and (iii) the order of overt pronoun and verb (pron+V, V+pron). We predict that participants perceive coordinated sentences with zero reference as the most acceptable and two separate sentences with zero reference as the least acceptable. It is also possible that word-order effects are present regarding the overt form. As a whole, our goal is to shed light on how different factors, at different linguistic levels, are at play when a speaker chooses the referential form.

References
A diachronic analysis of Nheengatu subordination constructions

Daven Hobbs
(University of New Mexico)

Keywords: Nheengatu, Tupi-Guarani, subordination, diachrony, usage-based

Nheengatu is a Tupi-Guarani language spoken by a few thousand people (precise figures are unknown) predominantly in the upper Rio Negro region of the Amazon rainforest. It is descended from Old Tupi, a language that was spoken along much of the Brazilian coastline at the time of the Europeans’ arrival into the New World in the late 15th century. Old Tupi as it was spoken during this time period is fairly well-documented; Nheengatu is thus unique among most other indigenous languages of the Americas in having documentation of its lineage in texts going back several centuries. This fact has made it easier to examine Nheengatu from a diachronic perspective. One can compare its structures to those of the language from which it descends, in order to make inferences about the processes of change responsible for any differences in structure that might be observed. The present paper takes this approach to investigate change in one grammatical domain in particular, namely that of subordination.

Three Nheengatu constructions—a complementation construction, a relative clause construction, and an adverbial clause construction—have been chosen for detailed diachronic analysis in this paper. These were chosen in particular because each one represents a different semantic subdomain within subordination, and each also developed out of a different (set of) source construction(s). These three constructions highlight the semantic range of this grammatical domain, as well as the heterogeneous sources and pathways of change through which it can become grammaticalized. The linguistic data used here come primarily from in-depth grammatical descriptions. For modern Nheengatu, the data come from Cruz (2011) and Navarro (2011). For Old Tupi, Navarro (1998) and Lemos Barbosa (1956) were used. Where informative, data from 19th century Nheengatu are also cited, which come from the collection of texts in Barbosa Rodrigues (1890).

This paper focuses on each of the three Nheengatu constructions in turn. Each construction is first compared synchronically to identifiable precursors found in Old Tupi. Then, a diachronic interpretation is developed concerning the processes of change by which each construction would have undergone grammaticalization. The development of the complementation construction is shown to instantiate a pathway from parataxis to syntax, from an unspecified semantic relation between events to true complementation. The development of the relative clause construction, in turn, illustrates a pathway from less to more grammatical within syntax, from interrogative to relativizer. Finally, the development of the adverbial clause construction is argued to showcase a particularly unique pathway from suffix to clitic, and from a general meaning of “futurity” to that of “purposive”. For each pathway of change, certain underlying cognitive mechanisms, such as form-function reanalysis (see Croft 2000), are suggested. The analyses developed here are couched within a usage-based and evolutionary understanding of language structure and change. This paper thus bolsters and elaborates
upon these frameworks, using data that have not yet been subject to detailed historical linguistic analysis.

References

Scope marking by intonation in negative clauses:
The case of Romanian

Emil Ionescu
(University of Bucharest)

Keywords: intonation, scope, negative concord, double negation, metalinguistic negation

In spoken Romanian the scope of the negative marker and negative words (N-words) in negative concord (NC) is determined by intonation. In this respect I present data which, to my knowledge, have been not exploited so far in a systematic way. The method I use involves basics of the theory of quantifiers and operators, more precisely aspects tight to their scope and scope interaction. A key distinction I also use is between descriptive and metalinguistic negation (Ducrot 1972, Horn 1989).

Negative clauses in Romanian contain either a negative marker (usually nu, ‘not’), or a (verb) negative marker plus at least one N-word (e.g. nimic, ‘nothing’). In the former case negation is (simple) sentential negation, in the latter one it is NC.

Spoken negative clauses display emphatic accent (EmphAcc). In simple sentential negation the accent affects either the negative marker, or some verb dependent. If the accent is carried by the negative marker, the scope of the negative marker is the sentence itself (Astăzi N-AM mâncat orez ‘Today I have NOT eaten rice’). If the accent is carried by some verb dependent, the scope of the negative marker is just the dependent under accent (Astăzi n-am mâncat OREZ ‘Today I have not eaten RICE’). In both cases, intonation determines the scope of the negative marker.

When a negative utterance is a NC, the EmphAcc is exclusively carried by the N-word. (Astăzi n-am mâncat NIMIC ‘Today I have not eaten NOTHING’). A NC with several N-words is a NC with each N-word under EmphAcc. The N-word is interpreted as a universal quantifier which scopes over negation.

The scope marker role played by the EmphAcc in NC becomes particularly obvious, when the accent is displaced on the negative marker (Astăzi N-AM mâncat nimic). The intonation thus obtained
cannot be associated any longer with the NC reading. This shows that the EmphAcc on the N-word(s) is a prosodic rule of the NC utterances.

When the negative marker is accented and N-words become non-emphatic, the utterance may be interpreted as a double negation (Astăzi N-AM mâncat nimic: ‘I DIDN’T eat nothing = I did eat something’). This reading apparently results from the scope interaction between the negative marker and the (unaccented) N-word: due to its EmphAcc, the negative marker takes scope over the unaccented N-word (just as in the case of simple sentence negation: Astăzi N-AM mâncat orez ‘Today I have NOT eaten rice’).

I hold, though, that double negation which results from this new intonation is nothing but an ordinary instance of metalinguistic negation (Ducrot 1972, Horn 1989). To this purpose, I use some well-known tests, such as reactivity, continuation, echoing and the distinction truth-assertability (Foolen 1991).

If my analysis is sound, two consequences follow for the status of the NC utterances in Romanian. Firstly, the availability of Romanian, as a NC language, for double negation (claimed in recent researches, Fălăuş 2007, Iordăchioaia 2010), turns out to have no clear base. Indeed, no proper (that is, descriptive) double negation, having the structure and the intonation mentioned above could exist. More precisely, the availability for double negation is stated at the price of not taking into account the prosodic rule of accenting all the N-words in a NC utterance.

Secondly, the data presented above render NC a phenomenon completely compatible with a compositional analysis: the negative marker has its normal negative content, whereas N-words (because of their preference for negative environments) are particular universal quantifiers, which scope over negation. This is essentially Giannakidou’s (1998) tenet about NC in Modern Greek. Giannakidou’s analysis may be adopted for Romanian, too. No resort to a more complex framework (i.e. polyadic quantifiers) is required.

References
Fălăuş, AnaMaria (2007), Le paradoxe de la double négation dans une langue à concordance négative stricte, in F. Floricic (ed.), La négation dans les langues romanes, Amsterdam, John Benjamins, 75-97.
Foolen, Ad (1991), Metalinguistic negation and pragmatic ambiguity. Some comments on a proposal by Laurence Horn, Pragmatics, 1.2., 217-237.
Horn, Laurence (1989), A Natural History of Negation, Stanford, CSLI Publications.

Linguistic interferences in the speech of Romanian children leaving in Slovenia

Ioana Jieanu

As a result of migration, over then 5 mil. of Romanian L1-Speakers now live outside of Romania and Moldova, that is why there are numerous different types of language contact, not only with the various
Romance languages, but also with languages which differ significantly in terms of genetic heritage and typology.

Related to this new spread of Romanian language around the world, and to its contact with different languages, I present in this study the linguistic effects of Romanian migration in Slovenia, starting from the characteristics of the languages involved in this process. The Romanian-Slovene contact analyses the bilingual or plurilingual community to which the children belong.

The study is based on a corpus conducted by questioning two children of 4 and 6 years old leaving in Slovenia, with Romanian parents, once a week, for 10 minutes each, during a 6 months period in 2019-2020 (500 hours of recordings), and random interviews with other five Romanian children living in Slovenia, belonging to mixed families.

The linguistic contact specifies the conditions of migration, the degree of influence of the official language on the Romanian language, and analyses the linguistic attitude of the Romanian children, which varies from the use of Slovene in most communication situations, the alternative use of two codes and the preference for communication in Romanian. The analyse of the corpus focusses on code switching (e.g. *Gremo pa na morje să inotăm*; *Za kaj spui aša; Mami, počakaj me că nu mai pot*), borrowing (e.g. *Nu zvârnești jucăriile; Najprej hai sâ ne bleidâm; Pojdi stran, pușastule!; Petra e bebeță!*), loans (e.g. *Copilul ăsta e slab* [sl. *slabo > engl. bad*]; *S-am speriat*), and word order (e.g. *Frig mi-e!; Al meu prieten*), which were identified in children’s communication.

**References**


**Motion verbs in music criticism: Their conceptual and communicative motivation**

Nina Julich-Warpakowski
(Leipzig University)

Keywords: conceptual metaphor, music, motion, keyword analysis, specialist discourse

Musical structure is commonly described in terms of motion: Melodies ascend and descend, motives return, and the music, harmonically, may arrive at a dominant seventh chord. It is the aim of this paper to investigate what motivates the conceptualisation of musical structure in terms of motion within the framework of Conceptual Metaphor Theory (Lakoff & Johnson 1999).

Given that music is a temporal art, Conceptual Metaphor Theory approaches suggest that the conceptualisation of music as motion through space is based on the conceptualisation of time as
motion through space (Johnson & Larson 2003, Cox 2016). Prototypical examples for the TIME IS MOTION metaphor include deictic uses of path verbs such as approach, arrive, come, go, or pass as in We are approaching Christmas (Moore 2014). Given the hypothesis that musical motion expressions are based on the TIME IS MOTION metaphor, it is expected that the majority of musical motion instances should mirror that pattern and be realised by deictically used path verbs.

For the analysis, a 6.7-million-word corpus of texts from the genre of music criticism was compiled. This domain-specific corpus consists of musical analyses of classical music and concert reviews retrieved from musicology journals and newspapers published in the UK and the US. A keyword analysis was conducted in order to identify key motion verbs in the music criticism corpus, and concordances for selected motion verbs were retrieved to analyse whether musical motion instances were used deictically or not.

The results reveal that path verbs prototypically associated with the TIME IS MOTION metaphor are negative keywords in the music criticism. Instead, many manner verbs are among the positive keywords. Furthermore, the concordance analysis indicates that only a minority of musical motion instances are used deictically.

The findings thus do not support the hypothesis that musical motion is primarily based on prototypical, deictic temporal motion metaphors. Instead, the frequent use of manner verbs is characteristic for the communicative function of review genres, in which manner verbs are used to assess the quality of a musical performance, and to express the subjective attitude of the reviewer (Caballero 2017). This use of manner verbs can be interpreted as being motivated by the Event Structure Metaphor sub-mapping ACTIONS ARE SELF-PROPELLED MOVEMENTS, in which manner of motion maps onto the manner of action (of a musician). Furthermore, the activity of music criticism is divorced in time from the actual musical experience which can explain the absence of deictic motion expressions. Instead, musical structure is described from a bird’s eye perspective in which musical development is conceptualised as motion irrespective of a deictic ego. Motion expressions referring to how the music develops over the course of a piece can be interpreted in terms of the Event Structure Metaphor sub-mapping PURPOSES ARE DESTINATIONS, or alternatively as instances of fictive motion. The results indicate that in addition to plausible conceptual motivations, it is the specific genre and communicative setting of music criticism that determines the particular use of metaphors for musical aspects.

References
The papal bull “Ineffabilis Deus” (1854) translated: presentation of a parallel corpus and analysis of 19th century Ibero-Romance varieties

Johannes Kabatek  
(University of Zurich)

Keywords: Parallel corpus, minority languages, Ibero-Romance, standardization, 19th century.

In 1854, Pope Pius IX proclaimed the bull “Ineffabilis Deus” about the immaculate conception of the virgin Mary, a Latin document of theological and philosophical argumentation with an extension of 4476 tokens in the original version. During the following years, the French priest Marie-Dominique Siré collected translations of the text into the languages of the world. He compiled more than 300 translations, among which 55 into European languages and varieties. Several of these languages did not count on elaborated written traditions in the 19th century, so that in these cases the text is a pioneering piece of elaborated prose. The whole collection of manuscripts was donated by Siré to the Pope and has been part of the Vatican library for more than 150 years without ever having been studied linguistically.

The contribution will consist in
a) the presentation of the project of an online parallel corpus of the text of the bull. This corpus will be built up step-by-step, starting with European languages and varieties. It will allow access to annotated transcriptions of the text as well as to the digitized original documents. The first step to be presented will consist of 20 European languages and varieties.
b) the presentation of four versions of the text in order to show how the corpus may be used for linguistic analysis. I will choose the Ibero-Romance varieties Catalan (Central Catalan as well as Catalan from Mallorca), Asturian and Galician. In all four cases, the traditions of written prose in the 19th century at the moment when the bull is translated are scarce. Only in one of the four cases, the Asturian translation, the community of speakers becomes aware of the existence of the text, whereas in the other cases the translations remain unknown.

I will analyse four dimensions of the texts (orthography, morphological unity, lexical variety and clause linkage) in. In all of these fields, the translations show a surprising degree of standardization: the orthography appears as rather systematic, morphology as rather unified, the lexicon offers a high degree of diversity and the clause-linking techniques show a wide range of aggregative as well as integrative techniques (Kabatek/Obrist/Vincis 2010).

On the other hand, the corpus allows for interlinguistic comparison of features of the respective languages in their diachronic state of the mid 19th century. I will compare some morphosyntactic features (clitic positions, differential object marking and possessive constructions) in order to show the potential of analysis the parallel corpus offers.

References
Ferrando i Francés, Antoni / Amorós, Miquel Nicolás (2011), Historia de la llengua catalana, Barcelona, Editorial UOC.


Siré, Marie-Dominique (1866), “Traduzione in tutte le lingue della Bolla Ineffabilis, colla quale Sua Santità il Papa Pio IX proclamò domma di fede l’Immacolata Concezione di Maria santissima”, in La Civiltà Cattolica, 17/7/IV, 729-737.


The Semantics and valency of verbs of the stem X (-stVCCVC-) in the Qur’an: A corpus-based analysis

Csilla Kász
(University of Kiel)

Keywords: Classical Arabic, statistical distribution of valency and semantical properties, root-and-pattern morphology, morpheme- and word-based derivation, corpus-based study

While the system of verbal derivational patterns in Arabic, usually referred to as stems I–XV (Arab. āwzān/ Hebr. binyamin), exhibits great formal regularity, its function is not clear-cut and a matter of ongoing study. The way grammars usually describe stems is to list frequent functions and illustrate them with few examples, often ignoring deviations from these functions. This method has already been found in the first grammars of Arabic (as for instance the Kitāb of Sibabawaih, cf. Jahn 1900) and pursued in more recent grammars of (Classical and Modern Standard) Arabic (cf. Wright 1967, Fischer 2006, Holes 2004, Badawi et al. 2004, Ryding 2005, and El-Ayoubi et al. 2010, to name few).

Aside from the grammars, a number of studies emerged, contributing to the description of individual derivational patterns (e.g. Leemhuis 1977 on stem II and IV in the Qur’an, Saad 1982 on transitive, causative and passive patterns, Zaborski 2007 on non-causatives of stem IV, Danks 2011 on stems III and IV, Ratcliffe 2005 on stem III in the Qur’an, and Granville 2018 on the identification of semantic relationships in the Arabic lexicon as whole). However a systematic corpus-based description of the verbal derivational patterns, that accounts also for irregularities and show the statistical distribution of their functions is still largely missing.

In my presentation I summarize the results of a recent corpus-based study on verbs of stem X that has been conducted within the scope of my PhD thesis on verbs of communication in the Qur’an. It includes 71 verb types (of stem X) with 369 tokens. The goal of the presentation is threefold:

First, to verify the already identified functions of the stem that have been described in the grammars of Arabic mentioned above as (i) reflexive to the causative or estimative of the stem IV (e.g. āhkama (IV) ‘make firm’ (causative) > īstakhkama (X) ‘make oneself firm’ (reflexive); akbara (IV) ‘consider sth. important’ (estimative) > īstakbara (X) ‘consider oneself important’ (reflexive)), (ii) estimative (e.g. hasuna (I) ‘be good’ > īstahsana (X) ‘consider sth. good’), (iii) desire y or ask...
for y’ whereas y is the meaning of the verb of stem I (e.g. ġafara (I) ‘forgive’ > istaġfara (X) ‘ask for forgiveness’).

Second, to identify, based on their statistical distribution, the core and peripheral functions of the stem.

And third, to address the ongoing discussion, whether derivation in Arabic is morpheme-based (i.e. the basis for the derivation are the root consonants) or word-based (i.e. the basis for the derivation are basic or already derived words) (cf. Ratcliffe 2005 and Granville 2018). My analysis confirms the view found among others in Ratcliffe (2006), that both approaches are relevant for the description of Arabic morphology.

References

Hindi-Urdu light verbs revisited

Liudmila Khokhlova & Boris Zakharin
(Moscow State University)
Keywords: Hindi-Urdu, light verb, telic verb, incremental theme, conative meaning

The compound verbs in Hindi-Urdu are of the Vv-type: they are formed by combinations of the main verb stem with finite forms of the corresponding light verb. The light verbs may double or modify the meaning of the base, add deictic features to its semantics, clarify the meaning of labile verbs etc. This paper is aimed at discarding certain linguistic myths concerning semantic-grammatical properties of compound formations with light verbs in Hindi-Urdu.

I. The inceptive or completive meanings of compounds in question are often ascribed to the grammatical and semantic properties of the light verbs (Singh 1990, Butt 1995, et al.). It will be shown in the paper that inceptive (versus completive) meaning of a compound verb depends on the feature of ‘telicity’ in the main verb. The compound verbs formed from non-telic stems usually imply the inceptive meaning (1), while those produced from telic stems may have completive sense (2).

(1) sītā haēs par-ā’
Sita laugh fall-AOR.F.SG
‘Sita started laughing.’

(2) laṛkā ghor-e se utar par-ā’
boy-M.SG horse-M.SG.OBL ABL get off fall-AOR.M.SG
‘The boy got off the horse.’

II. A general standpoint is that light verbs always signify a culmination of the event in its natural endpoint (Hook 1974; Liperovsky 1984; Nespital 1997; Montaut 2004, et al.). However, aorist forms from compound verbs may not denote the completion of the action in case the theme of the utterance is incremental (Dowty 1991).

(3) gītā ne gān-ā gā li-yā’
Gita ERG song-M.SG sing take-AOR.M.SG
par ākhirī lāīn bhāl ga-ā’
but last line.F.SG forget go-AOR.F.SG
‘Gita sang a song but she forgot the last line (of it).’

Three types of verbs with incremental theme denoting an incomplete action when combined with the light verbs will be discussed in the paper.

III. Contrary to some scholars’ opinion [Butt 1995, Singh 1990], not the light, but the main verb attributes the meaning of conscious choice to the actor, compare (4) and (5):

(4) laṛkā gir par-ā’
boy fall fall-AOR.M.SG
‘The boy fell’

(5) ciṭā laṛk-ē par jhapāṭ par-ā’
Panther boy-M.SG.OBL on pounce fall-AOR.M.SG
‘The panther pounced on the boy’

IV. There exists general opinion that compound verbs cannot be used in progressive aspect because of their implying either ‘inceptive’ or ‘completive’ meanings. However, this type of usage is allowed, but restricted to 1-st person verbal forms and denotes intention to perform the action:

(6) yah khān-ā bekār h-ai, is bār to
this food-M.SG bad be-3.PRES.SG this time Emph
khā le rah-ā h-āī’
eat take stay-PP.M.SG be-PRES.1.M.SG again NEG give-INF-M.SG
dubārā mat de-n-ā
‘This food is bad, now I am eating / will eat, but do not offer it next time!’
V. There is a general view that negation particles are not to be used with the compound verbs. The contexts where this usage is quite possible will be described in the paper.

References

Caritive constructions with *ka*-nouns in Tagalog

Sergei Klimenko
(Institute for Linguistic Studies, Russian Academy of Sciences)

Keywords: Tagalog, caritive, privative, abessive, corpus linguistics.

Caritive (abessive, privative) situations in Tagalog are encoded with a number of understudied periphrastic constructions. One of the three most common types of Tagalog caritive constructions involves the noun *kasama* (*ka-sama* ‘SHR-join’) ‘companion’ in combination with the negator *(hind)idi* as the predicate of an adjunct clause (1) (the other two most common caritive constructions in Tagalog are formed with the negative existential predicate *wala* and the noun *dala*, which is the root of the verb ‘to carry’ and also denotes an object that actor carries):

(1) …never=pa=niya=ng na-gawa na ma-tulog sa never=yet=3SG.ACT=LKP1FV.MOD-do[UV] LK AV.INF.STEM-sleep OBL iba=ng bahay na hindi kasama asawa=nya… other=LK house NEG companion spouse=3SG.ACT ‘…she has never slept in another house without her husband…’

Without negation, *kasama* occurs in a number of noncaritive constructions with a variety of functions. Apart from independent and attributive uses, these include comitative constructions (2), depictive secondary predicates (3), event-oriented adjuncts (4), and predicative complements (5) (Klimenko, forthcoming).

(2) K<um>ain=sila kasama si ali=ng Meding. <AV>eat[PFV]=3PL.NOM companion PRS.SG.NOM aunt=LK PN ‘They ate with aunt Meding.’

(3) Nag-mukmok=na=lamang=ako sa kwarto kasama ng AV.PFV.STEM-sulk=already=only=1SG.NOM OBL room companion GEN akin=g mga libro. 1SG.ACT=LK PL book ‘I just sulked in my room together with my books.’

(4) …p<in>ag-hirap-an=namin=yan maka-ipon
Kasama belongs to a class of nouns with the prefix *ka-* denoting “a person or a thing with whom the place, object quality or situation expressed by the root is shared” (De Vos 2011: 83). Other less frequently used nouns with the same prefix requiring physical or abstract presence of their arguments in the situation (as opposed to, for instance, *katrabaho* ‘workmate’) can also be used to build similar constructions. While *kasama* (146,579 tokens in the corpus) has the broadest meaning in this group of nouns, others refer to more specific situations of involvement. The most frequent ones are: *kabilang* ‘member’ (52,601), *kasabay* ‘simultaneous’ (16,943), *kausap* ‘interlocutor’ (13,050), *katabi* ‘adjacent’ (11,489), *kalahok* ‘participant’ (6,803), *kasali* ‘participant’ (4,306), *kasosyo* ‘partner’ (3,734), *kaharap* ‘somebody face to face with somebody else’ (3,262), *kaakitbat* ‘associate’ (3,027), *kasangkot* ‘involved’ (2,985).

This study will look into semantic and syntactic properties of different types of constructions with the ten *ka*-nouns listed above and their occurrence in caritive constructions based on data from the Tagalog Web 2019 (tlTenTen19) online corpus of over 230 million tokens (Jakubiček et al. 2013) at sketchengine.eu (Kilgarriff et al. 2004).

**Abbreviations:**
1, first person; 12, first and second person; ACT, actor form; AV, actor voice; DIST, distal; GEN, genitive; INF, infinitive; LK, linker; MED, medial; MOD, modal; NEG, negative; NOM, nominative; OBL, oblique; PFV, perfective; PL, plural; PN, personal name; PRS, personal; SG, singular; SHR, shared notion; STEM, stem-deriving prefix; TOP, topic marker; UV, undergoer voice.

Acknowledgements: This study was funded by the Russian Science Foundation (Project No. 18-78-10058 “Grammatical periphery in the languages of the world: A typological study of caritive”).

**References**

©
Detecting linguistic variation with geographic sampling

Ezequiel Koile & George Moroz
(National Research University Higher School of Economics)

Keywords: Linguistic geography, Variation, Sampling, Computational methods, Circassian languages

Geolectal variation is often present in settings where one language is spoken across a vast geographic area (Labov 1963). This can be found in phonological, morphosyntactic, and lexical features. It is often the case that linguists overlook this variation, assuming more homogeneous distributions (Dorian 2010). Let us consider the case of a geolectal continuum formed by a group of small villages whose dialects differ slightly with distance (Chambers & Trudgill 1998). For practical reasons, it is not always possible to collect fieldwork data from every single location in order to obtain this full pattern of variation. Instead, the most efficient would be to collect data from a selected group of these villages, but the question arises of which locations should be surveyed, in order to resemble the real distribution of linguistic features, or at least to detect the degree of different values existent for each of them. We propose and test a method for sampling different locations where a language is spoken, finding the optimal places to be included in a sample, with the goal of obtaining a distribution of typological features representative of the whole area.

For this goal, we use different clustering algorithms such as $k$-means and hierarchical clustering of locations based on their geographic distribution, and define our sample of locations on the basis of this clusterization (picking one location from each cluster), using the packages stats (R Core Team 2019) and lingtypology (Moroz 2017) in R. This assures us a geographically representative sample, less affected by spatial autocorrelation than a random selection. We test our methods against simulated data with different distributions of linguistic features, on various spatial configurations, and also against real phonological data from Circassian dialects (Northwest Caucasian). Our results show an efficiency higher than random sampling, both for detecting variation and for estimating its magnitude (see figure), which makes our method profitable to fieldworkers when designing their research.
Figure 1. Amount of different dialects detected with different methods and different sample sizes. On the x-axis, we plot the sample size in percentage of the total amount of villages (N=158). On the y-axis, we plot the amount of different dialects detected with each method. Hierarchical clustering is plotted in red, k-means in green, and random sampling in blue. It can be seen that for small sample sizes (up to 20 or 30%), clustering methods overperform random sampling, while the opposite is true for higher samples (80-90%). The latter is due to the fact that too many clusters (around 50 in this case, p=0.3) are not a good representation of the spatial distribution of these languages.

Data from 158 villages of speakers of 10 Circassian dialects. This plot is created with the ggplot2 package (Wickham 2016)

References
Using online language use data to explore morpho-syntactic language variation in English

Viktorija Kostadinova
(University of Amsterdam)

Keywords: prescriptivism, language variation, sociolinguistics, morphosyntax, online data

Studies exploring prescriptivism’s effects on language use have tended to use a precept vs practice approach (e.g. Auer and González-Díaz 2005, Anderwald 2012, 2014), where data from 18th- and 19th-century normative grammars (i.e. precept) is correlated with data from corpora (i.e. practice). These studies’ results range from showing that prescriptivism either has a short-term effect on language use (e.g. Auer and González-Díaz 2005) or no effect at all (e.g. Anderwald 2012, 2014) to indicating that prescriptivism may have substantially influenced the development, or rather demise, of some features (e.g. Molencki 2003, Laitinen 2009). The effects of prescriptivism remain insufficiently explored empirically, especially when it comes to present-day English and the linguistic behaviour of individual speakers (cf. Kroch & Small 1978). The reason for this is that prescriptive ideology tends to target morpho-syntactic features, and data on the use of these features in everyday language use is difficult to obtain using traditional sociolinguistic techniques. These challenges have been addressed by computational sociolinguistic studies, which leverage the ever-increasing availability of language use data in online communities (cf. Nguyen et al. 2016). This paper aims to explore the use of data from the online community Reddit for investigating prescriptively targeted features in present-day English (i.e. the split infinitive, number agreement, and pronominal forms in coordinated phrases) in order to answer the question: how can online language use data be used to explore morpho-syntactic language variation?

More specifically, the paper will (i) describe the variation in the use of prescriptively targeted features across time and across gender categories of Reddit users, and, in doing so, (ii) provide an example of how Reddit data can be used to explore sociolinguistic questions which are difficult to address through traditional methods. The general approach consists of establishing the variables of the prescriptively targeted features mentioned above, and tracking their variationist frequencies of occurrence in comments on Reddit. The Reddit data was obtained from pushshift.io, and analysed with Python. Only comments associated with available gender information about the user (available in the ‘author flair’ field in the dataset) were selected for the preliminary analysis and were annotated with the NLP Python library SpaCy, in order to facilitate extraction of the variables in question. Preliminary results obtained from a small subset of all available data suggest that while in general online language use data can indeed be fruitful in exploring cases of morpho-syntactic variation, this comes with a set of challenges (e.g. when it comes to annotation), which will be addressed in this paper.

References:


**Lineage-specific trends in the evolution of verbal negation**

Olga Krasnoukhova & Marc Tang

(University of Antwerp & University Lumière Lyon 2)

Keywords: verbal negation, negation pattern, evolution, phylogenetic methods

The evolutionary trends of verbal negation have, among others, been discussed with the ‘Jespersen Cycle’ (using Dahl’s 1979:88 terminology). The ‘classical’ version of this process involves the following steps (Jespersen 1917; van der Auwera 2009): a single clausal negator before the verb is joined by another marker after the verb, which is either negative too or becomes negative, and thus results in a construction that expresses a semantically simple negation twice (the so-called ‘doubling-stage’); the first negator can eventually disappear, leaving only the ‘new’ negator after the verb. This scenario is schematized in (1). The typological literature also reports on cases where the directionality of the process is reversed (e.g. van der Auwera & Vossen 2016; Vossen 2016 *passim*). In these languages, it is the postverbal negator that is joined by another marker placed before the verb, and this development may eventually lead to a preverbal single negation pattern (2).

\[
\begin{align*}
\text{(1)} & \quad \text{neg V} \rightarrow \text{neg V (neg)} \rightarrow \text{neg V neg} \rightarrow \text{V neg} \\
& \quad \text{Stage 1} \quad \text{Stage 2} \quad \text{Stage 3} \quad \text{Stage 4} \quad \text{Stage 5}
\end{align*}
\]

\[
\begin{align*}
\text{(2)} & \quad \text{neg V} \leftarrow \text{neg V (neg)} \leftarrow \text{neg V neg} \leftarrow \text{V neg} \\
& \quad \text{Stage 5} \quad \text{Stage 4} \quad \text{Stage 3} \quad \text{Stage 2} \quad \text{Stage 1}
\end{align*}
\]

We contribute by using phylogenetic methods to investigate the directionality and the transition patterns of negation marking in a selection of six large language families across the globe: Austronesian, Indo-European, Atlantic-Congo, Sino-Tibetan, Arawakan, and Tupian. Data on negation patterns comes from Dryer (2013) that is complemented by own data to ensure that each internal node
within a family is represented. We use the method of prior predictive simulation to generate a sample of 1000 phylogenetic trees for each family based on the constraints of the Glottolog tree (Hammarström et al 2018). Then, the phylogenetic method of Reverse Jump Hyper Prior (Gowri-Shankar & Rattray 2007) is applied to identify the likely directionality of change of negation and the transition patterns across the tree samples. The results partially corroborate the findings of Dunn et al (2011) by indicating that the transition patterns of negation diverge across language families, but only to a small extent.

By using the quantitative methods which have not been applied yet in research on evolution of negation patterns, the paper aims to contribute to our knowledge of diachrony of negation in general and the selected language families in particular.

References

Diachronic dynamics of a stem-derivational aspect system: how cluster analysis helps discover patterns

Joanna Wrzesień-Kwiatkowska & Björn Wiemer
(Johannes-Gutenberg University, Johannes-Gutenberg University – Mainz)

Keywords: Czech, aspect, stem derivation, cluster analysis, diachronic morphology

The talk demonstrates how cluster analysis can help discover semantic and frequential changes among patterns of stem derivation which are essential for the Czech aspect system.

The backbone of the aspect system of Czech (and other Slavic languages) builds on two productive patterns of stem-derivation (Breu 2000 and Wiemer/Seržant 2017): either an unprefixed verb stem (simplex, imperfective/IPFV1) is prefixed (see 1), or an already prefixed stem (perfective/PFV) is additionally suffixed (see 2) yielding another imperfective stem (IPFV2). If the two stems replace each
other under clearly definable conditions without changing the lexical meaning, the stems are not just synonyms, but so-called aspect (perfective—imperfective) pairs.

\[
\begin{array}{ccc}
\text{IPFV1} & \text{PFV} & \text{IPFV2} \\
(1) & \text{psá-t} & \Rightarrow & \text{na-psa-t} & \text{‘write’} \\
(2) & \text{pře-psa-t} & \Rightarrow & \text{pře-pis-ova-t} & \text{‘rewrite’}
\end{array}
\]

Pairs of type (1) arise if the prefix meaning overlaps with a semantic component of the simplex (Janda et al. 2013). If such pairs acquire another ipfv. stem (IPFV2) via suffixation, aspect triplets result (e.g., \(\text{dělit} \Rightarrow \text{roz-dělit} \Rightarrow \text{rozděl-ova-t} \text{‘separate’}\)). Triplets can “survive”, e.g. if IPFV1 and IPFV2 remain synonyms, but develop preferences for different subsets of functions associated with imperfective aspect. Otherwise triplets decay: either the lexical meanings of the imperfective stems dissimilate, or one of them becomes obsolete. Consequently, triplets are no homogeneous class; they vary considerably as for their (i) semantics and polysemy, (ii) frequency (comparing both IPFV1 vs IPFV2 from the same triplet and entire triplets), (iii) “age”, and (iv) persistence. We search for patterns in the interplay of these four parameters, in particular: do the persistence (or decay) of triplets and their meaning ranges correlate with their age, their prefixes, semantic features, or with frequency?

For this purpose we composed a database containing 651 potential triplets in Czech for 1750-2017, subdivided into six subperiods. Frequencies drawn from corpora (see References) are used for cluster analysis applied to these subperiods, for both IPFV2 and IPFV1. By employing time series clustering with non-Euclidean distance measures we propose a certain methodological transfer from engineering (Iglesias/Kastner 2013) and economy (Roelofsen 2018) to linguistic data. Contrary to previous cluster analyses (e.g., Divjak/Fieller 2014), diachronic change figures as a variable, and clusters are based on frequency changes, not (as in Jansegers/Gries 2019) on extensive annotation. The obtained clusters provide a starting point for qualitative analysis of selected typical and untypical representatives, with different prefixes and inherited vs borrowed (before or after 1750) roots from each cluster. Persistence rates and changes in meaning alternations and the lexical relation between IPFV1 and IPFV2 are established.

Preliminary observations suggest:

- Triplets attested before 1750 have persisted, which implies sufficient stability of the lexical relation between IPFV1 and IPFV2.
- There are more IPFV2 with initially low, but steadily growing frequency than initially highly frequent IPFV2 with a drastic reduction in the 20th century (38% vs 21% from 651).
- Deadjectival triplets behave differently: simplex stems appeared after the prefixed stems and more frequently tend to disappear again (e.g., \(\ddagger \text{slabít} \Leftarrow \text{oslabít} \Rightarrow \text{oslabovat} \text{‘weaken’}\)).

From our results we will draw conclusions concerning the stability of the aspect system.

References


Conditional clauses are traditionally defined as constructions exhibiting a cause-consequence pattern, in which “the truth of the proposition in the matrix clause is a consequence of the fulfilment of the condition in the conditional clause” (Quirk et al. 1985: 1088), as illustrated in (1). Prior research has demonstrated, however, that constructions introduced by the prototypical markers of conditionality (if in English and si in Spanish) can in fact express a wider variety of functions in discourse, ranging from expressions of politeness to metalinguistic comments, among others (Ford and Thompson 1986, Sweetser 1990, Ford 1997, Dancygier and Sweetser 2005, Warchal 2010, Brinton 2014, 2019, Lastres-López 2019, 2020), as in (2).

(1) So if I get the job, I’ll rent for a while until the house prices pick up (ICE-GB:S1A-019 #277:1:C)
(2) Siéntate aquí si quieres eh (C-ORAL-ROM efamcv05)
‘Sit down here if you want eh’

This paper explores conditional constructions in English and Spanish informal, face-to-face conversation, adopting a corpus-based and a contrastive approach. Our aim is to explore the uses and functions of these constructions in both languages, and the extent to which such functions show cross-linguistic similarities and/or divergences. Taking prototypical (cause-consequence) conditionals as a point of departure, we propose a categorialization of conditional constructions along the lines of the ideational, interpersonal, and textual metafunctions distinguished by Halliday and Matthiessen (2014).
This development from fully propositional meaning to interpersonal and textual functions is characteristic of processes of pragmatisalization (Claridge and Arnovick 2010: 179). In this suggested cline, ideational conditionals, as in (1), move beyond their prototypical conditional meaning to acquire “an essentially metacommunicative, discourse interactional meaning” (Frank-Job 2006: 397), as in (2). This path of pragmatisalization also throws light on constructions that originate from conditionals, such as cases of conditional insubordination (Mato-Míguez 2014, Kaltenböck 2016, Schwenter 2016, Lastres-López 2018), as in (3) and (4); or pragmatic markers introduced by conditional conjunctions (c.f. Brinton 2017, on English if I may say so; and Fuentes Rodríguez 2019, on Spanish si cabe).

(3) If you’ll just come next door (ICE-GB: S2A-054 #135:2:A)
(4) Si es que no me las voy a comer (C-ORAL-ROM efamcv07)
‘I’m not going to eat them’

The data analysed are extracted from the conversation datasets of the British component of the International Corpus of English (ICE-GB) (Nelson et al. 2002) and the Spanish component of the Integrated Reference Corpora for Spoken Romance Languages (C-ORAL-ROM) (Cresti and Moneglia 2005). Corpus findings confirm the diversity of discourse-pragmatic functions of conditionals in spoken English and Spanish. The results can also be interpreted in terms of decategorialization (Hopper 1991: 22), since pragmatisalized conditionals appear to lose some of the formal features characteristic of the ideational patterns. For example, conditionals in English typically contain a modal verb in the apodosis (Gabrielatos 2010, 2019), which is frequently absent in pragmatisalized conditionals. Likewise, despite initial position for the protasis being a language universal (Greenberg 1963), certain interpersonal conditionals favour a non-initial position for the protasis.

References

Fuentes Rodríguez, Catalina (2019), Combinaciones entre operadores argumentativos: El regulador escalar si cabe, Spanish in Context 16(2), 151-172.


Sweetser, Eve (1990), From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure, Cambridge: Cambridge University Press.

Typology of preverbs and transitivity in diachrony: Evidence from Greek and beyond

Nikolaos Lavdas & Leonid Kulikov
(National and Kapodistrian University of Athens, Ghent University)

Keywords: Indo-European, Greek, Vedic, reanalysis, preverbs

Indo-European preverbs can show features of a free morpheme, being, in particular, separable from the verb (“tmesis”; Watkins 1964), cf. (1a, b). In some other cases, however, they attest characteristics of non-separable bound morphemes. Previous studies have identified the transitive function of preverbs, as is the case of the Modern German applicative preverb be- (e.g. Michaelis & Ruppenhofer 2001). The aim of our paper is to demonstrate that (i) the direction of development of Greek preverbs testify to the existence of two types of preverbs, internal and external (Di Sciullo 1997; Di Sciullo & Slabakova 2005); and (ii) only internal preverbs (with idiosyncratic and non-compositional readings) can affect transitivity.

The diachrony of Greek provides several examples of a reanalysis of external preverbs into internal. Homeric Greek preverbs retain a concrete local or directional function and can assign case to a verbal argument, as, for instance, in (2), where the preverb en- assigns the dative to the NP nēusìn. In our study, we focus on the historical dimension of preverbs, mainly concentrating on the direction of change in preverbation in post-classical Greek. Thus, we analyze data on the compositional and idiomatic meaning of the verbal forms (cf. Ancient Greek sum-peri-phēromai together-round-carry: MEDIOPASS.PRS.1SG ‘I am carried round together’ or ‘I accommodate/adapt myself to circumstances’), paying special attention to the process of the loss of compositional meanings in post-classical Greek. Similar cases of transitivity of intransitive verbs accompanied by lexicalization are attested in Vedic, as in (3).

The observed diachronic tendency for the Greek preverbs is to develop an idiosyncratic (idiomatic) reading. Accordingly, external preverbs with compositional meanings tend to become internal, non-compositional preverbs, affecting telicity and transitivity. We also argue that parallel changes in the case of the fusion of preverbs and roots and the fusion of two preverbs (double preverbation; cf. the verb eks-apo-stéllō ‘send out, send away’) do not directly affect transitivity.

(1) a. tôn d’ aûte proséeipe `theà glaukôpis
   ART.ACC PTC again towards.spoke:3SG goddess:NOM gleaming-eyed:NOM
   Athēnē
   Athena:NOM
   ‘The goddess Athena with gleaming eyes addressed him again.’ (Hom. II. 1.206)

b. prôs d’ Eurûkleian éeipe
   towards PTC Eurykleia:ACC spoke:3SG
   ‘He addressed Eurykleia.’ (Hom.Od. 20.128) (Horrocks 1981: 41)

(2) pûr ém-pese nēusìn Akhaiôn
   fire:NOM in-fall:AOR.3SG ship:DAT.PL Achaean:GEN.PL
   ‘Fire fell into the ships of the Achaeans.’ (Hom. II. 16.112-113) (Viti 2008)

(3) upa-sthî-ya-te
   near-stand-PRES.PASS-3SG
   ‘[Agni] is worshipped.’ (KS 7.5:66.18-19 = KpS 5.4:53.14)
Acknowledgments
This research was supported by a research grant awarded to Leonid Kulikov from the European Union’s Marie Skłodowska-Curie program (grant nr. 665778) and from the NCN (POLONEZ grant nr. 2015/19/P/SH2/02028) during a research stay at Adam Mickiewicz University in Poznań.

References
Viti, Carlotta (2008), From space words to transitive markers: The case of ancient Greek en ‘in’. Transactions of the Philological Society 106(3), 375–413.

"The “-r- that characterizes the future” in Spanish: discourse functions of futurizing verb forms in Old Spanish poems"

Mary C. Lavissière
(University of Nantes)

Keywords: futurity, modality, Old Spanish, discourse functions, argumentation

As noted by Guillaume (1944), Romance tenses whose morphemes contain an –r– share the characteristic of referring to futurity. In Spanish, this marker extends across indicative and subjunctive modes (Table 1). Crucially, these modal systems saw a major diachronic change linked to futurity: the -ra form moved from an indicative form denoting past events to a subjunctive capable of referring to potential future events (Luquet, 1988). Variation among these forms, such as the -ra form and the conditional (Luquet, 1988), shows that they have clearly been associated in discourse. Studies about their discourse functions as a set, however, are lacking. This article focuses on these functions. We argue that one unifying criterion of these forms is that, thanks to their forward-referring potential, they are used for prospective shifts that serve rhetorical or dialogic functions.

To analyze these discourse functions, we adopt the point of view that discourse-pragmatics are vital to understanding how verbal categories relay meaning and how this meaning changes diachronically (Fleischman and Waugh, 1991). We present a fine-grain qualitative analysis of the Spanish verbs containing the -r- in preguntar poems (Chas Aguión, 2001) of the Cancionero de Baena. This poetry anthology was compiled in the 15th century, a critical period for the phenomena studied (Luquet, 1988). We chose poetic texts because of their focus on sound and their potential to associate several members of the -r- ensemble in a short passage. Our corpus includes 29,500 words and 320 tokens.

Our results demonstrate that these forms have two futurizing functions. First, they serve to advance the poet’s argumentation in poems containing philosophical debates. In these poems, the forms project the logical consequences of the author’s arguments (future indicative) or project a false future,
resulting from erroneous arguments (the conditional). Second, this set of forms is also used in the zones of interaction between locutor and interlocutor. The forms are concentrated in the initial question that launches the debate, in the opening of poems and at the end of poems, where the poet exhorts his or her interlocutor to respond.

The functions observed in our corpus demonstrate that the -r- was associated with prospective shifts in Spanish discourse. This association may have contributed to three major diachronic changes in Spanish: 1) the -ra form’s ability to evoke futurity (Luquet, 1988); 2) following the changes in the -ra form, the -re form’s becoming obsolete in Standard Spanish, yet continuing to exist in specialized genres such as legal texts (Eberenz, 1990, Lavissière, 2019); 3) the current growing obsolescence of the -se form, cantase [sing.SBJV.PST.3SG], which lacks the characteristic -r- (Rosemeyer and Schwenter, 2019).

<table>
<thead>
<tr>
<th>Tense</th>
<th>Mood</th>
<th>Example</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Future</td>
<td>Indicative</td>
<td>Cantará</td>
<td>He/She/You will sing</td>
</tr>
<tr>
<td>Conditional</td>
<td>Indicative</td>
<td>Cantaría</td>
<td>He/She/You would sing</td>
</tr>
<tr>
<td>-Ra form</td>
<td>Indicative (Medieval) / Subjunctive (Contemporary)</td>
<td>Cantara</td>
<td>He/She/You had sung / that he sing</td>
</tr>
<tr>
<td>-Re form</td>
<td>Subjunctive</td>
<td>Cantare</td>
<td>He/She/You will have sung</td>
</tr>
</tbody>
</table>

Table 1: Set of conjugated verb forms in Spanish that contain the characteristic -r-.

References
Baena, Juan Alfonso de (1993), Cancionero de Juan Alfonso de Baena in B. Dutton and J. González Cuenca (eds), Madrid, Espagne: Visor libros. (Original work published ca 1425).
Chas Aguión, Antonio (2001), Juan Alfonso de Baena y los diálogos poéticos de su cancionero. Baena: Ayuntamiento de Baena.
The Fourth Palatalization of Velars in Southwest Ukrainian: The Case of the Kryvorivnja Dialect

Oksana Lebedivna  
(National University of Kyiv-Mohyla Academy in Kyiv, Ukraine)

Keywords: Kryvorivnja dialect, 4th palatalization of velars, intrasyllabic harmony, flatness (rounding), i, i, e sound formants

The 4th palatalization of velars is a complex historical change of Common Slavic (CS) gy, ky, xy > gˈi, kˈi, xˈi in individual Slavic languages. In Ukrainian, CS y has been widely replaced by a phoneme i as in [fi]nuti ‘to perish’, [ki]pity ‘to boil’, [xi]tatsja ‘to stagger’. The Kryvorivnja dialect (K), one of the Hutsul dialects of Southwest Ukrainian (SWU), renders CS gy, ky, xy consistently as hˈi, kˈ(ˈ)i, xˈ(ˈ)i, hˈε, kˈ(ˈ)ε, xˈ(ˈ)ε with a transitional [ˈ] like sound after a sharpened velar. Not having been described for Hutsul before, forms in [ˈε] show, however, relatively high ranking in K.

In my paper, I examine the K reflexes of CS gy, ky, xy through the tendencies toward intrasyllabic harmony which may be reinterpreted with respect to the distinctive tonality feature of flatness (rounding), i.e., “high tonality” consonant allophone (palatalized) developed before vowel “with distinctive high tonality” (non-flat) and low tonality allophone (non-palatalized) developed before vowel with distinctive low tonality (flat) (see Timberlake 1978:726).

According to Flier (2007), the 4th palatalization of velars is remarkable for initial lacking a general prerequisite for sharpening; this is why the phonological motivation for the 4th palatalization was from consonant k > k· > kˈ to vowel y > ţ > i. Connected with the jer shift (early-mid 12th c. in SWU), the 4th palatalization could have occurred after a reconfiguration of the vowel system (Flier 2007; Timberlake 1978; Kasatkin 1968 for different dialects; see also Bethin 1998) or simultaneously with the shift of jers (for a chronology, see Shevelov 1979:237-245).

Velars came to be realized in K as sequences hˈi, kˈ(ˈ)i, xˈ(ˈ)i, hˈε, kˈ(ˈ)ε, xˈ(ˈ)ε. The requirement for this change in Hutsul is the overall lowering of front vowels. Calculation of sound formants in 11 different speeches on the Praat program (used for spectral analysis and measuring of sound formants) attested a close-front non-flat sound i (F1 383 Hz; F2 2318 Hz), near-close near-front non-flat sound a (F1 436; F2 2195), and an open-mid near-front non-flat vowel e (F1 528; F2 2032) as used in K: [fi]kajut ‘they scold’, du[xˈi] ‘spirits, ghosts’, fusɪ[kˈɛ] ‘headscarves’. Most K consonants are palatalized before an etymological non-flat vowel representing a tendency toward intrasyllabic harmony earlier introduced in CS: ro[bˈɛtˈɪ] ‘to do’ (U. ro[btr], CS *orbiti; ESUM 5:98-99) next to [pɛ]-šow ‘he went’ (U. [pʰi]-šow, CS *po-; ESUM 4:465).

The Praat analysis I provide concurs with Timberlake’s (1978) statement that the 4th palatalization of velars occurs as assimilation of high tonality of velars and vowels; in K this refers to flatness (rounding). Conceivably, Flier’s (2007) statement that a flattening system (round/unround) of SWU shows a switch from a gravity system (front/back) before the shift of jers is not therefore confirmed for K. Based on a tendency toward intrasyllabic harmony, the 4th palatalization of velars in K follows a basic pattern of the palatalizations of velars.

Acknowledgments: This research was funded by the Yukhymenko Family Endowment of the International Renaissance Foundation of the National University of Kyiv-Mohyla Academy.

References
Anti-cataphora effects, agree and phases

Paulina Łęska & Jacek Witkoś
(Adam Mickiewicz University, Poznan)

This study concerns cataphoric relations in English and Slavic (Polish). Backward pronominalisation (BP) is banned when pronouns c-command names but in English it is allowed when pronouns are within DPs that c-command names, as (3).

(1) *Hei believes that Johni is clever.

(2) *Oni wie, że Mariak mówi, że Jani jest zdolny.
   hei knows that Maria says that Johni is gifted

(3) [DP hisi latest movie] really disappointed Kusturicai.

In contrast, Slavic (including Polish) shows Anti-Cataphora Effects (ACE), stronger than BP, where c-command by the pronoun is less obvious (Despić 2013, 2015):

(4) *[njegovi najnoviji film] je zaista razočarao Kusturicai. (BCMS) his latest movie
   is really disappointed Kusturica

We address the following questions: (A)Does Condition C subsume the ACE? (B)Is it licensed by Agree or Move? [Ad. A]Willim (1989) and Witkoś (2008), express doubts as to whether the ACE is reducible to Condition C, as it depends on the depth of embedding and intervention (5-6), unlike (2):

(5) (?,?,?,?,?)[przyjazd [jego1 rodziców]] zupełnie Janka1 zaskoczył.
    his1 parents’ arrival surprised John1 completely

(6) ?[jej2 historie [o nim1]] wzburzyły Piotra1.

   her storiesNOM about himLOC exasperated PeterACC
Despić (2013, 2015) argues that the unacceptable status of (4) stems from Condition C. He claims that in contrast to English, in Slavic NP-lgs possessives are placed in an adjunct position from which they command outside NP in (8).

1. **English, Slavic DP-lgs:** [DP his [D’ 0 [NP movieN]]]

2. **Slavic NP-lgs:** [NP his [NP movieN]]

A different analysis appears in LaTerza (2016) and Franks (forthcoming) who argue for LF-raising of the possessives to account for (4) in Bugarian. We develop alternative proposal involving notion of ‘phase command’ defined in Bruening (2014):

3. X phase-commands Y iff (i) there is no ZP, ZP a phasal node, such that ZP dominates X but it does not dominate Y. Phasal nodes: CP, vP, NP (but not PP); (ii) X precedes Y.

4. **Condition C** (Minimise Restrictors): A definite description of the form the A (name) may not refer to a discourse referent in active set C if A could be dropped (become pronoun) without affecting either (i) the denotation of the description or (ii) its various pragmatic effects.

Phase command and Condition C in (10) account for core cases of the ACE in (4): possessives are adjuncts (unlike in English) and on definition (9) NP is not the first phase node dominating them; only CP is.

Furthermore, Bošković (2007) shows that phases (PIC) constrain Move but they need not constrain Agree (LD Agree in Chukchee). We combine (9-10) with the notion of PIC resistant Agree (both on the upward and downward leg of command). This move delivers several immediate bonuses. The PP in (6) is not treated as a phase and the pronoun commands from within it. Since intervention effects hold of Agree, (5-6) are accounted for. Finally, under PIC resistant Agree we can explain why Slavic DP-languages and Slavic NP-languages look alike for the Agree-based ACE while differ w.r.t. Left Branch Extraction.

**References:**


Diana Lewis

<not updated>
The choice of negation pattern in Võro and Seto

Liina Lindström, Maarja-Liisa Pilvik & Helen Plado
(University of Tartu; University of Tartu; & University of Helsinki, University of Tartu and Võro Institute)

Keywords: standard negation, word order, language variation, Uralic languages, South Estonian Seto

There is a typological tendency for standard negation markers to be preverbal, i.e. to occur before the main verb (Dahl 2010, Dryer 2013). Most of the Uralic languages also use preverbal negation, but several Finnic languages (e.g. Lude, Veps, Karelian, and Estonian Insular dialects) sporadically make use of postverbal negation as well. In this paper, we focus on South Estonian dialects Võro and Seto, where both pre- and postverbal negation (examples 1 and 2, resp.) as well as double negation (3) occur.

(1) ei vōiq jättäq vällä
NEG can leave:INF out
‘(one) can not leave (it) outside’

(2) ku marju saa as, syss...
when berry.PL..PRT get NEG then
‘when (we) did not get any berries, then…’

(3) inne ei tulõ vällä eiq ku keskpäivä
before NEG come out NEG than midday.PRT
‘(S)he doesn’t come out before midday’

Traditionally, Võro has been seen as a dialect of Estonian, and Seto has been treated as the easternmost subdialect of Võro. Although Seto and Võro, for the most part, coincide with regard to their phonology, morphology, and lexicon, there are crucial differences in syntax (e.g. Lindström & Uiboaed 2017, Lindström et al. 2019). Also in the case of negation, there is areal variation in negation patterns: postverbal negation is mainly used in Seto, but it is also common in Central and Eastern Võro (Iva 2007). Additionally, Lindström et al. (to appear) have demonstrated that postverbal negation is the dominant pattern in Seto, while in Võro, preverbal negation prevails. However, there is no comprehensive usage-based study on the distribution of the three above-mentioned negation types in Võro nor on the factors that trigger the choice of negation patterns.

In this study we look at the use and spread of preverbal, postverbal, and double negation in Võro and Seto dialects using the data in the Corpus of Estonian Dialects (http://www.murre.ut.ee/mkweb/). By applying quantitative variationist methods, we look for the variables that affect the choice between pre- and postverbal negation. The main research questions are: 1) Is there areal variation in using negation types as claimed earlier? 2) Does the choice of negation type depend on linguistic or extralinguistic factors?

Preliminary results demonstrate that areal differences do exist – postverbal negation is the default pattern in Seto and preverbal negation is preferred in Võro. The use of negation types in both varieties is primarily conditioned by strong persistence effects, i.e. tendencies to repeat linguistic patterns (Szmrecsanyi 2005) and the individual speaker rather than linguistic factors such as tense or mood. This implies that the negation system of Võro and Seto is predominantly affected by language contacts both in the societal and individual level rather than specific functional settings.

References
Lindström, Liina, Maarja-Liisa Pilvik, and Helen Plado (to appear), Variation in negation in Seto, in Olga Krasnoukhova, Mily Crevels, and Johan van der Auwera (eds), (to appear), Postverbal negation (Studies in Language).
Szmrecsanyi, Benedikt (2005), Language users as creatures of habit: A corpus-based analysis of persistence in spoken English, Corpus Linguistics and Linguistic Theory 1, 113–149.

The acquisition of VS structures in L2 Italian in a developmental perspective: A corpus-based study

Andrea Listanti, Jacopo Torregrossa & Liana Tronci
(University for Foreigners of Siena, University of Frankfurt, University for Foreigners of Siena)

Keywords: Second Language Acquisition, L2 Italian, postverbal subjects, syntax-discourse interface, corpus linguistics

The distribution of post-verbal subjects in Italian is governed by lexicon-syntax (with unaccusative verbs; Belletti 1988) and syntax-discourse (information-focus on the subject; Belletti 2004) interface conditions. In L2, phenomena related to the former interface are more easily attained (Sorace 2011). Accordingly, the performance of advanced L2-learners of Italian with VS differs depending on whether the subject follows an unaccusative verb or is focus-marked (Belletti et al. 2007; Lozano 2006 on L2-Spanish). However, it remains unexplored how VS emerges at lower levels of proficiency and whether the divide between lexicon-syntax and syntax-discourse interface phenomena is reflected at the developmental level (cf. Bettoni et al. 2009, Nuzzo 2015 for a similar perspective).

We conducted a study on the LIPS Corpus, that collects transcriptions of language assessment exams of learners with varied L1 and proficiency (A1-C2) (Vedovelli 2006). After dividing the transcriptions in units and isolating VS-structures (Table 1), we analysed them following a multi-layered annotation of verbs and subjects. We considered verb-type (unaccusatives, “piacere (like)-type” verbs with Theme and Experiencer DAT, unergatives and transitives) and semantic (agentivity) and information-structure features of the subject. To avoid arbitrariness in coding topic and focus, we relied on the inventory of topic- and focus-associated functions reported in Riester-Baumann 2013. The former includes givenness, while the latter contrast with a set of alternatives and occurrence with focus-operators.

As for the verb-type, Figure 1 reveals a decrease in the use of VS with V=“piacere-type” from B2 onwards and an increase in the use of VS with other types of unaccusatives from B1 to B2. There is also a steady increase in the number of VS with V=transitive. The analysis of the information structure
shows that the higher the proficiency, the greater is the tendency for postverbal subjects to be associated with focus (Figure 2, middle line), with the difference between B2 and C1 reaching significance ($\chi(1)=7.29, p=.005$). Most of these subjects are contextually given (top line), which is observed across all levels (with the lowest percentage – 55% - at B1). Finally, postverbal subjects are increasingly associated with agentivity across levels A1-C1, with a significant increase between B1 and B2 ($\chi(1)=5.31, p=.02$).

Overall, our results comply with the hypothesis of a developmental trajectory from lexicon-syntax to syntax-discourse interface phenomena. The semantic analysis suggests that post-verbal constituents are progressively assigned subject-status, being more and more associated with agentivity. This developmental pattern is also confirmed by the error analysis concerning verb-subject agreement (20% of errors in A1&A2, 14% in B1, 6% in C1 and 3% in C2). Likewise, the increasing occurrence of focused subjects (together with the extension of VS to unaccusatives, unergatives and transitives) denotes learners’ developing sensitivity to information-structure (with the preferred interpretation being contrastive focus). When looking at lower proficiency levels, the data seem to challenge the idea that “piacere-type” verbs emerge relatively late because of non-canonical alignment between thematic roles and grammatical functions (Bettoni et al. 2009). However, the observation that in these structures, postverbal constituents are not necessarily assigned subject-status (see data on subject-verb agreement errors) suggests that piacere-verbs function as “pivots” for a syntactic schema that is later extended to other verb-types.

<table>
<thead>
<tr>
<th>N. exams</th>
<th>A1</th>
<th>A2</th>
<th>B1</th>
<th>B2</th>
<th>C1</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>N. units (occurrence of a finite verb)</td>
<td>252</td>
<td>569</td>
<td>907</td>
<td>2387</td>
<td>2255</td>
<td>823</td>
</tr>
<tr>
<td>N. VS</td>
<td>26 (10,3%)</td>
<td>40 (7%)</td>
<td>61 (6,7%)</td>
<td>152 (6,4%)</td>
<td>150 (6,7%)</td>
<td>37 (4,5%)</td>
</tr>
</tbody>
</table>

**Table 1**: Total number of units and VS structures (per level)

**Figure 1**: Percentage of different types of verbs (unaccusatives, unaccusatives of “piacere-type”, unergatives and transitives) in VS-structures across proficiency levels.
Figure 2: Percentage of occurrence of the features “agentive” (first line from bottom to top), “focus” and “given” as associated with the subject phrase in VS-structures across proficiency levels.

References
Belletti, Adriana (1988), The Case of Unaccusatives, Linguistic Inquiry 19(1), 1-34.
Sorace, Antonella (2011), Pinning down the concept of interface in bilingualism, Linguistic Approaches to Bilingualism 1(1), 1-33.
A semantic map of adverbial conjunctions in Classical Hebrew

Christian Locatell
(Hebrew University of Jerusalem)

Keywords: Classical Hebrew, adverbial conjunctions, polysemy, semantic maps, grammaticalization

Adverbial conjunctions like “when, as, if, because,” are often multifunctional. This can be seen in the following examples of the English adverbial conjunctions “as” and “while.”

1.a “Do as I say, not as I do.” (comparative)
1.b “As I was walking, I found some money.” (temporal)
1.c “Can you clean out the refrigerator, as it is growing mold.” (causal)

2.a “You should sleep while I drive.” (temporal)
2.b “While I don’t like James Bond, do you want to go to the 007 movie?” (concessive)

This raises the question whether the polyfunctionality of adverbial conjunctions is structured and if so, how. Quantitative typological data on this word class has only become available relatively recently (esp. Kortmann 1997, 1998, 2001) and mostly for Indo-European languages.

The purpose of this paper is to address this lacuna by proposing for the first time a semantic map of multifunctional adverbial conjunctions in Classical Hebrew. Descriptive grammars typically provide a taxonomy of uses for each adverbial conjunction. Some studies deny any relationship between functions (e.g. Levinson and Zahn 2002:296) or deny the existence of different functions altogether (e.g. Follingstad 2001).

I heuristically employ Kortmann’s (1997) typological study of adverbial subordinators to organize the usage profiles of adverbial conjunctions from the corpus of the Hebrew Bible. It is shown that adverbial conjunctions in Classical Hebrew fill a contiguous range of semantic relations within the semantic map. However, there is a departure from the Zipfian Inverse Relation Hypothesis “which correlates an increasing degree of polyfunctionality with a decreasing degree of morphological complexity” (Kortmann 1997:113). Some Classical Hebrew adverbial conjunctions have variants differing in terms of morphologically complexity, e.g., ‘axar “after” and ‘axarey “after, because.” Unexpectedly, it is the morphologically more complex counterparts that are more polyfunctional.

I then discuss diachronic implications suggested by the semantic map regarding mechanisms and directions of development (Kortmann 1997:137–211; van der Auwera 2013). In terms of mechanisms, when particular uses adjacent on the semantic map are analyzed, they reveal plausible loci of diachronic extension. For example, I present tokens of the adverbial conjunction kî where it can only be interpreted temporally, where it can only be interpreted causally, and where it admits both interpretations. Unambiguous uses show that this is not simply a matter of vagueness, but of emergent grammar. In addition to inferential pressures (i.e. post hoc ergo propter hoc), I note the bridging contexts that would have facilitated its extension from temporal to causal conjunction, such as verbal form (temporal prefers to govern an imperfective, while causal does not) and position in relation to the main clause (temporal prefers initial position while causal prefers second position). In terms of directionality, besides the typological support for the change TEMPORAL > CAUSAL, I also present comparative Semitic evidence. For example, the longer textual history of Akkadian kî actually shows a proliferation of temporal uses before causal uses.

Thus, this study proposes a first semantic map of multifunctional adverbial conjunctions in Classical Hebrew and demonstrates the value of heuristically employing typological semantic maps.
References:


Incipient grammaticalization in synchronic language use – a study of epistemic adverbials in English

David Lorenz

Multiple nominal construals in Australian languages: Towards a typology

Dana Louagie

Passive predication in Portuguese: Alternation in Construction Grammar

Marcia dos Santos Machado Vieira
(Federal University of Rio de Janeiro)

**Keywords:** Construction Grammar, Portuguese, Variation, Passive predications, Associative links.

The aim of this poster is to provide, within the framework of Usage-Based Construction Grammar (Goldberg, 1995, 2019, and Traugott & Trousdale, 2013), a description of the alternation of some passive predication constructions that involve the uses of the verbs ser (to be) and ter (to have). We are
going to show their associative relations by similarity based on data licensed by such constructions in Portuguese written texts. The passive predication uses can be observed in sentences as:

- O tema teve destaque no periódico. *(the theme had a proeminent attention in the journal)*
- O tema foi destacado no periódico. *(the theme was highlighted in the journal)*

The corresponding active predication can be something like:

- O periódico destacou o assunto *(The journal highlighted the theme).*

In order to develop such comparability, a functional-cognitive perspective is related to a variationist sociolinguistic one and the restriction problem is faced according to a statistical perspective.

This study is related to a larger research about the discursive impersonalization predication constructions in Portuguese: focusing on their formal and functional attributes, as well as their relations of similarity and difference, in order to achieve generalizations about the network of predication constructions that license agentiveless constructs, impersonalized constructs. In order to achieve this goal, we investigate, among other predication constructions, the passive ones. We compare the passive sentences with inductive force participant (agent, force, origin, experiencer) expressed and also the ones without it. We consider collected data from two varieties of Portuguese (the Brazilian and European ones) and two discursive domains (the journalistic and the academic ones). Thus, we aim to present (i) the schematic formal-functional configuration of the passive constructions with support verb *ter* and the ones with the voice auxiliary verb *ser*, (ii) their productivity (token and type frequency) in the language varieties and discursive domains being examined, (iii) the network with the representation of the alternation in study, configuring a constructeme (Perek, 2015, and Cappelle, 2006).

This work is based on results from: a functional analysis of the state-of-affairs conceptualization design, a collexeme and collostructional anlaysis and also a multivariate analysis of the data of the microconstructions in alternation. We deal with the schematicity, compositionality, productivity parameters (Traugott & Trousdale, 2013) and also with contextuality. The intention is to describe, among other observations, that: (i) there are different passive constructions in written Portuguese varieties with different productivity according to some conditioning factors; (ii) some of them are frequently used as resources to conceptualize discursive impersonalized state-of-affairs; (iii) the meaning conceptualization involved in the constructs is determined by the lexemes (lexical constructions) and grammatical constructions which the two verbs in study combine (or not) with; (iv) semantic, discursive, pragmatic and social attributes influence their uses; (v) clues in the context and in the discursive environment can contribute to the access to the inductive force participant.

References

Cappelle, Bert (2006), Particle placement and the case for “allostructions”, in D. Schönefeld (ed), (2009). *Constructions all over: Case studies and theoretical implications* [special issue of Constructions].


DOM in the making: The emergence of Differential Object Marking in Aromanian varieties

Anastasia Makarova & Olivier Winistörfer
(University of Zurich, Switzerland)

Keywords: Language typology, Aromanian, Differential Object Marking, Romance languages, Grammaticalization

In this talk, we aim at describing the particular cases of Differential Object Marking (DOM) (Onea/Hole 2017, Bossong 1991), in several Aromanian varieties studied. While such DOM seems to have been grammaticalised in the Dacoromanian varieties from the 16th century on (over six centuries after the segregation of Dacoromanian and Aromanian) (Ciobanu, 2001), such patterns have been considered absent in all Aromanian varieties (Kramer 1991, Capidan 1932), or not systematic (Manzini/Savoia 2018, Markovik 2007). In the context of our recent fieldwork with questionnaires as well as semi-directed interviews in ethnically mixed villages in the districts of Struga, Resen, and Bitola in the Republic of North Macedonia, though, we observed patterns that indicate the existence and widespread usage of a DOM in the speech of native speakers of Aromanian from the villages Dolna Belica, Gorna Belica, and Malovishte. Despite many parallelisms, however, the behaviour of DOM does not seem to be homogenous in these varieties. Our informants from Gorna Belica categorically used the preposition *pi* (< Lat. *PER* ‘for, towards’) with pronouns referring to human entities, while systematically using the pronoun to express specificity with common as well as proper nouns designating beings, but not with nouns designating non-animates:

*I see the girl (one of many).’* (Aromanian from Gorna Belica, RaSt)

Like in Gorna Belica, our informants from Dolna Belica categorically used the preposition *pi* with pronouns referring to human entities, too, while also making use of the preposition to express specificity with both, animate and non-animate entities:

‘*I see this tree.’* (Aromanian from Dolna Belica, WiCa)

These examples suggest that in the Aromanian varieties from Dolna Belica and Gorna Belica, the Act-of-Speech-Participant Scale, Reference Scale, Animacy Scale, and Topicalisation Scale (Arkadiev, 2009) govern the distribution of DOM. Interestingly, we could not observe such patterns of Differential Object Marking in the adjacent varieties of Aromanian in Trnovo, Resen, and Nižepole. Thus, in our talk, we want to provide a thorough description of the patterns of Differential Object Marking in the Aromanian varieties of Gorna Belica and Dolna Belica, in order to shed light on possible processes resulting in the emergence of such a system in Aromanian and to propose typological explanations for the parallelisms not only within Aromanian, but with Dacoromanian varieties.

References
Arkadiev, P. M. (2009). Differential argument marking in two-term case systems and its implica-
tions for the general theory of case marking. In Differential subject marking (pp. 151–171). Dordrecht: Springer.


Ambitransitivity as an areal feature in the Balkans

Maxim Makartsev & Max Wahlström

(University of Helsinki & Institute for Slavic Studies, Carl von Ossietzky University of Oldenburg; Institute for Slavic Studies, Russian Academy of Sciences)

Keywords: ambitransitivity, Balkan linguistics, Macedonian dialectology, causativity, language contacts

This paper seeks to evaluate the reported tendency towards labile or ambitransitive verbs in the Balkan languages, thus far discussed only regarding standardized varieties (see, e.g., Aronson 2007: 18–20). Slavic languages typically decausativize transitive verbs through reflexivization, but the distinction between transitives and intransitives is maintained also through, for instance, suppletive transitive counterparts for intransitive roots. Both of these distinctions are prone to be lost in the Balkan Slavic languages vis-à-vis non-Balkan Slavic, consider Standard Macedonian examples 1 through 4:

1) Policija-ta ja zapre kola-ta.
   police-DEF ACC.F.SG stop.AOR.3SG car.F.SG-DEF
   ‘The police made the car stop.’

2) Kola-ta zapre.
   car-DEF stop.AOR.3SG
   ‘The car stopped.’

3) Lira-ta padna.
   lira-DEF fall.AOR.3SG
   ‘The Turkish lira fell.’

4) Toa ja padna lira-ta.
   this ACC.F.SG fall.AOR.3SG lira.F.SG-DEF
   ‘This made the Turkish lira fall.’
In Russian, the transitive verb in example 1, ‘to make stop’, *ostanovit*, would be decausitivized with a reflexive marker: *ostanovit*-sja. In examples (3 and 4, either the transitive verb could be reflexivized (cf. examples 5a vs. 5c) or a non-cognate verb would be used (5b).

(5a) *Lira obvali-l-a-s’.*
lira.NOM.SG make.fall-PST-F-REFL
‘The Turkish lira fell.’

(5b) *Lira upa-l-a.*
lira.NOM.SG fall-PST-F
‘The Turkish lira fell.’

(5c) *Èto obvali-l-o liru.*
this make.fall-PST-Nlira.ACC.SG
‘This made the Turkish lira fall.’

It has been shown that Bulgarian and Macedonian stand out from the other Slavic languages regarding their transitivization and detransitivization strategies (Grünthal & Nichols 2016), yet it has been argued that in comparison with Bulgarian, Macedonian would display even more ambitransitive verbs (Aronson 2007: 19). To assess these claims and to evaluate the areal distribution of the phenomenon, we sampled several Slavic and non-Slavic Balkan varieties, both standard and non-standard. We performed a survey based on a questionnaire, introduced by Nichols, Peterson, and Barnes (2004), that includes 18 verbal triads (non-causal continuous, bounded, and causal). Data was gathered from several standard languages (Albanian, Bulgarian, Croatian, Macedonian, Greek, Serbian) and local dialects in the Central Balkans (North-East Gheg Albanian, Tosk Albanian, several Macedonian and Aromanian dialects).

We demonstrate that, moving from the north to the south, the western Macedonian dialects display significantly more ambitransitive verbs. The same geographical distribution is paralleled in the Albanian dialects. It will be argued that this increase in ambitransitivity is an areal Central Balkan feature, which calls to evaluate it together with similarly patterning phenomena that have been attributed to Aromanian influence. Further, with the help of the non-standard data, we will discuss the type of P-lability involved (for terminology, see, e.g., Creissels 2014), and address the question of the directionality of change in this contact-induced phenomenon.

References

Paradigmatic uniformity under the pressure of intensive language contact
Michalis Marinis  
(University of Patras)

In Italiot’s nominal inflection—Italiot is the Greek system spoken in South Italy in two linguistic enclaves, in Puglia (Griko) and Calabria (Greko). For a more detailed description, see, for example, Ledgeway (2013)—all masculine nouns are today inflected according to a unified paradigm for all the cells of the plural number (table 1: IC1). In contrast, previously, the majority of this group of nouns formatted their plural paradigm according to two distinct inflectional classes (table 1: IC1 & IC2, Marinis 2019).

<table>
<thead>
<tr>
<th></th>
<th>ICi</th>
<th>ICii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular</td>
<td>-o</td>
<td>-α/ι</td>
</tr>
<tr>
<td>NOM.PL.</td>
<td>-i</td>
<td>-ες</td>
</tr>
<tr>
<td>ACC.PL.</td>
<td>-u</td>
<td></td>
</tr>
<tr>
<td>GEN.PL.</td>
<td>-o</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The two distinct inflectional classes (IC) for the plural paradigm of masculine nouns in Italiot.

This innovation could be treated as a case of direct borrowing of an inflectional suffix -i from the romance systems with which Italiot has been in a long term and intensive contact (Thomason 2001). According to this hypothesis, Italiot borrowed the inflectional suffix -i from romance varieties in contact (Salentino, Bovese, Calabrese, Standard Italian), where -i is the inflectional suffix combined with the most productive inflectional classes of masculine nouns.

Contrary to this hypothesis, in this paper, (a) I investigate the mechanisms (analogy, inter- and intra-paradigmatic levelling, reanalysis) through which the plural sub-paradigm of masculine nouns has been changed in Italiot nominal inflection. (b) I examine the nature of the causes leading to language change, showing that it has been achieved by both intra-linguistic and inter-linguistic factors.

In the former, I classify factors such as the tendency of the system to evolve to a state where each form corresponds to only one meaning/function (for example, Hock 1991, Hock & Joseph 1996). In the latter, I classify factors such as the intensive language contact of Italiot with romance systems.

I show that the leading cause of the language change is intra-linguistic, which is also evidenced by the presence of the phenomenon in other Greek dialects, also in those being in long-term isolation (for example, Tsakonian). At the same time, Italiot’s intense contact with Romance contributed to the prevalence of the phenomenon in the Italiot but not in the rest of Modern Greek dialects where the phenomenon is also scarcely present.

The data of this research comes from both the available written sources (for example, Rohlfs 1977, Katsoyannou 1995, Karanastasis 1997) and the 60-hours oral data available in the Gree.D database of the Laboratory of Modern Greek Dialects of the University of Patras. They have been complemented by data I collected personally during two fieldworks in the Greek-speaking areas of Southern Italy, the summer of 2016, and 2018.

References


---

**Integrating grammatical description, text collection and dictionary:**

**Language documentation and description for the digital age**

Florian Matter
(University of Bern)

Keywords: language documentation, digital linguistics, grammatical description, corpus linguistics, computer linguistics

The primary goal of GRAMR is to enable linguists to create interlinked digital representations of the three components of the “Boasian trilogy” – grammar, texts, and dictionary. It also incorporates elements of what has been called the “Himmelmannian” trilogy (Camp et al. 2018) – recordings, morphosyntactic annotation, and metadata. This is achieved with a web app written in the CLLD framework (Forkel et al. 2019), which allows for fast and user-friendly representation of various kinds of linguistic data. The application aims to close the gaps between linguistic documentation and description – the two processes are typically closely interwoven, but their respective end products often are not.

At the core of GRAMR is a collection of texts, which are sequenced into sentences and brought into the traditional interlinear example format, with a morpheme-by-morpheme glossing and a free translation, supplemented with audio. The second component is the grammatical description, which can contain interlinear examples taken directly from the text collection. This means that 1. examples can be looked up in the context in which they appear, 2. the audio can be played directly in the grammar, 3. segmentations and morpho-logical annotations are always up to date. The morphemes in the object line of interlinear examples are all clickable, leading to the respective dictionary entry; links to individual morphemes can also be inserted into the grammatical description. The dictionary is at the moment the least feature-rich component of gramr, mainly serving as a link between texts and the grammar. There is a dictionary entry for every morpheme, and entries display all examples in which the morpheme occurs, again with links to the corpus. Entries for grammatical morphemes discussed in the grammatical description can also have back references to the relevant sections.
Figure 1: Links between the three components of gramr; the dashed line represents op-tional, non-automatic links.

As of now, data entry is organized around two staple tools of field linguists: ELAN (The Language Archive 2018) and FLEX (Summer Institute of Linguistics 2019). Texts are transcribed and translated in ELAN, then exported to FLEX, where they are parsed into morphemes and annotated. The text collection is exported from FLEX, including ELAN annotation time and speaker information, which enables automatic extraction of audio snippets and linking examples to speakers. From the FLEX export, CLDF (Forkel et al. 2017) representations of the text collection and the dictionary are created. Speaker and text metadata are provided via CSV table files. The input for the grammatical description is provided as a collection of plain text (markdown) files, along with a CSV file detailing the structure and titles of sections. Along with traditional markdown elements like tables, all text in the app can contain GRAMR-specific markdown; for example ex:X-2 will render the example with the ID X-2; morph:X will provide a link to the morpheme with the ID X. This is all the linguistic input needed for creating the data-rich, holistic representation of linguistic data and analysis that is GRAMR. Online demo: http://gramr-demo.herokuapp.com/.

References

Grounding strategies in Lakurumau narratives
Lidia Federica Mazzitelli
(University of Cologne)
Keywords: discourse structuring; Oceanic; perfect; irrealis; iamitives

In this talk, I describe the use of the particle asang ‘new situation’ and the TAM marker daa ‘irrealis’ as grounding strategies in narrative texts in Lakurumau, a previously undescribed Oceanic language of New Ireland (Papua New Guinea). All data come from my own fieldwork (2017-2019).

Following Hopper (1979), I refer to the events that bring forward the actual story line as ‘foreground’ and to the events that support the foreground and do not themselves narrate the story as ‘background’. In Lakurumau, both can be in the unmarked realis, which can express both past and present reference and is the default choice for narratives. However, foreground events are frequently marked by the particle asang; background events may be marked by the irrealis daa. Asang-marked predicates bring forward the storyline while daa-marked predicates expand on it, encoding events seen as habitual or persistent and referring deictically to the present of the narration (the ‘speaker-now’; Fleischman 1972). In (1) the background (in square brackets) provides information about the magician-warrior Malaxon and his magic plant (a Cordyline australis).

(1) Orait nam Malaxon asang ka ramazik, di ramin asang a yaan aa xa xus a mu piu zana. Ka gonin asang naadi wana mu mos pena waan lo vezop. [Aa naan, nam Malaxon akam, a vuna roi moxom a waxaana si. Taim... kain taatai uruxaam... taim gu daa rapin maan a ron sim, ka daa tef tepin kam pana si.]

‘All right, Malaxon got up ASANG, they set ASANG a date and he told it to his warriors. He prepared ASANG them to go to the battle. [And him, that Malaxon, he used to hold a leaf of Cordyline australis. When... this kind of men... when you will throw your spear, he will fan it away with the Cordyline australis.]

I argue that the discourse functions of daa and asang can be directly derived from their semantics. The basic function of daa is to encode typical irrealis functions, including habituality and future tense reference: the latter is re-analysed as making reference to what the speaker is about to say later in the text. Asang expresses meanings typically related with already-, iamitive and perfect grams, such as an earliness implication, change-of-state with stative predicates and a ‘new situation’ with dynamic predicates (Krajinović 2019). Even though its functions resemble those of perfects, asang does not appear in the slot reserved for TAM markers, can co-occur with aspectual markers such as nga ‘progressive’ and it can modify nouns and adverbs, marking contrastive focus and unexpectedness. I analyse it therefore as a ‘new situation’ and ‘contrast’ marker; whether it is a iamitive (Olsson 2013) is an open question. Its function as a foregrounding device is consistent with this semantics: asang-marked events are seen as contrasting with the previous ones and thereby inherently marking a development in the story line. This function of perfect(-related) grams is found in other Oceanic languages, such as Tokelauan, where kua ‘perfect; new situation’ is also analysed as a foregrounding device (Hooper 1998).

References
SLE 2020 Book of Abstracts


Socio-stylistic aspects of diachronic variation: the case of Spanish asyndetic complementation between the 15th and the 18th century

Giulia Mazzola, Bert Cornillie & Malte Rosemeyer

<not updated>

Interrogative constructions in creoles and sign languages

Susanne Maria Michaelis
(University of Leipzig; MPI-SHH, Jena)

Keywords: Creole languages, sign languages, interrogative construction, typology

In this talk, I will explore potential commonalities in creoles and sign languages with regard to interrogative constructions. Several linguists have claimed that similar sociohistorical settings and cognitive constraints in creolization and sign language evolution have given rise to similar linguistic structures (e.g. Fischer 1978, 2015, Woodward 1978, Adone 2012). However, to be able to demonstrate significant commonalities between creoles and sign languages, it is not enough to only investigate features which are widespread in both types of languages: At the same time the features should be rare in non-creoles/non-sign languages world-wide.

Interrogative constructions are one of the best researched topics in sign languages (e.g. Morgan 2006; Šarac Kuhn & Wilbur 2006; Zeshan 2004, 2013; Cechetto 2012; Quer et al. 2017). My information on creoles comes from the *Atlas of pidgin and creole language structures* (APiCS, Michaelis et al. 2013), which covers 59 creoles. I will examine (i) polar (yes/no) questions and (ii) content (wh-) questions.

In polar questions, creoles and sign languages predominantly show an apparent commonality in that a polar question differs from a declarative sentence only by intonation (various non-manual markers, such as eyebrow raise or head forward position (Zeshan 2004: 19) being viewed as suprasegmental entities correlating to intonation in spoken languages; Sandler 1999). Jamaican, for example, has *Im av kyaar*? (rising intonation), lit. ‘He has car?’

At the same time, the intonation-only pattern seems to be rarer cross-linguistically (Dryer 2013). This would speak in favor of a shared preference for creoles and sign languages to feature intonation-only in polar questions. However, a closer reading of Dryer’s WALS chapter highlights the difficulty of qualitative comparison of data from different typological datasets: Given Dryer’s classification, WALS apparently underrepresents languages that mark polar questions intonationally, since minor strategies (e.g. question particles) in a given language are given priority over a co-existing intonation strategy in the overall typology. Therefore, interrogative intonation could be much more widespread world-wide and hence the status of interrogative intonation as a significant shared creole/ sign language feature remains to be proven by better matching typological datasets.

The picture in content questions is much clearer. Creoles and sign languages have different structures. In APiCS, two feature values are distinguished: fronted wh-phrases (‘What did you do?’)
and non-fronted wh-phrases (in situ lit. ‘You did what today?’ or some other non-initial position; Haspelmath 2013). Creoles show both patterns, with a strong preference for fronted wh-phrases. However, sign languages show more diverse patterns in the placement of the wh-phrase. Besides the possibility of fronting, many sign languages allow for a sentence-final position of the wh-phrase (CAKE EAT NOT WHO ‘Who did not eat the cake?’, Italian Sign Language (LIS), Cecchetto 2012). Interestingly, this rightward placement is unattested in creoles and is generally extremely rare in spoken languages (cf. Dryer 2013), and therefore seems to be a specific strategy of sign languages.

In conclusion: interrogative constructions do not lend support to the claim that creoles and sign languages share unusual commonalities.

References
Haspelmath, Martin (2013), Position of interrogative phrases in content questions, in Michaelis et al. (2013), 44-47.
Michaelis, Susanne Maria, Maurer, Philippe, Haspelmath, Martin and Huber, Magnus (eds) (2013), The atlas of pidgin and creole languages structures, Oxford: Oxford University Press.
Certainty, possibility, probability: Epistemic and strategic uses of Galician adverbs

Vítor Míguez
(Universidade de Santiago de Compostela & Instituto da Lingua Galega)

Keywords: adverbs, epistemic modality, Galician, mitigation, strengthening

In conversation analysis and similar frameworks, epistemic markers are studied in actual contexts of use, paying special attention to their interactional dimension (e.g. Kärkkäinen 2003). These works tend to assume that using expressions of (un)certainty either attenuates or reinforces the utterance, in other words, that epistemic expressions readily entail the presence of a mitigating or a strengthening discourse strategy. This paper claims that such standpoint is not adequate as it obscures the notions of epistemic modality and discourse strategy. Thus, its aim is to disentangle both concepts and show how they relate to each other on the basis of corpus data.

Epistemic modality is a functional domain concerned with the linguistic expression of likelihood, in particular, it is a qualification of a state of affairs as possible or more or less probable, as made by the speaker (cf. Nuyts 2001). Mitigation and strengthening are discourse strategies (hence linguistic phenomena) at the service of (im)politeness (a social phenomenon) used by speakers to decrease or increase illocutionary force, respectively.

Empirical studies of epistemic and evidential markers show the benefits of distinguishing strategic from non-strategic uses: mitigating uses of Spanish evidential expressions are rare and highly dependent on discourse genres (Estellés Arguedas & Albelda Marco 2017), whereas English certainly features strengthening as the most frequent use and epistemic modality as the most infrequent one (Byloo, Kastein, & Nuyts 2007).

This contribution studies the ‘post-manner’ uses of six Galician epistemic adverbials (certamente ‘certainly’, posiblemente ‘possibly’, probablemente ‘probably’, quizais ‘maybe’, se cadra ‘perhaps’, and seguramente ‘surely’) in essayistic, journalistic, and narrative prose. Samples of 100 occurrences of each adverb in each genre were taken from CORGA, the biggest corpus for present-day Galician. The resulting 1800 observations were coded for the relevant semantic-pragmatic categories, namely epistemic modality, mitigation, strengthening and other strategic uses, such as ‘tendentious’ and ‘ironic’ ones. The results reveal a clear difference between certainty, possibility, and probability expressions as regards strategic uses: certamente is primarily a strengthener and, to a lesser extent, an epistemic certainty marker; quizais and se cadra are markers of epistemic possibility and feature strategic uses, including mitigation, of little quantitative significance; posiblemente, probablemente and seguramente are close to being pure expressions of epistemic probability.

These results lead us to conclude that there exists an asymmetric relation between the epistemic value encoded by an expression and the discourse strategic goals it serves, insofar as a certainty marker is more likely to be used strategically than a possibility or, even more clearly, a probability marker. The basis of this divergence lies in the discursive status of (un)certainty: the default expression for certainty in language is an unqualified assertion, whereas the qualification of a state of affairs as uncertain requires a specific linguistic mark. Expressing epistemic certainty is, thus, anomalous by virtue of the maxim of quantity and poses a risk of threatening the face of interlocutors by asserting the cognitive superiority of the speaker. These issues are solved by using certainty markers as strengtheners, thus allowing for a range of pragmatically suitable strategies, such as showing agreement with interlocutors.

Acknowledgments: This work has received financial support from the European Union (European
Social Fund – ESF) and the Xunta de Galicia (grant numbers ED481A-2017/23 and ED431C 2017/34).

References


Temporal bias: a new type of bias for typologists to worry about

Steven Moran & Eitan Grossman

<not updated>

The speakers of minority languages are more multilingual

George Moroz and Nina Dobrushina
(HSE University, Moscow & HSE University, Moscow)

Keywords: multilingualism, language population size, Dagestan, L2, language evolution

Population size is often discussed as a factor which might have influenced patterns of language and cultural evolution (Wichmann & Holman 2009, Bowern 2010; Donohue & Nichols 2011; Nettle 2012; Bromham et al. 2015; Greenhill et al. 2018; Koplenig 2019; see Greenhill 2014 for an overview). One such factor is long-standing multilingualism. In this paper we advance the hypothesis that the number of different L2s mastered by speakers shows a negative correlation with the population size of their L1: the bigger the L1 population, the smaller the number of L2s mastered by its speakers.

The correlation between the size of language population and the level of multilingualism of its speakers is tested statistically on a large body of empirical data from Dagestan, Russia’s hotbed of language density and multilingualism. Linguists rarely have reliable sources on language population, especially for the periods when languages were vital (Greenhill 2014: 563). For Dagestan, the situation is better. Due to the digitalization of the Census of 1926, we have at our disposal reliable information about the population of Dagestanian villages before the urbanization and the rise of population caused by the First Demographic Transition. In combination with the fact that Dagestanian villages were almost exclusively monoethnic and monolingual, these data allow us to make credible estimations of the number of speakers of all languages and dialects of Dagestan before the changes of 20th century.
We can also cautiously assume that the size of village tended to be stable over time, because villages relied upon natural economy which, in its turn, heavily depended on the geographical location and thus could not change considerably.

The second type of data comes from a field study of Dagestani multilingualism (https://multidagestan.com/). Language repertoires of 4,032 people were collected and coded by using the method of retrospective family interviews during field research in 2011 - 2019. The data on multilingual repertoires covers only a small part of the villages for which we have population sizes, namely 54 villages speaking 29 different lects.

We then match population sizes of these 29 languages and the number of L2s spoken by the speakers of these languages from 54 villages, and run a Poisson mixed effects regression model that predicts the average number of second languages spoken by speakers from L1 communities of different size. The study confirms the hypothesis that the size of language population is negatively correlated with the multilingualism of the language community.

Extrapolation of our data suggests that differences in the numbers of L2’s are lost in population sizes higher than 400,000, the range other studies of linguistic correlates of population size usually deal with (see references above). The takeaway of this study is that small languages may teach statistical lessons that major languages cannot, and more attention from sociolinguistic typology is needed in the areas of high language density.

References
Koplenig Alexander. (2019). Language structure is influenced by the number of speakers but seemingly not by the proportion of non-native speakers. R. Soc. open sci. 6: 181274. http://dx.doi.org/10.1098/rsos.181274

“We were told women should not speak like that”: the fade-out of the Basque informal form of address hika among women

"
Garbiñe Bereziartua & Beñat Muguruza
(University of the Basque Country, UPV/EHU)

Keywords: forms of address, hika, gender, transmission, language prestige.

The social use of the Basque informal form of address hika has decreased in the last decades, but this decline has been much more dramatic among women, up to the point of becoming marginal. Verbal forms vary depending on the addressee’s gender, which means that the use of female forms of hika (i.e., noka) is on the verge of extinction, and therefore hika is turning into an exclusively masculine feature of the language (Echeverría, 2003). Hika has been and to a certain extent is still considered very informal and even disrespectful in some contexts. Labov (1972) stated that women tend to reject languages and varieties with less prestige, and many other studies have shown a similar pattern in different contexts (e.g., Gal, 1979; MacDonald, 1989).

This study looks at the social use of hika among different generations in the town of Azpeitia (Basque Country), with about 15,000 inhabitants, one of the strongholds of the Basque informal form of address in the 21st century. It also seeks to explore the specific reasons underlying the increasing gender gap on the social use of hika.

1,247 people answered an online questionnaire that inquired about their own use of hika. We designed two different versions of the questionnaire for two age groups: participants over 18 and secondary education students (age=12-18). Adult participants also responded narrative answers related to the reasons for the lower use of hika among women. We then conducted seven focus group discussions and three deep interviews with experts, where they discussed insights on this issue.

Our preliminary results show that women use the form hika less than men, and women are also addressed less in hika. This phenomenon does not seem to be diminishing in the near future, as our data show a well-defined trend: the younger the participants, the wider this gender gap. As to why this has happened, participants mentioned evident reasons such as lack of habit, lack of knowledge (of noka forms) and lack of transmission. Other reflections try to get to root of the issue, referring to the fact that women are told that it is rude or disrespectful to use hika, that men are afraid of sounding feminine, or that women are to be treated carefully and sweetly.

Our findings are in line with previous studies done concerning both hika in the Basque context: in the end, different factors discourage women from using the informal form of address. However, becoming aware of the factors that affect the choice of given forms of a language can be a first step towards reversing language or variety shift.

References
Echeverría, Begoña (2003), Language ideologies and practices in (en)gendering the Basque nation, Language in Society 32, 383-413.
MacDonald, Maryon (1989), We are not French: Language, culture and identity in Brittany, New York: Routledge.

No elbows on the table, and no colloquialisms either! How Flemish caregivers’ choices in Child-Directed Speech reflect broader linguistic ideologies
Keywords: variationist sociolinguistics, language acquisition, conversation analysis, pragmatics, child-directed speech

When competing language varieties co-occur, each variety’s use differs considerably by domain and social context, generating countless possibilities for sociolinguistic research. An especially interesting new avenue of study is how and why each variety is used in the home, particularly in Child-Directed Speech (CDS). CDS has the more general function of teaching children community norms, but it therefore also provides insight into which language ideologies parents adhere to (Ochs & Schieffelin 1984;7; Foulkes et al. 2005). In adopting this perspective, the aim of this paper is to determine how and to what extent language choices in the Flemish home reflect language ideologies in Flanders.

The language situation in Flanders forms an interesting laboratory as standardization of Dutch happened much later in Flanders than in the Netherlands because of how French dominated Belgian society. Belgian Dutch was only standardized in the twentieth century by adopting the standard from the Netherlands, leading to a period of hyperstandardization (Geeraerts 2001; Van Hoof & Jaspers 2013). Eventually the ideologies which privileged the exogenous norm lessened, as Flemings never found the foreign standard comfortable for everyday use, leading to the development of a local vernacular norm, Colloquial Belgian Dutch (CBD). However, the high prestige attached to the exogenous standard persists today (Grondelaers & Van Hout 2011).

We study this ongoing tension between exogenous standard and endogenous informal norm at the Flemish dinner table. The data consists of dinner conversations between 16 Belgian-Dutch caregivers (from the provinces Antwerp and Flemish Brabant) and their children (aged 2-8), and of metalinguistic interviews with caregivers about language attitudes. The data will be analyzed through quantitative variationist analyses, focusing on four sociolinguistic variables that alternate clearly between their Standard Dutch and CBD forms and reflect various degrees of salience (Lybaert 2014). The features in question are

1. second person singular pronouns gij/jij (you eat is ‘jij eet’ in SD and ‘gij eet’ in CBD)
2. diminutive –kel-je (little book is boekje in SD and boekskje in CBD)
3. word-final t-deletion on function words (that is dat in SD but da(t) in CBD)
4. adnominal flexion on definite and indefinite articles de/den and ne(n)/een (the apple is de appel in SD and den appel in CBD because appel is masculine).

To explain variation between Standard Dutch and CBD in the data, inferential statistical analyses examine how speakers’ and addressees’ ages and genders and the specific local contexts in which the utterances occur (operationalized by three discursive frames: relational/transactional/pedagogical) affect the use of standard and colloquial forms of these four variables. The analyses focus especially on how parent-directed speech differs from CDS.

The results show significant variation in the use of Standard Dutch and CBD within and between families, which reflects variation in language attitudes and ideologies, something that corresponds with the metalinguistic interviews and with previous research about ideologies regarding the use of standard and vernacular forms in CDS, both in Flanders and more broadly.

References

Preserving and Digitally Restoring Forgotten Stories from Romania’s National Phonogram Archives

Oana Niculescu
(Romanian Academy Institute of Linguistics “Iorgu Iordan – Al. Rosetti”, Bucharest)

Keywords: Romania’s cultural heritage, oral archive restoration, digitization

Designed as a “sound museum” (Șuteu 1958), Romania’s National Phonogram Archives (AFLR) showcases the country’s rich oral cultural heritage. Considered the largest and most diversified ethnolinguistic collection (Marin 1996, 2012), AFLR is representative of all Daco-Romanian subdialects. Ranging from the portrayal of various customs, religious practices and traditional beliefs, to short stories, folktales and songs, the Archive is a vast, multifaceted repository of the community’s collective memory and identity. Due to a lack of dialect literature (Vulpe 1978), these past narrations are of inestimable documentary value, serving as a link between sociolinguistics, dialectology, folklore studies and traditional culture. The systematic data collection (5 to 7 speakers from 4 generations) makes the recordings comparable in nature, allowing for an in-depth transdisciplinary analysis.

Through this presentation we aim to raise awareness concerning the protection, preservation and promotion of the stories gathered within the Archive. AFLR is currently kept at the Romanian Academy Institute of Linguistics “Iorgu Iordan – Al. Rosetti”. Either stored in an audio or text format (reel-to-reel tapes, dialectal texts, transcripts and monographs respectively) the material has undergone a limited digital restoration process (less than 300 tapes digitised from more than 1500 remaining). As a result of unsettling lack of equipment (only one tape recorder is still in working condition), most of the digital preservation projects have been orientated towards the audio retrieval of the reel-to-reel tapes Marin & Tiugan 2014, and Marin et al 2016). Nevertheless, there is also an urgent need to digitally restore all metadata files derived from the field recordings. It is this pressing, overlooked matter that our proposal addresses and intends to overcome.

During the dialectal inquiries (1960 – 2015), researchers organised the sociolinguistic information in two file types. The first one, entitled the speaker’s file, contains the date and place of the recording, the (sur)name, and, in some cases, the nickname of the informant, the age, education level, occupation

134
and possible travels outside the village (either work related or military service), details regarding family status, as well as other extralinguistic factors (e.g. recording conditions). Each file is accompanied by the informant’s photograph. The second metadata file consists of conversation topics, pronunciation variants, dialectal terminology, narrow phonetic transcriptions, glossaries or unique accounts delivered by the informants as to the age, origin or history of the region. These hand written documents form over 950 “village notebooks”. These “notebooks” are currently cleaned, properly stored, scanned and processed. The final product will be an open-access interactive platform of flipbooks and individual speaker’s files, optimally arranged for statistic and linguistic analyses. Alongside this unique instrument, we provide innovative services for human resource development within the Institute (such as audio processing and labelling in modern acoustic software, among other), aiming to improve current digital retrieval techniques and stimulate the entire recovery of the Archive.

Acknowledges
Grant de cercetare realizat cu sprijin financiar din Fondul Recurent al Donatorilor, aflat la dispoziția Academiei Române și gestionat prin Fundația „PATRIMONIU” GAR-UM-2019 I-1.5-3; 358c/15.10.2019.

References
Marin, Maria (1996), Arhiva fonografică a limbii române (După 40 de ani), Revista de lingvistică și știință literară, Chișinău, I, 41–46.
Marin, Maria, and Marilena Tiugan (2014), Harta sonoră a graiurilor și dialectelor limbii române, Bucharest: Romanian Academy Press (+ CD).
Marin, Marin, Carmen Ioana Radu, Daniela Răuțu, and Marilena Tiugan (2016) Harta sonoră a graiurilor dacoromâne vorbite în afara granițelor României, Bucharest: Romanian Academy Press.

The diachrony of the Danish IO and its promotion to subject in passive constructions

Peter Juul Nielsen
(University of Southern Denmark)

Keywords: indirect object, ditransitive passive, case, grammatical change, Danish

This paper presents an empirically based analysis of the connections between two diachronic developments in the grammar of Danish: The semantic specialization of the indirect object (IO) from broad Beneficiary meaning in Middle Danish to Recipient meaning only in Modern Danish (Heltoft 2014, cf. Barðdal 2007)
The emergence of the possibility of promotion of the IO of a ditransitive verb to nominative subject in passive constructions

An ongoing empirical study suggests that up until the 16th century, passive constructions did not allow IO promotion to nominative subject, and oblique forms of thematic IOs still dominate in texts of the late 17th century, cf. (1), but at this stage promotion of thematic IOs to nominative subjects begins to occur, cf. (2); both examples are from Køge Huskors (1674).

(1) **hende hav-er hver dag vær-et bud-en baade mad oc øll**

her.OBL have-PRS every day be-PRF.PTCP offer-PRF.PTCP both food and beer

’she has been offered food and beer every day’

(2) **Der **hand** **aff den gemen-e almue i enfoldighed blef**

when he.NOM by the common-DEF people in naivety become.PST

tilspur-t, hvorledis **hand** var ...

ask-PRF.PTCP how he.NOM was …

‘when he was asked by the common people in their naivety how he was …’

In present-day Modern Danish, IO-promotion in passives is fully established, and in a ditransitive passive construction, a thematic IO will occur as a subject in the nominative (as in (2)), cf. (3) from a 1998 newspaper article (KorpusDK).

(3) **hun er også netop blev-et tilbud-t et job som formand**

she.NOM is also just become-PRF.PTCP offer-PRF.PTCP a job as chairman

‘she has also just been offered a job as chairman’

Based on Henning Andersen’s (2008) concepts of reanalysis and regrammation and the theory of Connecting Grammaticalisation (Nørgård-Sørensen et al. 2011), the paper suggests that the emergence of IO promotion and the narrowing of the semantic potential of the IO should be analysed as connected. It is a consequence of the establishment of the IO as a content (logical) subject in a GET relation to the direct object (DO) in a secondary predication, cf. the ‘get’ implication in (4) (Diderichsen 1962: 170, Herslund 1986: 140, Hansen & Heltoft 2011: 131).

(4) **Bo[S] gav Lis[IO] en bog[DO]**

B. give.PST L. a book

‘Bo gave Lis a book’

→ **Lis[S] fik en bog[DO]**

L. get.PST a book

‘Lis got a book’

The pivotal role of the IO as a content subject will be related to the development from Middle Danish to Modern Danish of a fixed subject position before the position for objects and the development of the fixed order IO > DO, the latter reflecting the secondary subject-object relation between IO and DO. The empirical data for the analysis consists of IO constructions from a corpus of texts from various diachronic stages between 1300 and 2000.

References
Backgrounded motion events:  
A crosslinguistic study of associated motion in Mesoamerica

Mads Nielsen & Maria Messerschmidt  
(University of Copenhagen)

Keywords: associated motion, motion events, spatial grammar, linguistic diversity, Mesoamerica

Speakers of English are used to locating events in time by means of tense marking on the verb. Similarly, many languages of the world have grammatical markers that are used to locate events in space. Markers of location and direction are well-known, but not much attention has been given to grammatical markers of motion. An example of such a marker can be seen in (1).

Kaytetye (Koch 1984:27)

(1) *Ate nte athe-yene-ne.*
  shit 2SG.ERG excrete-go.and-IMP

‘You go and shit.’

The suffix -yene ‘go and’ is used to indicate that the event designated by the verb should take place after the addressee has moved away from the speaker. Note that -yene not only adds directional but also translational meaning (cf. Talmy 1975) to the utterance in (1). Elements like -yene that encode translational motion as background to another main event are referred to as associated motion (Koch 1984). Wilkins (1989) has shown that associated motion markers can be used to specify meaning in at least three dimensions: path, temporal relation between movement and main event, and moving participant. For two of these dimensions the following implicational scales have been proposed (cf. Levinson & Wilkins 2006, Guillaume 2016):

- prior motion > concurrent motion > subsequent motion
- movement of subject > movement of object
Both implications are confirmed empirically in the only major crosslinguistic survey of associated motion, namely Guillaume’s (2016) study of 66 Amazonian languages of which 44 were found to have associated motion affixes. Guillaume further observes that the same marker may be used for subject movement in both transitive and intransitive clauses, but that object movement is always specified with a separate marker, even in languages with ergative alignment. One reasonable explanation why associated motion favours this accusative-like pattern may be that most associated motion markers have probably grammaticalized from intransitive motion verbs whose sole, prototypically agentive, argument is more likely to have more things in common with the subject of a transitive clause rather than the object.

Associated motion, however, is not just an areal trait of Amazonia. In our study, we extend the description to Mesoamerica (cf. Campbell et al. 1986) by examining languages from all documented Mesoamerican families (10) distributed across the area. We find associated motion markers in 8 of the languages, all of which confirm the implication scales. Furthermore, we find evidence from two of the languages that the same marker can indeed be used for both subject and object movement. While this is at odds with crosslinguistic preferences, it may be accounted for in terms of the effect that applicative morphology has on alignment patterns.

Associated motion is not just relevant to typology but also to investigations of the relationship between language, cognition and culture (see Boeg Thomsen & Volhardt 2018). With our study, we aim to contribute to the typological description of associated motion and raise awareness, since despite being a fairly widespread phenomenon, this topic is still underdescribed in the literature.

References

Co-occurrence classes of consonants and the modular structure of consonant inventories

Dmitry Nikolaev
(University of Stockholm)
Keywords: phonological typology, consonant inventories, feature-economy principle, co-occurrence classes, dimensionality reduction

Considerable progress has been made in investigating the basic principles underlying the structure of phonological inventories. Different approaches, such as Quantal Theory (Stevens 1989), Geometric-Constraint Theory (Dunbar & Dupoux 2016), and Feature-Economy Principle (Clements 2003), have been proposed to account for the fact that languages tend to prefer some segments over others and to have particular combinations of segments (e.g., if a language has /d/ it is also likely to have /l/). What has been lacking, however, is systematic empirical investigation of particular groups of consonants that tend to appear together as a group, viz. to form stable subinventories. We conduct a statistical co-occurrence clusterisation study of consonant segments based on a sample of 2761 languages and show that although some of the empirical co-occurrence classes can be explained using previously proposed frameworks, others are idiosyncratic. We conclude that in order to better explain the existing typology of consonant inventories, theories need to account for their modular nature and provide motivation for the structure of major recurrent subinventories.

References

Complementizer selection and its variation in L2 Greek

Elissavet Nouchoutidou
(University of Athens)

Keywords: complementizers, matrix verbs, psych verbs, second language, Greek

The aim of this paper is to examine complementizer selection in Greek L2, focusing on psych verbs and the use of complement sentences to express the cause of a feeling. Complementizer selection can be challenging for L2 learners according to previous research (Theophanopoulou-Kontou et al. 1998, Varlokosta & Triantafyllidou 2003, Spyropoulos & Tsangalidis 2005, Roussou 2006), which is not surprising considering the system of sentential complementation in Greek. Complement clauses in Greek are introduced by various complementizers: oti / pos / pu (all of which could be the equivalent of ‘that’), an ‘if’ and na ‘to’. Each complementizer provides semantic information on the relationship between the matrix and the complement clause, bearing specific features that make it (im)compatible with the matrix verb. There are verbs that can be combined with only one complementizer (Ex. 1), while other verbs are compatible with two (or more), having a different meaning with each one (Ex. 2). Other factors, such as propositional operators in the matrix clause or the form of the matrix verb are important for the selection, as well.

(1) Thelo na/ *oti/ *pos/ *pu/ *an figho
want.1SG na/ *oti/ *pos/ *pu/ *an leave.1SG.
‘I want to leave.’

(2) a. O Janis lei oti tha fighi noris
ART.NOM Janis.NOM say.3SGoti will leave.3SG early
‘Janis says that he will leave early.’

b. O Janis NOM lei na fighume noris ART.NOM Janis.NOM say.3SGna leave.1PL early

‘Janis says we should leave early.’

This study focuses on matrix verbs compatible with at least two of the complementizers oti / pos, pu or na, discussing similarities and differences in native and non-native Greek speakers’ choices. A research was conducted using written questionnaires which were administered to approximately 300 adult learners of Greek L2, native speakers of various languages, studying Modern Greek at the University of Athens, and to a control group of native speakers of Greek. 50% of the L2 learners were at the elementary level, 30% at the intermediate level and 20% were advanced. Additional evidence was gathered from corpora of written productions (Iakovou et al. 2016).

The data show that the major differences between native and non-native speakers pertain to problems in selecting the right complementizer when choosing between oti/ pos and pu, to difficulties with psych and perception verbs and to the role of propositional operators. We will also demonstrate how the level of L2 knowledge and the native language of the L2 learners are significant factors in the case of matrix verbs that are compatible with more than one complementizers.

References


Varlokosta, Spyridoula, and Lida Triantafyllidou (2003), Language proficiency levels of L2 Greek, Athens: University of Athens.

Diachronic development of functions of se and sibi in Latin

Verónica Orqueda & Francisca Toro
(Pontificia Universidad Católica de Chile)

Keywords: Latin, Diachronic change, reflexives, Indo-European, deixis.

As already influential scholars suggested (Hahn 1963, Schmidt 1978, Shields 1998, and Puddu 2005, among others), *se might not have been a reflexive strategy in PIE, but instead slowly developed into that function throughout the history of Latin. However, a connection to an original non-reflexive function in Proto-Indo-European is still an unsolved problem: first, because most reconstructions still accept a PIE reflexive marker, and secondly, because the relationship between *só, *sí, *sé and *swé is still a matter of discrepancy.
The goal of this paper is two-fold. First, we aim to contribute with both qualitative and quantitative data to the hypothesis that Latin se develops from a non-reflexive element. Secondly, we will offer an alternative reconstruction that derives these forms from a deictic marker high in the features of Animacy and Topicality, that belongs to the paradigm of PIE *sō, *sī, *=se.

To support our hypotheses, we offer a rigorous semantic and syntactic classification of constructions that can work both for se and sibi, in order to get a more comprehensive perspective: reflexives, as in (1), ‘long-distance anaphorics’ (2), reciprocals, prepositional constructions, intransitives (3), reflexive possessives and self-beneficent (4):

(1) uobis placere potius quam sibi (Ter. Heaut., 52)
you-DAT.PL please-INF more than IND.REFL
‘(Young writers may be eager to) please you rather than themselves’

(2) dixit se facturam (Pl. Cas. 483)
said ANAPH do-FUT.PTC
‘(She) said (she) would do (it)’

(3) per nos odia se nostra expicient (Sen. Thy. 323)
through us-AC.PL hatred-N.PL INTR our-N.PL unfold
‘Through us our hatred unfolds’

(4) sibi consilium capiendum existimavit (Caes. Bell. Civ. 3.77.3.5)
SELFBEN resolution-AC.SG take-GER. considered
‘(He) considered that (he) should take a resolution’

The collected data is composed of 441 cases with se (acc./abl.; sg./pl.) plus 301 cases with sibi (dat. sg./pl.) found in four different authors in Archaic Latin (Plautus, Terentius, Naevius and Ennius). These data are compared to a smaller sample from four authors in Classical Latin (Caesar, Cicero, Livius, and Seneca). As well, the few existing cases with the accusatives sam (sg. f.), sas (pl. f.), sos (pl. m.), and dative sis (pl. m./n.) are analyzed, as these may have also belonged to the same paradigm at some historical point.

The comparison between the two historical periods let us observe that the number of reflexive cases slowly rises, while anaphoric cases decrease. This tendency can be interpreted as a grammaticalization path from free anaphors to clause-bound items, in line with Kiparsky’s (2012) claim of a hierarchical anaphoric domain in Ancient Greek (clause-bound reflexives < long-distance reflexives < discourse anaphora < pronominals), and is confirmed in later layers of the language, as shown in Cennamo et al. (2015). Therefore, this case-by-case study let us confirm that the reflexive must not be the original function and that, instead, it developed over time.

Acknowledgments: The research reported here is supported by the CONICYT-FONDECYT Research program, project No. 11170045.

References


Puddu, Nicoletta (2005), *Riflessivi e intensificatori: greco, latino e le altre lingue indoeuropee*. Pisa: ETS.


Subject clitic as phasehood diagnostic in Sanandaji Kurdish

Mojgan Osmani  
(Tarbiat Modares University)

Keywords: subject clitics; phasehood diagnostic; high applicative phrase; root modal phrase; progressive aspect

Many standard constituency tests, diagnostics, have been proposed in the literature to determine the phase boundaries in Chomsky’s term (2000 and 2001) of a given category. Examples of some diagnostics including extraction, ellipsis, quantifier raising, reconstruction, and wh-movement, among others. This paper investigates the positional distribution of subject clitics as phasehood diagnostic in Sanandaji Kurdish (SK) in the past transitive construction. The Kurdish dialect under discussion is called ‘Sanandaji’ or ‘Ardalani’ which is spoken in the Kurdish area in Sanandaj, Kurdistan, Iran. A subject clitic in SK is a pronominal clitic that obligatory doubled the subject and used to cross-reference the subject in a past transitive construction. Subject clitics appear on the direct object as shown in (1), on the indirect object in (2), the nonverbal part of the complex predicate (3), the verb (4), and the lower adverb in Cinque’s term (1999) in (5):

1) Eme kîtw-êkê-an=man da-Ø bê Sirwan.
   1PL book-DEF-PL=SCL.1PL give.PST-3SG to Sirwan
   ‘We gave the books to Sirwan.’

2) Ba Sirwan=yan da-Ø.
   To Sirwan=SCL.3PL give.PST-3SG
   ‘They gave it to Sirwan.’

3) Twaw=yan kîrd-Ø.
   Finish=SCL.3PL do.PST-3SG
   ‘They finished.’

4) xward=yan=man.
   eat.PST=OCL.3PL=SCL.1PL
   ‘We ate them.’

5) Aqlanê=yan entîxab kîrd-Ø.
   Wisely=SCL.3PL selection do.PST-3SG
   ‘They selected wisely.’

These examples show that clitics remain low in the VP domain. The standard claim is that vP and CP form the phase. Subject clitics attach to the right of the leftmost constituent of a lower domain in SK. The leftmost constituent of the defined lower domain in phase theory is the phasal edge. Subject clitics attach to the element at the phasal edge. Thus, the positional distribution of subject clitics is considered as phasehood diagnostic in the past transitive construction in SK. A problem with this view is, however, that, vP in the past transitive construction in Kurdish does not form a phase (Karimi 2013)
and 2017). He claims that since the past transitive vPs do not bear a full complement of uninterpretable φ-features, they do not form a phase.

We suggest that the high applicative head in the past transitive construction forms phase. In other words, we introduce the high applicative head as the low domain hosting subject clitics and subject clitics as phasehood diagnostic. Additionally, Phases are rarely explored in the structure of Kurdish sentences, and the few occasions that they have been considered, there is no consensus about the phasal boundary in the presence of aspectual and modal projection in the past transitive construction. This paper directly addresses this issue by claiming that the positional distribution of subject clitics at the progressive aspect, root modal phrase, and subjunctive mood show that these projections are part of the clause-internal phase, while epistemic modal and perfect aspect projection are not contained. Some conclusions are reached regarding the nature of the past transitive construction in SK, for which a new phase-based analysis is proposed.

References
Karimi, Yadgar (2013), Extending defective intervention effects. The linguistic Review, 30(1), 51-78.

Spatial distribution and properties of Differential Object Marking in Gascon

Tania Paciaroni
(University of Zurich)

Keywords: DOM, Gascon, language contact, microvariation, diachrony

The fact that Gascon displays differential object marking (DOM) is well known and has already been the subject of in-depth studies (see, among others, Rohlfs 1971, Séguy 1973, and Floricic 2018). Nevertheless, current spatial distribution and exact properties of DOM require further investigation.

In the Atlas Linguistique et Ethnographique de la Gascogne (ALG), map 2504 is dedicated to DOM, described there as geographically restricted to two Pyrenean areas with an empty corridor in between. The contrast in (1a-c) from Gavarnie (ALG, data point 697) is representative of the area displaying DOM: human objects (1a) occur in the marked accusative introduced by a, whereas the specific inanimate (1b) and the animate non-human objects (1c) occur in the unmarked form of the accusative:

(1) Gavarnie
   a. k ɛiˈi ˈβis a ˈtai ˈpaɪ ˈthat have.1SG  seen to your father(M)
   ‘I have seen your father’
   b. k ɛiˈi ˈβis era maiˈzu
The striking nature of this spatial distribution is certainly linked to the limits of the data collection method used, as Séguy himself points out (1973: 429–431), but it also invites to gather new data adopting up-to-date theoretical knowledge. Recent fieldwork investigations show variation and optionality both in some localities of the areas displaying DOM (2) and in the empty corridor (3).

(2) Gèdre (SiM, *1935)

that have.1SG seen DEF.M.SG your dog(M)

‘I have seen your dog’

(3) Cauterets (JLL, *1947)

a. that have.1SG seen to your father(M) to Jean-Louis

c. k ɛ̂₁ 'βis et 'twe 'ka

The evidence in (2)–(3) shows the relevance of further semantic properties reflecting the Individuation Hierarchy (see e.g. Cennamo 2003: 70–76) and pragmatic properties (see Iemmolo 2010 for the role of information structure; Ledgeway et al. 2019 for parallels in southern Italy).

This paper will (i) present new data collected during fieldwork in the Pyrenees since 2016; (ii) identify the relevant parameters of DOM; and (iii) show a diachronic trajectory for the recorded variation. These results will deepen our knowledge of (1) the internal variation of Gascon, (2) its behaviour in the contact zone (a) with Catalan and (b) with Languedoc French.

Acknowledgments:
The support of the University Research Priority Program (URPP) “Language and Space” of the University of Zurich is gratefully acknowledged.

References

Cennamo, Michela (2003), (In)transitivity and object marking: some current issues, in G. Fiorentino
Marked nominative alignment from reanalyzed relative clauses: towards an explanation of prefixes and suffixes in Northwest Kainji argument marking

Rebecca Paterson
(CNRS-LLACAN and SIL International)

Keywords: Kainji, Benue-Congo, marked nominative, case development, relative clause

Northwest Kainji (Benue-Congo) languages display either prefixed or suffixed noun class marking on nouns (Hoffmann 1967 [dri], Bendor-Samuel et al. 1973 [uth], and Smith 2007 [gel]). Affix placement is governed by the NP. Paterson (2019) argues that Ùt-Ma’in [gel] argument alignment patterns also depend on the structure of the NP: unmodified NPs display marked nominative alignment; NPs that contain modifiers display neutral alignment. From Ùt-Ma’in data, I argue that suffix class markers on unmodified subjects developed from the relative pronoun within a relative clause that has a relativized subject.

Noun class suffixes function as overt morphological nominative case for S/A arguments (1-2).
Noun class prefixes on P arguments (3) function as overt morphological accusative case.

S-NOM V
(1) sè [kɔ:t-jɔ] rwɔn ɔr-vastè
then guinea.fowl-c7 exit c5-last
‘Then a guinea fowl exited last.’ (Paterson 2019: 104)

A-NOM V P
(2) kɔnɔ [kɔ:t-jɔ] zɔ-t:è...ɔr-kjat ...
there guinea.fowl-c7 say-PFT C5-difficult
‘There a guinea fowl has said, “Difficult... “ ‘ (Paterson 2019: 104)

A V ACC-P
(3) ā=b hɔn [ŋ-kɔ:t]...
COND=2SG see c7-guinea.fowl
‘If you see a guinea fowl, ...’ (Paterson 2019: 104)
The unique form and function of the Ut-Ma’in nominative case marking and the prolific functions/locations of the accusative case marking are organized into a Type 2 Marked Nominative system (König 2008: 8, 158): (i) A and S are treated the same and simultaneously different from P, (ii) the accusative form is used as the citation form, and (iii) both case forms are morphologically marked; however, the accusative is used in a wide range of functions. The nominative, in contrast, is used in a very restricted set of functions. Overt case marking is only a feature of an NP argument if that NP argument has no modifier; when modifiers occur, no case distinction is made, i.e., the alignment system is neutral. The prefix-marked form of the accusative is the historical form of the noun word, matching the “expected” Niger-Congo prefixed noun class word form. The potentially innovative suffix-marked nominative form is the result of reanalyzing a relative clause structure as main clause syntax.

In an Ut-Ma’in relative clause, the relativized element is not expressed (4). When the relative clause expresses an event that is sequenced with the event conveyed by the main clause and the subject of the matrix clause is also the subject of the relative clause, the structure is semantically ripe for reanalysis. The relative pronoun, already marked for agreement with the class of the head noun, is reanalyzed as a subject marker. Compare tò ‘C6.REL’ in (4) with tò ‘C6’ in (5); they are tonally and positionally identical. The relativized subject is reanalyzed as a main clause subject; the verb of the clause is finite and interpreted as a main clause verb, (5).

\[
\begin{align*}
  & S_{\text{MATRIX}} \ [\text{REL}] \ S_{\text{REL}} \ V_{\text{REL}} \ ] \ V_{\text{MATRIX}} \\
  & \text{nět} \ [t-Ø] \ há:-n-é \ r-tākən \ ] \ f5η \ 5z5… \\
\end{align*}
\]

(4) person C6.REL.Ø go-DIST-FOC C5-beginning think.PST saying

‘People who came first thought that...’

S-NOM V

(5) nět-tó há:-n-é r-tākən

person-C6 go-DIST-FOC C5-beginning

‘People came first.’

Acknowledgements: Funding for this study was provided by the Center for the Study of Women in Society at the University of Oregon (2016) and the Firebird Foundation for Anthropological Research (2016). This project has also received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 758232).

References:
Progressive aspect in a Dutch-German-Limburgish contact variety

Nantke Pecht

(University of Maastricht)

Keywords: progressive aspect, syntax, semantics, language-dialect contact, Germanic contact variety

This talk presents a first analysis of progressive aspect in Cité Duits, a contact variety composed of features of Belgian Dutch, German and a Limburgish dialect. Cité Duits (lit. ‘mining district German’) developed among the locally-born children of immigrant miners of different European language backgrounds in the shared housing district of Eisden (B) in the 1930s. Nowadays, it is on the cusp of disappearing, with approximately ten speakers left, all of them men in their eighties. My contribution will be based on a corpus of audio data of spontaneous-like interactions (340 minutes) gathered by a method of sociolinguistic fieldwork (Labov 1972) in 2012/13 and 2015/16. By following the views postulated by Comrie (1976), Bybee and Dahl (1989) and Bybee et al. (1994), I consider aspect as a cross-linguistic category that exhibits a number of different realization strategies. An example from the data is the following:

(1) (0313_152448: 84.31 - 88.01)

\[ \begin{align*}
& \text{de} \quad \text{war(t)} \quad \text{television} \quad \text{an} \quad \text{gucke,} \\
& \text{flog} \quad \text{ihm} \quad \text{au(ch)} \quad \text{de} \quad \text{(au)Auge} \quad \text{kaputt.}
\end{align*} \]

‘He was watching television when his eye got hurt.’

In (1), the speaker reports a situation in the past where two events took place simultaneously. He first uses a progressive construction \((\text{war(t)} + \text{an} + \text{V-INF})\), indicating that the situation he referred to continued, and then switches to the simple past tense \((flog)\). Progressive aspect is expressed here by means of \(\text{an} + \text{V-INF} + \text{sein-FINITE} \) ‘at \(\text{V-INF be-FINITE}\)’.

The investigation of progressive constructions is particularly intriguing because the linguistic means in Cité Duits’ contact varieties differ to some extent but also overlap. Neither Belgian Dutch, Limburgish nor German have a single progressive construction. Although there are several ways to denote progressivity, they lack a ‘standard’ progressive marker. Besides, explicit marking of progressive aspect is optional in many contexts (Behrens, Flecken, and Carroll 2013: 95; Broekhuis, Corver, and Vos 2015: 150; Ebert 1996; Krause 2002; Tomas 2018: 53; Witt 2015: 5; Anthonissen, de Wit, and Mortelmans 2016: 9; van Pottelberge 2004). Nevertheless, it has been argued that Dutch has grammatical means to indicate progressive aspect, whereas German is said to have no grammaticalized aspectual markers (Flecken 2011: 62). Keeping these observations in mind, the question arises if and how progressive aspect is marked in Cité Duits.
In the first part of my talk, I will examine whether there is a progressive construction typical to Cité Duits. Based on frequency of occurrence, I will demonstrate that Cité Duits has one construction that is clearly the most widespread way to convey progressive meaning. In the second part of my talk, I will address the question of whether there are particular syntactic or semantic restrictions in the respective constructions that express ongoingness. Ultimately, I will discuss how these findings can be interpreted in terms of language contact.

Acknowledgments
This project is financed by the Netherlands Organization for Scientific Research (NWO, project number 322-70-008).

References

Multilinearity and multimodality in grammar:
The interaction of prosody, syntax and gestures in special questions expressing surprise and surprise-disapproval in a cross-cultural and cross-modal perspective
Erika Petrocchi  
(Ca’ Foscari University of Venice)

Keywords: surprise questions, syntax/pragmatics interface, prosody, sign language, gesture

I investigate counter-expectational surprise questions and surprise-disapproval questions, which are characterized by a special emotional interpretation that is obligatorily associated with a peculiar gestural and prosodic pattern. Consider (1) and the following context: Mary calls me on the phone and tells me that she has a new red dress to wear at tonight’s party. When I meet her at the party, I see that she has a blue gown. I am surprised and utter:

(1) Ma non era rosso?  
But wasn’t it red?  
(from Giorgi & Dal Farra 2018, ex. 4).

Consider (2) and the following context: I see Gianni with his best trousers kneeling in the dirt in the garden. I think that he will ruin his trousers. I disapprove of his activity and utter:

(2) Ma cosa fai?  
‘But what are you doing?!’  
(from Giorgi & Dal Farra 2018, ex. 6).

Researches conducted on Italian (Giorgi & Dal Farra 2019; Dal Farra & Giorgi 2018), German (Giorgi, Dal Farra & Hinterhölzl to appear) and Spanish (Furlan, MA Thesis 2019) show that the gestural, syntactic and prosodic patterns characterizing these structures display striking similarities cross-linguistically. These components turn out to be aligned in that the stroke of the gesture is simultaneously realized with the leftmost pitch of the sentence, usually found on the verbal form. The works mentioned above detected the presence of the same manual and non-manual gestural pattern in all the different languages studied. In order to study if these similarities are due to the cultural closeness of the languages considered, I run some experiments (repetition and elicitation tasks) studying the realization of these special questions (i.e. not information questions) in three culturally and geographically distant languages: Vietnamese, Korean and Japanese and in Italian Sign Language (LIS), the prosody of which is realized in the visual gestural-modality (non-manual components) paying attention to the relation between manual gestures and signs (Goldwin-Meadow & Brentari 2015). I detected a striking formal and functional regularity in the non-manual gestural pattern typically associated with surprise and surprise-disapproval – e.g. raised and furrowed eyebrows, forward head movement, nod, shake, widened and squinted eyes - in all the languages investigated. My results are coherent with recent findings in the literature. Moreover, in the signed modality, it is possible to find two of the three iterated manual gestures (i.e. not lexicalized items) found in Spanish and Italian along with signs. It is possible to formulate the working hypothesis that the non-manual components involved in the linguistic expression of surprise and disapproval might be universal, and that may be considered as part of Universal Grammar in that non-manual linguistic/bodily parameters seem to be regularized (sychronic alignment between syntax, prosody, and gestures) and show the same features cross-culturally and cross-modally. My study opens interesting possibilities of comparison between spoken and sign languages. Consider that sign language and gestures are rarely treated together in theoretical discussions on the human language faculty (Sandler et al. 2019).

References
Dal Farra, Chiara, Giorgi, Alessandra & Hinterhölzl Roland (Forthcoming), Syntax, Prosody and Gesture in expressing surprise, Paper delivered at the 40th annual Conference of the German
Lectal contamination: How lexical biases in linguistic variation may originate from lectal differences

Dirk Pijpops
(University of Leuven)

Keywords: lectal contamination, corpus, variation, lexical biases, usage-based

Grammatical variation is often driven by lectal differences, i.e. differences between regio-, dia-, ethno- or sociolects (e.g. Szmrecsanyi et al. 2016). For instance, the optional -s ending in the Dutch partitive genitives, as in (1)-(2), is far more often used in the Netherlandic regiolect of Dutch than in the Belgian regiolect (Pijpops & Van de Velde, 2018). Meanwhile, lexical variation also often exhibits lectal differences (Speelman et al., 2008). For instance, the adjective bijzonder ‘special’ as in (1) is more often used in the Netherlandic regiolect, while the adjective speciaal ‘special’ as in (2) is more popular in Belgium.

(1) *Is er iets bijzonder(s) dan?*  
Is there something special(-GEN) then  
‘Is there something special then?’

(2) *Is er iets speciaal(s) vanavond?*  
Is there something special(-GEN) tonight?  
‘Is there something special tonight?’

As a result, lexical items that are typical of Netherlandic Dutch, such as bijzonder ‘special’, will more often occur with the variant with -s ending, simply because they are more often used by people from the Netherlands. Now, there is of course language contact between the Netherlands and Belgium (van Agtmaal-Wobma et al., 2007). As such, Belgian language users will more often hear the variant with -s ending among partitive genitives with typically Netherlandic lexical items, e.g. *iets bijzonder(s) ‘something special’, than among partitive genitives with typically Belgian items, such as *iets speciaal(s) ‘something special’. If we then assume that language users are susceptible to lexical biases...
in the language usage around them (Bybee, 2013; Diessel, 2017), we should expect that Belgian language users themselves also start to prefer the variant with -s among partitive genitives such as *iets bijzonder(s)*, compared to partitive genitives such as *iets speciaal(s)*. In other words, lexical preferences of one lect would ‘contaminate’ the grammatical variation in another. This hypothesized effect is therefore named ‘lectal contamination’.

We have investigated lectal contamination in two ways. The first is corpus research. Here, it was found that partitive genitives with typically Netherlandic lexical items indeed appeared more often in the variant with -s ending, both among the language use of Belgians and Dutchmen, while controlling for other factors that are known to affect the use of the -s. The second is agent-based simulation (Beuls & Steels, 2013). Here, we built a computer simulation that incorporated the following assumptions, (i) there are two grammatical variants, with one group of agents initially preferring the one, and the other group relatively preferring the other; (ii) both groups of agents also prefer different lexical items, and these grammatical and lexical preferences are initially independent of one another; (iii) there is a limited degree of language contact between the groups; and (iv) the agents store concrete combinations of a lexical items with a grammatical variant in their memory, i.e. they are susceptible to lexical biases. The results of this simulation show that the effect of lectal contamination indeed emerges under these four assumptions.

**Acknowledgments**
I thank the University of Leuven for supporting this research through a BOF PDM-grant, as well as Freek Van de Velde for useful discussions regarding the nature of lectal contamination.

**References**


**Gender Agreement in NP and Clausal Domain: Theoretical and Typological Overview**
An asymmetric occurrence of agreement in different syntactic domains may be viewed as one of the points at issue of typologically oriented discussion of agreement. In this paper I make an attempt to outline and reconcile two diverging views on the primacy of agreement within a particular syntactic domain. Plenty of papers argue that, generally, agreement is characteristic of clausal domain rather than NP (Plank 1994, Nichols 1989, Matasović 2014). In terms of hierarchical relations clausal agreement implies NP-internal one for a given language. Meanwhile from the positions of canonical typology NP-internal agreement is said to be of more canonical nature than verb agreement with its subject (Corbett 2006: 21). Although I shall not address these two opinions as opposed, there is a certain need in more precise investigation of gender of all agreeing categories as it behaves more ‘canon-like’ and certainly goes against general typological expectations.

After brief discussion of previously proposed ideas I present evidence from genetically diverse sample that offers no support for universal correlations between presence of agreement in general (i.e. with feature underspecified) and particular syntactic domains: the category of gender clearly demonstrates a tighter bound with domain of NP than with clausal domain in all languages examined. This survey is based on the analysis of 59 languages with gender system attested. The sample of all gendered languages from the WALS’ list of 200 languages was extracted as an intersection of this list with feature 30A ‘Number of genders’ (Corbett 2013). It is expected that not only gender agreement in NP does not imply gender agreement in the domain of clause but there are also no languages which allow clausal gender agreement, but not NP-internal agreement.

In the next section I bring the discussed expectations into correlation with the results of this survey and to the possible extent compare these results with statistics presented in previous works on the subject.

Table 1. Gender agreement in 53 languages of our sample (This excludes 5 pronominal gender systems and Tagalog where agreement is only marginally expressed on Spanish loanwords in predicative position (Schachter and Otañes 1972:97).).

<table>
<thead>
<tr>
<th>Agreement pattern</th>
<th>Both domains 38</th>
<th>Only clause 2</th>
<th>Only NP 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>VA+; DA+; AA+</td>
<td>VA+; DA+; AA-</td>
<td>VA+; DA-; AA+</td>
<td>VA-; DA+; AA+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VA-; DA-; AA+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VA-; DA+; AA-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VA-; DA-; AA+</td>
</tr>
</tbody>
</table>

Languages 22 12 4 2 12 1 0

Given these findings, I argue that the hypothesis of gender agreement in clause implying gender agreement in NP is typologically confirmed. I do not claim it to be universal as there is a number of languages in every sample which do not conform to the proposed implication. However, statistically such cases are rare.

References

152
Multi-verb constructions in Ahamb (Vanuatu): Between serialisation and echo reference

Tihomir Rangelov & Julie Barbour
(University of Waikato)

Keywords: serial verbs, echo reference, sequential verb constructions, Oceanic, Vanuatu

Multi-verb constructions (MVCs) are a common feature of Oceanic languages. In the languages of Vanuatu, two common types of MVCs are regularly attested – serial verb constructions (SVCs) (often subclassified into Nuclear and Core SVCs) and echo referent constructions (ERCs) (Crowley 2002).

In Nuclear SVCs the two verbs share the same subject and object and there is no intervening morphology between the two verbs (1a). Core SVCs, on the other hand, exhibit a looser juncture between the verbs. They come in two types: the verbs can have the same subject (1b), or the first verb’s object can serve as the subject of the second verb (1c). Vanuatu languages typically exhibit SVO word order and have prefixed subject indexes (marking person, number and sometimes other grammatical information) and suffixed object indexes.

(1) (schematic representation as per Barbour & Dodd (2017))

<table>
<thead>
<tr>
<th>Type</th>
<th>Subject</th>
<th>SubjIndex-Verb1</th>
<th>Verb2</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear SVC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same Subject</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core SVC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ERCs typically display the same subject. The first verb carries a subject index, while all subsequent verbs have a prefixed echo referent morpheme that is not person/(number)-specific and points to the first verb’s subject index, as in (2).

(2) ERC

<table>
<thead>
<tr>
<th>Subject</th>
<th>SubjIndex-Verb1</th>
<th>EchoIndex-Verb2</th>
<th>EchoIndex-Verb3</th>
<th>...</th>
</tr>
</thead>
</table>
Ahamb is an endangered and previously undescribed language spoken by 1000 people in Vanuatu. This talk explores the form and function of MVCs in Ahamb in an attempt to position these data within the existing theoretical frameworks.

In Ahamb, there are both Nuclear SVCs (exemplified in (3)) and Switch Function Core SVCs (as in (4)).

(3) Na-jav por nmaru ili.
    1SG-cut split coconutDEM
    ‘I cut the coconut in two.’

(4) Na-rëng-i nga-roh aha.
    1SG-put-OBJ 3SG-be.located here
    ‘I leave it here.’

A distinctive feature of Ahamb is the presence of at least four series of subject indexes. The subject indexes in one of these four series are used exclusively on the second and subsequent verbs in constructions like the one in (5), usually involving sequential events:

(5) Na-van ne-rav mwog s-ag ne-varus van.
    1SG-go 1SG-take canoe POSS.GNR-1SG 1SG-paddle go
    ‘I went, took my canoe and paddled away.’

Examples as in (5) are neither prototypical Same Subject Core SVCs nor prototypical ERCs, but show similarities to both. Their form and function and their connection to related phenomena such as coordination and prosody are discussed in this talk, in order to establish where they sit within the existing typological and Oceanic frameworks on SVCs and ERCs (eg. Haspelmath 2016, Aikhenvald 2006, Bril & Ozanne-Rivierre 2004, Crowley 2002, Barbour & Dodd 2017, Ross 2017).

References

Barbour, Julie & Royce Dodd (2017), Serials and Echoes in the languages of Western Malekula, presented at the 12th Conference of the Association for Linguistic Typology (ALT 2017), Australian National University, Canberra, Australia.


Ross, Daniel (2017), Echo-subject marking in Vanuatu, presented at the 12th Conference of the Association for Linguistic Typology (ALT 2017), Australian National University, Canberra, Australia.

Grammaticalization of result clause constructions as degree modifiers: Evidence from Romanian and English
Keywords: construction grammar, second-order construction, degree modifier, result clause construction, resultative construction

Language embraces various means to render degree, also at the constructional level. We propose a corpus-based analysis of a pattern that has been less researched as source of grammaticalization for intensifiers, the finite result clause construction (RCX), expressing a primary predication in the main clause and a secondary predication in the result clause. For example, de mori ‘that one dies’ can occur in contemporary Romanian with a literal interpretation, where death is a potential result (1) Dacă nu ești atent, adversarul [RCX: te lovește [RCI: de mori]] (intended: If you’re not paying attention, your opponent will beat you to death), but also with a primarily intensifying function: (2) [RCX: Îmi doresc victoria [RCI: de mori]] (intended: I want victory so bad (that) I can taste it) or (3) [RCX: Râzi [RCI: de mori]] (intended: You laugh yourself to death).

What makes this construction particularly interesting for the purpose of our comparative, synchronic investigation is that, in Romanian, and also in English, it shows clear signs of grammaticalization as second-order or degree modifier construction, similarly to another syntactic pattern – the resultative. For example, one can express extreme amusement, hyperbolically, in English as (4) I [RCX: laugh so hard [RCI: (that) I split my sides]] or (5) I [resultative: laugh myself silly/sick/to death], etc. The resultative is widely discussed in the literature on Germanic languages, as source construction for intensifiers (among others, Hoeksema and Napoli 2008, 2019; See Claridge 2011, Margerie 2011, and Blanco-Suárez 2017 for supporting the admittance of the death-related resultative constructions in the intensifier set); however, it is proved to be only marginally attested in Romance (see Farkas 2009 for Romanian).

In defining degree resultatives as second-order constructions, Hoeksema and Napoli (2019) state that the original constructional meaning – i.e., a (potential or actual) result – is reinterpreted as an intensifier. Our analysis shows that, although syntactically different, RCXs can also function as second-order in the sense above, which adds to the variety of patterns attributed to these semantically motivated (ibid.) constructions. Moreover, it provides evidence that second-order RCXs are productive in the two languages investigated. To this end, we first conduct a comparative corpus analysis of the very frequent death-related secondary predicates de mori and to death, the first pertaining to a RCX, the second to a resultative construction. Then, we compare de mori with one of its contextual correspondents in English, (that) one can taste it – see (2), which is attested with good/close/strong/(want) sth. bad(ly), but it seems to have spread its current usage to everything causing extreme emotions/sensations (attested uses: so cold/event so magical (that) you can taste it). In both scenarios, we examine the collocational patterns on criteria such as a. semantic/morphological restrictions that characterize the class of collocates and b. type of meaning that prevails in contemporary language – which can be literal, subjective (based on mechanisms such as metaphor, hyperbolization/metonymic inferencing, where both literal and high degree interpretations coexist), or purely intensifying (where the secondary predicates cast aside literalness and contribute a merely intensifying interpretation). The latter criterion is judged on an automatically extracted list of the top frequent collocations found in the investigated corpora for a given secondary predicate. In the case of the RCXs, we also evaluate c. the degree of semantic and syntactic boundedness of the secondary predication to the construction (see also Rizea and Sailer 2019).

Our results indicate that, in contrast with to death, which is reported as a peripheral degree modifier whose most evident trait in synchrony is subjectivity (Margerie 2011: 132), de mori – which
collocates (almost) unrestrictedly, and mostly with positively connoted primary predicates – and (that) one can taste it register top frequent collocations which prove that the RCXs comprising them as secondary predicates manifest the highest degree of conventionalization on a purely intensifying meaning. We interpret this as a sign that de mori and (that) one can taste it are, currently, more advanced in the grammaticalization process than to death.

The investigation was carried out via Sketch Engine and KorAP corpus query and analysis tools, in four monolingual (RoTenTen16, EnTenTen15, CoRoLa, and Timestamped JSI web corpus 2014-2019) and two parallel corpora (OPUS2 English/Romanian).

References
CoRoLa – Corpus computational de referinta pentru limba romana contemporana [The Reference Corpus of the Contemporary Romanian Language].
Sketch Engine – http://www.sketchengine.eu

Small Talk in American and Ukrainian dating cultures: A cross-cultural analysis of female nonverbal communication

Oleksandra Romaniuk & Inna Stupak
(National University “Odessa Maritime Academy” & South Ukrainian National Pedagogical University named after K. D. Ushinsky)

Keywords: Small Talk, female nonverbal communication, American dating culture, Ukrainian dating culture, face-to-face encounters

Small Talk in face-to-face encounters is aimed at breaking the ice between the initiation partners. Knapp (1978) described Small Talk as “Big Talk”, highlighting its significance in establishing interpersonal relationships. Nonverbal cues, constituting 93 per cent of inferred meanings, are particularly helpful for potential partners to initiate conversation (Hall et al., 2009); they set the stage
for effective communication (Burgoon et al., 2010). However, nonverbal communication (NVC) varies across cultures (Andersen, 1988), leading to misunderstandings between the initiation partners from different cultural backgrounds.

The researchers tried to solve this problem by making a list of recommendations for successful Small Talk including nonverbal immediacy (DeFleur et al., 2014). However, only a few attempts were made to compare NVC across cultures and gender: German, Italian, and American gender behaviour (Shuter, 1977), German, English, French, Japan, and the Arab World (Hall, 1966). Likewise, there is no sufficient research for drawing a conclusion on male-female differences in NVC (Hall & Gunnery, 2013).

The literature review identified a research gap in the comparative research on Small Talk NVC in American and Ukrainian dating cultures (gender aspect). This motivates the study to provide insights into Small Talk NVC employed by American and Ukrainian females. The data were collected from the females ($N_{US}=55$ vs. $N_{UA}=55$) engaged in initial dyadic interactions. NVC, enacted by the bachelorettes of the dating shows The Bachelor US and The Bachelor UA (2012-2014) during face-to-face encounters, was assessed. The data came from a larger project on initial romantic relationships (Romaniuk, 2017), providing sufficient statistical power (> .80). NVC is less structured than verbal, non-linear, and spontaneous (Kouros & Cummings, 2011), which has strong ecological validity of the current research. Thus, interactional sociolinguistics was applied as one of the approaches to discourse analysis (Gumperz, 1982; Schiffrin, 2000); both qualitative interactive sociolinguistic and quantitative analyses were used (Berg, 2009). The nonverbal cues were compared utilising SPSS 23 to establish cross-cultural differences and similarities.

American ($N_{US}=446$) versus Ukrainian ($N_{UA}=371$) nonverbal cues were categorized according to Knapp et al. (2014) approach and compared: (1) touching behaviour (31.8 vs. 31.5); (2) eye behaviour (23.8 vs. 21); (3) speech-related gestures (18.6 vs. 17); (4) facial expression (13.5 vs. 15.6); (5) posture (12.3 vs. 14.8). The findings showed that Americans preferred other-focused touching behaviour (hug, hand-in-hand, kiss), eye behaviour (mutual gaze, gaze down), and speech-related behaviour (head tilt), while Ukrainians prioritised only eye behaviour. The findings also revealed the difference in the quality of these categories. Unlike American, Ukrainian females employed lip bites and face coverings; conversely, only Americans utilised Sh-gestures, hands on hips, hugs, caress and pats. This supports the hypothesis that American bachelorettes are quite outgoing and see no barriers to hugs, holding hand-in-hand, and kisses, while Ukrainians are more reserved as they preferred to maintain a formal style with a perfect stranger. These results indicated directions for further research.

References
Berg, Bruce (2009), Qualitative research methods, Boston, MA: Allyn & Bacon.


Romaniuk, Aleksandra (2017), Basic universal units and components of romantic discourse based on the dating show patterns of dyadic interaction, Analele Universităţii din Craiova, Seria Stiinte Filologice, Lingvistica 1-2, 370-380.


The influence of sentence type on the interpretation of Spanish future constructions

Malte Rosemeyer & María Sol Sansiñena
(Albert-Ludwigs-Universität Freiburg; KU Leuven & KU Leuven)

Keywords: Spanish; Pragmatics; Future tense; Corpus Linguistics; Sentence type

It is well known that Spanish futurizing morphology, whether synthetic or periphrastic, is frequently used not to express futurity but to formulate a hypothesis (Escandell-Vidal 2010; Escandell-Vidal 2014; Rodríguez Rosique 2019, among many others). According to Escandell-Vidal, in examples such as (1) the use of future tense encodes a procedural instruction to interpret the propositional content as the result of an inference by the speaker. As noted by many authors, in Peninsular Spanish the synthetic future tense (1a) is be more likely to express such epistemic reading than the periphrastic future (1b).

(1) [the doorbell rings]
   a. **Será** Juan.
      be.IND.FUT.3SG Juan
      ‘That must be Juan.’
   b. **Va a ser** Juan.
      go.IND.PRS.3SG to be.INF Juan
      ‘That must be Juan.’
The overwhelming majority of previous studies on the Spanish future tenses has focused on declarative sentences such as (1). Interestingly however, futurizing morphology can lead to rhetorical effects, a phenomenon that has not been studied thoroughly. Consider for instance, example (2) taken from the C-ORAL ROM (Cresti & Moneglia 2005), in which the speakers PAC and PAT are talking about a house built by some friends. PAT’s future-inflected interrogative is used as a rhetorical question to signal the irrelevance of PAC’s admiration of the house.

(2) PAC: con una terracita y todo? ‘With a small terrace and everything? ’
PAT: de todo ‘With everything’

PAC: joder eso menuda casa están haciendo [...] ‘Wow, what a great house they are building’

PAT: llevan dos o tres años con la casa (.) cómo estará la casa (.) impresionante ‘They have been working on the house for two or three years, how (else) can the house be? Impressive!’

This paper explores the relationship between futurizing morphology and sentence type on the basis of a quantitative analysis of about n=6,000 tokens of synthetic and periphrastic ‘future’ constructions in conversations in Spanish from Madrid, Buenos Aires and Santiago de Chile taken from the C-ORAL ROM and the COLA (Jørgensen & Eguía Padilla 2017). On the basis of a bottom-up classification of these tokens regarding their potential to express modal meanings, we demonstrate that because interrogatives imply a lower degree of epistemic authority by the speaker over the proposition (Heritage 2012; Enfield, Brown & De Ruiter 2012), polar and partial futurizing interrogatives are more likely to display modal meanings and associated rhetorical effects than futurizing declaratives. We also document substantial dialectal and diachronic variation in the use of futurizing morphology. This result points to a conventionalization of these form-function pairings for certain sentence types, as well as a clearer division of labor between synthetic and periphrastic future forms in the Spanish spoken by South American and Madrilenian adolescents.

References


Body and body-part configurations in Piaroa (Jodï-Sâliban)

Jorge Emilio Rosés Labrada
(University of Alberta)

Keywords: categorization, configurations, Sâliban, Piaroa, lexical typology

Body-part categorization has long been a concern of linguists. Previous studies have focused on universals of categorization (cf. Brown (1976), Andersen (1978), Palmer and Nicodemus (1985), and Schaefer (1985)) and on the semantic relations that obtain between different body-part terms (Wierzbicka 2007). The topic has also received attention from a cross-linguistic perspective recently (e.g. Enfield, Majid, and van Staden (2006)). However, terms for specific configurations of the body (e.g. crouching) or of particular body parts (e.g. fist, meaning “closed hand”) remain understudied. In this presentation, I focus on this area of the lexicon of Piaroa—a Jodï-Sâliban language spoken in Colombia and Venezuela—; all examples come from a five-hour corpus of naturalistic narrative data collected in the field.

There are numerous terms for body or body-part configurations in Piaroa. The examples below illustrate four of these terms in use. Examples (1) and (2) show terms for configurations of the entire body: kâ’ch’uy’ch’uy is used to describe the act of standing with both arms besides your body trying to look small while jareyũ describes being in a crouching position. The other two examples, however, exemplify terms for configurations of specific body parts—in both cases the eyes. In (3), the yuri, a type of lizard, is telling the man to keep his eyes closed and while there is no mention of the word for ‘eye(s)’, this is already encoded in the word âtodu. In (4), the Huirito—mischievous characters that live in the forest and antagonize humans—were looking at the man out of their corner of their eyes, hence the use of the verb root for ‘to be white’. The speaker also makes use of reduplication and length in this case.

(1) dajahuächö  kâ’ch’uy’ch’uy  kânhö
dahaw-aŋj-i  kâ’tû’s’ûs’û  kâ-ên-í
hide-?-NON.FIN  straight  stand-?-NON.FIN
topärü  kînûdô
tô-p-âr-û  kâ-în-û-do
see-CLS2-DUR-NON.FIN  stand-IMPF?-NON.FIN-EVID
‘hiding, standing straight, he stood looking on.’ (Babel019:63)

(2) jareyũ-nô  p-âŋ-ô  kajuahua’a  châhui’ô
jârêdômû-nî  p-ên-î  kâ-ahaw-âʔâ  ûîâwîʔ-î
crouching-?  sit-?-NON.FIN  3SG,F-cry-TAM1  continue-NON.FIN
‘in a crouching position, she cried, continuing’ (Babel021:272)

(3) âtodu  pâi’ô  riβip’ô
âtûdû  pêiʔ-î  rî-îp-îʔ-î
eyes.closed  turn-NON.FIN  stay-?-NON.FIN
‘stay put, with your eyes closed’ (Babel021:237)

(4) tti āriyā  twābābāāā / topinājōdo
th-iʔe-r-idžē  te-ēbē-bāːː  to-p-in-ētī-do
3PL-eye-CL:EYE-PL1  be.white-CL?:-RED  see-CLS2-IMPF?-CL:PL-EVID
‘they looked out of the corner of their eyes (with their eyes all white)’ (Babel025:85-86)

Note that in all four of these terms, the vowels are nasal; this is possibly linked to the association of nasality with small size (see (Krute 1989), (Rosés Labrada accepted)). Smallness is also conveyed in the translations offered by my consultants in Spanish, all of which have a diminutive suffix (e.g. paradito for ‘standing straight’ (1)). Additional terms include words for “bared teeth” and “spiky hair” as well as for “lying down with your arms beside your body”.

An in-depth investigation of this lexical domain in Piaroa does not only contribute to the description of the language—which remains seriously underdescribed—but also to our cross-linguistic understanding of this area of the lexicon. Further, it adds to recent advances in the area of lexical typology (e.g. Koptjevskaja-Tamm & Vanhove (2012)).

References


Koptjevskaja-Tamm, Maria, and Martine Vanhove, eds. 2012. “New Directions in Lexical Typology (Special Issue).” Linguistics 50 (3).


Le statut des prépositions issues de noms en Gbaya
Paulette Roulon-Doko
(LLACAN du CNRS; INALCO)
Keywords: preposition, grammaticalization, polysemy, meaning, Gbaya

Un certain nombre de noms ont en Gbaya, langue oubanguienne de RCA, la particularité d’avoir un sens notionnel produisant une très forte polysémie en particulier lorsqu’ils sont utilisés comme déterminés au sein d’un SN génitival à connectif tonal que j’appelle syntagme définitoire. Neuf de ces noms, dont sept pouvant désigner une partie du corps, peuvent être interprétés en position locative comme un syntagme prépositionnel. Mon analyse montrera les limites d’une telle interprétation pour certains d’entre eux pour lesquels le processus de grammaticalisation n’a pas abouti et qui, de ce fait, ne peuvent être considérés comme des prépositions à part entière. L’étude se fonde sur un important corpus de paroles spontanées récolté sur le terrain en R.C.A. entre 1970 et 2013, et les exemples sont plus spécifiquement tirés des 4h30 de corpus enregistré, transcrit et linguistiquement analysé.

L’examen de ces neuf noms à sens notionnel me conduit à proposer des statuts différents.

<table>
<thead>
<tr>
<th>Nom</th>
<th>sens notionnel</th>
<th>partie du corps</th>
<th>emploi seul, sans détermination</th>
</tr>
</thead>
<tbody>
<tr>
<td>nú</td>
<td>partie active</td>
<td>bouche</td>
<td>langage</td>
</tr>
<tr>
<td>tè</td>
<td>entité</td>
<td>corps</td>
<td></td>
</tr>
<tr>
<td>yik</td>
<td>surface plane</td>
<td>visage yeux</td>
<td></td>
</tr>
<tr>
<td>zù</td>
<td>sommet externe</td>
<td>tête rêve</td>
<td></td>
</tr>
<tr>
<td>zàŋ</td>
<td>étendue</td>
<td>ventre grosse</td>
<td></td>
</tr>
<tr>
<td>ndàyà</td>
<td>fondement</td>
<td>fesses</td>
<td></td>
</tr>
<tr>
<td>sèè</td>
<td>centre</td>
<td>foie</td>
<td></td>
</tr>
<tr>
<td>gün</td>
<td>base</td>
<td>0 base enterrée</td>
<td></td>
</tr>
<tr>
<td>?òò</td>
<td>restes</td>
<td>0 une même place</td>
<td></td>
</tr>
</tbody>
</table>

Certains de ces noms ont une valeur notionnelle qui peut faire émerger un sens spatial, d’autres non. L’absence totale de sens spatial dans le cas de nú « partie active » et de tè « entité » conduit à deux analyses différentes. Pour nú la multiplicité de l’expression de la fonctionnalité, lorsqu’elle correspond à une possible localisation reste entièrement contrainte par la nature du déterminant, ce qui ne permet pas l’émergence d’une préposition. Pour tè le sens notionnel unique d’« entité » devient, en dehors du rapport au corps, un moyen de circonscrire l’espace autour d’un déterminant pour en faire une localisation dont la nature peut être très variée, mais qui est nécessaire pour conférer à ce déterminant une fonction locative, car la construction directe n’est pas possible pour lui. Ce qui me conduit à retenir une préposition locative neutre tè « à » qui joue deux rôles, permettre tout d’abord à des noms d’accéder à une fonction locative et permettre aussi au nom verbal de certains verbes d’exprimer un gérondif. Enfin, ?òò « restes » qui n’a pas de valeur strictement locative produit la préposition « à la place de » qu’il s’agisse d’un remplacement ou d’une absence. Dans tous les autres cas, un sens spatial se déduit aisément du sens notionnel et, le plus souvent, la simple postposition d’un tel SN au groupe verbal suffit à l’exprimer, sans avoir besoin de poser l’existence d’une préposition. Seul, zù dont le sens notionnel « sommet extérieur » combine saillance et extériorité porte, dans certains de ces SN porte une valeur spécifique dégageant un espace particulier : l’espace « au-dessus » d’un support, quel qu’il soit. Ce qui me conduit à retenir zù comme préposition.

Directionality in N/V conversion: What do meaning and usage say?
Directionality in conversion remains largely an open question in today’s description, e.g. in cases such as smile\textit{N} / smile\textit{V}, at least when viewed as a word-formation process whereby new lexemes are created. This does not mean, however, that the issue has gone unnoticed, as shown by previous research on the topic (e.g. Marchand 1963 and 1964 for specific criteria on directionality in conversion, Clark & Clark 1979 for semantic patterns between nouns and verbs, and Bergenholtz & Mugdan 1979, Cetnarowska 1993, Don 1993, Štekauer 1996, Bauer & Valera 2005, Umbreit 2010, Bram 2011, or Lohmann 2017, for discussion of these or of related issues).

This paper treats such examples as cases of conversion, understood as an asymmetrical process (Sanders 1988: 155, Plank 2010: 82), in which two paronymically-related lexemes are involved (Cruse 1986, Valera & Ruz 2020). Other interpretations are possible where directionality is not a theoretical issue, e.g. word-class as underspecification (as in Whorf 1945 and, more recently, Farrell 2001).

Taking a lexico-semantic approach, this paper applies Marchand’s (1964) synchronic criteria of content for the establishment of directionality to a sample of formally identical simplex lexemes of the word-classes noun and verb. Unlike Bram (2011), this paper does not consider the chronological criterion of attestation for the identification of directionality. The aim is to test how well semantic criteria and frequency data of nouns and verbs may serve as criteria for the linguist to determine the directionality in these cases, if directionality is viewed, as argued by Plank (2010), as a matter of senses, not of lexemes. This means the following criteria are applied to all the senses attested in the \textit{British National Corpus (BNC)} for the sample described further below:

i. Semantic dependence
ii. Semantic pattern
iii. Semantic range
iv. Restrictions of usage
v. Frequency of occurrence

The data sample was obtained from the \textit{BNC} lemmatized frequency list, after automatic extraction of the lemmas recorded as both noun and verb. The list of noun/verb lemmas in the \textit{BNC} was manually cleaned for mistakes, and sampled to obtain a manageable sample of data for our study. Specifically, the sample consists of 50 lemmas within the frequency range 1–1000, as one word-class or the other. The above criteria were applied to the concordances where the lemmas are annotated as nouns and then to the concordances where they are annotated as verbs. For the criteria of semantic dependence, pattern, range and restriction of usage (classification as dialectal, colloquial, etc.), the paper relies on the information recorded in the \textit{Oxford English Dictionary (OED)}. As semantic definitions alone are typically signalled as vague or insufficient for the study and separation of word-classes, and in any case this diverts from Plank’s (2010), our paper lays emphasis on semantic classification according to the senses attested by the concordances of each sample entry. For the criteria of restriction of usage (distribution across \textit{BNC} source types) and frequency of occurrence, the paper relies on the information supplied by the \textit{BNC}, again, separated by senses instead of by lemmas.

References
Bauer, Laurie & Salvador Valera (2005), \textit{Approaches to conversion/zero-derivation}, Münster: Waxmann.
Bergenholtz, Henning & Mugdan, Joachim (1979), Ist Liebe primär? Über Ableitung und Wortarten,


Cetnarowska, Bożena (1993), The Syntax, Semantics and Derivation of Bare Nominalisations in English, Katowice: Universytet Śląski.


Don, Jan (1993), Morphological conversion, PhD dissertation, Utrecht University.


Lohmann, Arne (2017), Phonological properties of word classes and directionality in conversion, Word Structure 10(2), 204–34.

Marchand, Hans (1963), On a question of contrary analysis with derivationally connected but morphologically uncharacterized words, English Studies 44, 176–87.


Whorf, Benjamin L. (1945), Grammatical categories, Language 21(1), 1–11.

Frontable and unfrontable foci

Vieri Samek-Lodovici
(University College London)

Keywords: fronting, contrastive focus, corrective focus, right dislocation, Italian

Under Rizzi (1997, 2004), contrastive foci may always front left-peripherally. Unfrontable contrastive foci should not exist. Under Samek-Lodovici (2015), contrastive foci remain in-situ. They only front when the clause containing them is discourse-given and, as such, allowed to right-dislocate into rightmost position. Which analysis is empirically supported?

Bianchi et al.’s (2015) data favor the right-dislocation analysis. Sentences (1B-1D) are replies to (1A) focusing ‘MILANO’. Fronting is possible in (1B), where focus is followed by a discourse-given
clause. If the new information *fisica* (physics) is present, focus occurs in-situ (1C) but cannot be fronted (1D).

(1) A: Gianni studia a Roma. C: No. Studia fisica a MILANO.
   *John studies in Rome.* No. *(He)studies physics in Milan.*
   No. *In Milan, (he) studies.* No. *In Milan, (he) studies physics.*

The right-dislocation analysis accounts for these data. In (1B), right-dislocation of ‘*pro studia*’ (he studies) is licensed because the corresponding proposition ‘John studies somewhere’ is entailed by (1A) and hence discourse-given. Its right-dislocation determines focus-fronting in (1B). In (1D) right-dislocation of ‘*pro studia fisica*’ is unlicensed because ‘he studies physics somewhere’ is not entailed by (1A) and hence not discourse-given. Consequently, focus-fronting cannot occur. Rizzi’s analysis, instead, incorrectly predicts focus-fronting in (1D) to remain possible.

Bianchi (2013) and Bianchi and Bocci (2012) rescue Rizzi’s analysis by proposing that (1B) involves corrective focus and (1C) contrastive focus, and that only corrective foci front. They also maintain that the clause following corrective focus is always discourse-given because correcting propositions must match the corrected ones but for focus. Only contrastive foci, with no match requirement, can add new material.

My talk shows that their analysis is not tenable. If corrections require a match, then (1C), with its unmatched new material, should lack a corrective interpretation, contrary to observation. Furthermore, new data show that alternation (1) extends to confirmative foci, which lack a corrective import. Consider (2), where replies (2B)-(2D) focus ‘ROMA’. As before, focus fronts in (2B) where the rest of the clause is discourse-given, but not in (2D) where new information is added and right-dislocation blocked. Since focus is confirmative, the semantics of correction cannot account for this paradigm.

(2) A: Gianni studia a Roma. C: Si. Studia fisica a ROMA.
   *John studies in Rome.* Yes. *(He)studies physics in Rome.*
   B: Si. A ROMA, studia. D: *Si. A ROMA, studia fisica.
   Yes. *In Rome, (he) studies.* Yes. *In Rome, (he) studies physics.*

The key factor across (1)&(2) is the discourse-given status of the clause following focus. Fronting is possible in (1B&2B) because the clause is discourse-given, thus licensing its right-dislocation. In (1D&2D), instead, the added material leaves right-dislocation unlicensed, preventing fronting.

The above observations will be extended to additive focus alternations, enabling a unified analysis across corrective, confirmative, and additive foci. I will also show that the optionality of focus fronting follows from the optionality of right-dislocation.
References

Past Anterior in Xhosa as a Marker of Contradiction

Stefan Savić
(Rhodes University)

Keywords: Xhosa, TAM, Anterior, Perfect, contradiction

The interaction between TAM markers and information structure is a largely understudied area. In the present paper I investigate the cases in which the Past Anterior (Past Perfect) in Xhosa signals that the information about the event contradicts the hearer’s current knowledge. The Past Anterior in Xhosa contains a Past tense of the auxiliary verb -ba (‘be’) combined with the Anterior suffix -é (alternatively -ile). Its primary function is to refer to events that occurred before another event is well-documented in the literature on Xhosa and the other Nguni languages, e.g. in Bennie (1939), McLaren (1936), Louw & Jubase (1963), Pahl (1983), Einhorn & Siyengo (1993), Posthumus (1988). For example, in (1) the narrator’s event of arriving at his friend’s home represented by the Past Anterior form bendifiké (‘I had arrived’) temporally precedes the event(s) of said friend’s changing his shirt.

(1) Ú-y-é wa-yi-khulu=a le bendi-fik-é e-yi-nxib ile wa-z-a
SM.1-go-ANT SM.1.PSTCONS-OM.9-take=off-FV DEM.9.PROX
SM.1SG.be.PST-arrive-ANT SM.1.PTCP-OM.9-put=on-ANT.DISJ
SM.1.PSTCONS-come-FV
wa-nxib-a le.
SM.1.PSTCONS-put=on-FV DEM.9.PROX

He took off the one (shirt) he wore (had put on) when I arrived and put on the other one.

(Roux et al. 2001)

However, corpus data (Roux et al. 2001, Snyman et al. 2012, Eiselen & Puttkammer 2014) show that the Past Anterior often does not denote events which precede another past point of reference. Thus, in (2), an excerpt from a magazine article, a previously published information about a food product is
corrected: The Past Anterior besihalé (‘we wrote’) suggests that the published information about the product (which is part of the reader’s current knowledge) will be contradicted in the upcoming discourse.

(2) Ku-shicilelo lwe-thu lu-ka-Novemba=2006 besi-bhal-é
SM.17-11.issue POSS.11-1.PL POSS.11-1a-November=two=thousand=and=six
SM.1.PL.be.PST-write-ANT
nge-zi-dlo zee-ntsana sa-z-a sa-balul-a
INSTR-8-meal POSS.8-10.baby SM.1.PL.PSTCONS-come-FV SM.1.PL.PSTCONS-
mention
u-ku-ty-a kwee-ntsana kwa-kwa-Woolworths oku-tsha
RM.15-new
no-ku-phambili. S-enz-é i-mpazamo kw-elo njaku sa-th-i
5.article SM.1.PL-say=TV
oku ku-ty-a “ku-ne-zi-gcinakaliso”.
DEM.15.PROX 15-food-FV SM.15-COM-8-preservatives
Yi-mpazamo leyo.
COP.9-9.mistake DEM.9.MED
In our November 2006 issue, we wrote about baby food and we mentioned baby food from Woolworths, both the latest types and earlier types. We made a mistake in that article saying that “this food has preservatives”. This was a mistake.
(Snyman et al. 2012)

The Past Anterior may also indicate that the event (or its result) counters the addressee’s knowledge, statement or actions: In the elicited example (3), besithethé (‘we talked’) is appropriate as a form that reminds the hearer of the previously made agreement which she or he is about to breach.

(3) Besi-theth-ile nga-yo kaloku le nto!
SM.1.PL.be.PST-talk-ANT.DISJ INSTR-9 however
DEM.9.PROX 9.thing
But we talked about this!

The data suggest that the secondary function of the Past Anterior in Xhosa is to indicate that an event’s result contradicts the surrounding discourse or the interlocutor’s current state of knowledge. This function of the Past Anterior can be linked to its property in narrations (primary function) to denote results of events that are often cancelled (reversed) and that do not affect the flow of the main plot.

References:
Nominal causal constructions: Capturing variation across Slavic

Sergey Say
(Institute for linguistic studies, Russian Academy of Sciences)

Keywords: causal constructions, prepositions, Slavic, parallel corpora

Nominal causal constructions are constructions that express a causal relation between two events, in which the causing event is represented by a noun phrase. In Slavic, these noun phrases bear dependent markers, such as prepositions and/or case markers (cf. Ru заплакать от боли, Cz vzlykat bolestí, Bg разплача от болката ‘to burst into tears from the pain’).

Each Slavic language possesses an array of constructions that meet the definition above but differ in the intricacies of their semantic profile. The relevant factors have been thoroughly described for some Slavic languages, especially Russian (Iordanskaya, Mel’čuk 1996, Levontina 2003), but an analysis covering all branches of Slavic is lacking. The goal of this study is to compare the systems of nominal causal constructions across Slavic, to identify and quantify the differences between them, and to interpret the results in areal and diachronic perspectives.

The data were gathered using a parallel corpus (ParaSol). I compiled a database of 231 nominal causal contexts in 11 languages and annotated them for cognacy sets, such as e.g. Ru от — Br ад — Uk еди — Pl od etc. Based on these sets, it was possible to build a distance matrix for Slavic languages using relative Hamming distance (the ratio of contexts where languages employ non-cognate causal markers). This matrix is visualized in Fig. 1, using multidimensional scaling as implemented in cmacof package (de Leeuw, Mair 2009) for R (R Core Team 2019).

The profiles of Czech and Polish are clearly most divergent in Fig. 1. This deviant behaviour is mainly attributed to the fact that Czech tends to use its Instrumental case, and Polish, the z + genitive construction in those direct causal contexts where most other Slavic languages employ apudelative (‘from’) prepositions, such as Russian on etc. Apart from this, expressing direct causes (such as ‘cough from the smoke’, ‘moan with fear’ or ‘do something out of curiosity’) is fairly stable across Slavic. By contrast, expression of indirect causes (as in ‘he was late because of his brother’) is unstable: many different secondary markers have been recruited for that purpose in various languages. Similarities in this domain testify to small-scale genetic or contact-induced convergences, including calques.

These findings called for a method to measure similarities between systems independently of etymologies of markers. I propose such a distance metric based on Mutual Information (MI) and, ultimately, entropy. Applying this measure to my data showed that the factors governing the choice between causal constructions are diachronically more stable than the very markers involved. In Czech, for example, the distribution between competing constructions is very similar to that observed in other
Slavic languages, even though the basic direct causal marker is non cognate to functionally similar markers elsewhere. The one Slavic language where the very organization of nominal causal markers is divergent is Slovenian; here, the use of *zaradi* has blurred the basic distinction between direct vs. indirect causes.

Fig. 1. Visualization of differences between Slavic languages in terms cognacy of nominal causal markers

Acknowledgment: this study was supported by a grant from the Russian Science Foundation “Causal Constructions in World Languages (Semantics and Typology)”, grant No. 18-18-00472.

References


ParaSol: A Parallel Corpus of Slavic and other languages (http://parasolcorpus.org/)


Accommodation biases:

Loan adjectives in Dutch and Middle English
Marlieke Shaw  
(University of Leuven)

Keywords: historical linguistics, corpus linguistics, loanword accommodation biases, contact-induced syntactic change, loan adjectives

Loanwords are accommodated to the grammatical structure of their recipient language (Poplack et al. 2009). According to Wohlgemuth (2009), the dominant strategy to achieve that is direct insertion, which adds native inflections directly onto borrowed stems. However, in a previous case study on loan verbs we demonstrated that English-origin verbs in Present-Day Dutch, and French-origin verbs in Late Middle English (Late ME) are more prevalent in non-finite than in finite forms. For Dutch, it was additionally found that loan verbs favour inflectional categories that require no formal marking. Therefore, loan verbs may, even under direct insertion, be biased to certain native inflectional categories (cf. De Smet 2014 and Shaw 2019).

In the present paper, we look into accommodation biases, this time focusing on adjectives. We try to assess whether, like loan verbs, loan adjectives are preferred in any specific forms of the paradigm, and whether that applies to both English-origin adjectives in Dutch, and French-origin adjectives in ME. We carried out a corpus analysis, using a sample from the Corpus Gesproken Nederlands (Nederlandse Taalunie 2004) for Dutch and a fourteenth-century sample from the Helsinki Corpus (Rissanen et al. 1991) for Late ME. The attestations were manually annotated for their origin (Germanic vs. Romance), position (attributive vs. predicative), inflection (-e vs. zero), number and definiteness. For Dutch, we observe that loan adjectives are more prevalent in predicative than in attributive position, presumably since adjectives in predicative position are not inflected. In attributive position, where inflections alternate between zero and -e, zero inflection is the dominant form. The same inflectional bias is found for French loan adjectives in Late ME. Thus, for loan adjectives speakers rely increasingly on predicative forms (1) compared to attributive forms (2), where an -e is required:

(1) So _hydous_ was the _noyse_  
So hideous was the noise  
(c1390, PPCME2, Canterbury Tales, emphasis added)

(2) _many tempestes & many hidouse thondres & leytes_  
many tempests & many hideous thunderclaps & lightnings  
(c1400, PPCME2, Mandeville’s Travels, emphasis added)

It is remarkable, however, that the systematicity of the use of inflections in ME adjectives is highly dependent on the particular text in which they occur. Chaucer’s Parson’s Tale (c1390), for instance, is more systematic in its use of inflections than Trevisa’s Polychronicon (1387), where variation is across the board. Furthermore, we observe that French adjectives (e.g. noble) co-occur more often with French heads than their Germanic equivalents (e.g. highborn), suggesting that French adjectives tend to be integrated in larger French-origin phrasal units.

In this case study, we provide more evidence that direct insertion is an idealisation (De Smet 2014, and Shaw 2019), expanding our findings for inflectional biases in verbs to adjectives. Not only can this shed light on (strategies for and constraints on) loanword accommodation in different parts of speech, it can also give us a better understanding of contact-induced syntactic change.

This research was funded by grant G0D1418N from Fonds voor Wetenschappelijk Onderzoek (FWO) – Vlaanderen.
References

Multimodal double-voicing in Israeli Hebrew discourse

Leon Shor & Michal Marmorstein
(The Hebrew University of Jerusalem)

Keywords: double-voicing, multimodality, quotation, co-speech gestures, Israeli Hebrew

Co-speech gestures contribute in various forms to the meaning of the utterance of which they are a part (Kendon 2004). In this talk, we explore the use of the Quoting Gesture (QG)—typically realized by raising and flexing the index and middle fingers of both hands twice (Cirillo 2019)—in Israeli Hebrew discourse. Based on a collection of instances coming from television interviews, we analyze the formal realizations of QGs, their functions, and their temporal alignment and interplay with co-produced verbal segments.

While their gestalt is rather fixed, QGs present some variation in the number of hands used and flexes made by the index and middle fingers. Concomitantly produced with short nominal-phrases, QGs have an interpretive function in implicitly attributing these phrases to some other, concrete or imagined, speaker and thereby distancing the actual speaker from the cited words. QGs serve then as a stancetaking device which communicates a dissonant voicing or ‘contrastive perspectivation’ (Kotthoff 2002) of the utterance.

For example, in the cited excerpt from the TV program Soul Talk, the speaker describes her strong feelings as a child toward her divorcing parents due to the belief that her mother was abandoned by her father. Co-extensively with the verbal segment ke’ilu ha-nin-- nintešet ‘like the abandoned [wife]’ in
lines 7–8, the speaker produces the QG by raising and flexing the index and middle fingers of both hands once. The QG targets the adjectival-phrase *ha-nintešet* ‘the abandoned’ and demarcates it as someone else’s characterization, the mother in this case, with whom the speaker as an adult disaligns. Notably, the QG can operate in concert with other verbal resources, in this case the discourse maker *ke’ilu*, which indicates a lesser commitment of the speaker to the literal meaning of the phrase (Maschler 2009). The contrastive evaluation implicit in the use of the QG is given an explicit expression in further talk, when the speaker in line 9 makes the metalingual comment *be-merxaot hayom ani omeret et-ze* ‘I am putting quote marks around this today’.

Carried over from the graphic to the kinesic medium, we observe that the QG retains only a specific use of quoting marks—the scare quotes (Lampert 2013)—which involves relatively short and embedded segments. We suggest that this may have to do with the constraints of co-extensive production, as well as with the rich ecology of quoting or voicing markers in spoken discourse which include, besides the gestural, other verbal and prosodic resources.

**Example**

![https://www.youtube.com/watch?v=h3unNjL4BLQ&t=206s; 03:36-03:50](https://www.youtube.com/watch?v=h3unNjL4BLQ&t=206s; 03:36-03:50)

1. *hayu li harbe*,
   I had many
2. …*keasim o e*,
   anger issues or uh,
3. *txusot*,
   feelings of,
4. *em*,
   uhm,
5. *hizdahut meod gdola im ima šeli*,
   great empathy with my mother,
6. *betor yalda*,
   as a child,
7. ⟹ *še hi hayta % keilu ha-nin--*
   that she was like the aban—
   % flexes index and middle fingers of
   both hands once -->
8. ⟹ *# nintešet.%*
   abandoned (wife)
   ---> %
   # fig.1
9. *be-merxaot hayom ani omeret et-ze*,
   I’m putting quotes around that now,
System internal vs. sociolinguistic factors of variation in case systems

Kaius Sinnemäki & Francesca Di Garbo
(University of Helsinki)

Keywords: case systems, word order, population size, sociolinguistics, linguistic typology, hurdle models

We revisit earlier work on language evolution according to which the sociolinguistic context in which languages are learned and used may affect the development and loss of morphology. Lupyan & Dale (2010) argued that the number of cases depends inversely on population size, while Bentz & Winter (2013) argued that it depends inversely on the proportion of second language (= L2) speakers. However, it is also well-known that case systems are more likely to develop in verb-final than in SVO languages.

We evaluate the competing effects of word order, population size (n = 251), and proportion of L2 speakers (n = 80) on the number of cases. The data on cases come from Iggesen (2013), those for word order from Dryer (2013), and the sociolinguistic data largely from Lewis et al. (2016). We model sociolinguistic factors in competition with VO (verb object) vs. non-VO word order and use hurdle models (Fournier et al. 2012, Skaug et al. 2016) to address the high number of zero counts for number of cases. We thus focus separately on zero vs. non-zero counts (binary models) and then on the number of cases in languages with non-zero counts (truncated binomial models). Language family and geographical areas are modeled as random intercepts.

In the binary models, the probability of case is lower in VO compared to non-VO order in large languages, but the difference decreases as population size decreases (Fig. 1; A). The probability of
case is also influenced by VO order (Fig. 1; B) and the proportion of L2 speakers (Fig. 1; C). In the truncated negative binomial models, the number of cases is marginally affected by VO order (log est. = -0.3±0.2; p=0.053) but not by population size (log est. = -0.02±0.02; p=0.4) or the proportion of L2 speakers (Fig. 1; D).

The results suggest that sociolinguistic factors affect the probability of case but not the number of cases. Overall, measuring complexity as the number of distinctions in a linguistic domain does not seem a good predictor when evaluating sociolinguistic effects on language evolution (also Dahl 2019). We argue that the evolution of morphological complexity should be investigated beyond “inventory size”, studying how grammatical distinctions evolve as a response to both language-internal (here word order) and external factors (here demography). We expect that semantically transparent patterns of encoding emerge under the pressure of intense language contact (Kusters 2003, Trudgill 2011), so that, for instance, gender systems driven only faintly by semantics become clearly based on animacy distinctions (Verkerk & Di Garbo in preparation). Opposite developments, where complex language structures emerge or are reinforced through contact, cannot be excluded either. These contrasting tendencies towards contact-induced reduction, retention or bolstering of complexity could be the result of the interplay between internal pressures of change and the sociolinguistic contexts where variation and change occur.

Fig. 1. Marginal effect plots for the predictors in selected hurdle models.

Sinnemäki and Di Garbo have received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 805371).

References


**Slavic grammaticalization chains:**

*Selected issues of categorial and conceptual shifts in grammaticalization pathways for function words*

Piotr Sobotka & Magdalena Żabowska
(Institute of Slavic Studies of the Polish Academy of Sciences, Nicolaus Copernicus University)

Keywords: semantic change, syntactic change, grammaticalization chains, function words, Slavic languages

The term ‘grammaticalization chain’, introduced by Craig (1991) and refined by Heine (1992), refers to continuous changes whereby grammaticalized items emerge from autonomous or less grammaticalized items – e.g. content item > grammatical word > clitic > inflectional affix (as proposed in Hopper, and Traugott 2003: 7). Such pathways from autonomous items to grammatical words have not enjoyed as much scholarly attention in Slavic as, for instance, in Germanic, though there have been many detailed investigations of Slavic grammaticalization processes (cf. Nikolaeva 1985, Starodumova 2002, Wiemer 2011, Nomachi, Danylenko, and Piper 2014, Kleszczowa 2015, Wiemer 2015, Hansen, Letuchiy & Błaszczyk 2016).

The paper aims to present both the most frequent and the cross-linguistically unique patterns of Slavic grammaticalization in chosen group which result in function words, e.g. prepositions, conjunctions, intensifiers, approximators, particles, discourse markers. We explore grammaticalization pathways in categorial as well as conceptual domains. Analysis of Slavic data offers insight into the universality of the patterns in question at a conceptual level. Since the development of some Slavic function words differs from their equivalents in other languages, certain inventories of
grammaticalization chains can be distinguished. Another question is whether we can disjunctively classify these changes as purely semantic or purely syntactic.

Our proposed method offers a comprehensive semasiological approach to the development of Slavic grammatical word categories (such as pronouns or numerals) that serve as the basis for function words. It also brings an onomasiological analysis of the concepts (i.e. the specific values of grammatical categories) that serve as the links in grammaticalization chains.

We illustrate this by discussing grammatical category shifts according to the pattern such as (1) *pronoun > indefinite pronoun > approximator* and (2) *cardinal numeral > inverted cardinal numeral > approximator* with a postpositive fixed word order (cf. Rus. *Ona poluchila pjal’ rublej*. ‘She recieved five rubbles’ > *Ona poluchila rublej pjal’. ‘She recieved approximately five rubbles’). The first pattern can be implemented using different markers, cf. (1a) *V + ‘somewhere’ > ‘somewhere’* (Pol. *gdzieś*) + *PLACE > approximator* (Pol. *gdzieś*):

(1a) Pol. *Wyjchal gdzieś (ale nie wiem gdzie).*
he left somewhere (but I don’t know where)

or (1b) *V + ‘something’ > ‘something’* (Pol. *coś*) + *V\_FACT > approximator* (Pol. *coś*):

(1b) Pol. *Boli mnie coś.*
There’s hurting me something

resulting in semantically/conceptually distinct units of language. Next we examine the conceptual shifts of the lexis in question, cf.

```
approximation
< lack of knowledge < indeterminacy (as Pol. *gdzieś* ‘somewhere’)
< indeterminacy (cf. Lehmann 1992: 403f.) < number (as Rus.
< etc.
```

We outline a map of selected categorial and conceptual shifts based on a detailed analysis of Slavic data, comparing this with the claims of Kuteva et al. (2019) and showing the mutual relations between source and target categories. We propose an inventory of these categories and reconstruct their development and origin.

References


Micro-variation in a larger picture: Feature generalisation in Macedonian short pronouns

Barbara Sonnenhauser & Anastasia Makarova
(University of Zurich & University of Zurich)

Keywords: 3rd person pronouns, agreement features, object indexing, variation, Macedonian

The short forms of personal pronouns of Standard Macedonian distinguish SG and PL and, for 3SG, M/N vs. F. All forms may be used referentially, as *im* in (1). The ACC and DAT also function as direct and indirect objects indices, see *i* and *go* in (2).

(1)  
Im  rekov.  
DAT.3PL. say.AOR.1SG  
‘I told them.’ (makedonski.info)

(2)  
[svojot život],  ke  i  go,  posvetam  [na muzikata],  
my life.M.SG.DEF will DAT.F.3SG ACC.M.3SG dedicate.PRS.1SGDEP music.F.SG.DEF  
‘I will dedicate my life to music’ (makedonski.info)

In specific dialectal varieties, 3rd person short pronouns exhibit various degrees of gender and number generalisations. In (3), DAT.M/N.3SG *mu* anaphorically refers to *samovila*.F.SG. In (4), ACC.M/N.3SG *go* (phonological variants *gal/gu*) is used as F/M/N SG direct object index, DAT.M/N.3SG *mu* in (5) as index for F/M/N SG/PL indirect objects.
At first sight, these patterns seem straightforward contact phenomena, given that neighbouring Albanian and the Struga variety of Aromanian display this kind of feature generalisation as well, and that object indexing has been observed as being sensitive to contact situations up to the point of loss of feature distinctions on the pronominal index (see Fischer et al. 2019, 60 for Lima and Andean Spanish). Since object indices and short pronouns share the morphological material, both developments appear to be related to each other.

With the Indo-European languages of Europe preferring gender distinctions in 3rd person pronouns (e.g. Siewierska 2013) and dispreferring object indexing on the verb in their contemporary varieties, the developments in Macedonian appear remarkable. They might thus provide insight into larger historical developments, in that this dialectal micro-variation seems to result from the interaction of various conditions favouring in sum the mutual reinforcement of two otherwise volatile developments: object indexing and its generalisation into syntactic markers without any referential feature specifications, and the loss of gender/number distinction for referentially used short pronouns:

- Family: Preference for gender distinction in 3rd person pronouns; emergence / loss of object indexing in short-term cycles in Indo-European (see Dedio and Widmer 2020)
- Area: Stabilisation of patterns by synchronisation of developmental cycles in the Balkan context
- Contact: No gender distinction in 3rd person short pronouns in surrounding non-Slavic varieties
- Internal: Phonological developments in Macedonian dialects triggering the conflation of short pronominal forms (e.g. Vidoeski 1965)

The above data thus strongly support the assumption that considering the multiplicity of triggers involved in change helps to “understand the prehistory of seemingly independent but parallel developments in related languages” (Joseph 2013, 688).

In addition, facing the volatility of feature generalisation in Macedonian dialects (recent fieldwork suggests it to be vanishing again under the influence of the standard), this paper makes a case for considering also fine(st)-grained data. It is exactly such small-scale patterns that contribute to defining the larger picture based on which convenient linguistic abstractions are usually derived.

Acknowledgments: The research for this work was supported by the Swiss National Science Foundation (SNSF) grant Nr. 100015_176378/1. This support is gratefully acknowledged.
References
Vidoeski, Božidar (1965), Zamenskite formi vo makodonskite dijelekti [Pronominal forms in Macedonian dialects], *Makedonski jazik* XVI, 25–73.
Vidoeski, Božidar (2000), *Tekstovi od dialektite na makedonskiot jazik* [Macedonian dialect texts], Skopje: Institut za Makedonski Jazik “Krste Misirkov”.

Telic structure of visual, auditory and olfactory source-based perception verbs in Polish

Przemysław Staniewski & Adam Gołębiowski
(University of Wroclaw & University of Wroclaw)

Keywords: atelicity, telicity, perception, Polish, semantics-syntax interfaces

The paper focuses on the impact of the following parameters on the (a)telicity of source-based perception verbs in Polish: subject and complement semantics, nature of the sensory modality and the adopted perspective (source vs. experiencer). Depending on the configuration of the abovementioned parameters, the source-based verbs can be viewed as telic or atelic.

Although syntax and semantics of perception verbs (e.g. Gisborne 2010 and references therein) as well as the distinction between telic and atelic situations (e.g. Sasse 2002) have attracted a widespread attention, the phenomenon of (a)telicity in the perceptual domain has not been considered so far. Rogers (1974) and Viberg (1984, 2001) touched upon this problem only implicitly by classifying the experiencer-based verbs as states and (inchoative) achievements. The source-based verbs remain, however, unexplored.

We adopt the definition of telicity as presented in Wiemer & Seržant (2017: 249, cf. Wiemer & Laziński 1995; Dahl 1981) as “an inherent feature of a verb lexeme or a predicate that makes the denoted situation imply an inherent endpoint, regardless of whether this endpoint is realized or not”. For analytical and illustrative purposes, we employ Croft’s (2012) model for analysis of aspect and event structure. Compare the following examples:

(1) Widać dom-ø.
  to be seen.INF house-ACC.SG.M
  ‘A house can be seen’ – atelic

(2) Słychać strzelanin-ę.
  to be heard.INF gunfire-ACC.SG.F
  ‘Gunfire can be heard’ – telic

(3) Słychać pieśń-ø.
  to be heard.INF song-ACC.SG.F
  ‘A song can be heard’ – telic
Based on introspective and corpus data, we claim that the deployed parameters highly influence the telic structure of the analysed verbs, cf. (1) – (5). (2) – (4) demonstrate different (a)telic readings due to the occurring complements: (2) is telic with the implied endpoint, while (3) is telic with focusing on reaching the endpoint as the song is inherently temporally constrained. In (4), neither of the abovementioned endpoint readings is implied and therefore the sentence gains an atelic interpretation. In (5), the telic-atelic interpretation depends on the adopted perspective which relates to the nature of olfactory modality. Taking viewpoint of the source (which is in line with the syntactic structure) the sentence can be read as telic – the flowers give off a smell for an unbounded time interval. Therefore, the stimulus is independent from the experiencer. The experiencer cannot be, however, excluded from the perception event, even if it does not appear on the sentence level. Thus, if the flowers smell, the experiencer simultaneously smells them. This allows shifting to the experiencer perspective. Accordingly, (5) can be read as telic due to the nature of olfactory modality – the experiencer gets adapted to the smell and eventually ceases to perceive it after a short period of time. By contrast, in visual modality, cf. (1), the stimulus is persistent in time and perceivable regardless of perspective. This lends the sentence an atelic reading only.

References
Keywords: demonstratives, constructional variation, corpus study, multivariate analysis, Estonian

Introduction. Demonstrative pronouns (e.g. this) and adverbs (e.g. here) have gained much research interest, particularly from a typological perspective (Diessel 1999, Coventry et al. 2008, and Levinson et al. 2018). This research into demonstratives has shown that demonstrative choice (e.g. choosing between a proximal and distal term, such as English this and that) can be influenced by many factors (e.g. distance, salience and visibility of the referent), and also by discourse factors. However, little is known about what triggers the choice between demonstrative pronouns and adverbs, particularly if they can occur in similar syntactic positions. This is also the case of Estonian where both demonstrative pronouns inflected for spatial cases (example 1) and demonstrative adverbs (example 2) can be used as determiners to express seemingly similar meanings:

(1) Selles maja on kohvik.
\( \text{this-INE house-INE be.3SG café.NOM} \)

(2) Siin maja on kohvik.
\( \text{here.LOC house-INE be.3SG café.NOM} \)

‘There is a café in that house.’

Aim and research questions. The aim is to investigate Estonian noun phrases (NPs) in which a demonstrative pronoun or adverb occupy the determiner position. There are two research questions:

(i) What factors influence the choice of a demonstrative as a determiner?
(ii) Are there any differences in the functions of demonstrative pronouns as compared to adverbs as determiners?

Data. The data contains two main constructions: demPronNP (a demonstrative pronoun used in an NP; e.g. selles majas ‘in this house’) and demAdvNP (a demonstrative adverb used in an NP; e.g. siin majas lit. ‘in here house’). These two constructions have many sub-constructions, as both pronouns and adverbs have two terms (distal and proximal) in Estonian. Furthermore, all nouns in NPs can be inflected for six spatial cases (ablative, adessive, allative, elative, inessive, illative). The data represents all these sub-constructions and consists of 2400 clauses taken from the Estonian National Corpus 2017 via Sketch Engine.

Method. Conditional random forests and inference trees (Tagliamonte & Baayen 2012) are used to analyse the data. The dependent variable has two levels and shows whether a demonstrative pronoun (demPronNP) or an adverb (demAdvNP) occurs in the NP. The independent variables specify the meaning of the noun in NPs (concreteness, mobility, animacy and size of a potential referent), noun frequency, verb meaning and placement in a clause, and text type.

Results. Results indicate that demonstrative pronouns as determiners are inclined towards abstract nouns (including time; e.g. sel ajal ‘at this time’), and adverbs as determiners towards concrete nouns (including locations; e.g. seal poes ‘in that shop’). That being said, pronouns frequently combine with nouns of animate entities (e.g. sellele mehele ‘to this man’), whereas adverbs refer to animate entities only if expressed as a group of people (e.g. siin kambas ‘in this gang’). As such, demonstrative adverbs as determiners seem to bond more to spatial functions, whereas demonstrative pronouns are used in a wider range of contexts. Moreover, adverbs have a slight tendency to be used in informal and pronouns in formal texts.

References
How does a verb become a modal? Diachronic development of the newcomer necesitar ‘need to’ in the Spanish modal system

Miriam Thegel
(Uppsala University; KU Leuven)

Keywords: modality, Spanish modal auxiliaries, emergent modals, subjectification, semantic evolution

The Spanish modal auxiliaries have been a frequent object of study for several decades (e.g. Olbertz 1998, Fernández de Castro 1999, and Cornillie 2007), but more recently, there has been an increased interest in their diachronic development (e.g. Garachana Camarero 2017, and Zielinski & Espinosa Elorza 2018). In recent publications, the focus is mainly on the core members of the modal system, such as deber ‘must’, haber delque ‘have to’, poder ‘can’, and tener que ‘have to’, whereas peripheral members have fallen outside the scope. In this study, the diachronic development of the modal necesitar ‘need to’ is examined. This verb is mentioned only in passing (Gili Gaya 1961, and Fernández de Castro 1999), or not referred to at all (Olbertz 1998, and Garachana Camarero 2017).

The construction necesitar + infinitive can be considered a newcomer in the Spanish modal system. Necessitar was borrowed from Latin into Spanish in the Middle Ages (RAE: www.rae.es) as a lexical verb with the meaning ‘to oblige’. As indicated by data from Corpus del Español and CORDE, it was first used as a modal verb in the 16th century and increased notably in frequency during the 18th and the 19th centuries. When necesitar developed a modal use there were already several verbs in the system with meanings related to necessity, namely deber, haber delque and tener que (Garachana Camarero 2017). How come there was space for yet another modal auxiliary? Did necesitar add a meaning not already covered by the existing ones?

To answer these questions, randomly selected samples of deber, haber delque, necesitar and tener que have been extracted from Corpus del Español from the 16th century onwards to study the behavior of necesitar and its interaction with the other functional equivalents. A quantitative approach is used when describing the syntactic environment of necesitar, whereas a fine-grained qualitative analysis is carried out to classify the meanings of necesitar and other modals related to the notion of necessity. The framework used is inspired by perspectives combining cognitive and functional linguistics, such as Bybee & Torres Cacoullos (2009 and Byloo & Nuyts (2014).

An examination of cases of necesitar from 1500–1900 shows the increasing frequency of a particular schematic construction, namely necesitar X para Y ‘need (to) X to Y’, as in Necesitó de apoyarse sobre su hombre para sostenerse en pie ‘He needed to lean on his shoulder to stand on his feet’ (Davies, 18th century). The hypothesis is that this construction has been central for the semantic
development and establishment of *necesitar* in the modal system, as it added an intentional, subject-driven modal meaning that did not exist before.

When taking into account not only core members of the modal system but also emergent modals, new insights can be gained within the notion of necessity, which are relevant not only for Spanish modal research but also for research on modality and language change in general.

**Acknowledgements**

This study is part of the project “Survival of the fittest: semantic evolution in the Spanish modal system” (2019-00625), funded by the Swedish Research Council.

**References**


CORDE = Real Academia Española, *Corpus diacrónico del español*, http://www.rae.es

Davies, Mark: *Corpus del Español*, http://www.corpusdelespanol.org/


The use of Livonian personal pronoun *tāmāta* and demonstrative pronoun *se* in spoken language

Marili Tomingas
(University of Tartu)

Keywords: Livonian, pronouns, demonstratives, morphology, syntax

Livonian has two forms of 3rd person pronoun: long form *tāmā* and short form *ta*, meaning ‘he/she’. It is said that for all the Livonian personal pronouns short form is more common and long form is used rarely or only for emphasizing a referent in a sentence (Sjögren, Widemann 1861: 115, Viitso 2008: 332). However, this far Livonian personal pronouns have been described briefly only in dictionaries and grammar overviews and there is no empirical research yet about the exact morphosyntactic and semantic use of them.

In addition, Livonian demonstrative pronoun *se* ‘this’ may be used when referring to animate entity instead of 3rd person pronoun *tāmāta* (example 1). Oppositely, 3rd person pronoun may be used when referring to inanimate object (example 2).
This research focuses on analyzing and comparing the data on two Livonian pronouns: 3rd person pronoun tämālta and demonstrative pronoun se which can refer to both proximal and distal objects and is the main demonstrative pronoun used in Livonian. As pronouns are frequently used in spoken language and morphological forms and syntactic use in spoken language is especially diverse, the study is based on spontaneous spoken language data of Livonian recordings. The analyse material comes from Archive of Estonian Dialects and Kindred Languages (AEDKL). For the research, 13 recordings with 5 different Livonian speakers are analysed. The used recordings are recorded between the years 1986–2012 and are in the form of dialogues between researcher(s) and native Livonian speakers. In the research, only the data of native speakers is used.

In the presentation, I will analyse these pronouns morphologically and syntactically, the main research questions are: 1) How does the use of 3rd person pronoun long and short forms vary? 2) In which morphological forms are pronouns used in spoken language? 3) In which syntactic functions do the pronouns occur? For the research, 220-230 examples of both pronouns are collected and analysed. The analysis is both qualitative and quantitative, presenting morphological forms and syntactic functions in the data descriptively and how often do certain forms and functions occur in the data.

It is expected that short nominative form of 3rd person pronoun occur more often, but the long form is probably used only about persons and not inanimate objects (similarly to closely related language Estonian). The first results also show that both pronouns have some morphological forms in spoken language that differ from the forms mentioned in dictionary paradigms (e.g. Viitso, Ernštretš 2012).

References

Constraints on nominalizations:
Investigating the productivity domains of Italian -mento and -zione
Rossella Varvara  
(University of Florence)

Keywords: constraints on productivity, competition, deverbal nominalizations, derivational morphology, quantitative analysis

Despite the differences in terminologies and theories, scholars tend to agree that a pattern (or rule) of word formation usually applies to a subset of words, i.e. to a specific domain (Rainer 2005). In productive processes, it is possible to recognize features (also called constraints or restrictions) that identify the pattern’s domain. Investigating these features can shed light on the mechanisms governing word formation processes, especially in cases of competition.

Numerous works analyzed the domain of productivity for various affixes and languages (Aronoff 1976, van Marle 1986, Plag 1999, Gaeta 2004 among many others; see Rainer 2005 and Gaeta 2015 for an overview), but mainly from a synchronic point of view. Moreover, the use of quantitative analysis based on corpus data is not yet much widespread (Lignon 2013, Arndt-Lappe 2014, Varvara 2017, Bonami and Thuiller 2019).

This contribution analyzes the differences in base selection between two Italian event denoting suffixes, -mento (e.g. annegamento, “drowning”) and -zione (rotazione, “rotation”). These suffixes share the same evitative semantics, are both productive, and thus can be considered rivals in the formation of event nominalizations. However, it is still not clear how speakers choose among them in neologisms creation, if their domains of application differ and how these evolved diachronically.

Using a statistical approach (a logistic regression model with the suffix of the derivatives as response), I analyze the contribution of different features to the identification of the patterns domain. Specifically, I investigate the role of the number of characters, the frequency and the conjugation of the base verb, the presence of other affixes and the last character of the stem. The analysis is conducted on a dataset of nominalizations extracted from Midia, a diachronic balanced corpus explicitly built for morphological research and divided in 5 historical periods (Gaeta 2017). Results from the different centuries are thus compared in order to see the evolution of the suffix domain.

Preliminary results show that relative frequency, the conjugation of the base verb and the presence of the prefix a- significantly contribute to define the different domains in the last two centuries, confirming only partially previous synchronic findings (Gaeta 2004, Melloni 2007). The diachronic comparison will highlight when these tendencies started, and if other constraints led the choice of the event suffix in the past.

References

Gaeta, Livio (2004), Derivazione nominale deverbale, in M. Grossmann and F. Rainer (Eds.), La formazione delle parole in italiano, Tübingen: Max Niemeyer Verlag.
Gaeta, Livio (2015), Restrictions in word formation, in P. O. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), Word-formation: An international handbook of the languages of Europe (Vol. 2), Berlin: Mouton de Gruyter, 858-874
Plag, Ingo (1999), Morphological productivity, Berlin: Mouton de Gruyter.
Rainer, Franz (2005), Constraints on productivity, in P. Štekauer and R. Lieber (Eds.), Handbook of word-formation, Dordrecht: Springer, 335–352.

From noun classifiers to subordination:
Origins of relativizers and the complementizer Ńa in Tù’un na Ńuu Sá Mátxii Ntxè’è (Mixtecan, Otomanguean)

Albert Ventayol-Boada
(University of California, Santa Barbara)

Keywords: grammaticalization, Mixtec, nominal classification, noun classifiers, subordination
Grinevald (2000) proposes a continuum of nominal classification devices that runs from purely lexical to purely grammatical, with measure and class terms on the lexical end and gender and noun class systems on the grammatical end. This proposal allows Grinevald (2002) to generalize over attested grammaticalization patterns: class terms evolving into a classifier system, as attested in several languages of South East Asia (Bisang 1993, DeLancey 1986), and classifiers developing into a gender and noun class system, for example the Ngân’gityemmeri gender system (Southern Daly, Australia; Reid 1997) and the noun classes of Bantu languages (Givón 1970, Denny & Creider 1986).

In addition, there is evidence that classifiers are involved in the grammaticalization of other morphosyntactic categories. Song (2005) argues that possessive classifiers have grammaticalized into benefactive markers in Oceanic languages. Similarly, Epps (2008) shows that in Hup (Nadahup, Brazil) the source of the future marker can be traced back to an incipient noun classifier system of nominal origin. Finally, Rose (2019) demonstrates that classifiers are an intermediate step that explains the development of nouns into applicatives in Mojeño Trinitario (Arawak, Bolivia).

In Mesoamerica, noun classifier systems have been described for several languages in Mesoamerica (see Craig 1979, 1987, England 1983, Grinevald 2000, 2002 and Zavala 2000 for Mayan; and Costaoeuec & Swanton 2015, deLeón 1988, Macaulay 1996, Small 1990, Suárez 1983, Veerman-Leichsenring 2004 for Otomanguean). In this paper I present an analysis of the noun classifier system in Tù’un na Ńuu Sá Mátxixi Ntxè’è (Otomangean), a Mixtec variety spoken in the community of San Martín Duraznos in Oaxaca (Mexico). The analysis is based on 8 monologic narratives collected as part of the documentation of the language. This paper proposes a path of grammaticalization from different generic nouns in the language to noun classifiers, to third person pronouns, thence to relativization, and finally to complementization. In [1] =nã functions as a dependent pronoun, while in [2] the classifier and relativization functions can be observed. Examples [1] and [3] show some different functions for the non-human noun classifier =nã, i.e. pronominal,
(headless) relative clause in a cleft construction, and complementization. I argue that each step on the grammaticalization cline is functionally motivated and intermediate stages can be identified in synchronic uses. This study contributes to the overall understanding of noun classifier systems and their grammaticalization into other morphosyntactic categories beyond nominal classification.

[1] Ta`a sæ’a-na kii ña,

\begin{verbatim}
  taa ña sæ’a=na_i kii ña
  and REL.3GENR IPFV.do=3PL IPFV.COP COMP
\end{verbatim}

‘And what they do is that’


\begin{verbatim}
  nti=`i=na_i=máxìkà=ji nj sìká’a
  IPFV.search=3PL CLF.H.PL=music REL.PL POT.play
\end{verbatim}

‘they look for a band that will play.’ (SMD-0004-FiestaPatronal)


\begin{verbatim}
  sáá na kà’=ni=ji’i=án=ákivi súvá’=án=ín ña=lo’o
  then PRT POT.say=1SG with=3SG.F_k Q=IPFV.be.able POT.make=3SG.F_k one 3GENR_i small
  kivi kii’in Sándrá
  IPFV.be.able POT.grab Sandra
\end{verbatim}

‘I’m going to ask her if she can make one small [bag] for Sandra.’ (SMD-0012-VisitaMamá)

References


---

**The linguistic categorization of actions: an image-schematic approach**

Paola Vernillo & Rossella Varvara
(University of Florence)

Keywords: image schemas, action verbs, lexical closeness, action categorization, semantic representation

Action verbs may refer to ontologically and cognitively different action events. The English verb *to turn*, for instance, encodes both events in which an agent makes something move in a circle (e.g. *Mary turns the handle*), and events in which an agent changes direction along a path (e.g., *Mary turns left*). Viceversa, a given action event can be expressed by different verbs that only partially share their referential potential (e.g., *Mary turns/rotate the handle*; vs *Mary turns/*rotate left*).

In view of this, two main questions arise: 1) What does determine the extension of a given action verb, i.e. the verb capability to refer to different types of action events?: 2) Why can a given event be encoded by two semantically different action verbs (e.g., *Mary takes the book from the shelf/ Mary removes the book from the shelf*)?

We hypothesize that the semantic extension of an action verb is primarily determined by the cognitive similarity among the events the predicate refers to. In this context, we aim to make explicit the pivotal role that image schemas (Johnson 1987; Lakoff 1987; Gibbs & Colston 1995; Hampe 2005) play either in the cognitive and semantic construal of action events and in their linguistic expression (e.g., verb lexicon). Specifically, we support the idea that the identification of the image schemas characterizing action verbs semantics may help us to explain not only the predicates’ extension (Panunzi & Vernillo 2019) but also their semantic overlappings. Image schemas may provide us with the skeletal semantic information that is stored in the semantic core of an action verb, and that is also responsible for action concepts representation and categorization.

To test our hypothesis, we annotate the image-schematic components for a set of occurrences of 10 Italian action verbs. These occurrences are taken from the IMAGACT ontology (Moneglia 2014), a multilingual resource based on spoken corpora, and represent prototypical instances of action events. We then use the similarity values among action events computed by Gregori et al. (2019) as a benchmark for semantic and lexical closeness. In this work, Gregori and colleagues measure the
lexical closeness of two events on the basis of the number of predicates (in 13 languages) that can refer to both of them.

We expect that the semantic extension of an action verb presents an internally coherent image-schematic structure. Consequently, verbs that share common image-schematic components may also refer (at least partially) to the same action events. Moreover, we expect action events sharing the same image schemas to show high levels of lexical closeness. At a general level, we explore the relationship existing between cognitive and linguistic similarities in the categorization of actions.

References

The variation of Differential Object Marking in Spanish: Experimental data from four varieties and across six constructions

Albert Wall, Philipp Obrist, Senta Zeugin & Johannes Kabatek
(University of Vienna, University of Zurich, University of Zurich, University of Zurich)

Keywords: Differential object marking, elicitation experiment, animacy, secondary predication, variation

A comprehensive view of Differential Object Marking (DOM) in contemporary Spanish (Bossong 1985, Fábregas 2013), requires comparable datasets from different varieties. In order to address this gap, we conducted an elicitation experiment on six syntactic and semantic configurations sensitive to the variation of DOM. The experiment was replicated in four major centers (Madrid, Mexico City, Lima, the River Plate).

Challenging the well-known generalization that among direct objects (DOs), animate definites are always marked and animate indefinites only under specific configurations, whereas inanimates reject DOM, García García (2014) investigated the variable occurrence of DOM with indefinites in ADESSE (Vaamonde et al. 2010). He identified six potentially relevant configurations, many of which involve complex syntax (cf. Zdrojevski 2018): so-called reversible predicates require DOM even for inanimate DOs (1), as well as “double accusative” constructions, especially if the DO is not adjacent to the verb.
Inanimate DOs can receive DOM in AcI constructions (2), and even inanimate DOs of canonical transitive sentences show a robust rate of DOM (3). On the other hand, ditransitive sentences can block DOM even with animate DOs (4), and animate DOs of tener seem to reject DOM as well (5). The scant literature on regional variation refers primarily to canonical transitives, neglecting specific syntactic configurations.

(1) Reversible predicate
Un silencio temeroso siguió a la pregunta.
A fearful silence followed the question.

(2) AcI
Veo al agua caer muy rápidamente.
‘I see the water falling very quickly.’

(3) Canonical transitive, inanimate object
Abrazó a las columnas.
‘He embraced the columns.’

(4) Ditransitive
Pedro presentó su mujer a sus amigos. (García García 2014:53)
‘Pedro introduced his wife to his friends.’

(5) DO of tener
María tiene tres hijos.
‘María has three children.’

The experiment features a 2X2 Latin square design with relevant manipulations for each construction. For instance, with reversible predicates, animacy was manipulated and crossed for subject and object, meaning that each speaker produced at least four sentences per type, each under a different experimental condition. The items were distributed over 4 lists, and each participant saw the same number of items in each condition but never the same item in two conditions. Measures were repeated with construction types (1) and (4) for more robust results. Each participant produced 40 sentences, and we recorded 40 participants in each location, summing up to 6400 utterances.

Results demonstrate that categorical claims on DOM with reversible predicates and double accusative structures are too strong and that regional variation is considerable. Furthermore, comparison among varieties shows that inanimate DOs in canonical transitive sentences are twice as frequent in the Hispanic-American varieties under investigation (~11%) than in Madrid (~5%). We will discuss the results of all six constructions and hereby present a balanced and nuanced picture of the variation of DOM in four major varieties of Spanish.

References


**Multimodal packaging in action formation: Existential relative clause constructions in spoken Hebrew discourse**

Nikolaus Wildner

(University of Haifa)

Keywords: interactional linguistics, spoken Hebrew discourse, relative clauses, existential constructions, multimodal analysis

Employing the methodology of Interactional Linguistics (Couper-Kuhlen & Selting 2018), I explore the interactional functions and multimodal features of a complex syntactic sequence in spoken Hebrew discourse consisting of an existential predication and a subsequent relative clause (RC). As exemplified in ex. 1, the existential is expressed by the invariable “verboid” (Rosén 1977:107) *yesh* “there is”, whereas the RC is introduced by the invariable subordinating element *she-*:

Ex. 1

1 Alex: ...*yesh* l-anu doktorant xadash ba-ma’abada
   there_is to-1PL doctoral_student new in ART-laboratory
   ...we have a new doctoral student in the lab

   *she-bid’iyuk* ‘avar lagur sham,
   that-exactly move.PST.3M.SG live.INF there
   who just moved to live there,

Referring to these sequences as “existential relative clause constructions”, I address the question of “which discursive or social actions do speakers of Hebrew perform in interaction by means of these constructions?” For my study, I rely on video-recorded interactional data from the *Haifa Multimodal Corpus of Spoken Hebrew* (Maschler et al. 2020).

As a close look at the occurrences in my data reveals, distinct actions performed by existential relative clause constructions in Hebrew discourse feature in different sequential positions and come along with distinct non-verbal features, namely, prosody, gaze conduct and embodied behaviour. These distinctions become evident in comparisons between contrasting uses of the construction: Whereas ex. 1 exemplifies Lambrecht’s description that “the discourse function of this construction type is to introduce a new entity into a given discourse world and at the same time to express some piece of new information involving this entity” (2002:171), ex. 2 is employed to a different end:
Ex. 2
1 Hila: ‘ex hem kov’im le- ‘e mi holex le’an?’
how they determine.PRS.M.PL to uh who go.PRS.M.SG where_to how do they determine to uh who goes where?

2 Gaya:...yesh ‘eyze mishehi [mishi]
there_is some someone.F
...there is someone

she- ‘axra’it ‘al ze.
that-responsible on this.M.SG
who is responsible for that.

Also in this case, a new referent is introduced through the same construction. However, unlike ex. 1 after which the new referent “a new doctoral student” continues to be at centre stage throughout several subsequent intonation units, in ex. 2 Gaya asserts the existence of an underspecified referent (lit. “some someone”) in response to Hila’s question before abandoning it. That is, Gaya mentions here a referent in charge of a certain task only to satisfy Hila’s inquiry about some bureaucratic procedure, but not to develop it as a discourse topic. The functional contrast between these two discursive uses is, according to my analysis, not only expressed by the speakers’ lexical choices, but also by their bodily-visual behaviour and distinct recurrent prosodic patterns.

By discussing the different actions, their multimodal properties and the ways in which the latter map onto the former, I add further proof to the insight that constructions in discourse constitute integrated “multimodal packages for the production of action” (Hayashi 2005:47) whose full range of interactional functions is adequately detected by studying naturally occurring conversations.

References
Hayashi, Makoto (2005), Joint construction through language and the body: Notes on embodiment in coordinated participation in situated activities, Semiotica 156, 21-53.

The Balkan Sprachbund:
A multi-variate approach to the ‘prototypical’ case of language contact

Olivier Winistörfer
(University of Zurich, Switzerland)

Keywords: Language typology, Sprachbund, language areas, Romance languages, Slavic languages

In my talk, I want to combine qualitative and quantitative, multi-variate approaches in order to verify whether a set of independent, binary variables from the full spectrum of the linguistic system (phonology, morphology,
morphosyntax, syntax) support the claim of ‘typological’ proximity between the Romance and the Slavic languages in the Balkans discussed in previous studies.

Since the first mentions of parallelisms in the linguistic structures of the Romance and Slavic varieties in the Balkans by Jernej Kopitar (1829, 85-86) and the introduction of the term ‘Balkan Sprachbund’ by Nikolai Trubetzkoy (1930), there have been many cross-linguistic studies on the alleged language proximity of the (Romance and Slavic) varieties in the Balkans due to intense historic language contact. This led Friedman/Joseph to the conclusion that “[i]t is almost impossible to talk about the Balkans from a linguistic standpoint and not utter the term ‘Sprachbund’ […]” (Friedman & Joseph, 2017, 55). While there have been many qualitative studies on a restricted amount of parallelisms between languages in the Balkans (e.g. Tomić 2006, Thomason 2000, Sandfeld 1930, Kopitar 1829), there have been almost no attempts to describe the particular case of the Balkans from a quantitative point of view until now.

In my talk, I want to tackle this particular question with different typological datasets. After having presented the results from a quantitative pilot study I conducted with the data from the WALS (2013) database which seem to support the claim of a Balkan Sprachbund and having discussed the problems of this particular dataset, I propose a new approach to the question of the Balkan Sprachbund by combining qualitative and quantitative approaches already known from other fields of linguistic typology. On the basis of the variables presented in the World Atlas of Language Structures, WALS (2013), Autotyp (2013), as well as in Birchall (2014), and Haspelmath (2010), I introduce a set of independent, binary variables from phonology, morphology, morphosyntax, and syntax adapted to the particular case of the (Romance and Slavic) languages in the Balkans. In a next step, I classify the linguistic phenomena we find in the Romance (Aromanian, Dacoromanian, Meglenoromanian) and the Slavic varieties (Bulgarian, Macedonian, Serbian, Torlak) often considered members of the Balkan Sprachbund as well as their sister languages not located in the Balkans (French, Italian, Spanish from the Romance vs. Polish, Russian from the Slavic language family). The database is then used to verify whether a typological dataset and the quantitative analyses (Mann-Whitney test, Multi-Dimensional Scaling, NeighborNet) support the claims from previous studies suggesting that the varieties in the Balkans are linguistically ‘closer’ in comparison to the other branches of their language families.

This study can be considered an important contribution for future work on the Balkan Sprachbund and areal linguistics as it combines qualitative as well as quantitative approaches and as it considers non-standardised Romance and Slavic varieties in the Balkans, a group that has traditionally been neglected in quantitative studies, but can be of importance to the question of the Balkan Sprachbund. In addition, the linguistic diversity within the varieties can be represented in the set of variables due to the use of binary variables and lay the foundation for more such studies in the field of linguistic areas and language families.

References


The PIC in the spotlight: Agree without limits but Move with caution!

Jacek Witkoś
(AMU Poznań, Poland)

This presentation addresses the scope of the Phrase Impenetrability Condition (PIC, Chomsky 2001) for Agree and Move in the context of Long Distance Genitive of Negation (LDGoN) in Polish; Neg on main clause verb impacts case of direct object in infinitive:

(1) Jan nie każe Marii [PRO2 czytać mu*3/4 [listów Tomka3 do siebie1/2/3]] Jan NOM NEG told MariaDAT readINF himDAT lettersGEN TomGEN to self ‘Jan did not tell Maria to read him Tom’s letters to him/her/himself.’

(1) instantiates Agree between a probe (Neg) and a goal (the NP) based on a split-probe (Neg+v) action or case overwriting (Pesetsky 2013):

(2) [NegP NEG [vP v-told MariaDAT [CP [TP PRO T [vP readINF himDAT letters*ACC/GEN…]]]]

But if CP and vP constitute phases, the relation between Neg and letters violates the PIC. One solution is to extend the phase (Den Dikken 2007, Gallego 2010), to include NEG and the object (Blaszczyk 2001, Witkoś 2003). Then the phase is the smallest relevant convergent domain (Pesetsky 2013, Svenonius 2004). Alternatively, assume that PIC is selective:

(3) For Agree in Polish: (a) only a ‘content’ CP is a phase (C-Tfin, Cwh); (b) only NegP (negative vP) is a phase. (cf. Landau 2000 on CPinf).

Yet, liberal definitions of PIC lead to complications in a construction combining LDGoN and wh-movement. In (1) there are three interpretations for the reflexive, where siebie ‘self’ is bound by main clause subject Jan1, embedded clause subject PRO2, controlled by Maria2,DAT, and possessor Tomek3,GEN. Mu*3/4,DAT must be obviative w.r.t Tomek3,GEN, as it c-commands it. Yet, in (4) Tomek3/4,GEN can be coindexed with mu3,DAT, so an intermediate position is necessary for reconstruction (Fox 1999, Legate 2003, Lebeaux 2009):

(4) [[ilu listów Tomka3/4 do siebie1/2/4] Jan1 nie każe Mariii2 [CP PRO2 [vP l’ czytać mu3 t how many lettersGEN TomGEN to self Jan NOM NEG told MariaDAT readINF himDAT wczoraj]]]

‘How many of Tom’s letters to him/her/himself did Jan not tell Maria to read to him yesterday?’
This copy must be placed where mu3,DAT does not c-command Tomek3,GEN and at the same time PRO2 c-commands siebie1/2/3 to provide for the interpretation 2. This position is τ' at the edge of the embedded vP in (4), because wh-phrase need to move cyclically through phase edges and vP is a phase. But the LD Agree between NEG and the embedded direct object cannot tolerate this. (4) places contradictory requirements on this derivation. We conclude that (4) shows that Agree (selectively) disobeys PIC, while Move does not. This leads us to:

(5) PIC holds of Move but it does not hold of Agree. (Bošković 2007)

Bošković (2007: 613–614) points to languages that allow for Agree to reach into a finite CP; in Chukchee matrix v agrees with embedded object, which violates PIC:

(6) онан галгылу лңарқо-ний-ет [иңүң 0-рәтәмөң-нен-ат қора-т]
he regrets-3-PL that 3SG-lost-3-PL reindeer-3-PL ‘He regrets that he lost the reindeers.’

Polish LDGoN is a less extreme version of Chukchee LD Agree; it can reach only into infinitive complements. Ex. (4) lends further support to (6).

References

Effects of the Mandarin reflexivizing prefix zi- on clausal complements in its domain

Sally Wong, Eric Reuland & Martin Everaert
(Utrecht University)

Keywords: anaphoric dependency, aspect, temporal dependency, zi-verbs, Mandarin Chinese

The Mandarin prefix zi- not only reflexivizes simple transitive verbs, but also operates on verbs with full clausal complements, restricting the interpretation of anaphors and pronominals in this clause. For
instance in (1), the subject of the zi-verb is the only admissible binder of the anaphor zi-ji and the pronoun ta (see Chief 1997; Reuland, Wong & Everaert 2019):

(1)  Zhangsan, zi-yiwei  [Lisi, zhidaow [Wangwu, xinren zi-ji vryk | ta (vryk)discourse]].
Zhangsan REFLECTS think Lisi know Wangwu trust REFLECTS self | him/her/it
‘Zhangsan thinks that Lisi knows that Wangwu trusts him.’

We investigated further possible restrictions on clausal complements of zi-verbs, in the temporal and aspectual domain. 15 zi-verbs with clausal complements were compared with their unprefixed counterparts. Judgments were obtained from 15 native speakers of Mandarin. Applying the adverbial modification test using the adverbials zai yige xiaoshi zhinei ‘in an hour’ and da yige xiaoshi ‘for an hour’ (see e.g. Borik 2006), showed these verbs are all atelic. Some of these verbs were compatible with the perfectivity marker guo, supporting Borik’s result that perfectivity is not incompatible with atelicity. Next, temporal dependencies were investigated. For each choice of past, present or future of the matrix clause containing the zi-verb, it was determined whether the choice of the tense of the complement clause showed a restriction; the results were compared with those for related non-zi-verbs. The results were checked for plausibility effects. For instance, for a verb like dasuan ‘intend’, the time of the event denoted by the complement clause will necessarily be subsequent to the formulation of the intention. Such a restriction would be an effect of lexical semantics rather than of structure. After correction for such effects, it was found that zi-Vs are restricted to past tense and present tense. They don’t allow future tense, for reasons that don’t reflect lexical semantics and/or plausibility. In this they differ from their un-prefixed counterparts. The contrast is illustrated in (2).

(2)  (Premise: Ruguo gushi da sheng ‘If the stock market rose dramatically’)
    a. Lisi mingtian (hui) houhui zi-ji jintian mai le gupiao.
        Lisi tomorrow will regret REFLECTS self today sell ASP stocks.and.shares
        ‘Lisi will regret tomorrow that he sold the stocks and shares today.’
    b. *Lisi mingtian (hui) zi-huil jintian mai le gupiao.
        Lisi tomorrow will REFLECTS-regret today sell ASP stocks.and.shares
        Intended: ‘Lisi will regret tomorrow that he sold the stocks and shares today.’

This indicates that prefixation by zi- affects both argument structure and temporal structure.

References:

From ad hoc categorization to evaluation: the Romanian marker alde

Rodica Zafiu
(University of Bucharest and “Iorgu Iordan – Al. Rosetti” Institute of Linguistics)
Keywords: ad hoc categories, vagueness, evaluative particles, predeterminers, Romanian morphosyntax

As proved by recent typological inquiries (Mauri 2017, Mauri and Sanso 2018, etc.), languages display an extensive inventory of morphological and syntactic tools which are used for building ad hoc categories. In addition to general extenders, exemplification markers and other less specific linguistic strategies, colloquial Romanian uses a dedicated marker, the invariable form alde, which manifests a high level of versatility, due to the inherent context dependency of its functional class.

This presentation aims to analyze, in a comparative and typological setting, the semantic and functional values of alde, and to demonstrate that its diachronic changes are characteristic for a larger class of ad hoc categorization markers. Several constructions, some of which are preserved in regional varieties of Romanian, illustrate the change from marker of associative plurality to marker of ad hoc categorization and further to evaluation particle. Additional examples demonstrate that ad hoc categorization markers typically develop, through the pragmatic effects of vagueness, into ironical and deprecatory markers, a phenomenon that manifests itself in the case of antonomasia and collective suffixes as well (see for instance Magni 2018). The inherent vagueness of the ad hoc categorization has been extensively described for some typical strategies, especially for general extenders (Channel 1994: 119–156; Overstreet 2009).

The distribution examples of the different constructions and values of alde in contemporary (informal) Romanian are based on data extracted from two internet corpora (Romanian Web 2016 – roTenTen16 and CoRoLa). The predeterminer alde (whose origin is controversial, but probably related to the coalescence of al < Lat. ILLUM with the preposition de ‘of’, Giurgea 2012), is immediately anteposed to DPs with human reference (proper names, kin terms, and pronouns), either in singular or plural forms (1).

(1)  
a. alde Ion  
ALDEMARKER Ion  
b. alde frate-tău  
ALDEMARKER brother=your  
c. alde ăștia  
ALDEMARKER these ones

In its dialectal use, alde formerly functioned as a marker of associative plural (Daniel and Moravcsik 2005), which designated a kin group (Mărgărit 2001). The present colloquial use of this form is different: alde functions as a marker of similitative plural (with non-human referents), as a categorization trigger, as well as a comparative connective, and a marker of vague and usually ironical designation (Zafiu 2009). Alde is an ad hoc categorization trigger especially when the DP (indicating the exemplar from which a property has to be inferred) is multiple (coordinated), or in the plural form, but the same value may be actualized with a singular name as well. Many occurrences of alde are ambiguous between a (neutral or negative) reference to a semantic category and a subjective, slightly ironical reference to a specific individual (2).

(2)  
Ce să vorbeşti cu alde Mitică?
what SUBJ.MARKER speak.2SG with ALDEMARKER Mitică
‘What can you speak with (someone like) Mitică?’
We consider that the description of the multifunctional marker alde can contribute to better understanding the pragmasemantic potential of these types of markers, supporting the idea that the connection between ad hoc categorization and evaluation (more precisely, pejoration) is a regular one.

References
Mărgărît, Iulia (2001), Construcții cu alde în graiurile muntenești, Limba română 50 (5–6), 276–293.

Chaining clauses, serializing verbs? Ambiguities in the status of non-finite verbs in Muyu

Alexander Zahrer
(University of Münster)

Keywords: clause chaining; serial verb constructions; trans new guinea; non-finite verbs; switch reference

Muyu is an underdocumented lowland Ok language of West New Guinea spoken by perhaps 2,000 people in the Boven Digoel regency of Papua, Indonesia. Available resources comprise several wordlists in Voorhoeve et al. (1975) and Healey (1964), ethnographic material by Schoorl (1993) and a short grammar sketch by Drabbe (1950). Recent findings come from ongoing fieldwork as part of a PhD project. 9 hours of various genres have been recorded on video so far during 6 months in the field.

Muyu is a language with two types of chaining constructions: clause chaining and serial verb constructions (SVC). A similar system was described by Fedden (2011) in the related Ok language Mian. In this talk, I will argue that clause chains and SVCs in Muyu cannot be separated neatly in all cases due to ambiguities with the non-finite verb.

Verbs appear in one of three possible forms. They are distinguished by the inflectional morphology they take: finite verbs (occasional TAM marking + argument indexes), semi-finite verbs...
(same subject marker + argument indexes), and non-finite verbs (bare root + -e). Finite verbs occur in final clauses and spread their TAM features and subject left to all preceding verbs. Since semi-finite and non-finite verbs rely on subject continuity, finite verbs tend to signalise subject discontinuity in subsequent discourse. Consider the following examples:

(1) 
[\text{Ege } \text{angg-}en\text{-}go]_{\text{CL}} [\text{kido } \text{teb-}en\text{-}i]_{\text{CL}} [\text{yi } \text{kumungg-}an.]_{\text{CL}}

\begin{verbatim}
[here stay-3SG.M=PTC] [out move-SS-1SG] [3PL tell-1SG]
\end{verbatim}

\text{‘(The fish) stayed here so I came out and told them.’}

(2) 
[\text{Kawen-on-u}]_{\text{CL}} [\text{wan-e } \text{nin-e } \text{mun-un=}g\text{t}.]_{\text{CL}}

\begin{verbatim}
[climb-SS-3SG.F] [pick-NF hold-NF come-3SG.F=PTC]
\end{verbatim}

\text{‘She climbed and picked (the fruit) and brought (it).’}

(3) 
[\text{Kol-e } \text{ambip weneb-}on\text{-}up]_{\text{CL}} [\text{\textit{kawen-}e}]_{\text{CL}^2} [\text{epkat mo b-up}.]_{\text{CL}}

\begin{verbatim}
[leave-NF house go-SS-1PL] [climb-NF] [clothes only take:PL,O-1PL]
\end{verbatim}

\text{‘We went back home and entered (the house) and we just took the clothes.’}

Multiple non-finite verbs often concatenate as in the second clause of (2). They can be interpreted as serial verbs in a SVC based on the criterion of monoclausality (Haspelmath 2016, Aikhenvald 2018). On the other hand, semi-finite verbs like tebeni in (1) and kawenonu in (2) are means of marking non-finality in a chain of clauses. At first sight the division: semi-finite=clause chains, non-finite=serial verbs seems reasonable. However, non-finite verbs vary in their syntactic, semantic and prosodic integration to subsequent verbs. In certain cases, a non-finite verb seems to head its own clause. kawene in (3) does not integrate into the following clause but is linked more closely to the preceding clause with which it is sharing the argument ambip ‘house’. This ambiguity raises the question if a grammatical description of Muyu should strictly separate the two types of chaining constructions.

\textbf{Abbreviations}

\begin{tabular}{ll}
\text{CL} & \text{same subject} \\
\text{SS} & \text{non-finite} \\
\text{PTC} & \text{party} \\
\text{NF} & \text{CL} \\
\text{PL.O} & \text{object} \\
\end{tabular}

The research was supported by ELDP (Endangered Languages Documentation Programme) Individual Graduate Scholarship 0367.

\textbf{References}


Healey, A. (1964). \textit{The Ok Language Family in New Guinea}. ANU.


\[ \text{Resolving competition in English verbal argument structure: NPs vs PPs} \]

Eva Zehentner

\textlt{not updated>}

199
On the survival of Polish clusters in extemporaneous speech

Paulina Zydorowicz
(Adam Mickiewicz University in Poznań)

Keywords: phonotactics, morphonotactics, spontaneous speech, Polish, Corpus study

Polish possesses a rich phonotactic inventory. The complexity of the phonotactic system is visible in terms of 1) cluster size (up to 6 consonants), 2) the number of cluster types (over 2400), and 3) the phonological content of the cluster (sonority rises, reversals and plateaus). A considerable subset of the clusters may be triggered by morphology (Dressler and Dziubalska-Kołaczyk 2006). Cluster compliance with sonority (a traditional diagnostic of phonotactic markedness) amounts to 75%. Moreover, approximately 85% of word types contain at least one cluster. This means that running speech is highly saturated with clusters.

The aim of this contribution is to examine the production of clusters in extemporaneous speech of Polish speakers. The study focuses on two aspects of phonotactics in particular: the role of morphology and markedness in cluster production. It is predicted that morphologically-driven clusters might be reduced less frequently than phonologically motivated ones as they serve a morphological function. Markedness is measured by means of the Net Auditory Distance Principle (henceforth NAD, Dziubalska-Kołaczyk 2014). The NAD Principle formulates well-formedness conditions for clusters of various lengths and word positions. Cluster quality (preferred or dispreferred) is based on three criteria of consonant description: manner and place of articulation as well as the distinction between an obstruent and a sonorant in a sequence. The condition for a preferred word-initial or -final CC cluster states that the distance between the two neighbouring consonants ($C_1C_2$) must be greater than (or equal to) the distance between the vowel and the neighbouring consonant. Clusters which violate the NAD condition(s) are considered dispreferred and predicted to undergo simplification more frequently.

The empirical data were extracted from The Dictionary of phonetic variability of the contemporary Polish language (Madelska 2005), which comprises phonetic transcriptions of words from conversations with 30 Polish speakers. Clusters were tagged according criteria such as cluster position in a word, size, cluster goodness as measured by NAD, the presence of a morphological boundary in a cluster, the influence of the following vowel (the presence of homorganic environment), word length, word stress and word frequency. A mixed effects logistic regression model was fitted to the dataset with the lme4 package (Bates et al. 2015) in R (R Core Team 2018) to examine factors affecting cluster rendition with word and cluster as random effects, and cluster position, size, NAD status, morphological boundary, word length, homorganic environment and word frequency as fixed effects. The analysis confirmed statistical significance of all the variables. Word-initial clusters are modified more frequently than word-medial but less frequently than word-final sequences. Cluster modification rate increases alongside cluster length and word frequency. Preferred clusters undergo modification less frequently. Homorganicity of the (CC)CV sequence is strongly disfavoured. Cluster modification rate is proportional to word length. Unexpectedly, morphonotactic clusters are simplified more frequently than phonological clusters. It is hoped that the present study will constitute a valuable addition to the description of spoken phonotactics of Polish.
References


Madelska, Liliana (2005), Słownik wariantywności fonetycznej współczesnej polszczyzny [A dictionary of phonetic variability in contemporary Polish]. Kraków: Collegium Columbinum.

POSTER PRESENTATIONS
Negated quantifier phrase constructions: Typology and diachrony

Omri Amiraz
(The Hebrew University of Jerusalem)

Keywords: quantifiers, negation, ellipsis, typology, historical linguistics

This paper investigates the typology of negated quantifier phrase constructions in sentences such as *Not every cat enjoys being scratched*. The main finding is that negated quantifier constructions develop out of a particular type of ellipsis construction called *not*-stripping.

Definitions and scope:

(1) A NEGATED QUANTIFIER PHRASE CONSTRUCTION is a mono-clausal construction distinct from standard negation in which a negative marker is adjacent to a quantifier phrase.

(2) STANDARD NEGATION is a construction that is employed to negate main, declarative, unquantified, verbal clauses. (Miestamo 2005)

For example, standard negation is expressed in English by a construction in which the negative particle *not* follows the auxiliary or copular verb (3a). On the other hand, in a negated quantifier phrase construction, the negative particle *not* precedes the quantifier phrase (3b).

(3) a. John did not see Mary.  
   B. Not everyone saw Mary.

*Not*-stripping is an ellipsis construction found mostly in coordination where everything is deleted from the conjoined clause except for one constituent and the negative marker (Hankamer & Sag 1976: 409). The remaining constituent can be of various types, e.g., a proper name (4a) or a quantifier phrase (4b).

(4) a. Some patients hide out while waiting for their hair to grow, but not John.  
   b. Some of the students live in the dorms, but not all of them.

Materials and methodology: Negated quantifier phrase constructions are not often described in grammars. Therefore, this study requires collecting primary cross-linguistic data. For this purpose, I use historical and contemporary translations of the New Testament as a parallel corpus. The sample consists of 112 languages from around the world, and it contains a smaller sample of 43 languages from Europe, Western Asia and Northern Africa, for which diachronic data are also available.

Results: The results establish several novel linguistic universals, including the ones in (5) and (7).

(5) Universal 1A  
Every language that has a negated quantifier phrase construction also allows *not*-stripping with a remaining quantifier phrase, but not vice versa.

For example, Greek allows óxi óli ‘not all’ to occur in stripping (6a) but not in an independent clause (6b).

(6) Modern Greek (Spiros Filos translation, 1994)  
a. Esís íste katharí, allá óxi óli.  
you are clean butNEG all  
‘And you are clean, though not all of you.’ [John 13:10]  
b. *Óxi óli íste katharí.
.neg all you.are clean
(‘Not all of you are clean.’)

Universal 1B is demonstrated by Albanian, where the negative marker jo is used both in stripping and in negated quantifier phrase constructions, while standard negation is expressed by nuk.

(7) **Universal 1B**

If a language has a negated quantifier phrase construction, the negative marker that is used in this construction is the same marker that is used in *not*-stripping.

The most plausible explanation for the aforementioned universals is that negated quantifier phrase constructions develop out of *not*-stripping by syntactic reanalysis. I also discuss two potential counterexamples to Universal 1B (Portuguese and Malay) and propose diachronic scenarios that led to these exceptions.

**References**


**h-anticipation and hiatus resolution in Ancient Greek: how many rules?**

Roberto Batisti
(Università di Bologna)

Keywords: Ancient Greek, historical phonology, *h*-anticipation, hiatus resolution, phonological rules.

This paper investigates the relationship between two phonological tendencies in the history of Ancient Greek: the sporadic anticipation of word-internal *h* to a word-initial vowel or voiceless stop (Lejeune 1972: 95; Colvin 2006), and the resolution of hiatus between two adjacent vowels, enacted through several strategies including elision, gliding, coalescence, and diphthongization (De Haas 1988; Casali 2011).

I start by discussing two recent and innovative approaches by Jatteau (2016) and Sayeed (2019). Both authors distinguish two main varieties of *h*-anticipation:

1. **medial *h*** is copied to a word-initial vowel:
   
   *ευσό > *ευήο > *ευήο > ἡγό ‘singe’.

2. **medial *h*** causes aspiration of a word-initial voiceless stop:
   
   *προ- > *πʰρο- ‘watch’.

Jatteau keeps (1) (“report d’aspiration”) distinct from (2) (“coaléscence à travers *h*”) on the basis of chronology (historically fixed vs. living synchronical process), regularity (accidental vs. productive), target (vowels vs. vowels + stops), and interaction with syllabic structure (number of syllables preserved vs. fusion of two vocalic nuclei).

Differently, Sayeed argues that (1) (“Hauchumsprung”) and (2) (“phroura-rule”) – together with a hypothetical “Pre-Grassmann’s Law”, whereby medial *h* or *T* copied an internal aspiration to the beginning of a word, after the loss of intervocalic *h* but before the loss of *h* in medial sonorant clusters and regressive dissimilation of aspiration (= Grassmann’s Law).
I show that Jatteau’s approach avoids a problem with the chronology posited by Sayeed. In fact, (1) must be ordered before internal h-loss, as proved by h-anticipation in words with the structure #VhV- (*hieros > *hiheros > hieros ‘holy’); this becomes possible if PHROURA is considered a separate, much later rule. And in fact, PHROURA applies not only between prefix and stem as in (2), but also in the outputs of ‘crasis’, i.e. vowel coalescence across syntactically connected words, originating in spoken language and conditioned by high frequency of use (Fritz 2014; Pardal Padín 2015; Devine and Stephens 1994: 266–269):

(3) tò # hımation > tʰojmation ‘the garment’
(4) kai # hópōs > kʰápōs ‘and how’

PHROURA may accompany not only crasis but also different hiatus resolution strategies, such as elision (5) and synizesis (6):

(5) *tetra-hippos > té物业公司pos ‘with a yoke of four horses’
(6) *e-horaon > heʰrōn ‘I watched’

Analysis of the sources shows that the words or collocations in (2–6) share the prerequisites for ‘crasis’: they are living forms in Classical dialects, found in genres that reflect colloquial language or technical jargon, and have high frequency of use.

I will then discuss a further case, not treated by Sayeed and considered irregular by Jatteau:

(7) epi-horkeō > epʰiorkēō ‘I perjure’

I argue that in (7) the vowel -i- was subject to gliding (yet another hiatus resolution strategy), triggering the same shift of aspiration to the new onset that is assumed for PHROURA. In fact, this rule is simply a sub-case of PHROURA, with which it shares the same phono-morphological, chronological, and socio-pragmatical contexts.

References
Sayeed, Ollie (2019), Hauchumsprung and the historical phonology of Greek *h, Indo-European Linguistics 7, 164–175.
Discourse marker PALM-UP in Russian Sign Language (RSL)

Anastasia Bauer & Svetlana Burkova

The parrot next to the hamster (and) next to the bunny brings evidence for recursion in children

Adina Camelia Bleotu
(ISDS; University of Bucharest)

Keywords: language acquisition, recursion, PPs, coordination, contrast activation

The current paper shows, by means of a picture matching task (PMT), that Romanian children can understand recursion as early as age 4. Recursion has been considered the fundamental property of human language (Chomsky, Hauser & Fitch 2001), and a lot of effort has been invested by acquisitionists into bringing empirical support for its innateness (Perez et al. 2011, Roeper 2011, 2012, Hollebrandse & Roeper 2014 a.o). While many results are supportive of innateness, others are problematic, given the processing memory load problem affecting recursion experiments, as well as crosslinguistic variation. Previous research on Romanian (Sevcenco & Avram 2018) has shown on the basis of an ordering task that 5-year-olds interpret recursion as coordination. However, an act-out task only accesses the simpler default interpretation (coordination). In contrast, the current investigation starts from the hypothesis that, when given a better option (the recursive picture) than the default (the coordinative picture), children will recognize it as more adequate. Moreover, unlike the previous study, the current experiment exposes children to both recursion and coordination, activating the structural type contrast between the two. The subjects were 23 4 & 5-year-olds (Age range: 4-6;3, Mean: 4;9) and 45 adults. The method was a PMT, where subjects heard 16 randomized sentences (4 structure types x 4 animal combinations) containing a recursive (common noun/ proper name)/ coordinative structure (comma/ conjunction) such as in Table 1, and they had to pick the matching picture out of two (recursive/ coordinative) (see Examples and Figure 1). The results show that children are sensitive to the recursion/ coordination difference, and, at an even more fine-grained level, to the difference between recursion with common nouns/ PNs/ coordination with comma/ coordination with conjunction. A mixed-effects logistic regression was performed by means of a model using Interpretation as Variable, Structure type and Age Group as Fixed Effects, Items and Participants as Random Effects, and recursion with common nouns as the baseline, and it revealed significant differences depending on structure type and Age Group (p < .001, see Figure 2). This shows that (in)definiteness, the conjunction marker are triggers for distinguishing structure types. Account: The PMT shows subjects prefer recursive readings for recursive structures in a context where the coordination-recursion contrast is activated. The contrast activation account is supported by evidence from other experiments e.g. experiments on scalar implicatures, where the presence of the alternative (“all”) activates the <some, all> scale, i.e. subjects produce more implicatures when the test contains items with “all” alongside “some” (e.g. Foppolo, Guasti & Chierchia 2012). Thus, given the appropriate experimental context, children as young as 4 can be shown to understand recursion and distinguish it from coordination.
Acknowledgments: This paper is the result of work done during the ongoing grant “A Colouring Approach to How Children Understand Recursion in Romanian” (Supervisor: Larisa Avram), provided by The Interdisciplinary School of Doctoral Studies, University of Bucharest. Many thanks go to the educators and children at Waldorf Kindergarten, No. 11 “Ion Heliaide Rădulescu” School, Bucureşti, for their cooperation and help, as well as to Tom Roeper for his suggestions.
References

Tone reversal in central Neo-Štokavian
Dorde Božović
(University of Belgrade Faculty of Philology)

Keywords: Serbo-Croatian, prosodic typology, tone-stress interaction, vowel reduction

Neo-Štokavian is a dialect of Serbo-Croatian with a relatively complex pitch accent system, based on the interaction of two tones (H and L), vowel length, and ictus (i.e. stress). Any syllable may bear a H tone, but only one such tone is allowed per word and a word must have at least one H tone. The ictus co-occurs with H if the syllable bearing it is word-initial, thus producing a falling contour tone. Otherwise, due to stress retraction that has operated in Neo-Štokavian, the ictus is found on the syllable immediately preceding H, thus producing a rising contour tone. As a consequence of these stress-assignment rules, final syllables in polysyllabic words are never stressed, falling pitch is allowed on initial syllables only, and monosyllabic words may only have falling pitch. Rising pitch on the other hand may occur on any syllable other than the last. In addition, the stressed and any following unstressed syllable may be long (i.e. bi-moraic) or short. For historical reasons that have produced this pitch system, no unstressed syllable preceding the stressed one may be long. Standard Serbo-Croatian is based on this prosodic system (Lehiste & Ivić 1986, Inkelas & Zec 1988), as found in eastern Neo-Štokavian (spoken in East Herzegovina).

Central Neo-Štokavian dialects (as spoken in Bosnia and the surrounding areas), however, may differ prosodically from standard Serbo-Croatian in a number of innovative ways. First of all, they are characterized by a strong vowel reduction affecting unstressed short syllables. When disyllabic words with a rising pitch on the initial syllable and a short unstressed vowel in the second syllable are affected by reduction, they surface as monosyllables with a rising pitch (e.g. kònji ‘horses’ > kònj, dòdî ‘come-IMP’ > dòd), a situation dispreferred in standard Serbo-Croatian. In addition, in a number of polysyllabic words they have an innovative rising pitch on the initial syllable instead of the etymological falling one (e.g. mákja ‘mother’ for the etymological májka, kárta ‘card’ for the etymological kárta, etc.), thus effectively manifesting tone reversal.

In this paper, relying on Stratal OT (Kiparsky 2000, for Serbo-Croatian cf. Zec & Zsiga 2009), I will discuss the possible causes of this reversal in the context of diachronic evolution of the Neo-
Štokavian prosodic system. It seems that Neo-Štokavian is in a transition from a pitch system to stress system. As a transitory idiom, it manifests features of both.

In addition, as tone reversal generally represents a rather rare and interesting phenomenon from the point of view of phonological theory, I will then proceed to discuss the possible insights that central Neo-Štokavian may provide for the dynamics of tone-stress interaction and the typology of tonal and pitch systems. As low tone seems to attract stress in central Neo-Štokavian, it doesn’t fit easily into general typology of tone-stress interactions (de Lacy 2002). Again, it features as a transitory idiom between different types.

References


The acquisition of Italian accusative and dative clitic pronouns in restructuring contexts

Sara Cerutti, Valentina Scarda & Francesca Volpato
(Ca’ Foscari University of Venice)

Keywords: accusative clitic pronouns, dative clitic pronouns, restructuring verbs, Italian, language acquisition

This study investigates the acquisition of 1st, 2nd and 3rd person singular accusative and dative clitic pronouns in Italian. We tested 87 typically developing children (age 3;6–9;5) with a repetition task of sentences containing restructuring verbs, to observe the effects of person, case features and clitic positions on accuracy.

In Romance languages, 3rd person pronouns are acquired later than 1st and 2nd person pronouns, and are more difficult to process (Coene, Avram 2011; Tuller et al. 2011). In Italian dative clitics are better preserved than accusatives, more frequently omitted (Caprin, Guasti 2009). No data have been systematically collected on the acquisition of restructuring. Antelmi (1997) and Bernardini, van der Weijer (2017) found an asymmetry between proclisis and enclisis in monolingual and bilingual acquisition of Italian; Bennati, Matteini (2005) tested clitic pronouns in L2 learners. Proclisis requires longer dependencies and the functional analysis of restructuring verbs, entering mono-clausal structures (Rizzi 1978; Cinque 2004).

For this study, we designed a repetition test to investigate the production of complex structures that would be difficult to elicit (Devescovi, Caselli 2001, 2007; Del Puppo et al. 2016).

The test includes 49 sentences containing clitic pronouns and restructuring verbs, and 6 control sentences (length: 15-16 syllables). Trials are divided as follows: 28 sentences with 1st, 2nd, 3rd (masculine and feminine) person singular accusative clitics; 21 sentences with 1st, 2nd, 3rd (masculine) person singular dative clitics. Each sentence contains one or two restructuring verbs (modal and
motion) and a clitic in one of the two (proclitic or enclitic) or three (proclitic, enclitic-intermediate, or enclitic-final) available positions (examples (1) and (2)).

Children were divided into six groups: 11 children age 3;6-4;6 (G1, mean 3;9); 16 children age 4;9-5;6 (G2, mean 5); 15 children age 5;7-6;2 (G3, mean 5;10); 15 children age 6;6-7;3 (G4, mean 6;11); 15 children age 7;6-8;5 (G5, mean 7;11); 15 children age 8;4-9;5 (G6, mean 9). A control group of 16 adults (G7, age 20-28, mean 24) was also tested.

Case and person features did not influence accuracy. Children performed significantly better in the repetition of 2-verb-sentences and in sentences with proclitic pronouns, hence preferring monoclausalisity. The most frequent error was clitic misplacement (mean: 11%): the clitic was moved most frequently from the enclitic to the proclitic position in 2-verb-sentences, or from the final to the intermediate enclitic position in 3-verb-sentences. This can be explained with Jakubowicz’s (2011) Derivational Complexity Hypothesis. Enclisis on the intermediate verb implies the analysis of the highest verb as lexical (Cardinaletti, Shlonsky 2004). Both types of structures entered by restructuring verbs seem available to children. Adults almost performed at ceiling (mean: 98,4%), sporadically moving the pronoun from the intermediate to the final enclitic position.

(1) Ti vengo a trovare domenica nel pomeriggio.
Vengo a trovarli domenica nel pomeriggio.

 I will come see you Sunday in the afternoon.

(2) Mi deve passare a prendere dopo la lezione.
Deve passarmi a prendere dopo la lezione.
Deve passare a prendermi dopo la lezione.

He has to pick me up after class.

References
The vowel system and the linking processes construct the transition-based percepts

Man-ni Chu
(Fu Jen University)

Keywords: vowel system, linking processes, transition-based percepts, noise-masking experiment, place of articulation

The goal of this study was to examine which factors more significantly affect one’s transition-based percepts. Seventy-seven listeners, 21 American English, 19 Dutch, 20 French and 17 Polish natives participated in this noise-masking (0db, -20db) experiment, where the stimuli were CaC]. The Polish has eight vowels in its system, while French has 17, Dutch has 16, and General American English has 15. Several mixed-effect models were performed to indicate that only coda, group, and their interactions, significantly affect the recognition of the Place of Articulations (POA). Across four codas, shown in Figure 1, where the grey bars indicate the stimuli with -20db noise-masking, and the black ones indicate the stimuli with no masking. English and Dutch groups outperformed French and Polish ones. This implies that if the less numbers of vowels within the language’s inventory, the less robustness for natives to practice their percepts in formant cues, resulting in the least accuracy of the identification of POA. This explains why the least accuracy of the Polish group, but it is not the case of the French group. Probably the linking process of the vowel-initial words prevents French natives from perceiving the coda. Although English shows few particular cases of the linking processes, it seems that the linking process allows a gradient of instead of a clear-cut boundary. One should note that once the language (e.g., Polish) does not have a large number of vowels in its inventory, the lack or existence of the linking processes has no effects at this stage. In the end, since the similar patterns of the codas (p>t>k~=?) at two levels of noise-masking, that formant cues are still easier to recover in poor listening conditions, supported by Wright (2001). To conclude, when one identify unreleased final stops, not only his vowel system, but also his habitual treating the integrity of syllable structure ensures transition-based percepts.

References

Figure 1. The percentage of correct identification among groups.
Keywords: Sámi languages, syncretism, spatial relations, location, source

In this paper, we will take a look on the conflation of location and source of motion in the eastern Sámi languages, namely Northern, Inari, Skolt, Kildin, and Ter Sámi. In Proto-Sámi, as in many other Uralic languages, the spatial case system had a three-way distinction coding LOCATION, SOURCE, and GOAL each with its own case (Aikio & Ylikoski forthcoming: 21). However, this distinction is lost in the modern eastern Sámi languages due to phonological processes, and now they have only two spatial cases, one coding GOAL (illative) and one coding both the LOCATION and SOURCE (locative). This kind of syncretism is extremely rare in the World’s languages (Creissels 2006: 19–23, 2009: 615).

There are indications that, when a three-way distinction in spatial cases is disrupted, the language tries to repair it by some means. Two examples from Uralic languages distantly related to Sámi languages and each other can be mentioned. In Veps, a Finnic language, the distinction between location and source case was lost similarly to the merger that happened in the Sámi languages. Modern Veps has agglutinated the postposition päi ‘towards’ to the case ending to form a new case expressing SOURCE, thus restoring the three-way distinction (Grüenthal 2015: 74–78). However, the inessive case, which is the successor of the syncretic case, carries some functions that are typical to the SOURCE case, and the restoration of the three-way distinction is thus complete only in the spatial domain (Grüenthal 2015: 75–76). Similarly in Mari the old spatial case coding source lost its productivity, and the language recruited the postposition гэг (nowadays without any other semantic content) to function as a marker of source, so that the three-way distinction was spared (Alhoniemi 1977).

It has been shown that Northern Sámi has also recruited a postposition, namely eret ‘away’ to function in conjunction with the locative case to convey the source function. This postposition works like its Veps counterpart though, and it can be used only in spatial contexts (Aikio & Ylikoski forthcoming: 20). In the other Sámi languages this kind of reparation is not mentioned (see Feist (2015: 245–250) for Skolt, Kert (1971: 163) for Kildin, and (Tereshkin 2002: 102–103) for Ter Sámi). However, preliminary fieldwork done among Skolt Sámi speakers seems to indicate that there are various strategies for distinguishing LOCATION and SOURCE.

The aim of this paper is to examine the use of the locative case in the eastern Sámi languages to find out what kind of strategies, if any, they use to distinguish the functions of location and source from each other. This group of languages is well-suited for this kind of study, because they are closely
related, but seem to react differently to the case syncretism. By examining the different restoration strategies these languages employ, new light can be shed on the typologically rare phenomenon of syncretism in coding LOCATION and SOURCE.

References

Subordination strategies in Paraguayan Guaraní

Dmitry Gerasimov

While Paraguayan Guaraní is probably the most studied indigenous language of South America, existing literature is heavily skewed towards its unique socio-linguistic situation, with many morphosyntactic topics still left largely unexplored. [Estigarribia 2017: 79] identifies clausal syntax as one of the areas where the need for further work is especially strong. The present talk surveys subordination strategies in the language, based on published texts and own fieldwork.

The study involves 27 constructions of Paraguayan Guaraní, identified, grouped and coded in such a way as to ensure compatibility of findings with [Dixon & Aikhenvald (eds.) 2009; van Gijn 2014; van Gijn et al. 2015]. They cover all major types of semantic relations recurrently encoded by subordinate clauses in the languages of the world.

Paraguayan Guaraní is shown to employ a wide range of event-linking constructions, varying in the degree of syntactic integration: from finite subordinate clauses introduced by free conjunctions to participant and event nominalizations to serial verb constructions, verbal compounds and derivational affixes. Virtually every major subordination strategy discussed in [van Gijn et al. 2011] is present in the language. Most notably, despite prolonged contact with Spanish, modern Paraguayan Guaraní still possesses most of the constructions types identified in [van Gijn et al. 2015] as characteristic of the Tupian languages. In the domain of relativation, at least three of the four nominalizing markers
reconstructed for Proto-Tupí [Rodrigues & Cabral 2012] have been retained, although their reflexes have lost in productivity due to expansion of the Proto-Tupí-Guarani general nominalizer.

Of special theoretical and typological interest are causal and circumstantial clauses in which the semantic relation between clauses is encoded by a case suffix attached directly to a verb stem that bears no markers of nominalization/relativization, cf. (1). In [van Gijn 2014, 2019] the use of case or adpositional marking to indicate the relation of the dependent clause to the proposition expressed in the main clause is taken as a sign of nominalization. However, the dependent clauses in question show no other sign of demotion as far as argument encoding and TAM marking is concerned; such “zero-nominalized” verbs also cannot appear in other nominal contexts. We tentatively assume that markers like -gui ABL in (1) are best treated as transcategorial operators able to transform constituents of different categories into adjuncts (more or less like the de is employed in Mandarin to transform almost any category into an adnominal modifier, cf. [Paul 2012] and references therein).

The same Ablative suffix -gui is used to mark standards of comparison. However, Paraguayan Guaraní comparative clauses contrast with reason clauses in that in the former some nominalizing device must obligatorily be present before Ablative (if the standard is non-nominal), as in (2). [Sudo 2015] proposed an analysis of variation in comparative clauses under which the use of nominalizing morphology in languages with fixed case comparatives follows from morphosyntactic properties of the standard marker, namely, their subcategorization for a strictly nominal complement. The fact that in Paraguayan Guaraní non-nominal standards still must be nominalized before attaching -gui, although the latter can combine with bare verbs, undermines this proposal and constitutes an argument in favor of a semantic motivation of the observed variation.

**Examples**

(1) Ñane-retã i-py’a-guasú-va-pe-nte o-henãi, 1PL.INCL-country 3-liver-big-REL-LOC-RESTR 3A-звать ha pé-inã pei-me peã pende-rehe o-jeroviá-guí. and DEM-DUR 2PL.A-be 2PL 2PL-OBL 3A-believe-ABL ‘Our country calls only the brave, and here you are because she believes in you’

(2a) Ro-hayhu-ve chu-pe a-hayhu.*(há)-gui. 1SG.A/2SG-love-CMPR 3-LOC 1SG.A-love-NMZ-ABL ‘I love you more than him/her’.

(2b) Che-mba’ê-ve ne-mba’ê.*(rã-launch-gui. 1SG-thing-CMPR 2SG-thing-COND-ATTR-ABL ‘This is mine rather than yours’.

**Abbreviations**


**References**


Subjectivity in different registers and genres: the case of Estonian particles

Tiit Hennoste, Külli Habicht, Helle Metslang & Külli Prillop
(University of Tartu)

Keywords: subjectivity, registers, particles, language variation, Estonian language

The topic of our presentation is the use of particles which express varying degrees of probability and attitudes in different registers and genres. We concentrate on three epistemic particles (vist ‘probably, possibly’, ilmselt ‘obviously, evidently’, tegelikult ‘really, in fact’) and three value judgment particles (õnneks ‘fortunately’ and kahjuks, paraku ‘unfortunately”).

Our study belongs to the broader topic of the expression of subjectivity and intersubjectivity in language (Narrog 2017; Traugott 2010). There are only a few studies of subjectivity/intersubjectivity which include a register-based dimension and those works concentrate on a few genres and registers only (e.g. Baumgarten, Du Bois & House 2012). Our analysis of several different registers and genres attempts to fill this gap.

We have two research questions:

a) How do registers cluster on the basis of usage frequency of different particles?
b) Which contextual features define the differences in the frequency of particles?

Our data come from the Estonian Language Corpora (https://cl.ut.ee/korpused/) of written, spoken, and computer-mediated texts and conversations. Registers analyzed are: academic prose; print media (daily newspapers); narrative fiction; spoken interaction (ordinary and institutional); computer mediated communication (instant messaging and comments to news and opinion stories in Internet media portals). Our approach is centered on usage-based and corpus-driven studies of discourse and registers. We use cluster analysis as a method.
Our results reveal that the analyzed registers form five clusters distinguished by the frequency of different particles: 1) ordinary interaction (ordinary spoken interaction and instant messaging); 2) institutional spoken interaction; 3) narrative fiction; 4) Internet comments; 5) print media and academic prose. The contextual features which define the difference of the frequency of different particles are dialogicity/monologicity, spontaneity/editedness, personality/impersonality (neutrality), ordinary/institutional interaction.

Analysis shows that the key features distinguishing clusters are dialogicity/monologicity and spontaneity/editedness and key distinguishing markers are *vist* ‘probably’ and *ilmelt* ‘evidently’. *Vist* is a marker of uncertainty and/or low truth value. *Ilmelt* expresses truthfulness based on an argument, experience, or perception.

Clusters are grouped into three groups which form a scale. At the first end of the scale are spontaneous dialogues (1st and 2nd cluster). This group is distinguished from other clusters by high frequency of *vist* and low frequency of *ilmelt*. Within this group, the institutional interaction is characterized by lower frequency of *vist* and higher frequency of *ilmelt* than ordinary dialogue.

At the other end of the scale is 5th cluster: edited monologues which are distinguished from other clusters additionally by impersonality. This cluster is characterized by very low frequency of *vist* and high frequency of *ilmelt*.

The third group consists of narrative fiction (3rd cluster) and Internet comments (4th cluster) and form an intermediate group distinguished by the frequency of *vist* and *ilmelt*. Texts in those clusters contain both dialogue and monologue parts, comments are semi-spontaneous and fictional dialogue mimic often spontaneity.

The research has been supported by the Centre of Excellence in Estonian Studies (European Union, European Regional Development Fund) and by the Estonian Research Council (grant PRG341).

References
Baumgarten, Nicole, Du Bois, Inke, and House, Juliane (eds), (2012), *Subjectivity in Language and Discourse*, Brill.

Towards canonical marking of core arguments in Romanian?
The case of *plâcea*
Mihaela Ilioaia  
(FWO Flanders, Ghent University)

Keywords: subject, transitive, non-canonical, Romanian, SAE

This paper deals with the Romanian construction in (1), in which the psych verb *plăcea* ‘please, like’ shows two different configurations, one with a dative experiencer (1a), the other with a nominative experiencer (1b).

```
(1) a. Îmi place de vecin-a mea  
    me.DAT please.3SG of neighbor-the my  
    ‘I like my neighbor’

b. Colegi-i mei mă plac  
    colleagues-the.NOM my me.ACC like.3PL  
    ‘My colleagues like me’
```

Recent studies suggest that Romanian, just like Spanish, expands its non-canonical marking of core arguments (Bossong 1998, Van Peteghem 2016, 2017). For this reason, Bossong (1998) and Haspelmath (2001), highlighting the tendency of typical SAE languages towards canonical marking of core arguments, oppose Romanian to other Romance languages, such as French, Spanish, or Italian, and group it with East-Slavic and Baltic languages. Indeed, both in Romanian and in Spanish, the dative domain expands itself and attracts, besides psychological verbs, several other verb classes, such as verbs of happening, cognitive verbs, and even existential verbs (Melis & Flores 2013).

Bearing this in mind, besides this expansion of non-canonical marking of core arguments one may expect a simultaneous regression of canonical marking in Romanian. Yet, the transitive construction in (1b) provides evidence against this. By means of a diachronic corpus study, the evolution of the two constructions with *plăcea* ‘please, like’ in (1a-b) is examined from their first attestation until present-day Romanian. For the purpose of this case study, the Romanian Web Corpus RoTenTen16 is used for the 21st century, alongside with a self-made corpus for the period between the 16th and the 20th century of Romanian, which contains approximately eight million words. Interestingly, whereas a simple query in the corpus for the 21st century reveals that this construction occurs rather frequently in this period, the corpus for the older periods of Romanian did not return any example of the transitive structure in (1b), before the 20th century. This result points toward the possibility that the construction in (1b) represents an innovation in Romanian, most probably under the pressure of the transitive syntax specific to SAE languages. It is argued in this paper, based on the case-study of *plăcea* ‘please, like’, that the tendency of SAE languages toward a transitive syntax, and hence, toward canonical marking of core arguments, contrary to all evidence, causes specific oblique structures to develop into canonical, transitive structures in Romanian. The question is raised whether this is a generalized tendency in Romanian or a unique phenomenon limited to the verb under scrutiny. It is also suggested that a more comprehensive study – involving a larger number of predicates – can contribute to a better understanding of the position of Romanian with respect to SAE languages.

References
Vowel duration in hearing aided, cochlear implanted, and normally hearing children

Asma Izadi Bidani, Nafiseh Taghva & Vahideh Abolhasanizadeh
(Shahid Bahonar university of Kerman, University of Shiraz & Shahid Bahonar university of Kerman)

Key words: GENDER, VOICE, Hearing aided, Cochlear implanted, Duration

Each language consists of different vowels with different lengths. In old Persian, vowels were divided into two groups of long [ɒ, i, u] and short [æ, e, o] (Samare, 2012). But in modern Persian the length opposition is realized only in certain limited environments on the surface (Doustdar Toosarvandani, 2004). The investigations of some recent researches contradict this length-based pairing (Aronow, McHugh, and Molnar, 2017). Linguists believe that different factors affect vowel length. Disability is one of these factors (Neumeyer, Harrington, and Draxler, 2010) besides place of vowels in the vowel space (Myers, 2005) and vowel context (De Lacy, 1998; Durvasula and Lou, 2012; J. Raphael, F. Dorman, and Geffner, 1980). Gender is also a factor affecting duration of vowels (Simpson and Ericsdotter, 2003; Ericsdotter and Ericsson, 2001).

The present research aims to measure duration of six Persian vowels. As Persian vowels are long or short, they will be produced with their intrinsic duration by the speakers. So, vowels were put in CVC framed meaningless words in order to investigate the effect of plosives on them. The words were repeated by hearing aided (HA), cochlear implanted (CI), and normally hearing (NH) children with DISABILITY (HAs and CIs), GENDER (females and males), and VOICE (voiceless and voiced) as factors. Each word consisted of a vowel [æ, ɒ, e, i, o, u] in the center and one of the plosives [p, b, t, d, k, g] in the onset and coda, [pæp, pep, pop, pɒp, pip, pup] are examples of used words in this study. After recording and analyzing the sounds using a Shure microphone and PRAAT software, repeated measure ANOVA test was used to evaluate the effect of the mentioned factors on duration of vowels. The results revealed that the effect of DISABILITY and VOICE on mean duration of vowels were significant. The results of a Post Hoc Bonferroni test showed that NHs produced vowels with the longest mean duration, but vowels produced by CIs had the shortest mean duration among three groups of participants. The mean duration of long vowels in HAs and CIs was more than the mean duration of short vowels in these groups of participants. Moreover, long vowels did not have longer mean duration comparing to short ones in NHs. The investigations also proved that vowels in CVC
framed words with voiceless plosives had longer mean duration comparing to those in the same patterned words with voiced plosives. It is also important to mention that long vowels in CVC framed words, with voiced or voiceless plosives, had longer mean duration comparing to short ones. Therefore, hearing problems and vowel context affect the mean duration of vowels and the mean duration of long vowels is not always more than the mean duration of short ones.

References:
Durvasula, Karthik and Lou, Qian (2012), Voicing, aspiration, and vowel duration in Hindi, *164th meeting of the acoustical society of America Kansas city*, Missouri.
Simpson, Adrian P. and Ericsdotter, Christine (2003), Sex-specific durational differences in English and Swedish, *15th ICPHs*, Barcelona.

**Variation and change in the Basque dialect of Itsasu (France): A case study based on recordings from two different decades**

Irantzu Epelde & Oroitz Jauregi
(CNRS, IKER UMR-5478 & University of the Basque Country)

Key words: Basque; language variation and change; dialect levelling; language contact

Basque village dialects of France have low prestige. Although dialectal speech varieties are hardly heard in public domains, they must have some covert prestige in local communities and might be much more widespread in people’s homes than most Basque linguists are aware of, living in a society with a strong monoglossic ideology (Coyos 1999) and no tradition of variationist sociolinguistic studies. Still, both the number of dialect speakers and the number of dialectal features in the speech of Basque villagers is decreasing, though not as fast as predicted earlier (cf. Peillen 1994, Davant 1996).

Two languages coexist in the Basque Country of France: French, the official language, which is present in all areas of use, and Basque, the non-official language and more and more excluded from the areas in which it has traditionally been present (family, community, social areas such as religion).
Basque is a minority, threatened language in the Basque Country of France living under strong pressure from French, and the vitality of the Basque-speaking community is rather weak, compared to other areas.

The variation and gradual loss of dialectal traits in Basque has been object of only a couple of earlier studies in the Spanish Basque Country (Haddican 2005, Lujanbio 2016). They both compare phonological features in the speech of several generations in one village (Oiartzun, in Gipuzkoa, in the first case; Goizueta, in Navarre, in the latter). As far as the French Basque Country is concerned, this type of studies are non-existent. We have been lucky to obtain sound recordings of the dialect of Itsasu, a traditional village in the province of Labourd, from two different decades, covering the period 1971-2019. Our own recordings (from 2000-2002) were obtained in varying social settings. Currently, they are all available in the repository Basque Oral Archive and in INA (Institut National de l’Audiovisuel, Délégation Pyrénées-Toulouse).

In the paper we will present to which degree these recordings are suitable for a study of variation and change in the use of dialectal features, not only in the speech community in general, but also in the speech of individual speakers. We will consider 10 phonological features, 10 morphological features and 10 syntactic features in the speech of 6 native Itsasuan informants (180 items in total), through two sets of data: i) naturalistic data (spontaneous observation data from 22 hours of recordings), and ii) semi-experimental data (auditory judgment task). We will also present the results of the comparisons, showing which dialect features are more stable than others. A case study of the speech of one of the speakers, who was recorded in different periods and under various social conditions, suggests that changes in social setting had more effect on her speech than a time-lapse of 20 years.

References
Haddican, Bill (2005), *Aspects of Language Variation and Change in Contemporary Basque*, NYU Dissertation.

Stativity and markedness in the adjective/adverb interface

Sandra Jiménez-Pareja & Salvador Valera
(University of Granada)

Keywords: adjective, adverb, interface, stativity, markedness

In a paper written in 1984, Göran Kjellmer explained how -ly suffixation is blocked by stative adjectives, because a dynamic interpretation of the resulting -ly word is not possible, hence great: greatly but not big:* bigly (cf. also Quirk, Greenbaum, Leech & Svartvik 1985, and Bauer, Lieber & Plag 2013).
Based on a corpus of 52,203 occurrences extracted by lemma from the *British National Corpus* (BNC) and the *Corpus of Contemporary American English* (COCA), this paper discusses a type of -*ly* affixation that appears to be counterevidence to Kjellmer (1984), namely -*ly* suffixation in stative adjectival bases:

(1) She snatched up the receiver and sat *wetly* on the bed  [+Predicative]  [–Adverbial]
(2) Her parents watched *worriedly* through the window  [+Predicative]  [–Adverbial]

These have been described as subject-relatedness, i.e. subject-oriented adverbs that retain only the predicative reading because their adjectival bases do not convey the typically circumstantial, adverbial meaning associated with the suffix -*ly* (Díaz-Negrillo 2014, and Valera 2014). Subject-relatedness explains why stative adjectives may take -*ly* suffixation and, in doing so, they limit the possibility of -*ly* suffixation to the formal, not the semantic level, because they do not convey adverbial meaning. However, subject-relatedness also raises the question of which adjectives display this behaviour and when, as not all adjectives do so:

(3) 50 soldiers gathered *sleepily* around a speaker  [+Adverbial]  [+Predicative]
(4) The three scarved women […] stared *proudly* ahead  [+Predicative]  [+Adverbial]
(5) I walked *quickly*, head down.  [–Predicative]
(6) She spoke *clearly*, but never raised her voice  [–Predicative]  [+Adverbial]

Regarding the first question, this paper researches the semantic profile of the -*ly* words of examples like (1) and (2), and to their distribution in the BNC. The results show a range of cases in which *wetly* and similar examples occur with various types of subjects (e.g. AGENT, AFFECTED, or FORCE) and verbs (e.g. DYNAMIC and STANCE) and where their adverbial interpretation is not possible, because the adjectival bases denote a property, mental or physical, that is not subject to control by the subject they refer to. The feature ±CONTROL has been considered in the distinction between stative and non-stative verbs and adjectives (Ljung 1975; cf. also Lakoff 1966, Dixon 1977, and Lyons 1977). The adjectival base of *wetly* has the feature –CONTROL, and is not affected by the subjects or the predicates that precede it, so *wet* is stative but still it allows -*ly* suffixation.

Regarding the second question, i.e. when this behaviour is possible, the corpus data evidence the relevance of register in view of the number of examples of the category *Fiction* in the BNC. This is in line with the argument that subject-relatedness is domain-conditioned (Bauer 2014, and Jiménez-Pareja & Valera 2018).

These cases illustrate a mismatch between the morphology and the semantics of these -*ly* words: they have the form of an adverb, but the suffix does not affect their meaning, i.e. the suffixed forms retain the meaning of their adjectival bases.

**References**

Bauer, Laurie (2014), ‘What is the plural of mouse?’ and other unhelpful questions for morphologists. Plenary lecture delivered at the 47th Annual Meeting of the Societas Linguistica Europea, Adam Mickiewicz University, Poznań, Poland, 11th-14th September 2014.


---

**Grammaticalization in Somali and the shaping of prosodic types**

Nina Hagen Kaldhol & Sverre Stausland Johnsen

(University of California San Diego & University of Oslo)

Keywords: Somali, grammaticalization, tone, accent, prosodic types

We demonstrate how grammaticalization may cause the restructuring of prosodic systems, leading to the development of new prosodic types. Illustrated with data from Northern Somali (NS), we show how a restricted tone system has developed into one of morphological tone through (a), grammaticalization of independent words to bound forms, and (b), a subsequent reanalysis of tone patterns as properties of morphological constructions rather than prosodic domains. The High tone (H) in Somali is usually assigned to the final or penultimate mora/vowel of words, with the location determined by grammatical features, e.g. gender (1) (Hyman 1981). However, some constructions have different tone patterns (2)–(3).

(1) **One H, final/penultimate**
   - *inantinán* ‘boy/girl’
   - *qaalin/qaalin* ‘camel (M/F)’
   - *damēer/dameer* ‘donkey (M/F)’

(2) **One H, non-demarcative**
   - *cūnayaa* ‘I am eating’
   - *karinayaa* ‘I am cooking’
   - *isticma layaa* ‘I am using’

(3) **Two Hs**
   - *cūnayó* ‘I am not eating’
   - *gīrīgān* ‘house.dem’
   - *gīrīgāya* ‘house.poss’

**Aims:** First, building on Moreno (1955), we analyze the progressive forms of verbs (*cūnayaa*) as grammaticalized forms of a former periphrastic construction (*cūni hayaa*). As a result of grammaticalization, the H is no longer restricted to the final or penultimate mora of words, hence its former demarcative property has been weakened. Furthermore, as the negative forms of verbs have High-toned suffixes, the grammaticalization of negative progressive *cūni hayó* to *cūnayó* has weakened the
former *culminative* property (maximum one per word) of the H. Second, we demonstrate that a similar development has taken place in the nominal system: Based on evidence from related dialects which have not undergone these changes (authors’ fieldwork; Lamberti 1986; Moreno 1955), we sketch a series of diachronic changes which have caused the determiners in NS to become bound: For example, a Noun-DEF DEM construction has first contracted to Noun DEM (4a). Word-internal sandhi alternations (e.g. voicing of [k] intervocically) have then been extended to Noun-DEM (4b) in analogy with Noun-DEF: gūrī+ka → gūriga ‘the house’ (cf. gūrī kalē ‘another house’). Similar developments have affected possessives as well.

(4) **Proposed developments**

a. *Contraction*

<table>
<thead>
<tr>
<th>house-DEF</th>
<th>DEM</th>
<th>house DEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>gūri-ga kán</td>
<td>⇒</td>
<td>gūri kán</td>
</tr>
</tbody>
</table>

b. *Extension of word-internal sandhi*

<table>
<thead>
<tr>
<th>house</th>
<th>DEM</th>
<th>house.DEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>gūri kán</td>
<td>⇒</td>
<td>gūrigán house</td>
</tr>
<tr>
<td>house DEM</td>
<td>⇒</td>
<td>house.DEM</td>
</tr>
</tbody>
</table>

We propose that grammaticalization has caused the H to lose its “accentual” properties, and a new system has developed: at a previous stage, tone was tied to metrical structure/prosodic domains, but the synchronic system is one of *morphological tone* in which *tonal schemas* are associated with morphological constructions. We argue that the “prosodic word” (PWord) is not a useful unit in accounting for the distribution of the H synchronically in Somali, contra recent analyses proposing that e.g. gūrigān consists of two PWords due to the presence of two Hs (Downing and Nilsson 2019; Green and Morrison 2016;2018). Finally, as this synchronic state results from recent diachronic developments in NS, with the original constructions still existing in other dialects, we argue that the Somali dialect continuum constitutes a rich area for the study of microvariation in prosodic systems and the development and shaping of prosodic types.

**References**


---

**Mirative meanings of predicative constructions in Latvian**

Andra Kalnača & Ilze Lokmane

(University of Latvia, Riga)
Keywords: mirativity, evidentiality, modality, predicative constructions, mirative constructions

Mirativity subsumes sudden discovery, surprise, and unpreparedness of the mind of the speaker (or the addressee) (Aikhenvald 2004, 2012). Thus, mirativity as a semantic category expresses the speaker’s attitude toward the proposition, not some objectively perceivable aspect of it (DeLancey 2001, 2012).

While Latvian does not mark mirativity morphologically, there are several predicative constructions where mirative meanings are present as either primary or secondary meanings. These constructions are partly lexicalized in that they involve specific grammatical forms of the verb and specific lexemes, mostly, pronouns. The main objective of this study, thus, is to provide, for the first time, a comprehensive analysis of mirative constructions in Latvian.

1. Evidential constructions

Latvian has a special reportive evidential marker – the oblique form of the verb (e.g., nitina, Grigorjevs 2013). As pointed out by Aikhenvald (2012), “the non-firsthand evidential may acquire a mirative connotation if something happens contrary to the speaker’s expectation and much to their regret.” (see also Rett, Murray 2013) In Latvian, this happens in exclamatory sentences involving the verb būt ‘be’ in the simple present oblique as a copula to a nominal predicate, i.e. esot, and a subject expressed by the demonstrative pronoun tas ‘this/that’, referring to a previously mentioned contextual item.

(1) Un grīda!
   Tas esot parkets!
   this.NOM.SG be.COP.OBL.PRS parquet_floor.NOM.SG

   Melns kā darva!
   ‘And the floor. Is this called “parquet floor”? Pitch black, it is.’ (A. Eglītis)

Sentences of this type express defeated expectations of sorts, with an added negative evaluation. Thus, it is possible to speak of three concurrent meanings here – evidentiality (somewhat weakened), epistemicity (disbelief, irony) and mirativity (unexpectedness, mostly unpleasant).

2. Modal infinitive constructions

Rhetorical questions beginning with the lexicalised construction kas tur ko ‘what’s the point’ express the modal meaning of lack of necessity and, simultaneously, surprise and contrariness to what is expected:

(2) Kas tur ko laipot?
   what.NOM there what.ACC manoeuver.INF

   ‘Why be roundabout?’ (LVK2018)

In sentences of this type, mirativity is an overtone to modality.

3. Modal predicative constructions

Other exclamatory constructions encoding mirativity include rhetorical questions beginning with a question word and expressing unexpectedness and surprise:

(3) Kur tas redžēts –
    where this.NOM.SG seen.PTCP.PASS.NOM.SG
    sievieši ar ieročiem!
    woman.NOM.PL with weapon.INS.PL

   ‘Who’s ever seen such a thing – women with weapons!’ (LVK2018)
Elliptical predicative constructions beginning with the word group *kas par* ‘what a …!’ and referring to something described in context are another case in point:

(4) *Ir laimējies pasēdēt blakus pilotiem turbo lidmašīnās.*

*Kas par baudu!*

what.NOM PREP delight.ACC.SG

‘I have had the luck to sit next to pilots in turbo planes. What a delight!’ (LVK2018)

In sum, mirativity in Latvian occurs in primary and secondary meanings of conventionalized and partly lexicalized constructions used in modal exclamatory or interrogative (rhetorical) sentences (on mirative sentences in general see Mocini 2014, Simeonova 2015). While these constructions have been discussed in Latvian grammars before (e.g., Nītiņa, Grigorjevs 2013) they have never been analyzed as a complex of mirative meanings with its own place in the universal typological system of mirativity.

The examples used in this study come from a variety of sources – the Balanced Corpus of Modern Latvian (LVK2018), literary prose and the press.

References


Anastasia Kharlamova  
(Institute for Linguistic Studies, Russian Academy of Sciences, Saint-Petersburg)

Keywords: phonetics, Balkan Studies, language contact, Albanian language, Aromanian language

This presentation is dedicated to analyzing the results of language contact between Albanian and Aromanian in the town of Selenica (Southern Albania). Selenica has been a predominantly Aromanian settlement for more than three hundred years; though Albanian is spoken by nearly everyone as the country’s state language, Aromanian is still preserved not just among the older residents, but by the younger generations as well.

Our goal is to determine the results of phonetic-level contact between the two languages. Based on our previous studies of Albanian-Aromanian interference, our working hypothesis has been that not only does Albanian naturally influence Aromanian, as the dominant language, but the reverse process also occurs. The currently analyzed materials from Selenica, gathered during our four expeditions to the town in 2018 and 2019, demonstrate that such is, indeed, the case.

There is clear evidence of Albanian influence in our informants’ speech. Most interestingly, the Albanian velarized [ɫ] is not just found in Albanian words that have it originally (such as Arom. [t’ilgul] ‘sound, voice type’ < Alb. tingull [tingul]), which is expected. This consonant is also in free distribution with [l], no matter whether the informant is speaking Albanian or Aromanian, so that it occurs:

- in Albanian words that originally have [l] – for example, the word for ‘basil’, borzilok [borzilok] in standard Albanian, can be pronounced either with [l] ([bɔrз’ilok]) or with [ɫ] ([bɔrз’ilok])
- in Greek loanwords ([elaða] ‘Greece’)
- in indigenous Romance words [mul] ‘many’; in fact, this pronunciation is far more frequent in our materials than the “standard” [mul], which can be found, for example, in the Aromanian texts from Greece in Bara et al. (2005)

At the same time, though, we can also see major Aromanian influence on our informants’ Albanian speech. The most important of Aromanian features that gets imported into Albanian is, in our opinion, palatalization. Palatalization is active in Romance languages – see Sawicka (1997) for a detailed analysis of the topic – but is not present at the current stage of standard Albanian. However, our informants’ speech demonstrates quite clear evidence of it: for example, the Albanian vjet [vjet] ‘year-PL’ gets pronounced as [v’jetɛ]: not just the first consonant gets palatalized, but the affricate as well, as we believe that postalveolar affricates, due to them mostly occurring in Turkish loanwords, are still foreign for the phonetic system of Aromanian speakers.

Therefore, we can see that the phonetic interference between the two idioms spoken in Selenica is active in both directions, despite Aromanian being nominally a minority language.

Acknowledgments
Funded by Russian Science Fund grant 19-18-00244.

References
Bara, Maria, Kahl, Thede, and Sobolev, Andrey (2005), Die südaromunische Mundart von Turia (Pindos), München: Biblion Verlag.

**V-V compounds and Adjunct Modification in Japanese**

Hideki Kishimoto  
(Kobe University)

Keywords: syntactic V-V compounds; raising; control; adjunct; Japanese

Compound verbs, which are formed by combining two verbs into one morphological unit, are found in many Asian languages. Japanese, too, abounds with compound verbs. It is generally agreed that compound verbs are divided into lexical V-V compounds, which behave as single words in syntactic terms despite their complex morphology, and syntactic V-V compounds, which involve complementation, i.e., the first verb is embedded under the second verb syntactically (see e.g. Kageyama 1993; Koizumi 1999). Syntactic V-V compounds are further divided into raising and control types, which can be distinguished, for instance, according to whether they allow clausal idioms to appear (Carnie 2006).

(1)  
\[
\begin{align*}
a. & \text{Kono-mise-de kankodori-ga naki-kake-ta.} \\
& \text{this-store-at cuckoo-NOM sing-ABOUT:TO-PST} \\
& \text{‘This store was about to get fewer customers.’} \\
\end{align*}
\]

\[
\begin{align*}
b. & \ast \text{Kono-mise-de kankodori-ga naki-oe-ta.} \\
& \text{this-store-at cuckoo-NOM sing-FINISH-PST} \\
& \text{‘The cuckoo finished singing at this store.’} \\
\end{align*}
\]

Given that (1a), but not (1b), can have an idiomatic interpretation, it is often claimed that the compound verb in (1a) involves raising, whereas the compound verb in (1b) involve control. Thus, the following configurations are often posited for (1a) and (1b).

(2)  
\[
\begin{align*}
a. & \text{[TP SBJ [vP2 [vP1 SBJ V1]-V2]T] (Raising)} \\
b. & \text{[TP SBJ [vP2 SBJ [vP1 PRO V1]-V2]T] (Control)} \\
\end{align*}
\]

In raising V-V compounds, the subjects are generated in vP1 projected from the lower main verb. On the other hand, control V-V compounds generate their subjects in vP2, which control PRO. While the structures for raising and control V-V compounds are often stipulated on conceptual grounds, it is shown that their configurational structures are verified empirically via adjunct modification. I argue in this paper that while the subject of control V-V compounds is generated in the upper clause, PRO is not projected because the embedded clause does not comprise a TP projection licensing null Case, contrary to the traditional view.

Empirical evidence comes from the fact that in control V-V compounds, unlike raising V-V compounds, agent-oriented adverbs (like *kenmei-ni* ‘hard’) are not allowed to appear in the embedded clause.

(3)  
\[
\begin{align*}
a. & \text{Ken-ga [(kenmei-ni) yama-miti-o aruki]-oe-ta.} \\
& \text{Ken-NOM hard mountain-path-ACC walk-FINISH-PST} \\
& \text{‘Ken finished walking the mountain path (hard).’} \\
b. & \text{Ken-ga [(kenmei-ni) yama-miti-o aruki]-kake-ta.} \\
\end{align*}
\]
Ken-NOM hard mountain-path-ACC walk-BE.ABOUT.TO-PST
‘Ken finished walking the mountain path (hard).’

Note that PRO can be readily targeted by agent-oriented adverbs.

(4) [PRO kenmei-ni aruku] kotoo-wa ii koto da.
    hard walk that-TOP good thing COP
    ‘It is a good thing to walk hard.’

Since adjuncts can be added to their modifying targets, if present, the data suggest that subjects of raising V-V compounds initially occur in the embedded clause, and raised to the matrix subject position, as depicted in (2a), but that neither subjects or controlled PROs are not present in the embedded clause of control V-V compounds, as in (5).

(5) [TP SBJ [vP2 SBJ [vP1 V1]-V2 ]T] (Control)

I suggest that in raising constructions, V1 is responsible for assigning theta roles, whereas the subject of control V-V compounds is assigned a theta role by V2 due to the transmission of theta roles from V1 to V2.

References

Control vs. complex predication: infinitival constructions in Sanskrit

John J. Lowe, Adriana Molina-Muñoz & Antonia Ruppel
(University of Oxford)

Keywords: Sanskrit, Control, Complex predication, Infinitives, Corpus linguistics

A number of verbs in Sanskrit can govern infinitival complement clauses, which at least superficially appear to involve control or raising:

(1) a. Devadattaḥ kaṭaṃ kartuṃ śaknoti
    D-NOM mat-ACC make-INF can-3SG
    ‘Devadatta can make a mat.’
(1) b. Devadattaḥ kaṭaṃ kartuṃ jānāti
    D-NOM mat-ACC make-INF know-3SG
    ‘Devadatta knows how to make a mat.’
(1) c. Devadattaḥ kaṭaṃ kartuṃ icchatī
    D-NOM mat-ACC make-INF want-3SG
    ‘Devadatta wants to make a mat.’

These constructions have received interest as potential examples of complex predicates, i.e. syntactically complex constructions which nevertheless function as single predicates in terms of
clausal argument structure. Evidence for this includes passivization possibilities (as in 2, discussed below). However, while some work has focused on the facts implied by the grammatical rules of Pāṇini’s Aṣṭādhyāyī, there has been no detailed attempt to uncover the proper syntactic analysis of these constructions by applying established syntactic criteria to actually occurring data.

(2)a. kaṭaṃ kartuṃ śakyate
   mat.ACC make.INF can.PASS.3SG
   ‘A mat can be made.’

(2)b. kaṭaḥ kartuṃ śakyate
   mat.NOM make.INF can.PASS.3SG
   ‘A mat can be made.’

The indigenous Sanskrit grammatical tradition, beginning with Pāṇini’s Aṣṭādhyāyī, acknowledges differences between verbs of desiring (as in 1c) and other verbs, but the rules concerned do not give a clear picture that could easily be connected with modern concepts such as control or complex predication. Deshpane (1980) discusses the treatment of these constructions in the indigenous grammatical tradition, and proposes that both types involve “compound” verbal constructions (based on evidence such as passivization patterns). Kiparsky (2002) likewise focuses only on the statements of the indigenous grammatical tradition (in fact, only on Pāṇini), and claims that with verbs of desiring we are dealing with control, but that with śak ‘can’ (and by implication other verbs classed with śak by Pāṇini, including jñā ‘know’), we are dealing with a “verb union process” resulting in a “single predicate for purposes of the grammar’s [argument structure] licensing constraints.” However, the grammaticality judgements claimed by Kiparsky (2002) differ considerably from those claimed by Deshpane (1980). Moreover, neither author considers in any detail what data is actually found in Sanskrit texts. A further complication is that although constructions with śak ‘can’ are universally taken to be the most like a complex predicate, it has been shown that the descendant of this verb in Hindi/Urdu, sak, is a control verb (Bhatt et al. 2011). Both Deshpande (1980) and Kiparsky (2002) argue that the passive type in (2b) provides evidence for ‘verbal union’, i.e. a complex predicate. However, we show that it is equally possible to explain the difference between (2a) and (2b) in terms of different types of control/raising structures. Yet whether we are really dealing with control/raising or complex predication depends on more than this. We present the results of a large-scale corpus search, in which we have extracted all the relevant constructions from an electronic corpus of Sanskrit running to over 5 million words. Not only patterns of passivization, but also evidence from binding, negation and adverb scope, allow us to establish the most appropriate analysis for each verb.

References

The morphosyntax of clause combining in Australian languages: a typology
Ellison Luk & Jean-Christophe Verstraete  
(University of Leuven, University of Leuven)

Keywords: clause-combining, conjunctions, finiteness, Australian, adverbial clauses

The domain of clause combining in Australian languages has often been described as morphosyntactically restricted, with relatively few specialised means to mark subordination (e.g. Dixon 2002: 86-87). Accordingly, the bulk of typological work in this domain has focused on a limited number of phenomena, like the syntactic status of so-called ‘adjoined relative clauses’ (Hale 1976, Nordlinger 2006), switch reference (Austin 1981) and complementising case (Dench & Evans 1988). This paper develops a more comprehensive typology of the morphosyntax of clause-combining in Australian languages, using a sample of 50 languages. We show that (i) specialised morphosyntactic marking is not rare, and the range of resources available is more diverse than is covered by the current typological literature, (ii) not all constructions treated as subordinate can be confidently labelled as such, and (iii) specialised morphosyntactic marking tends to be limited to a restricted semantic subdomain, with other interclausal relations relying on the co-optation of intra-clausal resources (like TAM-marking).

On the basis of our sample, we define a basic typology of five macro-types of morpho-syntactic marking, with most languages in our sample having at least three of these. We show that well-known types like dependent verb forms are quite diverse, showing various combinations of clausal reduction (e.g. ‘semi-participial forms’ in Wirangu, Hercus 1999: 119-129) and nominal-like behaviour (e.g. ‘nominalisations’ in Gooniyandi, McGregor 1990: 396-411). More importantly, there are also less-often discussed types like distributed marking, where the ‘arraying’ of markers across the two clauses defines the interclausal link (as in Nyulnyul (1), where both clauses are inflected similarly [IRR+IMP ‘past unrealized’] to produce a past conditional construction), and a broad range of conjunctions, often thought to be atypical in Australia but found in at least half of the languages in our sample.

(1) Nyulnyul (Nyulnyulan)

\textit{nga-la-kan \hspace{1cm} derby-ung \hspace{1cm} i-li-jabal-an-ngay}

1MIN.NOM-IRR-carry-IMP \hspace{1cm} Derby- ALL1 \hspace{1cm} 3NOM-IRR-ask-IMP-MIN.ACC

‘I’d have taken him to Derby if he had asked me.’ (McGregor 2011: 658)

Based on this database of morphosyntactic phenomena, we take up Nordlinger’s (2006: 26) call to develop a more fine-grained analysis of the morphosyntax of clausecombining. Our alternative typology achieves this by validating the phenomena against basic behavioural criteria such as clausal ordering potential, sites of inflectional reduction, and the coding of dependents. We demonstrate that language-internal distinctions may be obscured by \textit{a priori} descriptors such as ‘subordination’ and ‘non-finite’.

Finally, our analysis shows that while there is a broad range of morphosyntactic devices in the sample, specialised means are limited to specific semantic subdomains of clause-combining. For example, purposive and contemporaneous relations tend to have morphosyntactically specialised marking, while causal and contrastive relations tend not to (compare also Givón 1980). These semantic restrictions may actually be what underlie the classic idea that Australian languages are morphosyntactically restricted in clause-combining. For other semantic subdomains, clause combining appears to rely on recruiting existing clause-internal resources (compare Merlan 1981, McGregor 1988, and Verstraete 2010), for instance modal marking being co-opted for suspended propositions.

\textbf{References}


“Illya Makarchuk
(National Research University Higher School of Economics, Moscow)

Keywords: lexical aspect, Aktionsart, diminutive, scalarity, evaluative morphology

This paper describes a family of aspectual derivations that I will call *verbal attenuative*. Verbal attenuatives denote events that are *incomplete, deficient* or in other way *lesser* than the base. To my knowledge, there has not previously been any typological overview of such an aspectual meaning other than two sketches of a typology by Volkov & Ignatenko (2017) and Tatevosov (2003).

One class of verbal attenuatives marks events of lower intensity than normal and often comes along with a meaning of pluractionality. Below is an example of French verbs in *-oter/-iller* (Stosic, Amiot 2014).

(1) *neig-er* ‘to snow’ → *neige-ot-er* ‘to snow lightly’
(2) *mord-re* ‘to bite’ → *mord-ill-er* ‘to nibble’
(3) *travaill-er* ‘to work’ → *travaill-ot-er* ‘to work half-heartedly’

Another class is events that are brief, transient or partial. Below is an example of Hill Mari Attenuative (my own fieldwork data). In (4) the duration of the event is short, in (5) the duration of the result state is short, and in (6) the result state is partial. Another example of this class is a well-known case of Slavic Delimitative in *po* (see e.g. Kagan 2016).
The third class is instantaneous, momentaneous events. Below is an example of Moksha Semelfactive (my own fieldwork data). In (7) a derivative means a single act of a internally plural event, in (8) an attenuated brief action, and in (9) an inchoative. A known case of Slavic Semelfactive in -nu is also of this class (see e. g. Makarova, Janda 2009).

(7) tərnat- ‘to shiver’ → tərnat-kst- ‘to shiver once’
(8) raka- ‘to laugh’ → raka-kst- ‘to laugh a little bit’
(9) urakəd- ‘to wail’ → urakəd-kst- ‘to start wailing’

In the talk, I will provide examples from a convenience sample of languages and discuss other cases of the domain in question. I argue that the three classes of verbal attenuative exemplify different types of measuring an event. Derivations like (7-9) are event-minimizing, that is taking the minimal part of an event. Derivations like (4-6) are event-portioning, taking a portion of an event. Derivations like (1-3) are different from previous two: they do not take a part of an event but compare the event to some norm. In my presentation I will discuss these classes in depth and will try to provide ways to distinguish between them.

Acknowledgements
The research has been supported by RFBR, grant № 19-012-00627.

References
Volkov, Oleg and Daria Ignatenko (2017), Pokazateli nepolnoty situatsii, ili Limitativno-semelfaktivnaya derivatsiya v yazykakh mira [Markers of incomplete situation, or Delimitative-semelfactive derivation throughout languages of the world]. Paper presented at the 14th Conference on typology and grammar for young scholars, November 2017, St. Petersburg.

A semantic map for progressive and the notion of proto-process
Keywords: progressive, imperfective drift, proto-process, actionality, expressive.

Typological studies dedicated to PROGRESSIVE (e.g. Dahl (ed.) (2000: §4) and Vafaeian (2018)) have not yet provided extended analyses of accumulated data. The semantic map model has yet to be applied to the progressives.

Based on the data from 36 progressive markers in the languages of Europe (Bertinetto et al. (2000) i.a.) I’ve plotted a semantic map for PROGRESSIVE.

The map consists of 15 functions defined relative to sentences of the PROGQ questionnaire (Bertinetto et al. 2000). The set of functions is subdivided into three groups.

Actional functions (green) consist in changing the actional class of the verb. For example, Spanish Progressive used in this function freely co-occurs with the perfective Simple Past (1). If the former were semantically imperfective, this should have been impossible. Therefore, its function in such cases is only to reassign the verb to the atelic actional class <P, P> (Tatevosov 2002).

(1) estuvo leyendo todo el día
was:SP:3SG read:GER all the day
‘S/he spent the whole day reading’. (Bertinetto et al. 2000: 535)

Expressive functions (lilac) consist in expressing (rather than describing) the speaker’s beliefs and feelings (Kroeger 2019: §2.6) with regards to the situation.

 Aspectual functions (red or unmarked) imply inclusion of the Topic Time in the Time of Situation (Klein 1994) and do not imply expressive meaning.

It is desirable to provide a unifying semantics for PROGRESSIVE as a cross-linguistic category-type, from which the observable variation in language-particular progressives can be derived.
I assume that the progressives which only have the “agentive” function impose the following constraints on the predicate that they are applied to and on the situation described: (i) agentive subject; (ii) perceptual accessibility to the speaker; (iii) lack of truth-value relative to moments; (iv) temporariness; (v) restrictedness to one occasion (Cusic 1981). If a progressive has both the “agentive” and the “absentive” functions, this progressive requires that all these constraints except (ii) be obeyed. Etc.

This set of constraints is encapsulated in the notion of Proto-Process. The semantics for PROGRESSIVE can be formulated as inclusion of TT in the Time of a Situation corresponding to the Proto-Process. The core differences between progressives amount to different levels of strictness of the correspondence condition.

The notion of Proto-Process is diachronically relevant. Only the aspectual functions are derivable via this notion and — coincidentally — only these functions are retained in imperfectives. Thus, PROG imperfective drift (Bertinetto et al. 2000) can be modeled as an incremental weakening of the Proto-Process correspondence condition.

In the talk I will discuss in more detail the strengths and limitations of the current methodology.

Glosses: GER — gerund; 3SG — 3 person singular; SP — simple past.

Acknowledgements: The study was implemented in the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University) in 2020.

References


G. Vafaeian. Progressives in use and contact: A descriptive, areal and typological study with special focus on selected Iranian languages. PhD dissertation, Department of Linguistics, Stockholm University, Stockholm: 2018.

On some non-canonical adjectival past participles in Romanian

Carmen Mirzea Vasile
(University of Bucharest)

Keywords: adjectival past participles, Romanian participle, deverbal adjectives, non-canonical participles, modal value of participles

The general literature on the adjectival past participles in Romanian discusses, in the first place, the adjectival passive past participles (1a), and then the adjectival active past participles (1b-e) (see Pană Dindelegan 2013: 222-233). Recently, the adjectival active past participles have received an extensive
description (Niculescu&Mîrzea Vasile 2017, 2019); the main findings of the recent surveys of old Romanian and Daco-Romanian varieties are as follows: (i) compared to modern standard Romanian, non-standard (and old) Romanian displays a larger inventory of participial adjectives with active value; (ii) the participles of verbs expressing a (non-motional) controlled process can be adjectivized (1d), thus, unaccusativity is not a strong factor determining the possibility of adjectivization of participles; (iii) adjectival participles can preserve the direct object of their transitive source verb (1e).

(1) a. carte 
book.DEF.F.SG read.PPLE.F.SG ‘the book read’
b. bărbați
man.M.PL drunk.PPLE.M.PL ‘drunk men’
c. fete plecate în străinătate
girl.F.PL left.PPLE.F.PL in foreign-countries ‘girls who have left abroad’
d. cu zâmbită (...) față (DVS.1682-6: 420’)
with smiled.PPLE.F.SG face.F.SG ‘with a smiling (...) face’
e. băiat făcut armata (TDM.I.Olt, în Marin 1991: 51)
boy.M.SG done.PPLE.M.SGarmy.DEF.ACC ‘a boy who has completed the military service’

The adjectival participle in (1a) is a canonical one, whilst the adjectival participles in (1d,e) are atypical and those in (1b,c) are, at least, non-core.

The present paper represents an empirical-descriptive approach to some marginal adjectival past participles in non-standard and old Romanian. More precisely, on the basis of textual (and lexicographic) sources, we examine some types of adjectival past participles that are not canonical (cf., a.o., Albrespit 2007, Alexiadou & Schäfer 2013) as far as the underlying arguments modified (2a,b) and their modal value (2c,d) are concerned. So far, these types of Romanian adjectival past participles have not been thoroughly discussed in the literature.

(2) a. apă (...) negriță
water.F.SG unspoken.PPLE.F.SG ‘untouched/unspoilt/unspoken (fresh) (...) water’
b. minune mirată (DVS.1682–6: 416’)
miracle.F.SG amazed.PPLE.F.SG ‘amazing miracle’
c. subt chipure văzute și atinse (Eustratie Log., 1645, in DLR)
under form.N.PL seen.PPLE.F(=N).PL and touched.PPLE.F(=N).PL ‘in a visible and perceivable form’
d. troiță (...) neîmpărțită (CL.1570: 29’’)
trinity.F.SG undivided.PPLE.F.SG ‘(...) undivisible trinity’

Firstly, we present some underlying semantic-syntactic configurations with verbs that yield unusual adjectival participles (cf. Shagal 2019). Besides the adjectival past participles with an active value (oriented towards the Agent of an transitive verb (1b) and towards the Agent (1b) or the Patient of an intransitive verb), already discussed in the literature (Niculescu&Mîrzea Vasile 2017, 2019;
Dragomirescu 2016 and the references therein), we describe other types of non-canonical adjectival past participles, that are modifying (i) an underlying locative (2a), (ii) an argumental PP (2b), and (iii) an indirect object in Dative. We will pay a special attention to the variation displayed by these underlying structures (dative/accusative variation, Locative PP/Locative DO, etc., see also Pană Dindelegan 2016: 72-89), which implies that it is not always possible to determine with certainty which underlying argument the participle modifies.

Secondly, we look at the **adjectival past participles with a modal value** (2c,d) (for general details, see Rainer 1999; Oltra-Massuet 2014; Martin&Pitteroff&Pross 2016; Fábregas&Marín 2017). Whereas in contemporary standard Romanian there exist few negative adjectival participles which can convey a sense of potential-elative (3), in old Romanian texts, adjectives of this type (2c,d) are much more numerous (in our sources, more than 70, most of them negative ones) and more varied. These adjectives are mentioned in passing in Brăescu&Nicula (2019) and Mîrzea Vasile (2012).

(3) **frumusețe nespusă**

beauty.F.SG unspoken.PPLE.F.SG

‘unspeakable beauty’

Our second goal is thus to provide more data on these modal adjectives (a large inventory; the exact modal value; the relation with synonymous affixal word-formation and syntactic patterns, etc.).

---

**References**

Albrespit, Jean (2007), Atypical passives, *Études anglaises* 60 (4), 466–482.


Grapheme size is processed like stress: Experimental evidence from Chinese script

James Myers
(National Chung Cheng University, Taiwan)

Keywords: phonology, psycholinguistics, Chinese, writing systems, regular vs. irregular

Despite obvious functional and social differences between writing systems and spoken/signed languages, they share many nontrivial formal and psychological similarities (see, e.g., Evertz 2018, Myers 2019). The present study demonstrates yet another similarity: like stress in spoken languages, the processing of grapheme size in Chinese characters depends on its positional regularity.

Studies on a variety of languages (e.g., Cutler and Clifton 1984, Cutler and Van Donselaar 2001, Domahs et al. 2013) have found an asymmetrical effect of mis-stressing on spoken word recognition. For example, Cutler and Clifton (1984) found that when English speakers make lexical decisions on words without vowel reduction, they recognize regular (stress-initial) disyllabic words (nútmeg) more quickly than incorrectly stressed variants (nutmég), whereas mis-stressing irregular (stress-final) words (typhóon as týphoon) has little effect (Figure 1a). This asymmetry suggests that expectations about regular stress rules can override lexical exceptionality.

The analogue to prosody in Chinese script relates to character-internal position-based sizes. Individual characters are written from left to right and top to bottom, so as with final lengthening in spoken and signed languages (Beckman and Edwards 1990, Sandler 1993), character strokes and constituents are generally larger at the right and bottom (Myers 2019). In particular, semantic radicals, which are mostly small, strongly favour the left edge, with the phonetic component to its right (see Figure 2).

To look for possible parallels with the stress results, we selected 50 regular and 50 irregular characters (in the traditional script still used in Taiwan), matched in token frequency, stroke number, and constituent type frequency, and for each created an incorrectly shaped variant (enlarging the semantic radical and shrinking the phonetic component) using Wenlin (https://wenlin.com/). Wenlin was also used to recombine the constituents (kept in their lexical positions) into 100 nonlexical characters, along with their incorrect shape variants.
A lexical decision experiment with 40 traditional Chinese readers confirmed that incorrectly shaped characters were recognized more slowly. Although this was true regardless of shape regularity, responses to regular characters showed a significantly weaker influence of character frequency, as if judgments for them were initiated pre-lexically. In an acceptability judgment experiment with just the nonlexical characters on an additional 42 readers, regular nonce characters (with left-edge semantic radicals) showed a significantly stronger effect of shape correctness than irregular ones (Figure 1b), as if readers mentally regularized the latter.

Our results thus replicate those seen in the stress experiments, albeit more weakly (the overriding of lexical access by shape regularity may be limited in characters, which represent morphemes and thus tend to have higher token frequencies than spoken words, which include morphologically complex ones). These results thus join many other recent experimental and corpus-based findings in revealing nontrivial similarities between writing and spoken/signed language, though the underlying causes still require further exploration.

Fig 1a. Reaction times for (in)correct stress in irregular vs. regular English words

Fig 1b. Acceptance rates for (in)correct shape in irregular vs. regular Chinese characters

Fig. 2. Regular (left) and irregular (right) character shapes

References


Phonological complexity and suprasegmental variables in contact situations

Ricardo Napoleão de Souza & Kaius Sinnemäki
(University of Helsinki)

Keywords: phonology, suprasegmentals, areal linguistics, typology, linguistic complexity

Studies in language contact have mostly focused on morphosyntactic structures, resulting in phonological variables rarely being addressed beyond changes to segment inventories. Despite the wealth of data in these investigations, the overall lack of information regarding the suprasegmental domain (e.g. syllable structure, tone systems, etc.) creates a major gap in our understanding of the role of phonological structure in language contact situations. On the other hand, the limited data on suprasegmental phenomena that do exist tend to emphasize how contact brings about structural simplification, as in syllable patterns (Kang 2011). The current paper evaluates the occurrence of contact-induced phonological change on the suprasegmental domain using a geographically balanced sample. Specifically, it addresses cases of reported increases in phonological complexity, using syllable structure as well as tone systems, lexical stress systems, and vowel and nasal harmony phenomena as variables.

Method. We gathered a convenience sample using the sampling method described in Maddieson et al. (2014-2019) and Maddieson et al. (2013). This method uses geographical area as well as genealogical affiliation to draw six different groupings: (1) Africa; (2) Central & South America; (3) Europe, South, & West Asia; (4) East & South-East Asia; (5) North America; and (6) Oceania. For each area, we then searched the contact literature and/or areal surveys (e.g. Salmons 1992, Matisoff 2001, Matras 2009, Foley 2010, Mithun 2010, Epps 2019) for data on contact-induced changes in the variables of interest. Individual reference grammars and articles provided additional data. The final sample contained 41 languages belonging to 25 language families from all pre-determined areas.

Only changes that led to increases in complexity are discussed here. Complexity was operationalized as an absolute measure (e.g. the number of tones in a language), but also following Dahl (2004), so that the longer the description of a phenomenon, the more complex it is. For instance, we considered the development of unpredictable stress placement due to contact as an increase in suprasegmental complexity compared to fixed stress. We further coded the data were also coded genetic affiliation, and donor language(s).

Results. Contact-induced increases in complexity in the suprasegmental domain occurred in all areas of the world. Of these, changes to syllable structure occurred in 53 percent of our sample - somewhat unsurprising since all languages have syllable structure - followed by changes to stress systems (27 percent), tone systems (13 percent), and nasal harmony systems (7 percent). These may reflect
bibliographic bias and overall deficiency in the reporting of suprasegmental data rather than actual prevalence of the phenomena. As such, no statistical tests were run on the current data.

**Discussion.** Data from 41 languages from all across the world suggest that contact may also lead to complexification of phonological structure. Our results further indicate that the suprasegmental domain provides fertile ground for inspecting contact-influenced increases in linguistic complexity. Overall, the data reviewed here highlight the relevance of phonological structure as a variable in language contact investigations, which have been mostly preoccupied with morphosyntactic structures.

**References**


Salmons, Joe (1992), *Accentual Change and Language Contact: A Comparative Survey and a Case Study of Northern Europe.* Routledge.

---

**Exploring non-culmination in Bantu**

Bastian Persohn  
(University of Hamburg, Rhodes University)

Keywords: Bantu, semantics, aspect, actionality, non-culmination

This paper examines two related aspectual phenomena: non-culminating accomplishments (NCAs) and non-culminative readings of implied-result verbs.

Concerning NCAs, the combination of Vendler’s (1957) accomplishments and the perfective aspect is typically interpreted as the obtainment of the inherent endpoint. This reading, is, however, defeasible in many languages. In absence of the non-contradictory continuation, the first clause in (1) would be understood as conveying the construction of a complete house (numbers indicate noun classes).
(1) Xhosa (Bantu; xho)

\[
\text{Nd-akh-e indlu kodwa a-nd-a-yi-gqiba.}
\]

SBJ1sg-build-PFV 9.house but NEG-SBJ1sg-CONSEC-OBJ9-finish

‘I built a house, but I did not finish it.’

With implied-result verbs the defeasible culmination pertains to the attainment of a resultant state on part of the object. This is illustrated in (2). With many verbs, we can distinguish between a construal of a partial change, as in (2a), and the absence of any change, as in (2b)

(2) Nyakyusa (Bantu; nyy)

(2a) Mwakyoma a-ŋ-dingseenie Mwalokama mbaatiko,

Mwakyoma SBJ1-sbj expl-PFV Mwalokama 10.rule

leelo a-ka-si-kola syosa

now/but SBJ1-NEG.PFV-OBJ10-grasp/PFV 10.all

‘Mwakyoma explained Mwalukama the rules, but Mwalukama did not grasp them all.’

(2b) … leelo Mwalukama a-ka-syangania=po

now/but Mwalukama SBJ1-NEG.PFV-understand=PART

‘… but Mwalukama did not understand anything.’

While both phenomena have been documented in a geographically and genetically diverse array of languages, and while some cross-linguistic commonalities as well as axes of variation have been identified (Martin 2019; Martin et al. 2016; Demirdache and Martin 2015), they have so far not been studied for any continental African language. This paper therefore presents an explorative approach to the two phenomena in two Bantu languages: Xhosa (South Africa) and Nyakyusa (Tanzania), with data coming from targeted elicitation during fieldwork.

It was found that in both Xhosa and Nyakyusa the use of the perfective aspect with accomplishment predicates requires the relevant (sub-)eventuality to be terminated, even though the two language differ significantly as to how far the eventuality needs to have unfolded.

In Xhosa, verbs of consumption and destruction, as well as transitive change-of-state verbs allow for an NCA construal from the earliest increments of the eventuality on. With verbs of creation and performance, on the other hand, an NCA construal requires that the created object resemble the entity described by the object noun phrase; a similar pattern has been described for Hindi (Singh 1994).

In Nyakyusa, felicity of NCA construals turn out to be less dependent on predicate class. However, NCAs in this language compete in interesting ways with the employment of a verbal partitive marker; the latter is illustrated in (3).

(3) Nyakyusa

\[
n-nw-ile=po ufinga jango
\]

SBJ1sg-drink-PFV=PART 9.wine 9.POSS1sg

‘I drank of my wine.’

As for implied-result verbs, Xhosa and Nyakyusa conform to a widespread pattern that zero-change construals are more readily available with subjects having agentive properties (or that stand in for an implicit agent), whereas partial-change construals do not require any agentivity.

These and further findings bear on our understanding of the interaction between the lexical/phrasal dimension of aspectuality and morphosyntactic aspectual operators (see Sasse 2002), both in the Bantu languages (see Nurse 2008) and beyond.
References

Syntactic difficulties in bilingual high-school students with Italian L2.
Two case studies of syntactic training.

Elisa Piccoli, Anna Cardinaletti & Francesca Volpato
(University Ca’ Foscari of Venice)

Keywords: bilingualism, Italian relative clauses, syntax, syntactic movement, training

Overview and goals. This paper investigates the linguistic competence of bilingual high-school students with Italian L2 compared to monolingual Italian age peers in the repetition and production of complex syntactic structures. Based on studies on school-age children (Prévost et al. 2010), this study investigates whether difficulties in mastering syntactic structures derived by syntactic movement persist in high-school students. Furthermore, this study aims at discussing the improvement and generalization effects promoted by syntactic training (Thompson et al. 2003, 2007, Levy & Friedmann 2009, and D’Ortenzio 2019).

Participants. G1 consists of 8 bilingual high-school students (age range: 14-20). Three of them were born in Italy from migrant families and five of them moved to Italy when they were children (age range: 3-5). G2 consists of 30 Italian age peers.

Materials. Two oral tasks were administered. The sentence-repetition task includes 33 experimental sentences, both simple (left-dislocation sentences) and complex (long-distance subject and object wh-questions, clefts, oblique relatives) and 16 filler sentences of the same length (Del Puppo et al. 2016). The production task contains 20 trials eliciting subject, object, and oblique relative clauses. Two bilingual participants were also assessed after training using these same materials.

Assessment of linguistic skills (before training). G1 produced more ungrammatical sentences than G2 in both tasks and more sentences typical of sloppy registers in the production task. Overall, bilingual individuals have difficulties in mastering complex structures derived by syntactic movement. Focusing on the repetition of oblique relative clauses, G1 and G2 did not however differ. This may be due to the frequency at which they are used at school or in formal contexts. Monolinguals and bilinguals are exposed to this language variety to the same extent.
Individual analyses show that the participants that were exposed to Italian after age 4 are significantly less accurate than those exposed earlier (p < .05). The students born in Italy from migrant families showed native-like performance. This study confirms 4 years as the cut-off point for native-like performance (Meisel 2009, and Unsworth et al. 2014).

Two case studies. Two bilingual students (G1) were administered a syntactic training based on Levy and Friedmann (2009) and focused on relative clauses, the most complex structures. After the training, their performances significantly improved in both tasks. They also improved in the repetition of untrained sentences: clefts, wh-questions. Generalization effects were found in similar studies on other populations and show that training generalizes from more over less complex structures of the same syntactic type (Thompson et al. 2003, 2007, Levy & Friedmann 2009, and D’Ortenzio 2019).

Issues for future research. Persistent difficulties with structures of the formal register lead us to reflect upon the language varieties monolingual and bilingual students deal with at school and to develop protocols of syntactic training helping them improve their linguistic competence in formal Italian.

References


Approximative Adverbs in Mandarin Chinese and Scalarity

Shu-Ing Shyu

<not updated>
On the evolution of TAM categories and lexically defined aspect: the case of Pluperfect and Future Perfect in Slavic

Dmitri Sitchinava
(Russian Language Institute and Higher School of Economics, Moscow)

Keywords: pluperfect, future perfect, lexically defined aspect, Slavic languages, non-compositionality

The Slavic languages are known to have, historically, both lexically defined aspect (the so-called “Slavic-type aspect: Perfective vs. Imperfective) and TAM categories such as Aorist, Imperfect, Perfect, Pluperfect and Future Perfect. These properties can combine in different ways (such as Aorist Perfective and Imperfect Perfective in Old East Slavic or Modern Bulgarian). The theoretical importance of their co-occurrence is relevant for the discussion of the status of aspect and of whether Perfect-like forms are temporal or aspectual in nature. It has been claimed (cf. Benacchio 2002) that the TAM system of Slavic was reduced to three tenses due to the rise of the lexically defined aspect.

A theoretically important case in point is the correlation between the Pluperfect and Future Perfect, the so-called absolute-relative tenses, and the lexically defined aspect in Slavic. Berger (2012, for Czech) and Barentsen (2015, for Bulgarian) claim that the changes in the co-occurrence of Pluperfect with both aspects are explained by the relative status of the former, and that the aspect supersedes Pluperfect as a marker of temporal sequence.

However Pluperfect, as described by many authors since Dahl 1985, and Future Perfect, as Pen’kova 2018 shows, both exhibit non-compositional polysemy. Such uses as cancelled result, frame past, initial point of narration, and experiential for Pluperfect, or irreal, evidential, and conditional for Future Perfect, are widespread typologically.

As the sources of our study we used massively parallel texts and monolingual corpora of Slavic languages (including historical corpora of East Slavic such as TOROT or RNC) and translation questionnaires for native speakers of Slavic languages.

It appears that in Russian co-occurrence of Pluperfect (and its successor, the bylo particle) with Imperfective (IPF) rises after 1450 and decreases by 1700s, with a further shrinking by 1930s. Czech (studied by Berger) largely follows that pattern, eventually losing Pluperfect other than in Subjunctive. But in Belarusian and Ukrainian the Imperfective rate remains stable after the 18th century, whereas the development of aspect in East Slavic was much the same. In modern South Slavic languages, even with a rich TAM system, Pluperfect Imperfective is rare as compared to the previous diachronic stage.

At the same time the Russian form, is shown to evolve into the direction of modality or uncompleted action, and the BCMS forms, towards the expression of cancelled result, losing other uses. Future Perfect is known to change, in Polish, Rusyn and Slovene, and in the Western Ukrainian dialects, to simple Future. With the exception of Slovene, the perfective forms in Future were also lost.

We suggest that these changes were driven by the semantic evolution of Pluperfect and Future Perfect towards restricted non-compositional non-relative uses, and by a decline of uses that favored the disregarded aspect (e.g. frame past or experiential for imperfectives in Pluperfect, and cause or temporal precedence for perfectives in Future Perfect). This change was dictated neither by the semantics of temporal precedence nor by the evolution of the Slavic lexically defined aspect (perfective vs. imperfective) per se.

References
The development of linguistic complexity: Evidence from recent changes in Present Day English

Vasileios Symeonidis

On the morphosyntactic status of the particle *de* in the V-*de* constructions in Mandarin Chinese

Jen Ting

(National Taiwan Normal University)

Keywords: morphosyntactic, V-*de* construction, Chinese, local dislocation, postsyntactic

This paper aims to provide a morphosyntactic analysis of the particle *de* in the V-*de* constructions in Mandarin Chinese. Despite the extensive attention received by syntactic behaviors of the V-*de* constructions, the issue regarding the category of the morpheme *de* has often been dismissed by simply glossing it as part of the V-*de* sequence (see Huang 1988, 1992, and Lin 2000). We argue that *de* merges with the V host at PF in both resultative (1a) and descriptive (1b) constructions. We first show that the V host and *de* must be discrete entities at PF. Assuming that right node raising (or RNR) effects are derived by PF deletion (Cheng and Ting 2015), the fact that the “*de* + result XP” sequence can be an RNR target indicates that *de* must be elided together with the result XP as in (2) (cf. Lasnik’s 1995 argument for tense affixes in English merging with their V hosts at PF). We then argue that the merging of *de* and its V host at PF is sensitive to hierarchical structure in resultative sentences but is sensitive to linear adjacency in descriptive sentences, based on the puzzling asymmetry between (3a) and (4a): once the offensive postverbal NP in the descriptive construction is moved away by topicalization in (4b), the sentence becomes well-formed. Sensitivity to the hierarchical structure for the merging of *de* and the V host is exactly what underlies the difference between tense morphemes and comparative/superlative morpheme –*er*/*est* in English. As argued by Embick and Noyer (2001), the postsyntactic movement of the superlative morpheme is blocked by the intervening adverb in (5b). This claim, however, is challenged by Kiparsky (2005) and Emonds (2009). The contrast in (4) exhibited by the descriptive *de* will lend further support for the kind of affix hopping that is sensitive to adjacent constituents instead of syntactic headedness. Results of this study will shed light on the
morphosyntax of V-de constructions in Chinese and also reveal theoretical implications for both syntax and morphology.

(1) a. Xiaozhang pao-de hen lei.
   Xiaozhang run-DE very tired
   ‘Xiaozhang ran to such an extent that he got tired.’
   b. Xiaozhang pao-de hen kuai.
   Xiaozhang run-DE very fast
   ‘Xiaoming ran fast.’

(2) Laozhang zuotian gongzuo de hen lei; er Laoli zuotian ye gongzuo
   Laozhang yesterday work DE very tired and Laoli yesterday also work de hen lei.
   DE very tired
   ‘Laozhang worked to such an extent that he was tired. Laoli did, too.’

(3) a. Laozhang da-de Laoli tou-po-xie-liu.
   Laozhang beat-DE Laoli head-break-blood-flow
   ‘Laozhang beat Laoli so much that Laoli was badly hurt.’
   b. Laoli, Laozhang da-de tou-po-xie-liu.
   Laoli Laozhang beat-DE head-break-blood-flow
   ‘Laoli, Laozhang beat so much that Laoli was badly hurt.’

(4) a. *Laozhang da-de lanqiu hen hao.
   Laozhang beat-DE basketball very good
   ‘Laozhang plays basketball very well.’
   b. Lanqiu, Laozhang da-de hen hao.
   basketball Laozhang beat-DE very good
   ‘Basketball, Laozhang plays very well.’

(5) a. Mary is the most amazingly smart person …
   b. *Mary is the t amazingly smart-est person …

References
Emonds, Joseph (2009), Inflectional merger = PF merge of bound morphemes, Lecture given at University of Newcastle.
Lin, Huei-Ling (2000), V2 as the main predicate in Chinese descriptive constructions, in Zheng-sheng
How cyclic is the negative existential cycle in Indo-European?

Annemarie Verkerk, Luigi Talamo & Shahar Shirtz
(Universität des Saarlandes, Universität des Saarlandes, independent researcher)

Keywords: typology, diachrony, negation, Indo-European, phylogenetic comparative methods

This paper examines the Negative Existential Cycle (henceforth NEC, Croft 1991) in a broad sample of over 100 Indo-European languages, testing its validity using both analytic and quantitative methods. The NEC is a typological hypothesis on the diachronic relationship between different types of negative existential constructions and their relation to standard verbal negation. In (1), for example, the existential predicate is negated by the standard verbal negation marker while in (2), a special negative existential marker, distinct from verbal negation, is used.

<table>
<thead>
<tr>
<th>Language</th>
<th>Text</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sicilian (Italic, Italian Romance)</td>
<td>i iatt-I sevvatic-I nun exist-unu</td>
<td>‘There are no wild cats’</td>
</tr>
<tr>
<td>Modern Irish (Celtic, Goidelic)</td>
<td>Níil cait fiáin ann</td>
<td>‘There are no wild cats.’</td>
</tr>
</tbody>
</table>

Croft (1991) recognizes six negative existential construction types and posits a unidirectional, cyclic, diachronic pathway connecting them. Recent work (Veselinova 2013, 2014, 2015, 2016) finds that the changes in negative existentials sometimes do not take the form of a cycle: the six stages of the NEC do not necessarily follow each other, languages can have several constructions of different types, and the stages are not equally stable. The NEC also interacts with other cycles through which negation markers arise (van Gelderen to appear).

The current paper asks whether the attested changes in the expression of negative existentials are sufficiently cyclical to allow for the NEC to arise, or whether other types of changes are attested frequently enough in order to say that the NEC does not explain the patterns in the expression of negative existence. To answer this question, we have gathered data on negative existentials and standard negation in over 100 contemporary and extinct Indo-European languages, building further on Verkerk & Shirtz (to appear).

First, we reconstruct the history of negative existential construction types in Indo-European using a parsimony approach and show that both cyclical and non-cyclical changes between construction types are required to accurately describe changes in negative existentials (in line with Veselinova 2014, 2016). We show that this is in many ways similar to a “traditional” analytic approach to morphosyntactic reconstruction (following, e.g., Barðdal & Gildea 2015).

Then, we use phylogenetic comparative analyses (Pagel, Meade, and Barker 2004) to test whether Croft’s (1991) NEC explains the attested distribution of negative existential construction types better than alternative models. We discuss the requirements these methods impose on sample size and the amount of attested variation, and the difficulties in modeling when languages have several negative
existential construction types. We test Croft’s NEC proposal against alternative models such as the reverse direction of the NEC, and models based on pathways attested in our first analysis and in Veselinova 2014, 2016. Combining initial results from both analyses, we suggest that Croft’s NEC is indeed attested in Indo-European, but does not fully explain the diachronic distribution of construction types.

References

Noun Incorporation: Early Saturation as well as Early Restriction

Niina Ning Zhang
(National Chung Cheng University)

Keywords: agent-oriented resultative, incorporation, restriction, saturation, Mandarin

This talk identifies a new type of (pseudo) noun incorporation (NI). Two types of V-V resultatives in Mandarin Chinese are compared: the patient-oriented resultative (PR) and the agent-oriented one (AR).

(1) a. Yaoyao chewan-po-le na shuang xie. (PR)
   Yaoyao wear-broken-PRF that pair shoe
   ‘Yaoyao wore that pair of shoes so much that it got broken.’

b. Yaoyao chewan-ni-le na shuang xie. (AR)
   Yaoyao wear-tired.of-PRF that pair shoe
   ‘Yaoyao wore that pair of shoes so much that she got tired of it.’

The semantic grouping of the first verb (V1) and the post-verbal nominal of an AR may have idiomatic readings and selection effects. But unlike in PR, V1 in AR shows a weak reachability which
is parallel to that of an incorporated nominal in the familiar type of NI: no quality modification and no anaphoric linking to a proform (Mithun 1984).

V1 can be modified by a manner adverbial in a PR, but not in an AR.

(2) a. Ta hen xiaoxinde ga-jian-le naxie zhaopian. (PR)
   he very carefully wipe-clean-PRF those photo
   ‘He carefully wiped those photos clean.’

   b. Ta (*hen xiaoxinde) ca-ni-le naxie zhaopian. (AR)
   he very carefully wipe-tired-of-PRF those photo
   ‘He had wiped those photos so much that he was fed up with them.’

V1 can be the antecedent of the proform use of nong ‘make’ in a PR, but not in an AR.

(3) a. Yaoyao wan-huai-le ziji-de wanju, Lili ye (PR)
   Yaoyao play-broken-PRF self-DE toy Lili also
   zhe yang nong-huai-le ziji-de wanju.
   make-broken-PRF self-DE toy
   ‘Yaoyao played her toy such that it got broken, and so did Lili.’

   b. Yaoyao wan-ni-le ziji-de wanju, *Lili ye zhe yang
   Yaoyao play-bored-PRF self-DE toy Lili also so
   nong-ni-le ziji-de wanju. (AR)
   make-bored-PRF self-DE toy
   Intended: ‘Yaoyao played her toy such that she got bored, and so did Lili.’

I argue that in an AR, V1 and the post-verbal nominal undergo a type of NI parallel to the familiar type, i.e., in (1b), na shuang xie is incorporated into chuan. In such NI, it is the verb that exhibits root-like properties. Such properties are viewed as a consequence of an early combination of a verb and a nominal, before each of them takes an event-argument, in a neo-Davisonian semantic perspective. The early combination is early restriction in NI examples such as strawberry-picking (e.g., Chung & Ladusaw 2004; Dayal 2011), but early saturation in the type of NI reported here.

No root-like properties are found in the V1 of PR, and thus no NI is seen in PR.

An incorporated nominal can be a phrase (Massam 2001), and so is the post-verbal nominal in AR. Such a nominal and the incorporating verb can be separated in their surface positions (Dobrovie-Sorin et al. 2006; Dayal 2011), and so are the nominal and V1 in AR. An incorporating verb is never a verb of change of state (Johns 2007), e.g., freeze, so is V1 of a resultative, which must denote an activity. All of these are compatible with the newly identified type of NI.

References
WORKSHOP PRESENTATIONS
WORKSHOP 1

The grammaticalization of manner expressions into complementizers or quotatives

Caroline Gentens & Kasper Boye
(University of Stockholm & University of Copenhagen)

Keywords: manner expression; grammaticalization; complementizer; quotative; mistaken belief

Saxena (1995) drew attention to the cross-linguistic polyfunctionality and grammaticalization of manner deictics such as thus or so into complementizers or quotative markers. She suggested a pathway of development along which manner expressions first develop a function as quotative markers, which then gives them the potential to acquire the same range of functions as verbs of speaking that grammaticalized into complementizers. Saxena further proposes that this development proceeds unidirectionally along the following implicational hierarchy (including other functional extensions of manner expressions):

(1) \( \text{(manner expression} < \text{)} \) \text{direct quote marker, complementizer} < \text{purpose, reason marker} < \text{conditional marker} < \text{comparative marker} \)

In later work, the specific aptitude for manner expressions to develop into quotatives as in (2) or complementizers as in (3) has been confirmed to be widespread (Heine & Kuteva 2002: 274, 258; Güldemann 2008: 317-328; König 2015; Boye & Kehayov 2016; Treis & Vanhove 2017). These studies show that the quotative stage posited in (1) is not a necessary part of the development, and that similar developments into the domain of complex sentences apply to manner demonstratives (e.g. so), manner question words (e.g. how), manner nouns (e.g. way), manner affixes, and to closely related simulative verbs (e.g. resemble) or simulative prepositions (e.g. like), which we therefore all include under our umbrella term of ‘manner expressions’.

(2) For so seyde Salomon, þe wise: “Pe mon þat her wel deþ, he cumeþ þar he lyen foþ on his lynes ende”
   “For so said Salomon the wise: “The man who does well here, he obtains recompense at his life’s end.”” (cataphoric quotative marker, OED, Middle English)

(3) He told me how (≈ that, ≠ the manner in which) John might never return to his home country.

The grammaticalization of manner expression remains understudied, however. This has been attributed to their semantic and syntactic versatility (e.g. König 2015: 39-40), and to the problems associated with making more fine-grained distinctions between markers of manner ‘proper’, similarity, comparison, instrument, and means (Kortmann 1997: 81, 84, 146; Hengeveld & Mackenzie 2008: 263-265). In this workshop, we want to draw on the recent surge in relevant descriptive studies to refine existing distinctive criteria, to document the ways in which manner expressions acquire grammatical functions in complex sentences, and to examine their relevance to the cross-linguistic study of the semantics of complementation.

This last question relates to the functional specialization of the resulting expression that introduces a complement clause in the broad sense. Some grammaticalized manner expressions are
specialized for non-presupposed complement types in the domain of direct and indirect speech and thought reporting, for instance into quotative, i.e. direct speech-related, uses (e.g. *so* in (2) and *be like*), or into indirect speech-related uses (e.g. *so* as a propositional anaphor in *He said so*). Others specialize for so-called presupposed, ‘factive’, propositions (e.g. *how* in (3)) rather than illocutions (Gentens 2016, Legate 2010, Nye 2013, Boye & Kehayov 2016). Yet another type occurs specifically with counterfactual (CTRFCT), so-called ‘mistaken belief’ complements as in (4) (e.g. Evans 1995, 2003). We would like the participants of the workshop to explore the question of how the different specializations relate to each other - in terms of areal distribution, possible mutual exclusiveness, and/or traces of semantic or syntactic persistence with respect to the source construction.

(4) niya nguthaliya-th, maraka kalka-th (Kayardild, Tangkic)
3sgNOM pretend-ACT(UAL) CTRFCT be sick-ACT(UAL)

‘He pretended he was sick.’ Evans (1995: 379)

The goal of this workshop is thus to bring together papers documenting the synchronic and diachronic multifunctionality of manner expressions that have come to function as propositional complementizers or as quotatives, both in English and beyond. We would like the contributions to explore one or more of the following questions:

- which empirical criteria, tests, or markers allow you to distinguish between closely related uses of a manner expression (e.g. manner and simulative uses) or between closely related complement types (e.g. between presupposed, interrogative, and exclamative *how*-complements, cf. Nye 2013)?
- how do different uses of the same marker relate to the distinction between states of affairs, propositions, and illocutions (see e.g. Vendler 1967: 122-146, Lyons 1977, Boye 2012)?
- which semantic specialization does the element introducing a complement clause have? Does it specialize for so-called ‘factive’ propositions as in (3), or rather for non-presupposed complements in the directly and indirectly reported domain (cf. *so* in (2)), or more specifically for counterfactual, so-called ‘mistaken belief’ contexts (e.g. the marker roughly paraphrasable as *as if* in (4))?
- does this semantic specialization show in a restriction on the set of main clause predicates the complement type co-occurs with? Is the marker also used in cases of (semi-)insubordination?
- how does this semantic clause type relate to the larger system of complementation in the language?
- which source construction(s)/mechanism(s) of change underlie the development of the complementizer/quotative use? Which other uses does the marker have outside of the domain of complex sentences?

References


Lyons, John (1977), *Semantics*, vol. II. Cambridge: CUP.

**Manner expressions in Basque: Dialectal and cross-linguistic comparison, and some grammaticalization paths**

Borja Ariztimuño-Lopez
(University of the Basque Country)

Keywords: Basque, manner expressions, grammaticalization, internal reconstruction, cross-dialectal comparison

In this paper I claim that the Basque complementizer suffix -ela (2) is the result of the grammaticalization of the –distal– adverbial demonstrative of manner hala ‘thus’ (1), and not a merge of two suffixes, as has been traditionally claimed (Michelena 1972: 105, n. 37).

(1) **hala** esan zidaten
    thus say:PTCP AUX
    ‘They said that to me’ (lit. ‘thus they said to me’)

(2) **Etorriko** zirela esan zidaten
    come.PRSP AUX.PST.COMP say:PTCP AUX.PST
    ‘They said to me that they were coming’

Compare the Basque examples with the following sentences from Khwe (*apud* Kuteva et al. 2019: 398):
(3) **Tá** xám kx’úí.
    thus lion speak
    ‘Thus speaks the lion.’
I also propose that the relational form (-elako) might have been directly grammaticalized from the derived adverbial demonstrative halako ‘such (a), (any) like that’.

Previous scholars have proposed that the complementizer suffix was the outcome of the merging of two different suffixes: the first one would be the subordinator suffix -(e)n or the genitive case suffix -(r)en, and the second one has been thought to be the ancient suffix of manner -la (de Rijk 1995) or the allative case suffix -ra (Terence H. Wilbur, apud Michelena 1972: 105, n. 37). Both views assume that a) the subordinator -en underlies all the subordinated verb forms and b) that it has something to do with either the genitive- ablative -rean or the inessive -n case suffixes.

Nevertheless, and besides other morphological and diachronic issues, those analyses lack a crucial aspect of any reconstruction, and perhaps even more important in the case of a language isolate like Basque: the dialectal comparison. In doing that, we see that there are both diachronically and diatopically coherent reasons to pose that the variant -ala should be regarded as the former one: it is recessive, and appears in both western and eastern varieties, leaving the central dialects as the innovative ones.

It remains open for discussion if the -la suffix itself is related to the abovementioned -ra allative and/or the dialectal variant -ala, which cannot be explained by pure phonological evolution.

References


Keywords: spoken Polish, quotatives, manner expressions, demonstratives, deictics

This paper addresses the understudied quotative uses of two manner/quality demonstratives in colloquial conversational Polish: the adverbial (manner) tak ‘thus, like so’ and the adjectival (quality) tak-i (-al-ie/etc.) ‘like this’. Both are used in quotative ‘demonstrations’, where they introduce quoted direct speech, targeting illocutions rather than complement clause propositions in indirect speech. It this sense, they are true quotative markers in the narrow sense. Examples (1) and (2) illustrate this use
of tak and tak-i, respectively. The study draws on data from the Spokes corpus of contemporary conversational Polish (http://spokes.clarin-pl.eu/).

(1) i my tak „No to może tam by się udało nas?”
   ‘And we (were) like this “Well, maybe you could be (moved) there?””

(2) i on wtedy taki „Nie, no co ty, grzecznie się bawiłśmy”
   ‘And he (was) like this “No, not at all, we were behaving ourselves””

The quoted material may also represent creatively constructed discourse representing thought, attitude, gesture, facial expression, physical activity or sound effect. This last use is illustrated in (3):

(3) cze wracać ale jeszcze się odwraca i tak „hau hau”
   ‘(S)he wants to come back but (s)he manages to look around and (goes) like this “Woof woof”

While tak in combination with a speech verb would be a canonical quotative frame (cf. say thus/so: “Quote”), the attributive tak-i, otherwise a quality deictic, may also be paired with ‘say’ into a quotative frame; this is only possible in colloquial speech. Further, spoken data indicate another common strategy, i.e. the verbless construction (Conj)+NP+tak/tak-i, cf. (1–3). I argue that this construction is an extension of (Conj)+NP in (4) – a broader strategy of null predicates of action (e.g. hitting, speaking, movement) common in Slavic (Weiss 2013). Note that example (4) may be extended to no to my tak/takie...; cf. also the use of the conjunction i in (1–3).

(4) no to my „Hmm, no to chodźmy”
   ‘So we (were like) “Hmm, well, let’s go then”

Tak and tak-i are thus optional elements and are independent of overt predicate verbs (although they may be combined with one); they may also be stacked with other quotatives. The variants with overt verbs are treated as stackings, i.e. ‘say’ and tak/tak-i operate independently and ‘redundantly’ as two signals of quotation.

Tying in with the theme of the workshop, I argue that tak/tak-i have grammaticalized from demonstratives into quotatives, which is common cross-linguistically (Buchstaller and Van Alphen 2012, König 2015). Their source constructions are the broader types of manner/quality deictics in their exophoric (gestural), anaphoric and cataphoric uses, cf.:

(5) taki samochód ‘a car like this’ (+gesture)
   ‘such a car (=previously mentioned)
   taki samochód, który... ‘such a car which...’ (cataphoric)

In quotation, tak/tak-i have developed cataphoric uses whereby speakers demonstrate and enact reported discourse, attitude, physical action, etc.

Tak/tak-i also display the polyfunctionality typical of such grammaticalized items. Semantic extensions and discourse uses to be discussed include hedging, approximation, booster of dramatization and sense of immediacy.

References
The grammaticalization of manner expressions into complementizers: 
Insights from Semitic languages

Rodrigo Hernáiz
(The Open University, UK)

Keywords: grammaticalization; complementation; manner expressions; Semitic; Akkadian

Complementation strategies in ancient and modern Semitic languages include the use of complementizers formed from a common element k- (Modern Hebrew kî, Akkadian kīma, Ugaritic k(y), Ge’ez kama). These conjunctions, which typically cover a wider semantic range beyond complementation, are generally assumed to derive ultimately from a common Semitic comparative particle (Deutscher 2000).

Nevertheless, the steps involved in the eventual readaptation of a manner expression into a marker of sentential complementation are still not sufficiently clear. A reassessment of the documented evidence is necessary to address the following questions: What can the oldest record of Semitic languages inform us about this grammaticalization process? Can an examination of the distributional patterns of manner expressions in early Semitic languages reveal any semantic or syntactic footprint from source constructions?

Despite attempts to describe this process of grammaticalization, the issue remains unsettled. For the oldest substantial Semitic written record, Akkadian, it has been proposed that a causative meaning of kīma is the source of complementizer kīma, via reanalysis of pre-existing causal clauses (1) as factive complement constructions (2):

(1) He said/spoke to the governor because (kīma) the barley was not collected
(2) He said/spoke to the governor that (kīma) the barley was not collected. (Deutscher 2000)

The same idea permeates in Zuckermann’s (2006) analysis of complementation in Hebrew. Streck (2002), on the other hand, argues that the comparative meaning is the direct source of Akkadian complementizer kīma, suggesting reanalysis from bridging context of the type:

(3) I told him how I had travelled to Egypt years ago
(4) I told him that I had travelled to Egypt years ago (Streck 2002)

I present the results of a quantitative and contextualized analysis of the oldest corpus of evidence of complementation in Semitic, to assess the type, frequency and functional distribution of complementation in a synchronic and diachronic perspective. The study of patterns of distribution of manner expressions and complementizers focuses on the Old Akkadian (OAkk), Old Babylonian (OB) and Old Assyrian varieties of Akkadian. The study was based on data completeness for OAkk (record published until 2018), and on data representativeness for the far more extensive OB documentation (ACCOB corpus). The findings are also contrasted against information from texts and grammars of other ancient Semitic languages to provide a clearer picture of the cross-linguistic extension of the phenomenon.

258
The research data suggests that k-complementizers might have grammaticalized at an earlier date than previously claimed, without neither intermediate quotative nor causal stage. The latter claim is in line with cross-linguistic observations where adverbial causatives are not frequently proven to be the direct source of object complementation markers (Boye & Kehayov 2016; Kuteva et al. 2019; Schmidtke-Bode 2014). Moreover, the earliest distribution of complement kīma-clauses in Akkadian do not illustrate contexts where reanalysis like the above can be inferred, but reveals an early occurrence in noun-complement clauses and a significant association with verbs of knowledge in pragmatically marked utterances.

References


Factive manner complementizers in English: A diachronic case study

Mikko Höglund & Caroline Gentens
(University of Stockholm)

Keywords: manner expressions, complementizers, factivity, diachronic development

In Present-day English, complex NPs with the noun way can introduce a factive (i.e. non-reported) propositional complement as in (1). In these contexts, the complex NPs with way do not necessarily invoke a manner reading which fills an adverbial gap in the following clause (cf. How has he involved the rhythms of the branches __ while keeping the forms in their place?); rather, they are functionally similar to clauses introduced by the fact that (Legate 2010).

(1) I liked Anthony Eyton’s ‘Oak Tree’, and I specially admired the way that he has involved the rhythms of the near branches with the top of the oak in the middle-distance while keeping the forms in their place. (OED n. rhythm, 1965)

In other uses, the noun way is the antecedent of a relative clause, in which it functions as a path (2), or manner (3), expression.
Historically, the first uses of *way* + finite clause involve path expressions (2ab) mainly with predicates expressing some kind of motion (e.g. *tread*, *retire*). Complex path NPs with an adverbial function in the main clause as in (2b) set the way for a second use, in which *way* serves as an adverbial manner expression also within the relative clause in (3). Only in a third and later stage can the sense of manner modification to the relative clause be lost (1).

In this talk, we present a corpus-based study of the semantic and distributional changes affecting the *way*-*that*-clauses. The results of the analysis are based on extractions for the noun *way* from corpora for the Early Modern (EEBO) and Late Modern (CLMET3.0, CEAL) English period. The historical development attested for *way* sheds light on grammaticalization on two different levels: more specifically on the grammaticalization patterns of *way* (see e.g. Tabor & Traugott 1998 for *anyway*), and on a more general level on the mechanisms underlying a cross-linguistic tendency for path and manner expressions originating in lexical items to grammaticalize into clause linkers such as conjunctions introducing complement clauses (Boye & Kehayov 2016, Güldemann 2008, Hopper & Traugott 2003 [1993]). Moreover, it provides a possible explanation for the semantic specialization of the complementizer-like use of *way*, i.e. for the fact that it is used to introduce factive, but not reported complements in PDE, cf. ?*He said the way that he balanced the rhythms of the branches in the painting.*

References

Data sources
CEAL: Corpus of Early American Literature.
CLMET 3.0: Corpus of Late Modern English Texts version 3.0.
EEBO: Early English Books Online.
The Basque marker *bait-*: from a manner expression to subordinator

Dorota Krajewska
(University of the Basque Country UPV/EHU)

Keywords: Basque, diachrony, syntax, complement clauses, subordination

In Basque, the prefix *bait-* marks subordination, but it also appears in exclamatives and independent clauses. In subordinate clauses, it co-occurs with clause-initial conjunctions in reason (1), manner or result clauses or with pronouns in relative clauses. On its own, it is used in relative (2), reason, result and complement clauses. Only a limited group of verbs take *bait-* complements: factive predicates, such as *pena/domaia da* ‘it’s pity’, *erokeria da* ‘it’s crazy’ or *harrizeko da* ‘it’s surprising’ (3), and predicates of happening, such as *gertatu* ‘happen’ (4).

(1) *Erra-ten da mortale-a, zeren ill-zen bait-tu arima*
   say-IPFV AUX.3SG mortal-DET because kill-IPFV bait-AUX.3SG>3SG soul
   ‘It is called mortal because it kills the soul.’ (Beriaín, 17th century)

(2) *Axeri-a, bait-ta animale oro-ren fiñ-en-a, …*
   fox-DET bait-be.3SG animal all-GEN fine-SUPER-DET
   ‘The fox, which is the finest of all animals, …’ (Tartas, 17th century)

(3) *Pena da egiazko hobendun-ek ez bait-tute hain laster*
   pity be.3SG true culprit-ERG.PL NEG bait-AUX.3PL>3SG so fast
   aitor-tzen confess-IPFV
   ‘It’s pity that real culprits do not confess so fast.’ (Larzabal, 20th century)

(4) *Gertha zedin bada hil bait-zedin paubre-a*
   happen AUX.PST.3SG so die bait-AUX.PST.3SG poor-DET
   ‘So it happened that the poor died.’ (Leizarraga, 16th century)

A question which has intrigued many Basque scholars (e.g. Lafon 1966) is whether *bait-* is diachronically related to *bai* ‘yes’. I will argue that while it is difficult to derive the functions of *bait-* from the affirmative *bai*, it is easier to link them to the manner expression use of *bai*. This paper, thus, attempts to reconstruct the development of the subordinator *bait-* in general, and the complementiser function in particular. It is based on two types of data: a corpus study of the usage of the marker *bait-* and the particle *bai* in the oldest Basque texts and internal reconstruction.

According to Trask (1997) and Lakarra (2013) *bai* comes from a form of the verb *edin* ‘be, become’ meaning ‘it is (so)’. I will show that traces of the manner expression function of *bai* are attested in the oldest texts, where the particle is used as standard marker in equative and comparative constructions. Given the tendency of manner expressions to renew (König 2015), it seems plausible that *bait-* was initially *bai* reinforced with the conjunction *eta* ‘and’, and that the subordinator developed from there.

As regards complement clauses, I will look into the range of predicates taking *bait-* complements and whether this function directly relates to the manner expression or rather constitute a further extension of the subordinator. Finally, I will reflect on the competition between *bait-* and the more common complementation suffix *-la*, which, incidentally, is also related to a manner expression (*hala* ‘thus’, as proposed by Ariztimuño 2013).

References
Manner expressions in Finno-Ugric: their use in quotative constructions and beyond

Eda-Riin Leego & Denys Teptiuk
(University of Tartu)

Keywords: manner expressions, similitative adpositions, manner demonstratives, quotative indexes, perception and epistemic verbs

In quotative constructions of Finno-Ugric languages, manner expressions (MEs), like similitative adpositions and manner demonstratives, either occur with reportative verbs like ‘say’, ‘think’, cf. (1), or are used as grammaticalized elements with equational verbs (2) and as single quote-introducers (3) (see Teptiuk 2019a, b). Although it is hypothesized that cognitive verbs behave similarly to reportative verbs in complex sentences (Palmer 1986: 135; GüldeMann 2008: 4ff.), cf. (4), the use of MEs with epistemic (know, guess) or perceptive (see, hear) verbs has not been investigated so far.

(1) Udmurt (Blog subcorpus)
Nošt soiž šoraz učke no jua
and DEM.3SG at.ILL3SG stare.PRS3SG and ask.PRS3SG
kad’: “Kytyń bon Mańi?”
like where PTCL PN
‘And (s)he stares at him/her and as if asks: “Where is Mani?”’

(2) Estonian (New media subcorpus)
tegelt pärnus on nii et vabandust aga ma
basically PN.INE be.PRS.3SG so COMP sorry but 1SG
unustasin oma pileti koju…
forget.PST.1SG own ticket.PART home.ILL
‘Basically in Pärnu it’s so (that) sorry, but I forgot my ticket at home…’

(3) Finnish (Internet communications, korp.csc.fi)
Vähän niinku “haluaisin uskonnon opetusta, mutta
a.bit like want.COND.1SG religion.GEN education.PART but
A bit like “I’d like to get a religious education, but where can I get it from?”.

(4) John said/thought/knew/guessed/concluded that Mary would come (Palmer 1986: 135).

Teptiuk (2019a, b) reports that in quotative constructions simulative adpositions indicate approximateness or non-factuality of the quote, cf. (1), (3). As for manner demonstratives, their use is primarily associated with endophoric function. However, besides reproduction of factual and fictional discourse (for the latter, see ex. 2), there are also cases where manner demonstratives can be used as mimetic markers (see Güldemann 2008: 287ff.), introducing the discourse-insertion of representational gestures and movements rendered into words (see Teptiuk, submitted). With both types of markers, the fact-assertiveness of propositions is usually lowered, and quotations appear as subjective representations of someone’s utterances and thoughts.

In this study, we expand previous accounts and investigate the use of MEs that were previously observed in quotative constructions for their combination with epistemic and perceptive verbs, and check if they show any restrictions here. The choice of verbs is not accidental. As for reports of perception, they are “inherently more simulative than reports of speech” (Meyerhoff 2002: 352). Therefore, one can expect the combination of perception verbs with MEs quite naturally in contexts where speakers aim to evoke additional connotations towards something that they hear or see. In case of epistemic verbs, we hypothesize that the combination with MEs contributes to the expression of propositions similarly to the use of MEs in quotative constructions (see above). Thus, instead of expressing assertiveness to proposition denoting epistemic processes, the reporter might present them as approximate, non- or counter-factual.

For this purpose, we look into five languages, representing three branches and two areas of Finno-Ugric (Estonian, Finnish, Erzya, Udmurt, Komi). The data come from different corpora (see e.g. Arkhangelskiy 2019; Blokland et al. 2016) and social network sites.

References

Arkhangelskiy, Timofey (2019), Corpora of social media in minority Uralic languages, Proceedings of the fifth Workshop on Computational Linguistics for Uralic Languages, Tartu, Estonia, January 7–8, 2019, 125-140.


Güldemann, Tom (2008), Quotative indexes in African languages: a synchronic and diachronic survey. (= Empirical Approaches to Language Typology 34), Berlin: Mouton de Gruyter.


Teptiuk, Denys (2019a), Quotative indexes in Finno-Ugric (Komi, Udmurt, Finnish and Estonian), PhD dissertation, University of Tartu.

Teptiuk, Denys (2019b), New quotatives in Finnish and Estonian, FUM 42, 207-249.

Teptiuk, Denys (submitted), Manner deictics as quotative indexes in Finno-Ugric, to appear in Å. Næss, A. Margetts, and Y. Treis (eds), Demonstratives in discourse ( = Topics at the Grammar-Discourse Interface), Berlin: Language Science Press.
From deverbal noun to complementizer: The case of the Turkish verbal noun -(y)Iş
Julian Rentzsch

Clausal complements (CCs) are most commonly constructed by means of verbal nouns in Turkish, i.e. bound complementizers which are suffixed to the stem of the predicate of the complement clause and produce nominal forms that accept possessive suffixes, case suffixes, etc. Frequent verbal noun markers in complementizer function are, among other things, -DIK and -mA, items relatively well described and characterized as factive vs. non-factive (Kornfilt 2007: 315) or assertive vs. non-assertive (Csató 2010: 114, 116–117).

While these two items have a documented history of several centuries in Turkish, there is also an item -(y)Iş, equally capable of forming CCs, but described as "more restricted in its functions" by Göksel & Kerslake (2005: 427–430). This item has developed from a deverbal noun suffix, i.e. a derivational marker which is said either to encode the manner of performing an action (giyinş ‘the way of dressing’) or to form abstract nouns (yüriş ‘walk’; Gencan 1975: 213–214). As a complementizer, -(y)Iş usually expresses either the manner in which the action is performed (ex. 1) or a single instance of an event, a usage which makes the event countable (ex. 2; cf. Göksel & Kerslake 2005: 428).

However, there are also occurrences of -(y)Iş as a complementizer that cannot be explained in either way, and which are difficult to describe. In some instances, -(y)Iş can be replaced by -mA (ex. 3), in others by -DIK (ex. 4), and in some instances by both (ex. 5). The vacillating use of -(y)Iş can be motivated by the fact that it is an emerging complementizer and not yet fully paradigmaticized. Note that different from how-clauses in English, Turkish -(y)Iş-clauses may occur in positions not assigned case (cf. Legate 2010: 124), i.e. in subject slots (ex. 6) and in predicative noun slots (ex. 7). The proposed paper attempts to classify the uses of -(y)Iş in Turkish CCs. It tests commutability with the non-factual item -mA on the one hand and the factual item -DIK on the other, and tries to motivate the uses of -(y)Iş as a complementizer in light of its origin.

Examples
(1) Adamın gözümün içine bakarak kahkaha at-tuş-i hoşuma gitmiyordu.
   ‘I didn’t like the way the man was laughing while he was looking into my eyes.’

(2) Zaten bu hikâyede gerçekliginden emin olduğum tek şey odyu; kalbimin her at-tuş-inda, her soluk al-tuş-imda, gözlerimi her açıp kapa-yuş-imda karnımda anbean büyüyekte olan bu bebek.
   ‘Anyhow, this was the only thing in this story which I was sure about: The baby, which was growing in my womb with every beat of my heart, with every breath I took, with every opening and closing of my eyes.’

(3) Bazen saygıya benzeyen bir ifade oluyor yüzünde, çogu zaman hiçbir şey anlamıyor ol-tuş-un (ol-ma-nun/*ol-duğ-unun) acıklı boşluğu.
   ‘Sometimes a complexion resembling respect appears in his face, but often the tragic emptiness of understanding nothing.’

(4) Paris’e ilk gid-tuş-imde (git-tiğ-imde/*git-me-mde) on gün kalmıştım.
   ‘When I had visited Paris for the first time, I had remained there for ten days.’

   ‘The strange thing was that I couldn’t see any nuns. There were only men.’
It was strange that everybody had an occupation.

The basic problem is that we are misunderstanding some things.

References

Diachronic evolution of the subordinator **kak** in Russian

Natalia Serdobolskaya & Irina Kobozeva
(Institute of Linguistics, Russian Academy of Sciences; The Pushkin State Russian Language Institute & Lomonosov Moscow State University)

Keywords: complementation, adverbial clauses, manner expression, subordination, question word

Manner question words may grammaticalize into complementizers (Heine, Kuteva 2002; Güldemann 2008; Boye, Kehayov 2016). Many languages show polysemy of manner question words and eventive complementizers (Boye, Sørensen 2016). We show that this polysemy is not necessarily a result of grammaticalization of manner expressions, rather a narrowing of functions of a subordinator with a very general use.

Modern Russian **kak** ‘how’ is both a question word (1) and used in subordination (2) in adverbial clauses of comparison (2a), eventive (2b) and manner complements (2c), temporal clauses (2d) (Shvedova 1980; Arutyunova 1988; Kobozeva, Inkova 2018). It may also introduce clauses of reason (2e); though, only with the copula (with an archaic nuance). The question arises, which are source constructions and which developed later. Specifically, is the complementizer use a newly developed construction or has it emerged together with comparative clauses? We analyze several large historical documents of Old Russian, starting from first available manuscripts dating 11th century and providing exact frequencies of various functions of **kakъ**/**kako**.

Complement clauses with **kakъ**/**kako** are attested since the 11th century (Shevel’ova 2009), when they could not only encode events, but also propositions (3) and polar indirect questions. Hence, in complementation we do not observe an enlargement of uses, but a narrowing of functions. The same is true for adverbial clauses: in OR **kakъ**/**kako** could introduce temporal, conditional and reason clauses (Bulakhovskiy 1958), while in MR it is only used in temporal clauses (in reason clauses it is restricted).

Thus, the evolution of the subordinator **kak** involves a narrowing of its functional domain. The **kakъ**/**kako** in OR is a “contextual” subordinator, whose semantics is vague and largely determined by
context. Then it underwent a gradual loss of its functions, including “propositional complements”, “reason” and “condition”. In MR *kak* is used in subordination only as illustrated.

Therefore, coindexation of manner and event in complement clauses can be explained not only by grammaticalization of manner expressions, but also as a result of the specification of meaning of "contextual" subordinators with a large range of uses.

(1) **Kak** projti k metro?
   ‘How to get to the metro?’

(2) a. On svoi knigi pishet, kak rabochij detali shtampuet. (Kobozeva, Inkova 2018)
   ‘He writes his books like a worker makes machine components.’

   b. YA videl, kak poezd proshel mimo, ne ostanavlivayas’. (Arutyunova 1988: 116)
   ‘I saw the train pass without stopping.’

   c. YA znayu, kak ona varit shchi.
   ‘I know how she makes the soup.’

   d. Kak priedesh’, pozvoni.
   ‘Call me as soon as you have arrived.’

   e. YA govoryu eto tebe ne so zla, a po dushe… kak ty est’ mój syn… (Kobozeva, Inkova 2018)
   ‘I am telling you this not because I’m angry, but from my heart… because you are my son…’

(3) про местаткъ память на радослава како казало бъло на местатке 3·г. гривено про женън то татка
   ‘Debt letter about Radoslav and Mestatka saying that he (Radoslav) willed to receive 7 grivnas for a theft by Mestatka from Radoslav’s wife’.

References
Boye, Kasper & Peter Kehayov (eds.) (2016), *Complementizer Semantics in European Languages*, Berlin: Mouton de Gruyter.
Sheveleva, Maria N. (2009), «Soglasovanie vremen» v yazyke drevnerusskih letopisej (k voprosu o formirovanii otnositel’nogo upotrebleniya vremen i kosvennoj rechi v russkom yazyke), in Russkij yazyk v nauchnom osveshchenii, № 2 (18), 144–174.

What do manner expressions add to the semantics of mistaken-belief constructions? An Australian survey
Keywords: mistaken-belief, similitatives, reported speech, Australian languages

Especially since Wilkins (1986) drew attention to the existence of mistaken-belief expressions in Mparntwe Arrernte, authors have attested these all over the Australian continent. Mistaken-belief expressions are mono-sentential expressions signalling ‘x (= not the speaker) thinks/says/believes p, but I (= the speaker) know that not-p’, or, that ‘the speaker mistakenly believed that the semantic content of that constituent was true with respect to the total proposition, but now realizes s/he was wrong’ (Wilkins 1986:588). Consider (1):

(1) Mparntwe Arrernte (Pama-Nyungan, Central Desert)
   Unte mane re-nhe apmwerrke ine-ke-kathene
   2sgA money 3sg-0 yesterday get-pc-KATHENE
   ‘I wrongly assumed you had picked up the money yesterday’ (Wilkins 1986:588)

Explicitly marked mistaken-belief expressions in Australian languages can roughly be divided into three types: (I) the Mparntwe Arrernte-type, as in (1) in which the mistaken belief meaning is signalled by a particle of morphological marker (here the enclitic =kathene). A second type is found in languages with dedicated syntactic constructions for mistaken belief (II) as in the Ungarinyin example (2), in which the specific placement of a matrix clause in combination with a modal element brings out a mistaken belief interpretation:

(2) Ungarinyin (Worrorran)
   kanda=karra a-ma-Ø wardulu
   n.PROX=EPIST 3msg-say/think/do-PRS close
   ‘He mistakenly thinks they are close to here’ (Spronck 2015)

A third type (III) is represented by the Jaru example in (3), in which a semblative adverbial element (or, in other languages, markers that diachronically derive from semitative/similative manner expressions) is used to signal mistaken belief.

(3) Jaru (Pama-Nyungan, Ngumpin-Yapa)
   gulanga gurnygga pinang-an-Ø
   as though dead stay-CONT-PAST
   ‘It looked as if (or I thought) it was dead (but in fact it was alive)’ (Tsunoda 1981:206)

The observation that mistaken-belief expressions can be marked with rather distinct, but cross-linguistically re-occurring markers raises the question: what do these markers semantically have in common?

In this paper I aim to answer this question, specifically in relation to the semitative markers (as in 3) and reported speech/thought constructions (as in 2), by demonstrating on the basis of a sample of 40 Australian languages that semitative-based mistaken-belief expressions as in (3) are common, and do not deviate semantically from mistaken-belief expressions that do not contain semitatives. This leads to the suggestion that the function that motivates the use of markers such as gulanga ‘as though’ in (3) can be related to the conventional semantics of reported speech/thought, specifically the meaning components Spronck & Nikitina (2019) identify: evidentiality, modality and ‘demonstration’ (in the sense of Clark & Gerrig 1990). This proposal may also explain why reported speech/thought constructions can grammaticalise into semitatives.
The grammaticalization of simulative morphemes and manner demonstratives in Ethiopian languages

Yvonne Treis
(LLACAN, CNRS/INALCO)

Keywords: grammaticalization; simulative; manner demonstratives; complementation; Afroasiatic

Based on data from grammars and text collections, my paper is a study of the distinct grammaticalization paths of manner demonstratives and simulative morphemes in a sample of 20 languages of the Ethiopian Linguistic Area belonging to the Cushitic, Omotic and Semitic branches of Afroasiatic. Under “simulative morpheme” I subsume ‘manner’-nouns/nominalisers and grammaticalised ‘like’ suffixes/enclitics; these non-deictic sources of complementisers are to be distinguished from manner demonstratives (‘like this’) whose core function is to refer to a contextually salient manner inherent in the current speech situation (visible or gesturally demonstrated) or described in the preceding discourse (anaphoric).

Treis (2017) has shown that simulative morphemes have developed into complementisers (and adverbial clause markers) in about 40 Ethiopian languages. It could be excluded that the grammaticalization process proceeded via an intermediate quotative stage, because none of the languages examined uses non-deictic similatives as quotatives synchronically. In contrast, manner demonstratives, which were not considered in Treis (2017), are widely used as quotatives in direct speech-related uses, as shown, e.g., for Kambaata (Cushitic) in Nikitina & Treis (forthcoming). Similative morphemes and manner demonstratives can thus be assumed to have mutually exclusive extended functions.

While Treis (2017) only examines whether similatives mark a(ny type of) complement clause in a language, the present paper aims at a more fine-grained study of the semantic specialisation of simulative-based complementisers and the division of labour between different complementation strategies in Ethiopian languages; cf. Treis (2010: 25) on Kambaata. It is shown that simulative-based complementisers of the sample languages are used with a restricted set of main clause predicates, i.e. utterance (‘ask, tell’) – see (1), propositional attitude (‘believe’), knowledge (‘come to know’) and perception verbs (‘hear’, ‘see’) as well as certain manipulative verbs (‘do’, ‘cause’). With utterance verbs, simulative-based complementisers mark indirect speech reports (declarative, interrogative, directive complements), whereas manner demonstratives introduce or end direct speech reports – see (2) – and refer back to preceding representations of speech as propositional anaphors.

KAMBAATA (CUSHITIC, ETHIOPIA)

References

Yvonne Treis (LLACAN, CNRS/INALCO)
Complement clause marked by 

(2) \( \text{...} \) \textit{abb-ishsh} \textit{maal-

\textit{siis-ano-si=} \textit{ga} exceed-

\textit{CAUS}1.3M.PFV.CONV surprise-

\textit{CAUS}2.3M.IPV.REL.3M.OBJ=SIM \textit{jaal} \textit{a} \textit{ha-n-s} \textit{kul-áno} \textit{ikke} friend-PL-DAT-LNK-3M.Poss tell-3M.PFV PAST ‘He used to tell his friends that it amazed him (...).’

Speech report with following manner demonstrative

(2) ‘\textit{Xumm-á hós-s?}’ \textit{hittig-úta y-i=} \textit{ ké}’ peace-M.ACC pass_the_day-2S.PFV.CONV like_this-F.ACC say-3M.PFV.CONV=SEQ \textit{marbaat-á-s af-

\textit{siishsh-o} lamp-M.ACC seize-CAUS2.3M.PFV ‘He said ‘(...) Have you spent the day in peace?’ and then lighted his lamp.’

References

Nikitina, Tatiana & Yvonne Treis (forthcoming), The use of manner demonstratives in discourse: A contrastive study of Wan (Mande) and Kambaata (Cushitic), in Ä. Næss, A. Margetts, and Y. Treis (eds), \textit{Demonstratives in Discourse} (Topics at the Grammar-Discourse Interface), Berlin: Language Science Press.

Treis, Yvonne (2010), Purpose-encoding strategies in Kambaata, \textit{Afrika und Übersee} 91, 1–38.


Manner and mood: Complementizers in southern Italy

Nigel Vincent
(The University of Manchester)

Keywords: Salentino, Sicilian, Latin, cartography, (ir)realis

Southern Italian dialects display a variety of complementizers introducing finite clauses (Ledgeway 2016). Two, namely \textit{ka/ca} < Latin \textit{QUA} ‘because’ and \textit{ke/che} < \textit{QUID} ‘what’, have generally agreed etymologies. A third — \textit{ku/cu} — is traditionally considered to derive from \textit{QUOD} ‘because’, thus parallelling the semantic trajectory of \textit{ka/ca}. An argument against this view is the fact that in dialects like Salentino these two items contrast: \textit{ca} introduces propositional complements — \textit{disse ca} ... ‘he said that ...’, \textit{crisciu ca} ... ‘I think that ...’ — and \textit{cu} irrealis complements: \textit{ogghiu cu} ... ‘I want that ...’ (Calabrese 1993). Moreover, in Salentino the complementizer contrast has extended to relatives, where in (1) the relative identifies an actual person and in (2) a desired person who may not in fact exist (Simone 2002):

(1) \textit{ha bbinutu lu vappône} have.3SG.PRS come.PSTPRT DEF boy ca a ffâttu stu lavòru COMP have.3SG.PRS.IND do.PSTPRT this work
‘The boy came who did this job’

(2) sta tfèrcu nu vappòne
be.PRS look.PRS.1SG INDEF boy

cu mme fàttsa stu lavòru
COMP me do.PRS.SUBJITIVE this work

‘I’m looking for a boy who can do this job for me’

Such data reinforce Groothuis’s (2018) conclusion that the source of cu is in fact the manner expression QUO (MODO) ‘in which way’. This in turn opens up a link to other dialects in which MODO alone yields the irrealis complementizer mu/mi in a group of dialects in southern Calabria and north-eastern Sicily as in (3) and (4):

(3) ìddù annau mi si curca
he go.PST.3SG COMP REFL.3SG lie down

‘he went to lie down’

(4) forse vole lo Signoremì edifico altro monasterio
perhaps want.PRS.3SG DEF Lord COMP build.PRS.1SG other monastery

‘Perhaps the Lord wants me to build another monastery’

(Legenda di Beata Eustachia, Sornicola 1992)

As (4) demonstrates, this item has been attested since the 15th century, while in modern usage it is characterised by three distinctive properties, namely:
— it has shifted to become an infinitival marker as evidenced by its new position following the embedded subject and any verbal negation (Ledgeway 1998);
— hence, unlike ca and che, it does not enter into recomplementation structures;
— it displays insubordination effects as in (5), cited from Ganfi (2018):

(5) a) non mi ci cerchi a machina a idda
NEG COMP there ask.IMPER DEF car to her

‘Don’t ask her for the car!’

b) m’u dici tu!
COMP-itsay.PRS.2SG you

‘if you say so!’

In the light of such data, my paper aims to do three things:
a) embed the modern phenomena in their historical context and review the arguments for a contact effect with Greek (Rohlfs 1972 vs Ledgeway 2006);
b) develop a manner > mood trajectory parallel to the causal > propositional one;
c) explore the way this manner and mood relation can be integrated into an alternative to the cartographic account built on Rizzi’s (1997) concept of a split-CP which has dominated the literature on these constructions to date.

References
Polish jakoby: an exotic simulative-reportive doughnut?

Tracing the pathway and conditions of its rise

Björn Wiemer
(Johannes-Gutenberg University, Mainz)

Keywords: reportivity, comparison, manner, complementizer, Polish

Of simulative origin, Polish jakoby derives from the connective jako ‘how, that’ univerbated with the irrealis enclitic by. From the earliest attested stages (late 14th century) into the 17th century, jakoby was used as comparison marker and as subordinator of manner or purpose clauses. The former use persisted, the latter was ousted. After the 16th century jakoby further developed into a reportive marker either as particle (1) or complementizer (2). As complementizer jakoby changed its preference from verbal to nominal attachment sites (3); cf. Wiemer (2015).

(1) Jesteś jakoby lekarzem.
‘Reportedly, you are a physician.’

(2) Wspomniano mi, jakoby miał taki zamiar.
‘I was told that, allegedly, he had such an intention.’

(3) Uwierzył plotkom, jakoby P.
‘He believed gossip (saying) that P.’
This semantic and functional history of jakoby is at variance with Saxena’s (1995) implicational hierarchy adduced under (1) in the CfP of this workshop, but it largely confirms Gipper’s (2018) general pathway of similitative markers:

(4)  (i) similarity (∋ irreal) > (ii) visual/perceptual evidence > (iii) inference+uncertainty

Nonetheless, certain aspects differ from Gipper’s studies or require clarification. First, jakoby has run through more than the pathway in (4) acquiring a predominantly reportive function; it thus “stripped off” functions in (4)’s middle part, but retained the initial stage. Second, it likewise lost its manner and purpose uses. Third, it is scarcely attested in quotative contexts and after pretence-CTPs. Moreover, fourth, although jakoby was employed as a subordinator from its earliest attestations, it is difficult to pinpoint when it started being used as a complementizer.

Against the diachronic and areal backdrop of jakoby’s functional development (Wiemer 2015, 2018: 313-328) the talk will provide a token-based account derived from the PNC and two diachronic corpora (see References) to cover several periods. The particular issues to be discussed are:

(a) How to distinguish complementizer from particle use, particularly in earlier stages. Basically, the problem amounts to identifying the suppression of independent illocutions in finite (i.e. balanced) clauses headed by jakoby.
(b) Why manner and purpose uses coincided systematically before they dropped out.
(c) Whether, and why, loss of morphological segmentability of jako+by (‘how/that’+IRR) correlates with increasing use as evidential marker.

I will argue that both manner-purpose and evidential use were favored by predominant employment in finite clauses, and that support was given by frequent co-occurrence with modal auxiliaries, which created conditions for modal/evidential concord. However, manner-purpose usage developed along another path than evidential usage, before the former disappeared. Moreover, the “drop-out” of stage (ii) from (4) may prove illusory, since it is scarcely attested for jakoby and a direct link between stages (i) and (iii) (and thence to reportive use) can be established: comparison may have implicit scope over propositions, which opens up an immediate road into epistemic and evidential marking.

An account of these assumptions and the expected results will contribute to filling a gap left for reportive complementizers by Boye et al. (2015: 8-10), Kehayov/Boye (2016: 828f.) and other contributions to Boye/Kehayov (2016).

References


**Corpora**

**PNC:** *Polish National Corpus* (NKJP): [http://nkjp.pl/](http://nkjp.pl/)

**KorBa:** Electronic corpus of Polish texts from the 17th and 18th centuries (up to 1772) [https://korba.edu.pl/query_corpus/](https://korba.edu.pl/query_corpus/)

Corpus of 12 mln tokens (texts from 1750-1917), not yet publicly available
WORKSHOP 2

Towards a diachronic typology of middle voice systems

Guglielmo Inglese & Andrea Sansò
(KU Leuven, University of Insubria)

Keywords: diachronic typology, valency, middle voice, reflexivity, verbal voice

Description and aims:
The middle voice is a grammatical domain whose complexity has long puzzled scholars, as it lies at the cross-roads between the grammatical domain of voice phenomena and the structure of the lexicon. In spite of the existing typological work on the topic, there is no systematic diachronic typology of this domain, so that we only have an impressionistic understanding of how middle voice systems emerge and develop in the world’s languages. This workshop aims at filling in this gap, by combining the study of individual languages with a cross-linguistic view on middle voice systems, both in a synchronic and diachronic perspective (see below for the specific research questions and possible topics).

Cross-linguistic research on the middle voice has shown that from a synchronic standpoint middle voice systems (MVSs) typically feature a split distribution, to the extent that in individual languages middle markers (MMs) can act as valency changing markers with some verbs but also display an obligatory lexically specific distribution with others (see e.g. Kemmer 1993, Manney 2000, Kazenin 2001, Kaufmann 2007, Calude 2017, Zúñiga & Kittilä 2019, Inglese forthc.). These two classes of verbs can be referred to as oppositional and non-oppositional middles.

Typical functions of oppositional middles include valency changing operations like anticausative, passive, reflexive, reciprocal, impersonal, and involuntary agent constructions (cf. Dixon & Aikhenvald 2000, Kulikov 2010, Zúñiga & Kittilä 2019). Non-oppositional middles show a consistent distribution across languages, as they tend to occur with a specific cluster of situation types: grooming verbs, verbs of position and of change in body posture, verbs of (non-)translational motions, verbs of speech, cognition, and emotion, and spontaneous events (cf. Kemmer 1993). Kemmer (1993) was among the first scholars to suggest that this distribution across grammatical functions and lexical classes is not random. Generalizing over the situation types covered by language-specific MMs, Kemmer reached the conclusion that the middle domain can be cross-linguistically characterized as a “coherent but relatively diffused category that comprises a set of loosely linked semantic subdomains” (ibid.: 238). She further argues that the reason why languages group all these situations under the same coding is that they all share the common functional property of low degree of elaboration of events, whereby middle situations feature participants and sub-events that are not fully distinguishable.

In spite of the existing synchronic typology, diachronic studies of the middle remain few, to the effect that nowadays there is no comprehensive diachronic typology of MVSs. Existing studies focus almost exclusively on the diachrony of individual valency changing functions (e.g. passives [Haspelmath 1990, Wiemer 2011], anticausative [Haspelmath 1993], reflexives [König & Siemund 2000, Schladian 2001], reciprocals [Heine & Miyashita 2008], antipassives [Sansò 2017]). What is lacking in most of these studies, however, is a careful consideration of how oppositional, i.e. grammatical, functions of MMs historically relate to non-oppositional, i.e. lexical ones. The mainstream view on the development of MMs maintains that there is a unidirectional path from oppositional to non-oppositional functions, with reflexive markers constituting one of the more
frequent source constructions of MMs (cf. Kemmer 1993, König & Siemund 2000, Heine & Kuteva 2002, Haspelmath 2003, Kaufmann 2007). However, the reverse scenario, i.e. MMs that start from a lexical distribution and only later develop oppositional functions, including the reflexive, has been proposed for a few cases (cf. Dom et al. 2016 on Bantu, Luraghi forthc. and Inglese forthc. on Hittite and Proto-Indo-European).

These and many other findings call for a rethinking of the diachronic typology of the middle voice domain, in order to explore the full range of possible sources and mechanisms of language change that may lead to the emergence of MVSs, and to understand the extent to which the synchronic variation that one observes among MVSs cross-linguistically is determined by the history of these systems in the first place.

The aim of the workshop is to bring together scholars working on the middle voice and related phenomena in a typological and diachronic perspective in (1) languages (language families) with well-documented history (e.g. Semitic, Indo-European) and (2) languages for which less historical evidence is available, but which can still provide us with some valuable data on the basis of comparative evidence (e.g. Bantu, Oceanic). We welcome abstracts dealing both with language-specific middle voice systems as well as those dealing with more general typological questions or relevant implications for the theory of language change.

We welcome contributions that address the following TOPICS (the list is not exhaustive):

- documentation of previously undescribed MVSs, especially in language families other than the ones represented in Kemmer (1993);
- the possible inventory of situation types associated with MMs cross-linguistically, with a focus on non-oppositional middles;
- the development of individual MMs, either in individual languages or in language families, with a focus on the historical relationship between oppositional and non-oppositional functions;
- corpus studies on the historical development and spread of MMs in specific languages;
- the synchronic and diachronic relationship between MMs and competing constructions (e.g. dedicated markers for oppositional functions, such as reflexive or passive markers) in individual languages;
- the role of phylogenetic vs. areal biases in the development of MVSs;
- comparative studies on the reconstruction of the original function of MMs;
- possible sources and processes of language change that may give rise to MMs;
- the role and interplay of grammaticalization and lexicalization processes in the emergence of MVSs;
- synchronic vs. diachronic explanations of the configuration of MVSs in individual languages;

Acknowledgements
The workshop is organized as part of the project Towards a Diachronic Typology of the Middle Voice run by Guglielmo Inglese and funded by the FWO (Research Foundation Flanders).

References

Dom, Sebastian, Leonid Kulikov & Koen Bostoen. 2016. The middle as a voice category in Bantu: setting the stage for further research. Lingua Posnaniensis 58 (2): 129-149.


Is Yuki -il a middle voice suffix?

Uldis Balodis

<not updated>

From oppositional to non-oppositional: Middle-marked verbs with no valency change in Spanish

Carlota de Benito Moreno
(University of Zürich)

Keywords: middle-markers, reflexive verbs, aspectual se, Spanish, analogy

The distinction between middle-markers (MMs) associated to a valency change (oppositional MMs) and lexically obligatory MMs (non-oppositional MMs) does not fully capture the complex situation of some middle systems. In Spanish, for instance, the MM can (but not must) be attached to some intransitive and transitive verbs without affecting their valency. While in some contexts the presence of the MM might imply a semantic difference (as in (1)), this is not always the case (as in (2)). Based on examples like (1), some authors have tried to link the presence of the MM to aspectual constraints related to telicity (Nishida 1994, Zagona 1996, De Miguel & Fernández Lagunilla 2000, Sanz & Laka 2002, Sánchez López 2002, among many others). More recently, variationist studies of specific verbs have suggested that pragmatic notions such as counter-expectation determine the presence or absence of the MM (Aaron 2003, 2004, Aaron & Torres Cacoullos 2005, Torres Cacoullos & Schwenter 2008).

(1) a. Juan fue.
   John go.PST.3SG
   ‘John went.’

   b. Juan se fue
   John MM.3 go.PST.3SG
   ‘John left.’

(2) Juan (se) vino.
   John MM.3 come.PST.3SG
   ‘John came.’

A more holistic variationist approach, however, shows important lexical differences in the frequency with which intransitive and transitive verbs take the MM, suggesting that they are at different stages of the spread of the MM – while in some cases its presence is quite rare (estar(se) ‘to be’, saber(se) ‘to know’), in other cases it is almost obligatory (reír(se) ‘to laugh’, imaginar(se) ‘to imagine’). Accordingly, the study of these verbs can illustrate at least one possible development from oppositional MMs to non-oppositional MMs.

The goal of this study is to provide an explanation of how such a development can take place. It mostly relies on a quantitative study of the most frequently middle-marked verbs with no valency change, based on data from semi-directed interviews with speakers of different dialects of Peninsular
Spanish. Although the data are synchronic, the comparison between varieties, verbs and contexts sheds light on the spread of the MM from valency-changing functions to other contexts. Moreover, data from previous historical studies (Martín Zorraquino 1993, 1998, 2002, Bogard 2006) will be used to check the validity of the synchronic results.

The data show that the MM spreads analogically to verbs semantically similar to those in oppositional middles (many of which belong in the semantic groups identified by Kemmer (1993)). The effects of the analogy are also syntagmatic, since middle-marked verbs with no valency change appear most frequently in contexts that are reminiscent of oppositional MMS with valency change (in terms of animacy and control of the subject as well as the presence of an external cause). The MM can spread gradually to less similar contexts and, in a last step, it might become obligatory (i.e. a non-oppositional MM).

References


De Miguel, Elena & Marina Fernández Lagunilla (2000), El operador aspectual se, Revista española de lingüística 30 (1), 13–43.


Martín Zorraquino, María Antonia (1993), Ir e irse en el Cantar de Mío Cid, Aragón en la Edad Media 10–11, 575–578.


Syntactic parenthesis and prosodic integration: some clues to understand discourse markers production and use

Maria Cristina Lo Baido
(University of Bergamo)

Keywords: Italian, discourse markers, parenthesis, prosody, spoken discourse

We intend to analyse a class of verbal discourse markers in Italian such as *penso* (‘I think’) addressing the following question: how are such DMs produced and used in spoken discourse? By studying their behaviour on the syntactic, distributional and prosodic level (439 occurrences taken from the KIParla Corpus, Mauri et al. 2019), we will make some considerations on their role in language production. Given that the DMs at issue are typical parentheticals - and parenthesis implies an interruption - we will study to what extent such disruption is prosodically mirrored and how the DMs’ possible prosodic integration can be related to their functions as indexicals strategies operating on spoken communication organization (Diewald 2011).

In the sample of 439 DMs, we will specifically study the profile (Praat - Boersma 2001) of 150 DMs in their immediate environment, by spotting the occurrence of silent pauses (interruptions of the sound signal lasting more than 200 milliseconds, Candéa 2000, Crible 2018) on the level of the DMs boundaries. Verbal parentheticals display various behaviours:

1) voi avrete ((0.4s)) e::h non so ((0.9s)) dei foglietti (KIParla, BOD1005)
   ‘you will find eh I don’t know some instructions’

2) può contare allora presumo su un’ottima conoscenza della lingua inglese (Lip, NA12)
   ‘So you can rely I presume on an excellent proficiency of English’

The marker *non so*, in (1) indicates the speaker’s hesitation, which it is reflected by the filled and silent pauses, and the vowel lengthening (Bortfield et al. 2001). Conversely, the similar marker *presumo*, in (2), facilitates the speaker’s production, through the expression of both the content level (host) and the more functional meaning (Degand and Evers Vermeul 2015: 78) of evidentiality given by the prosodically integrated DM, despite the lack of hierarchization (Haselow 2016). In cases like (2), despite the syntactic parenthesis, the DM production does not (seem to) pose problems for the speaker, as demonstrated by its prosodic integration. Furthermore, such integrated items may arise from complex predicates (such as strong assertive or semifactive predicates, presumably used intentionally), rather than from generic items conveying hesitation. Beside their prosodic integration, the occurrence of DMs in medial position (180/439) contributes to give evidence for their positive role in helping the speaker for the production of elusive meanings with varying scope during the on-line discourse production.

Prosodic integration is a clue indicating the way parentheticals are produced, namely, as items eventually integrated to the speech flow conveying different functional meanings, which contextualize the hosts meaning. When they are integrated and they do not occur with markers of disfluency (as in 1), they are more likely to work as resources. The prosodic integration of
parentheticals is not a new issue (see Dehé and Kavalova 2007, Dehé and Wichmann 2010a,b, Simon 2004, inter al.). Therefore, this contribution offers additional evidence for the integration of verbal DMs, by correlating the latter with the status of DMs as a linguistic category (and resource) not exceptional with respect to their hosts, except for the performing of functional meaning(s) on discourse level.

References


Dehé, Nicole and Anne Wichmann (2010a), Sentence-initial I think (that) and I believe (that): Prosodic evidence for use as main clause, comment clause and discourse marker, *Studies in Language* 34:1, 36–74.


Simon, Anne Catherine (2004), *La structuration prosodique du discours en français* (Sciences pour la communication; 74), Bern: Peter Lang.

Tracing the grammaticalization of a middle voice marker:

A diachronic study of *ke* (可) constructions in Chinese

Yiwei Fang, Xianhui Ye, Song Lu & Foongha Yap (Chinese University of HongKong, Shenzhen)

Keywords: middle voice, modality, diachronic study, grammaticalization, *ke* construction

280
This study investigates how the Chinese character *ke* (可) develops into a middle voice marker. Previous studies such as Wang (2011) have shown that modal *ke* (可) could already be used as a middle voice marker with a potential reading in pre-Qin Chinese (i.e., before 221 BCE). As noted in Xue (2015; see also Yutaka, 2005, among others), facilitative readings of *ke* (可) and *keyi* (可以) middle constructions emerged more recently in Modern Chinese. However, in the absence of diachronic studies, how modal *ke* developed into a middle voice marker and how *ke* and *keyi* constructions are related to each other remain unclear. This study traces the development of *ke* from lexical verb to middle voice marker based on data from the Guodian Manuscripts (Warring States period) and texts from the Ancient Chinese sub-corpus of the State Language Commission Corpus.

As shown in Figure 1, *ke* was initially used as a verb meaning ‘approve’ during the Shang period (17th-12th century BCE). During the Zhou period (11th-3rd century BCE) *ke* extended its usage to verbs meaning ‘accord with’, as well as an evaluative adjective meaning ‘good’ or ‘favorable’. Deontic modal uses of *ke* expressing ability emerged during the Zhou period, ostensibly extending from its use as a favorable adjective in [NP *ke* ye] construction (e.g. Wu, 2018). Ability readings of *ke* was initially used in [NP<sub>agent</sub> *ke* (‘can’) V] constructions to express appraisal and extended into expressing general ability. During the Zhou period *ke* was also attested as a deontic modal with permission readings and as an epistemic modal with possibility readings. Approximative adverbial uses, as in (1), emerged during the Warring States period. [NP<sub>patient</sub> *ke* V] constructions were also attested during the Zhou period. [NP<sub>patient</sub> *ke* V] constructions with ability readings and permission readings could further develop into potential middles, as shown in (2) and (3) respectively. By the Spring-&-Autumn period (8th-5th century BCE). [NP<sub>patient</sub> *ke* V] constructions with possibility readings also developed potential middle voice uses, as in (4).

During the Zhou period, *ke* also often combined with utility verb *yi* ‘use’ (以) to form a *keyi* (可以) construction meaning ‘can use (it) to V’. During the Spring-&-Autumn period, *keyi* developed into a disyllabic epistemic modal *keyi* with an ability/possibility reading, as in (5), with a facilitative reading emerging in Modern Chinese, as in (6).

Consistent with Kemmer’s (1993) crosslinguistic observation, our analysis reveals that the emergence of potential middle voice uses of *ke* involves a process of subjectification, whereby the speaker expresses his/her evaluation of the inherent property of the referent in subject position. Our findings further reveal that argument ellipsis in Chinese, in particular agent omission, can give rise to decausativized constructions, which play an important role in the reanalysis of modal uses of *ke* as middle voice markers. This study contributes to our understanding of the grammaticalization of Chinese middle voice markers, with implications for the relationship between modality and middle voice.

Examples

(1) *ke si qian ren* ‘(The army) *ke* (‘is about’ < ‘could be’) four thousand people’
(2) *wo xin fei xi, bu ke juan ye.* ‘My heart is not a mat and it is not rollable.’
(3) *er gui ke yong xiang* ‘Two containers are usable for offering sacrifices.’
(4) *shi … bu ke ru* ‘gentlemen … are not insulting’
(5) *you guo zhi mu, keyi chang jiu.* ‘if you possess the dao, you can last and endure’
(6) *che ke yi kai hen kuai.* ‘The car can drive very fast.’
Figure 1. Semantic extensions of ke constructions in Chinese

References

‘Ke’

Greek middle voice across millennia

Silvia Luraghi & Dionysios Mertyris
(University of Pavia & Ulster University)

Keywords: middle voice, Greek, diachrony, anticausative, passive
Greek is the only Indo-European language that has preserved the morphological middle inherited from Proto-Indo-European. As Greek is among the earliest documented Indo-European languages, it offers a unique opportunity for research on the diachrony of voice. In spite of the relatively easy accessibility of Greek data, quantitative studies comparing various language stages are missing. In our paper, we plan to present a corpus study based on a diachronic corpus ranging from Homeric up to M(odern) G(reek), and show how the various functions of the middle voice and their distribution varied over time, both in terms of token and type frequency. Our hypothesis is that the basic function of the middle voice changed over time along with the development of voice opposition for ancient media and activa tantum. Available descriptive studies highlight both differences and similarities between the semantics of the A(ncient) G(reek) and of the MG middle: while the (anti)causative alternation is mostly indicated by voice alternation in AG (Allan 2003; Sausa 2016), MG largely relies on verbal lability (Nichols et al. 2004; Alexiadou & Anagnostopoulou 2004), even though anticausative remains among the meanings of the MG middle (Manney 2000). Passive is often described in reference handbooks as the main function of the MG middle, with reflexive and reciprocal uses showing a strong decrease (Holton et al. 1997: 211-216). A preliminary investigation unveils a number of tendencies.

(a) Decrease in the token frequency of media tantum: in Homer (8th c. BCE) media tantum account for 15.9% total occurrences of the 200 most frequent verbs, while in a corpus of literary MG from the 1980s they cover 8.5% (Stamatiou 2017, henceforth LitMG). Lemma frequency is also lower, as it accounts for 38.2% of lemmas in a sample of 355 verbs from Homer (data from Romagno 2010) and only 5.4% in a sample of 333 verbs from the Historia Imperatorum Turcorum (HIT, late 16th-17th c. CE). This tendency emerges at an early stage, as the number of oppositional middles is on the rise from Homeric Greek onward (Schwyzer 1959).

(b) Even though the MG middle largely preserves the semantics of the AG middle, token frequency shows a different distribution of the various meanings. Among the 50 most frequent verbs in Homer, out of 15,818 occurrences middle tokens account for 4050, of which 2209 are oppositional middles. Among these occurrences, 861 are from verbs that do not show any semantic differences between voices. We restricted the observations to verbs with (i) passive (ii) reflexive/reciprocal/autobeneficiary (iii) anticausative middles, and compared them with the verb samples form HIT and from LitMG. Comparison yields the results in Table 1 that confirm Holton et al.’s (1997) description.

<table>
<thead>
<tr>
<th></th>
<th>PASSIVE</th>
<th>REFL/REC/AUTOB</th>
<th>ANTICAUSATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homer</td>
<td>19.6%</td>
<td>38.3%</td>
<td>42.1%</td>
</tr>
<tr>
<td>HIT</td>
<td>22.5%</td>
<td>33</td>
<td>44.5%</td>
</tr>
<tr>
<td>LitMG</td>
<td>63.9%</td>
<td>12.1%</td>
<td>24%</td>
</tr>
</tbody>
</table>

Table 1. Token frequency of middle voice functions over time

In our paper we will discuss the trends based on an in-depth description of these and other data extracted from text.

References


languages. *Linguistic Typology* 8.2: 149-211.
Romagno, Domenica. 2010. Anticausativi, passivi, riflessivi: considerazioni sul medio oppositivo. In:
Putzu, I., Paulis, G., Nieddu, G.F. and Cuzzolin, P. (eds.), *La morfologia del greco tra tipologia
e diacronia*. Milano: FrancoAngeli, 430-441.
Sausa, Eleonora. 2016. Basic valency orientation in Homeric Greek. *Folia Linguistica Historica* 37:
205-23.

**Trans-Himalayan “middle voice” and the case of Macro-Tani languages: Functions, origins and
categorical status**

Yankee Modi & Mark Post

<not updated>

**Semantic developments of the middle in Cushitic**

Maarten Mous

(Leiden University)

Key words: Middle, Cushitic, Semantics, autobenefactive, imperfective

Cushitic languages have a middle suffix in *t*. For Cushitic languages, middles are expressed on the
verb not inflectionally but derivationally. In (1) an example from Iraqw is given. In (1a) the agent is
the first person singular, expressed on the verb, and the patient, the ankle, is object; in the sentence
with the derived middle, (1b), the agent is now conceived to be the body (part) itself, expressed as
subject. Lexicalizations of middle derivations occur, e.g. *bu‘uut* ‘be sufficient’ which originated in a
middle derivation from *buu* ‘pay’ but no longer has any semantic link to ‘pay’ (Mous 1993).

(1a) ya’e-r-‘ée’ a-ga tunquláa‘
    leg-F-my O.F-PF sprain:1SG
    ‘I sprained my ankle.’

(1b) ya’e-r-‘ée’ aa tunqulu‘-út
    leg-F-my S3:PF sprain-MIDDLE:3F
    ‘My ankle sprained.’

The most noticeable aspect of Cushitic middles is the development of a productive autobenefactive
meaning. My claim is that this is a contact phenomenon. Once speakers realise by some knowledge of
another language that a (cognate) middle suffix can express the sense of doing the action for one’s own benefit, the applicability of this sense is appreciated and this sense becomes productive in the receiving language. In my paper I intend to develop this in detail. More generally I discuss why and how certain senses catch on. For example, another potentially productive sense, the facilitative use, which is closely connected to the spontaneous action middle (Kemmer 1993:148), is not or rarely present in Cushitic languages.

Elsewhere (Mous 2004) I have argued that the core sense of the Cushitic derivational middle is body-relatedness. This is reflected in recurrent lexicalization patterns. Cushitic languages show some additional semantic fields compared to those proposed by Kemmer (1993).

Some lexicalization patterns have grammatical effects: In South Cushitic, some (lexically specified) imperfectives are formed by a middle; the nomen agentis derivation involves either a causative or a middle derivation; and pairs of middle-causative deponent verbs develop as if an inflectional middle could emerge.

I show that the presence of other markers such as reciprocal and reflexive pronouns prevent the middle markers from developing those senses.

References
Hayward, Richard J. (1975), Middle voice verb forms in Eastern Cushitic, Transactions of the Philological Society 1975, 203-244.
Mous, Maarten (1993), A grammar of Iraqw (Cushitic Language Studies, 9), Hamburg: Helmut Buske.
Mous, Maarten and Martha Qorro (2000), Middle voice in Iraqw, in Kahigi (eds.) (2000), Lughaz Tanzania / Languages of Tanzania: Studies dedicated to the memory of Prof. Clement Maganga, Leiden: CNWS, 141-159.

The middle voice in symmetrical voice languages: Toward a diachronic typology

Naonori Nagaya
(The University of Tokyo)

Keywords: middle voice, diachrony, typology, Austronesian, symmetrical voice

Austronesian languages of the Philippines and western Indonesia, such as Tagalog, Balinese, and Indonesian, have been characterized as symmetrical voice languages (Himmelmann 2005; cf. Zúñiga and Kittilä 2019). The defining characteristic of these languages is the presence of at least two voice alternations marked on the verb, neither of which is clearly the basic form (Himmelmann 2005: 112). This system is represented, for example, by the Actor-Voice (AV) construction in (1a) and the Locative Voice (LV) construction in (1b) in Tagalog. The Actor NP is in the nominative (or ‘subject’) relation in the former, while the non-Actor NP is so marked in the latter.
Tagalog (Nagaya 2012:62)

a. *Nag-bihis* si Ricky.
   AV.RL-dress P.NOM Ricky
   ‘Ricky dressed.’

b. *B<in>i>his-an* ni Ricky ang apo niya.
   dress<RL>-LV P.GEN Ricky NOM grandchild 3SG.GEN
   ‘Ricky dressed his grandchild.’

In this paper, I investigate synchronic variation in middle marking among symmetric voice languages of the Philippines and Indonesia, seeking for a possible diachronic typology. There are four claims I argue in this paper. First, in symmetrical voice languages, the AV marker is often recruited for middle marking. As in the Tagalog example above, the AV construction in (1a) represents a grooming middle situation. This interpretation is not available in the LV construction in (1b). Note that, in symmetrical voice languages like Tagalog, non-middle verbs also receive morphological marking (e.g., the LV marker in the case of (1b)).

Second, the middle marking with the AV marker is widespread across symmetrical voice languages, covering both oppositional and non-oppositional middles. Although it is only recently that this voice category has caught attention from scholars working on these languages, there has been an increasing number of studies featuring the AV middle marking (Shibatani and Artawa 2007, Nagaya 2009, Beavers and Udayana under revision).

Third, there is variation in the inventory of middle situation types associated with the AV middle marker in symmetrical voice languages. For example, in Tagalog, AV constructions express a wide range of non-oppositional middles, including grooming actions, changes in body posture, and (non-)translational motions (Kemmer 1993, 1994), as well as oppositional middles, such as anticausatives and reciprocals. This is not always the case with other symmetrical voice languages.

Lastly, the source construction of the AV middle marker is agent nominalization. It is now widely accepted (but still controversial) that Austronesian voice morphology was once nominalization morphology (Starosta, Pawley, and Reid 1982, Kaufman 2009). Under this hypothesis, the AV middle marker is assumed to have been a marker of agent nominalization. This paper proposes that the AV middle marking started as an agent focusing strategy. The two strategies share the function of indicating “a process centering in the subject [=agent], the subject [=agent] being inside the process” (Benveniste 1971: 148).

To conclude, symmetrical voice languages in the Philippines and western Indonesia provide an interesting case of synchronic variation in the encoding of middle situations and offer a possible scenario of language change that may give rise to middle marking.

References

Beavers, John and I Nyoman Udayana (Under revision), Middle voice as generalized argument suppression: The case from Indonesian, The University of Texas at Austin and Udayana University.


Nagaya, Naonori (2009), The middle voice in Tagalog, Journal of the Southeast Asian Linguistics Society 1, 159–188.

Nagaya Naonori (2012), On the syntactic transitivity of Tagalog Actor-Focus constructions, NINJAL Research Papers 4, 49–76.


The oppositional and non-oppositional middle functions of the reflexive prefix in the Tanzanian Bantu languages Hehe, Nilamba and Nyaturu

Lengson Ngwasi
(University of Gothenburg)

Keywords: Reflexive, Reciprocal, Middle, Oppositional middle, non-oppositional middle, Grammaticalization

This paper has two goals. First, it presents field-work data showing that in three North-Eastern Bantu languages Hehe (G62), Nilamba (F31) and Nyaturu (F32) the reflexive prefix is polysemous. It is used to encode oppositional functions such as reflexive, reciprocal and anticausative. That is, unlike many Bantu languages, these languages have not kept the Proto-Bantu distinctions between the reflexive prefix *(j)i, the reciprocal suffix *-an, and the anticausative (neuter) suffix *-ik (cf. Dom, Kulikov, and Bostoen (2016); (Meeussen, 1967); Polak (1983); Schadeberg and Bostoen (2019)). Instead, they behave like most South-Western Bantu languages (Zones H, K, and R (Bostoen, Forthcoming) in losing distinct affixes for reflexive, reciprocal and the anticausative.

The following data from Hehe in (1) and (2) respectively demonstrate that the reflexive prefix i- is a productive reflexive and reciprocal marker; similar data can be found in Nilamba and Nyaturu:

1. a - ka - i - on - ile
   3SG-PST-REFL-see-PFV
   S/he saw her/himself’

2. va - ka - i - on - ile
   3PL-PST-REFL-see-PFV
   ‘They saw each other/themselves’

The reflexes of the Proto-Bantu reciprocal suffix *-an only occurs in few natural reciprocal constructions where in most cases it is fossilized. For example, the natural reciprocal event ‘to meet’
has a fossilized reciprocal suffix -an in all three languages. i.e. ku-i-taang’ana (Hehe), ku-i-taangana (Nilamba), and u-khaangana (Nyaturu).

Moreover, the reflexive prefix i- can also appear with some verbs which fit well into the situation types of middle voice (non-oppositional middle) by Kemmer (1993). For example, spontaneous events, emotional verbs (media tantum) etc., as shown by the data from Hehe in (3) and (4) respectively; similar data can be found in Nilamba and Nyaturu.

3. i- ki-tabu  ki-ku i- guts-ag-a wunofu
   aug-CL7-book 7.SM-PRS-REFL-sell-HAB-FV well
   ‘The book sells well’

4. a - ka -i[lap] - ite
   3SG-PST-REFL[swear]-PFV
   ‘S/he swore’

Second, I will argue that the development of non-reflexive functions for the reflexive prefix can be explained as the result of two grammaticalization processes: (i) reflexive > reciprocal (cf. Heine and Narrog (2009), (ii) reflexive > middle (non-oppositional middle voice) (cf. Heine and Kuteva (2002). This proposal is in contrast to Heine (2000) grammaticalization chain: body, head > emphatic reflexive > reflexive > reciprocal > Middle > Passive, which implies that reflexive markers must pass through the reciprocal stage before encoding other functions. Heine (2000) grammaticalization chain also cannot account for other Bantu languages such as Swahili (G42) and Zulu (S42) where reflexive prefixes encode middle functions, without encoding reciprocal function. Thus, positing two grammaticalization processes, instead of one grammaticalization chain as Heine (2000) hypothesizes, makes it possible to account for the grammaticalization of reflexive prefixes both in Bantu languages where they are used to encode reflexive, reciprocal and middle, and where they encode middle functions, without encoding reciprocal function.

References

Dom, Sebastian, Kulikov, Leonid, & Bostoen, Koen. (2016). The middle as a voice category in Bantu: Setting the stage for further research. Lingua Posnaniensis, 58(2), 129-149.


The origin of the middle voice and the rise of labile syntax in Bribri

Sara Pacchiarotti & Leonid Kulikov
(Ghent University)

Keywords: middle voice, labile verbs, intransitivisation, morphosyntactic mismatches, diachrony, Chibchan

In this paper, we show that the Bribri (Chibchan, Costa Rica) middle voice suffix -r derives passive voice from active transitive and agentive intransitive verbs (cf. (1-2)), as well as anticausative verbs from nominal and adjectival roots.

(1a) ie’ tô balo’ ya’ tâi
    3SG:PRX:H ERG chicha drink:PFV.REM much
    ‘He drank a lot of chicha.’

(1b) balo’ ya-r tâi
    chicha drink-MID:IPFV much
    ‘A lot of chicha is drunk.’

(2a) ie’ tsê’ tâi
    3SG:PRX:H sing:PFV.REM much
    ‘He sang a lot.’

(2b) tsê-n-ê têi
    sing-MID-PFV much
    ‘There was a lot of singing/It was sung a lot.’

Applying the methodology of internal reconstruction and drawing upon diachronic typological generalizations in the domain of voices and valency-changing categories, we investigate the origin of the middle, not uncommon in other Amerindian languages (see e.g. Beck 2000, Estrada Fernández 2004, Thompson 2004), but traceable to sources different from the one we propose here. We argue that in Bribri, the anticausative, originally derived by adding the intransitivizing -r/-n suffix to the causative stems in -k (as, for instance, in akê ‘ripe’ – akê-k ‘to ripen smth.’ – akê-n-ûk ‘to become ripe’, or bâ ‘hot’ – bû-k ‘to make hot’ – bû-n-ûk ‘to become hot’; cf. (3)), became the main diachronic source of the middle voice.

(3) tsuru’ bû-n-ê
    chocolate hot-MID-PFV
    ‘The chocolate (became hot and) is hot (now).’

Specifically, the anticausative suffix was extended to all transitive verbs to form passives. This functional extension of the anticausative morpheme could have been triggered by an important development reconstructable for causative verbs: the causative suffix -ûk (ultimately going back to the verbal root ‘make’), homophonous with the infinitive ending -ûk of other verbs, was reanalyzed as an infinitive ending and eventually disappeared in the conjugation forms of the anticausative verbs, thus leaving no trace of causative morphology whatsoever.

We will further demonstrate that a small group of middle verbs lacking an active counterpart without the suffix -r (media tantum verbs) have become labile, occurring in both intransitive and transitive constructions with no formal (supra)segmental change (cf. (4)). These verbs create a small “island of lability” within the syntactic system of a language which overtly marks transitivity oppositions.
The emergence of the transitive uses of the media tantum verbs (originally only used in intransitive constructions) must be due to the analogy with the system of syntactic oppositions of transitive verbs. This diachronic scenario of the rise of lability, hitherto unnoticed in the typological literature, must be of particular relevance for diachronic typology of media tantum verbs as well as for diachronic typology of lability.

Acknowledgments
This research was supported by a research grant awarded to Leonid Kulikov from the European Union’s Marie Skłodowska-Curie program (grant nr. 665778) and from the NCN (POLONEZ grant nr. 2015/19/P/SH2/02028) during a research stay at Adam Mickiewicz University in Poznań.

References

Diachrony of Spanish Analytic Middle Constructions

Cristina Sánchez López & Margot Vivanco
(Universidad Complutense de Madrid & Universidad de Castilla La Mancha)

Keywords: change verbs, reflexive, causative, IL-predicate, SL-predicate

This study deals with the role and interplay of grammaticalization and lexicalization processes in the emergence of midde voice systems, focusing on the diachronic development of a sub-system of analytic change of state constructions in Spanish. These are built by a pseudo-copula, i.e, a light verb expressing change, and a non-verbal predicate, expressing result. They convey spontaneous (1b, 2b) or externally caused (1a, 2a) change events. Analytic middles, like lexical middles, give rise to oppositional middles, so they alternate between a causative-transitive use (1a, 2a) and a middle-unaccusative one (1b, 2b). The latter is always obligatorily marked by the clitic se (anticausative alternation, 1b), except from the suppletive alternation found in the pair dejar ‘leave’/quedar(se) ‘remain’ (2), where se is optional:

(1) a. El frío puso enferma a la niña.
   The cold put sick ACC the girl
   ‘The cold made the girl sick’
b. Ella se puso enferma
   She REFL put sick
   ‘She got sick’

(2) a. El accidente dejó ciega a la niña.
   The accident left blind ACC the girl
   ‘The accident made the girl blind’

b. La niña (se) quedó ciega.
   The girl REFL remained blind
   ‘The girl went blind’

Among other light verbs of change like muto and uerto, Latin had a suppletive pair, unaccusative fio (3a), and causative facio (3b), which, in the passive form, produced an anticausative variant (3c) (cf. Bassols 1948; Martín Puente 2010, 2013; Martín Puente & Santos Guzmán 2002; Pinkster 2015):

(3) a. pectus quoque robora fiunt (Ov., Met. XI, 82)
   breast.NOM.PL too oak-wood.NOM.PL become.3PL
   ‘The breast becomes oak wood too’

b. fac proserpentem bestiam me (Pla., As., 695)
   make.IMP slithering.ACC best.ACC me.ACC
   ‘Make me the slithering beast’

c. uirga aurea facta est (Ov., Met. XI,109).
   Twig.NOM.F.SG gold.NOM done.F.SG is
   ‘The twig became gold’

At least three processes are involved in the evolution from Latin to Old Spanish:

a) Reduction from the suppletive alternation of fio/facio to the anticausative alternation of fazer (Spanish hacer ‘make’) (4):

(4) Por quebrantar sus carnes fácense hermitas. (SD, 1236)
   For break their flesh make3PL.REFL hermits
   ‘In order to tear their flesh, they become hermits’

b) Grammaticalization of full verbs into light verbs of change. The items undergoing this process belong to well-established semantic classes: verbs of movement, like tornar ‘(re-)turn’ (cf. Suárez Fernández 1999); venir ‘come’, caer ‘fall’, volver ‘(re-)turn’, convertir ‘become’; verbs of lack of movement, such as fincar, remanir and quedar (all meaning ‘remain’); and verbs of caused movement, like poner ‘put’ and dejar ‘leave’ (5). These verbs allow optional descriptive secondary predicates (5a); the change into light verbs occurs when such predicates become obligatory and express the resulting property or state, either in the middle structure (5b) or in the causative one (5c):

(5) a. tornóse espantado, que non bolvió cabeza a ninguna cosa.
   turned.3SG.REFL horrified, that not turned.3SG head to no thing
   ‘He turned away in horror, and he didn’t turn his head for any thing’ (CD 1251)

b. Maguer que era blanco, negro se fue tornando. (LAL 1240-1250)
   Although that was white black REFL was turning
   ‘Although it was white, it was turning black’

c. auédeslo contrastado pensat cómo lo tornedes alegre (LAP 1240)
   have.2SG.him saddened think.2SG how him turn.2SG glad
   ‘You have saddened him; think of how making him happy’

c) The most striking characteristic of Spanish light verbs of change is that they select the result phrase they combine with depending on whether this expresses a Stage-Level or an Individual-Level
predicate. This restriction started with the causative variants of *hacer* ‘make’, which chose I-L predicates (6a), and *dejar* ‘leave’, which preferred the S-L ones (6b), but it did not start affecting the unaccusative variants until the 16th century. The whole sub-system of analytic middles gradually organized around the I-L/S-L dichotomy.

(6) a. Dixo la mugier: “Quien te *fyzo* rey?” (*FUM*, 1200)
   Said the woman who you made king
   ‘The woman said: “Who made you king?”’

   b. comió con [ellos], et *dexo*los confortados en la su Fe (*FR* 1254)
   ate.3SG with-them and left.3SG.them comforted in the their Faith
   ‘He dined with his disciples and made them feel comforted in their Faith’

To conclude, analytic middles use the mechanisms already available in the language –the anticausative and the suppletive alternation– but, crucially, they exploit, in a novel way, the semantic nature of the *state* in a change of *state* event. Taking into account that in the Middle Ages *ser* (‘be-I-L’) is the only copula, the newly created light verbs of change distribute around the S-L/I-L distinction, before the pair *ser* (‘be-I-L’) / *estar* (‘be-S-L’) even exists. This issue has not been addressed in Spanish diachronic linguistics before; therefore, this study does not only contribute to our knowledge of the evolution of middles, but also to our understanding of the evolution of reflexives, anticausatives and the IL/SL dichotomy.

References


Corpus

From agent-oriented verbalizer to middle marker: The diachrony of the middle voice in Malayo-Sumbawan

Andrea Sansò & Dawid Gajewski

<not updated>
The reflexive and middle voice in Hittite: A diachronic analysis

Emily Smith
(University of Chicago)

Keywords: Hittite, Indo-European, historical linguistics, reflexive, middle voice

Hittite is an Indo-European language that exhibits active- and middle-voice (MV) verbal conjugations and an enclitic reflexive marker in ḫa (1). The reflexive marker in ḫa has been observed to perform a number of semantic and grammatical functions in addition to indicating true reflexivity (Boley 1993; Carruba 1969; Cotticelli Kurra and Rizza 2011, 2013; Hoffner 1969, 1973, 1996). One of the more recent studies on the functions of reflexivity and middle voice in Hittite (Cotticelli Kurra and Rizza 2013) suggested that within the Indo-European language family, possibly beginning in late PIE, the reflexive particle was used in conjunction with some middle-conjugation verbs. Later (e.g., in Hittite) the use of ḫa spread to active conjugations, and eventually (e.g., in Romance languages) replaced the middle-voice verbal conjugations. However, a diachronic comparison within Hittite, conducted as part of my dissertation research, shows an opposing development: in the oldest Hittite texts, the reflexive only occurs with active-voice verbs, and has a completely different function than the middle voice. Additionally, in OH ḫa occurs in both valency-increasing (2) and valency-decreasing (3) constructions, but does not yet seem to indicate lexical distinctions (e.g., New Hittite kiš- ‘to occur, happen’ vs. ḫa kiš- ‘to become’). It is only in post-Old Hittite texts that ḫa begins to occur with certain middle-voice verbs (4), often with no apparent difference in meaning from their active counterparts (cf. 1). In other words, it seems that in the earliest records of the Hittite language, ḫa and the middle voice occurred in complementary distribution (Inglese forthc. reached similar conclusions independently). While this runs counter to what Cotticelli Kurra and Rizza predict, it also raises another question: does this lack of overlap mean that verbs that originally occurred in MV were somehow incompatible with reflexive meanings? In this project, I address this question through a corpus-based, diachronically organized reassessment of the use of ḫa in Old, Middle, and New Hittite, in order to shed further light on the domains of both the reflexive and MV.

Examples

1) EGIS-ŠU
ma-ḫa-kan
weten
ārri

‘Afterward she washes herself with water.’ (KUB 7.53 ii 25)

2) nu-şwa-LUGAL-ūnna
(labaran
halzir)

‘And they also called me king Labarna.’ (KUB 29.3 i 3–5)

3) nu-šw
et
ekue[r]

‘They ate and drank (with ḫa); they drank off a whole p.-vessel (without ḫa), and they were/became sated.’ (KUB 17.5 i 10–12)

4) nu-ṣwā

kēdantaw[eten]

arrattaru
‘With this water he must wash himself.’ (KUB 59.66 iii 7–8)

References

The Origin and Divergence of Middle Voice in Kuki-Chin Languages

Kenneth Van Bik
(California State University, Fullerton, CA)

Keywords: Middle Voice; Reflexive; Proto-Kuki-Chin; Kuki-Chin Languages; Proto-Typical Reflexives.

This paper compares and contrasts middle voice in six Kuki-Chin (KC) languages: Three from southern Chin, Daai (So-Hartmann 2009), Hyow (Zakaria 2017), and Khumi (Peterson 2013); two from central Chin, Haku Lai (Smith 1999) and Mizo (Chhangte 1986); and one from “Old Kuki” Chin, Sorbung (Mortensen & Keogh 2011), and claims that, on the one hand, it is possible to reconstruct a proto middle prefix *ŋ*- at the Proto-Kuki-Chin (PKC) level, as shown in (1).

(1) Reflexes of PKC middle prefix *ŋ*-

<table>
<thead>
<tr>
<th>PKC MID Prefix</th>
<th>Daai (So-Hartmann 2009)</th>
<th>Khumi (Peterson 2017)</th>
<th>Sorbung (Mortensen &amp; Keogh 2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ŋ</em>-</td>
<td>ng-</td>
<td>(a)ŋ-</td>
<td>ng-</td>
</tr>
</tbody>
</table>
On the other hand, there are some KC languages, e.g., Hyow, Lai, and Mizo, which do not have the reflexes of this PKC middle prefix *ŋ-. That is, Hyow has a reflexive suffix -êy, shown in (2), that is unrelated to the reflexive prefixes ii- and in- found in Lai and Mizo respectively, shown in (3) and (4).

(2) Hyow (Zakaria 2017)

\[ ə ù-hmuʔ-ho \; éy-mik=ní \; hmu-êy \; éy \]
yes 3A-see.II-PM 3SG.POSS-eyed=FOC close.eyed=MID ANAPH.DEM

\[ y/k-khɔɔe-hnɔ-ní \; ponge=shɔk-khɔ \]
hear-PM 3SG.POSS-ear=FOC embrace-CAUS-PM

(3) Hakha Lai (Smith 1997)

\[ kan-hmāay \; kan-ii-phiaʔ \]
1PL.POSS-face 1PL.SBJ-MID-wash
‘We washed our faces’

(4) Mizo (Chhangte 1986)

\[ a^/-in-meet³ \]
3SG.SUB-MID-shave
‘he is shaving’
(The number in the data in Mizo are tone markers)

Even though the middle voice certainly originates from the reflexive construction in Hyow, Lai, and Mizo (cf. Kemmer 1993), this change seems to result from independent development in each language, probably after losing the PKC middle prefix *ŋ-. In addition, the development of middle voice has different pattern in individual KC languages. For example, Hakha Lai has a fully-fledged middle voice (see (3) and (6)) above whereas its dialectal sister Falam Chin shows only a reflexive construction, as shown in (4).

(4) Falam Chin (King 2010)

\[ Mang \; cu \; amah \; le \; amah \; a-at-aw \; pang \]
Mang TOP 3SG.PRON and 3SG.PRON 3SG.SBJ-cut.1-RFL accidentally

‘Mang cut himself accidentally’

Finally, LaPolla (1995) proposes three stages of middle voice development: prototypical reflexives, less prototypical reflexives, and middle situations. This paper further proposes to separate the KC middle voice in two categories: primary middle voice e.g., grooming middle, exemplified with data from Hyow in (5) (also found in all the languages studied so far) and secondary one, e.g., intensive middle (used for emphatic purpose), found only in Lai, as illustrated in (6).

(5) Hyow (Zakaria 2017)

\[ éyóò \; ú-lúkí \; hángháng \; kháʔ- éy-ài=tù \]
then 3SG.POSS-head all shave.II-MID-DEP=R.EVID

‘Then he shaved all his head himself’

(6) Hakha Lai (Tluangneh 2002)

\[ kán-rőol \; nánʔii-ày \; nán-ii-ày \]
1SG.POS-food 2PL-MID-eat.II 2PL-MID-eat.II

\[ iʔ \; kán-lēeŋ \; riʔ \; ɲat \; üu \]
CONJ 1PL.O-visit.I still SUGG HORT

‘you guys ate and ate our food, so you have to visit us for a while’
(SUGG stands for Suggestive)
In conclusion, this paper hypothesizes that there was a PKC middle prefix *ŋ*, reflected by Daai, Khumi, and Sorbung whereas this PKC morpheme has been lost in languages such as Lai, Hyow, and Mizo which develop middle voice independently from reflexive morphemes.

My research in Kuki-Chin languages has been supported by an NSF Grant (BCS-1911385).

References

**Persistence and variation in Turkic deponent verbs**

Matthew Zaslansky
(University of California, San Diego)

Keywords: Turkic, deponency, morphology, middle voice, reflexive

Though deponency—the morphological pattern in which middle-marked (MM) verbs exist without constrative unmarked counterparts (Kemmer 1993: 22)—is widespread throughout the Turkic family, this fact has typically only received limited attention. Like many languages with deponent verbs (famously Latin), many verbs in Turkic exhibit synchronic competition between the deponent forms and unmarked forms. Kemmer (1993: 152, 239) hypothesized that this sort of competition yields either semantic differentiation of the competing forms, or that the unmarked form might fall out of use. Turkic deponents provide ample comparative data to test such hypotheses. On the basis of comparative evidence from a survey of cognate verbs from grammars and with novel data from 22 language varieties, I argue that competing forms in Turkic diathesis paradigms do not tend to be differentiated semantically, but rather persist in (mostly) free variation.
The most common pattern for deponents in the family involves the Turkic MM -(I)n marking the canonically unmarked active stem, creating a morphological alternation with the causative suffix -t. An example of deponency in Azerbaijani (Oghuz Turkic) is given in the diathesis paradigms in Table 1 below, with the relevant middle markers shaded.

<table>
<thead>
<tr>
<th></th>
<th>Default class (tr.)</th>
<th>Deponent class (tr.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>döj-  -mæc ‘hit-INF’</td>
<td>*öjræ-  -mæc ‘learn/teach-INF’</td>
</tr>
<tr>
<td>Active</td>
<td>döj   -mæc ‘to beat / hit’</td>
<td>öjræ -n -mæc ‘to learn’</td>
</tr>
<tr>
<td>Passive</td>
<td>döj   -ül -mæc ‘to be beaten / hit’</td>
<td>öjræ -n -il -mæc ‘to be learned’</td>
</tr>
<tr>
<td>Reflexive</td>
<td>döj   -ün -mæc ‘to throb (beat oneself)’</td>
<td>———</td>
</tr>
<tr>
<td>Reciprocal</td>
<td>döj   -üf -mæc ‘to fight’</td>
<td>öjræ -t -mæc ‘to teach’</td>
</tr>
</tbody>
</table>

Table 1: Infinitival diathesis paradigms for döj- ‘beat’ and *öjræ- ‘teach, learn’

Deponent verbs exhibit accusative syntax when transitive, but there are many intransitive deponents as well. While many deponent n~t alternating verbs like LEARN/TEACH are inherited in modern Turkic languages (Erdal 1991), new verbs are also frequently added to this class, especially denominal change of state verbs.

Contrary to the expectation that the lack of transitive form would facilitate the use of a MM in the active (Kemmer 1993: 108), variation in Turkic active forms (e.g. Meshgīn Turki ögræš/ögræn- ‘learn’ (*ögræ-, cf. ögret- ‘teach’) tends to involve the innovation of active forms with reciprocal markers (-Iş) despite the pre-existence of active forms with reflexive markers -(I)n). Similar variation between deponent active forms stretches as far back as the Orkhon inscriptions (Akhmetov 1978).

This suggests that even where the reflexive is not productive as an oppositional MM, it is nonetheless involved in maintaining paradigmatic contrasts. This is clearest in those cases where the expansion of the MM -(I)n to various active stems has created variation in other diathesis categories, namely the causative, which may either be based i) on an active stem where -(I)n alternates with -t in the causative, or ii) on an active stem in which the MM -(I)n is reanalyzed as an unsegmentable part of the stem, in which case the default causative marker (-DIR) is used.

References
‘The grammar of thinking: Comparing reported thought and reported speech across languages’

Daniela Casartelli, Silvio Cruschina, Pekka Posio & Stef Spronck  
(University of Helsinki)

Keywords: Reported Speech, Reported Thought, Evidentiality, Grammaticalization, Pragmaticalization

Speakers may resort to a variety of linguistic strategies to report communicative acts (‘reported speech’ = RS) or mental states (‘reported thought’ = RT). While RS has received relatively much attention (Buchstaller & Alphen 2012; Güldemann & Roncador 2002; Janssen & Wurff 1996), RT is not often explicitly discussed. Authors either group RT together with RS as a type of ‘inner speech’ (Vygotsky 1987; Vološinov 1973), or treat it as a completely separate phenomenon. This is the case for formal semantic approaches to quotation and propositional attitudes, where RS more naturally fits the former category and RT the latter or and also for functionalist analyses of quotative meaning (Boye 2012).

Both positions have merit. From a syntactic viewpoint, structures involving RS and RT are roughly equivalent (Palmer 1986: 135; Spronck & Nikitina 2019). Moreover, both RS predicates (e.g. say, tell) and RT predicates (e.g. think) behave as bridge verbs allowing for a number of syntactic phenomena, including extraction across wh-questions, embedded V2 in Germanic languages, and complementizer deletion (Cocchi & Poletto 2002; Dor 2005; Erteschik-Shir 1973; Vikner 1995; see also Salvesen & Walkden 2017). The structural resemblance is even greater in languages that do not make a lexical distinction between ‘say’ and ‘think’ at all, examples of which have been found on nearly all continents (Güldemann 2008; Larson 1978; Reesink 1993; Rumsey 1990; Saxena 1988; Spronck 2015).

Nevertheless, predicates of RS and RT may select different complementizers (see, e.g., Ledgeway 2005) or different moods in the embedded clause (see, e.g., Laca 2013), and behave differently with respect to other phenomena such as negation raising (Horn 1989). Functional, corpus-based analyses of RS/RT have revealed differences in preferred syntactic patterns with respect to word order, expression and omission of arguments and complementizers, and different preferences for hypotactic vs. paratactic constructions in colloquial speech (e.g., Posio & Pešková, to appear). Formalist accounts highlight differences in the syntactic structure of RS/RT complement clauses: the complement of RT predicates seems to lack an independent illocutionary force, while RS predicates select for clauses with a full structure licensing root phenomena (Heycock 2006; Hooper & Thompson 1973). Semantically, this difference has been related to the fact that only the complement of RS predicates constitutes an independent speech act. Differences also emerge in the grammaticalization/pragmaticalization processes out of predicates of RS/RT that yield different types of evidential and epistemic structures and markers such semi-grammaticalized constructions, parenthetical expressions, evidential or epistemic adverb(ial)s (cf. English methinks, Spanish dizque, Greek lei), discourse markers and modal particles, and grammatical elements (e.g. complementizers) (Cruschina 2015; Cruschina & Remberger 2008; Thompson & Mulac 1991; Posio 2014; Wiemer 2018, and references therein).
The aim of the proposed workshop is to bring together linguists from different frameworks working on the relationship between RS and RT, and on their different linguistic manifestations. On the one hand, it continues the trend towards the joint analysis of categories of cognition and perspective, which in recent years has gained a strong presence in typology, syntax, semantics, and pragmatics. On the other, it also seeks to assess whether RT constitutes a separate category, RS/quotation or (other) attitudinal expressions. We seek to bring together scholars working on the analogies and differences between the two domains of RS and RT, embracing the investigation of their semantic, pragmatic and morpho-syntactic properties, as well as of the historical linguistic processes that lead to the emergence of RS (evidential) and RT (epistemic) markers. Ultimately, we aim to contribute to the understanding of the notions RS and RT, their crosslinguistic extent and their limitations, and to foster debate across theoretical divides and approaches.

The questions addressed in the workshop include (but are not limited to):

- Which morphosyntactic properties characterize RS and RT in individual languages or in a crosslinguistic sample?
- Should RT be treated as a distinct category? Which tests and criteria can be used to distinguish between the two domains both syntactically and semantically?
- Does complementation under RS/RT predicates involve different semantic and syntactic units and objects? Does complementation under RS/RT predicates involve different semantic and syntactic units and objects? Do the differences that have been observed in complement clauses align with RT/RS or rather with the distinction between assertive and non-assertive predicates?
- Can we identify patterns of polysemy between concepts of ‘saying’ and ‘thinking’?
- What differences and similarities in the expression of RS and RT can be found in corpus-based studies? Are such tendencies specific to individual languages or varieties or rather extend across languages?
- Are RT and RS subject to similar pragmatic constraints?
- What is the relationship between the different linguistic manifestations of RS and RT, including both full predicates and grammatical elements?
- Does RT differ from RS with respect to processes such as grammaticalization, pragmatization or also acquisition?
- Semantically, how does RT relate to quotation on the one hand and propositional attitudes on the other? Can RT be interpreted as ‘inner speech’?

We accept submissions that contribute to the description, discussion, and analysis of these and other issues concerning RS and RT in any language or in a typological/comparative perspective. We welcome contributions from all frameworks and approaches, including synchronic, diachronic, data-driven, corpora, discourse, typological, and theoretical analyses.

References


On inferential hearsay readings in European languages

Bert Cornillie
(KU Leuven)

Keywords: indirect evidentiality - inferential – hearsay – European languages

The most frequently quoted typologies of evidentiality clearly distinguish between inferential readings and hearsay/quotative readings. Within the group of indirect evidence, Willett (1988) differentiates between the two evidential readings in terms of “Reported (second-hand/third-hand/folklore)” and “Inferring (from results/reasoning)”. Chafe (1986) distinguishes between “induction (mode of knowing) from evidence (source of knowledge)” and “hearsay (mode of knowing) from language (source of knowledge)”.

Yet, both seminal typologies do not include inferences from reported evidence. Izquierdo Alegría (2016) convincingly argues that inferences are not a type of evidence, nor a source of information/knowledge, but they are a mode of access in complementary distribution with direct access and hearsay access and that the inferential mode of access can be combined with a sensory, discourse or cognitive basis. Hence, Izquierdo Alegría’s (2016) proposal includes the possibility of inferences from third-hand information. That is, by means of an inferential marker speakers can express linguistically that they have made an inferential conclusion from some piece of discourse pronounced by somebody else.

It is achieved knowledge that polysynthetic languages often have a grammatical paradigm with clearly distinguished evidential values, whereas European languages lack such a set of expressions. The traditional typologies of evidentiality were built on the basis of data from the former languages, but they cannot easily be applied to European languages, which are characterized by indeterminate evidential expressions that, for instance, convey both inference and hearsay readings, depending on the specific contexts of use (see Cornillie 2009, 2018). I claim that that inferential hearsay readings have been completely forgotten in the literature. In the Spanish example with al parecer ‘seemingly’ in (1), the speaker makes a conclusion first and then refers to the coparticipant’s words.

(1) … para que tenga conocimiento de los incumplimientos, al parecer, según sus palabras, reiterados, de la parte demandada, es decir, del portero, del emplea… del empleado o ex-empleado de la finca urbana. (CORLEC. PJUR005C. Lawyer consultancy on the radio regarding usurpation of a key)
‘So that he has knowledge of the breaches, apparently, according to your words, repeated, by the sued part, that is to say, by the doorkeeper, the worker or ex-worker employed at the urban estate’

On the basis of conversational data from Spanish, this paper explores qualitatively how such inferential hearsay readings can be disentangled from hearsay proper, i.e. an evidential value without inferential dimension and from inferentials based on sensory and cognitive evidence. It will be shown that, paradoxically, explicit contextual cues referring to someone else’s discourse indicate that other words serve as a basis for the speaker to make an inference. More generally, it is argued that in languages without a grammatical evidential paradigm specific particles favor the hearsay dimension (Wiemer 2010), whereas adverbials often combine both inferential and hearsay readings, leading to inferential hearsay readings in certain contexts. In addition to the Spanish data, evidence will be presented from the literature on Baltic, Germanic, Romance and Slavic evidentials.

References

Thinking out loud? Je pense, je me dis and j’étais là in French talk-in-interaction

Sophia Fiedler
(University of Neuchâtel)

Keywords: Interactional Linguistics, Conversation Analysis, Direct Reported Speech, I think, multimodality

The most common complement-taking predicates (CTP) for making publicly available thoughts in everyday talk seems to be to think (French: penser). However, conversational data show that there are also other constructions accomplishing a similar task in French talk-in-interaction: the reflexive form je me dis in its present and in its past forms (je me disais and je me suis dit) and the formulaic construction j’étais là. I argue that these constructions are used by interlocutors for the organization of social interaction to accomplish specific actions. Je pense, for instance, mostly projects an opinion, as
in *Je pense c’est trop lourd* whereas *j’étais là* is mostly followed by a strong emotional stance expressed mimo-gesturally. *Je me dis* can be used by speakers to close a sequence of disagreement.

Whereas research about *to think*, especially in English, is very dense (Kärkkäinen 2003, Thompson/Mulac 1991), the French constructions have not been studied in a systematic way concerning oral data (but see Andersen 2007 on *je pense*). Especially when it comes to an interactional perspective, which will be adopted here, research about the selected constructions is still rare. Most research about CTPs making available thought, especially *I think*, focuses on the phenomenon of grammaticalization (Hopper/Traugott 2003). Grammaticalized forms share characteristics like the omission of the complement, morpho-phonological reduction or variability in position. This paper, however, wants to go beyond the conception of these constructions having grammaticalized or being pragmatic markers used for mitigation or hedging (Pekarek Doehler 2016). I argue, in line with Pekarek Doehler (2016), that in addition to pragmatic functions, those constructions play a crucial role concerning the infrastructure of social interaction. They are not only used to accomplish the action of making publicly available thought, which puts them functionally in line with DRT. They also contribute to the organization of social interaction. Concentrating on the analytical unit of turns – “one party talks at a time” (Sacks/Schegloff/Jefferson 1974: 699) and sequences – i.e. the larger unit where one action is accomplished – I propose that those constructions are used for interactional purposes, e.g. to close a sequence of disagreement or to project a stance.

Analysing 18 hours of French video recordings, transcribed according to the Jeffersonian conventions of Conversation Analysis (2004) this study examines the conversational actions speakers accomplish by using the abovementioned constructions. Within the analyses, I will also take into account the interlocutors’ non-verbal conduct, i.e. gaze (Rossano 2013) or gestures, transcribed according to the conventions elaborated by Mondada (2014). The present study thus contributes to already existing research on DRS and DRT in two ways. It adds an interactional perspective (Couper-Kuhlen/Selting 2018) thus completing typological and conversation analytic studies (Holt/Clift 2007) and it takes into account the interlocutors’ multimodal conduct thereby showing that language and the body often come in a ‘package’ (cf. Hayashi 2005).

The present study is being carried out with the generous support of the Swiss National Science Foundation, grant no. 100012_178819, as part of the research project *The Emergent Grammar of Clause-combining in Social Interaction: A Cross-linguistic Perspective (the French Component)*.

References


---

**Reported speech contexts and complementizers across languages**

Alessandra Giorgi

(Ca’ Foscari University of Venice)

**Introduction:** In this work I show that certain properties of reported speech contexts, typically introduced by saying verbs, are connected to the characteristics of the complementizer introducing them. In particular, the properties I discuss are the Double Access Reading (DAR) and Indicative vs Subjunctive selection in Italian, and *shifting* contexts in Armenian and Hindi.

**The data:** the so-called Double Access Reading (DAR) In Italian arises with communication verbs, which normally (the precise meaning of normally will be specified in the work), trigger the indicative. In Italian, and in many other languages, such as English, the following sentence has a typical interpretation:

(1) Gianni ha detto che Maria è incinta

Gianni said that Maria is pregnant

Maria’s pregnancy must overlap both with Gianni’s temporal location at the time of his saying, and with the speaker’s temporal location. Subjunctive complement clauses, or indicative ones with the imperfect *era* (was), do not exhibit the DAR.

The other important property of reported speech contexts concerns the distribution of first person forms in certain languages. Armenian, and Hindi as well, admits for an embedded first person to refer to a non-first person superordinate subject, with all the indexical items coherently interpreted (Armenian, from Giorgi & Haroutyunian, 2018, ex.43), i.e. shifted form the speaker to the subject of the main clause:

(2) Erkušabti Hakob-n inj as-ac’ t’e vale mekn-elu em

Monday Hakob me said that tomorrow leave-FUT.PTCP AUX .1SG

**Literally:** “On Monday Hakob told me that I will leave tomorrow”

**Meaning:** “On Monday Hakob told me that he (Hakob) was going to leave the day after Monday”
This only happens with the complementizer *t'e* and not with the complementizer *wor*. Interestingly, the complementizer *wor* is used both with the indicative and the subjunctive and does not trigger the DAR, whereas *t'e* only selects the indicative and triggers the DAR.

**The hypothesis:** The hypothesis is that the properties above, i.e. the DAR across languages, which in Italian is connected to the indicative-subjunctive choice, and person shifting in Armenian (and Hindi) are due to the characteristics of the complementizer introducing them. Capitalizing on Giorgi’s (2010) work, I will argue that in the Complementizer introducing the complement clause of saying verbs is present an indexical component, which is not there in complementizers introducing complement clauses of other kinds of verbs. This indexical component, realized in the specifier position of the complementizer, can be conceptualized as a *null* demonstrative pointing at the speaker. I will show that only certain complementizers exhibit this property and trigger the characteristics above. The null determiner is responsible for the peculiar anchoring with the DAR and the first person shift in Armenian and Hindi. Note also that is a well-known fact that, even if in Italian the complementizer for saying, thinking and wishing contexts is always *che* (that), in many other languages they differ, as in Salentianian (Calabrese, 1993) and Greek (Roussou, 2010).

**References**


Giorgi, A. 2010, *About the Speaker*, OUP.


**Reporting, perspective-taking and epistemic stance:**

The polyfunctionality of the Upper Napo Kichwa reportative construction

Karolina Grzech

(University of Stockholm)

Keywords: reported speech, epistemic stance, Quechua, Kichwa, discourse

In most Quechuan languages, reported speech and thought is marked by the reportative evidential =shi. However, Upper Napo Kichwa (UNK) does not exhibit a reportative evidential, but marks reports with verbs of speech (1) or thought (2) and the marker =mi.

(1) *Marcos shamu-shka=mi*’ *ni-sha=mi rima-nun.*

NAME come-ANT =M1 say-COR=MI talk-3PL

‘They say “Marcos has come back.”’

(2) *Chi churi-wa. ſũka iva-qa-ni churi-wa=mi.*

D.DEM son-DIM 1SG think-PST-1 son-DIM=MI

‘This little boy…I thought “[he is the farmer’s] little son.”’
The reportative use of =mi in UNK stands out from other Quechuan languages, where =mi marks direct evidentiality (Weber 1986, Floyd 1997, Faller 2002). In UNK, =mi encodes *epistemic authority*: ‘the relative right to know or claim’ (Stivers et al. 2011: 11).

While most Quechuan languages belong to the Andean linguistic area, Upper Napo Kichwa is spoken in the Amazon. It is also typical of the Amazonian area in that it lacks an indirect reported speech construction (Stenzel & Franchetto 2017). Thus, employing =mi with verbs of speech and thought is the only morphosyntactic means of marking reported speech (RS) and reported thought (RT) in UNK.

This talk analyses the RS and RT constructions in the language. The research questions I aim to answer are the following:

1) What are the morphosyntactic properties RS and RT constructions in Upper Napo Kichwa?
2) How is the reportative use of =mi compatible with its epistemic meaning?
3) How do the RS and RT constructions in UNK compare to those used in adjacent languages?

These questions are investigated on the basis of a 13h audio-visual corpus of spoken UNK, comprising a variety of genres and containing data from ca. 40 speakers.

Preliminary suggests that in the RS constructions ni- (‘say’) functions as a complementiser, rather than a verb of speech. This is exemplified in (1), where ni- co-occurs with rima- (‘talk’), a verb retaining its original semantics and functioning as the head of the main clause. RT can be introduced by =mi and iya- (‘think’), as shown in (2). However, iya- does not co-occur with semantically equivalent verbs of thinking. This suggests that, unlike ni- (‘say’), iya- (‘think’) function as a verb, rather than a complementiser. Other verbs of speech are also not attested in the complementiser function. In the first part of the talk, I discuss the differences between these verbs, and compare RS and RT complements with other complement-taking predicates.

Despite the differences in their grammatical status, RS and RT constructions have a common role in UNK discourse: they introduce parallel perspectives. The co-existence of multiple perspectives is a salient feature of discourse across Amazonia (Uzendsoki & Calapucha-Tapuy 2014), as is the existence of evidential and epistemic marking. In the second part of the talk, I compare the Upper Napo RT and RS constructions with those in adjacent languages, so as to confirm the hypothesis that the UNK reportative constructions are more typical of the adjacent Amazonian languages that of the Quechuan language family.

References


©
Reporting on ‘thinking’ in Spanish and Portuguese and the role of the subject pronoun

Anja Hennemann
(University of Potsdam)

Keywords: mitigation, evidentiality, epistemic modality, subject pronoun, cognitive verbs

Spanish and Portuguese belong to the so-called pro-drop languages. In these languages the use of the pronoun is not obligatory (because the pronouns are grammatically/pragmatically inferable). It is therefore particularly interesting to examine the optional use of the pronoun ‘I’ in comparison to its non-use in the context of the speaker’s thinking.

Adopting quantitative and qualitative perspectives, I am concerned with the performativuse of the verbs sp. creer and pt. achar (‘think’), i.e. when these verbs of thinking encode the subject’s cognitive (epistemic/evidential) attitude. The data are retrieved from the CdE (https://www.corpusdelespanol.org/) and CdP (https://www.corpusdoportugues.org/) and reveal that:

1) While the construction [subject pronoun (yo/eu)+verb] is usually followed by the complementizer que (“that”), expressing a relatively strong commitment concerning [p],

(a) sp. Yo creo que eso es bueno para todos. (CdE:eldesmarque.com)
   ‘I think that this is good for everyone’
(b) pt. Eu acho que é um preço muito bom. (CdP:SempreUPdate Notícia)
   ‘I think that it is a very good price’

2) the construction [verb+subject pronoun] is usually found in parenthetic (c) and right-peripheral (d) sentence positions, 3) often accompanied by questions requiring back-channel behavior on behalf of the interlocutor like in (e) and (f), thus serving an intersubjective function:

(c) pt. Deve ganhar mais alguns GPs até ao fim do ano, acho eu, mas chegar ao mundial é muito improvável. (CdP:AutoSpor)
   ‘[He] should win a few more GPs by the end of the year, I think [think I], but winning the Championship is very unlikely’
(d) pt. A Maria está muito consciente disso, acho eu. (CdP:Nova Gent)
   ‘Mary is very aware of that, I think [think I]’
(e) sp. […] no es com…, no es el… el Bonnie and Clyde auténtico, creo yo, ¿no? por los convencionalismos de la sociedad española […] (CdE:Habla Culta:Madrid:M4)
   ‘[…] it is not like[…], it is not the … the authentic Bonnie and Clyde, I think [think I], right? Because of the conventions of the Spanish society […]’
(f) pt. […] fica feio… para isso existe a rubrica Apple acho eu, não? (CdP:Pplware)
   ‘[…] it is ugly… for that there’s the Apple section I think [think I], right?’


(g) sp. […] es él del pantalón rayado, creo. Bueno. Todos elegantísimos, elegantísimos; la fiesta más increíble, eso sí. (CdE:habla Culta: Bogotá:M42)
   ‘[…] it’s the one with the striped pants, [I] think. Well. All of them very elegant, very elegant, the most unbelievable party, this [is] for sure.’
(h) pt. [...] eu devo descartar a possibilidade de ser assexual (acho).
(CdP:a2forum.forumeiros.com)
'[…] I should rule out the possibility of being asexual ([I] think).

5) Hence, I show semantic-functional differences between reporting on the speaker’s thinking with and without the subject pronoun, also respecting the position of the pronoun.

References
Aijón Oliva, Miguel Ángel and María José Serrano (2010), El hablante en su discurso: Expresión y omisión del sujeto de creo, Oralia 13, 7–38.
Schneider, Stefan (2007), Reduced parenthetical clauses as mitigators: A corpus study of spoken French, Italian and Spanish, Amsterdam / Philadelphia: John Benjamins.

Thought Presentation Embedded in Speech Presentation in Thai News Reports
Prapatsorn Tiratanti
(Thammasat University)

Keywords: thought presentation, speech presentation, language of the news, news reporting, Thai

A previous corpus-based study of discourse presentation in English shows that thought presentation is rarely found in the press, occurring only 5.66% of the corpus (Semino & Short, 2004, p.59). However, the data from my earlier research of speech presentation in Thai news reports in newspapers and online news (Tiratanti & Tungvioonayakit, 2019) has revealed that, in the Thai language, thought presentation is often found in the news articles. This study is conducting a follow-up investigation into how reporters use thought presentation in news reporting. The objective of this current study is to describe how thought presentation is used in news articles in the Thai language.

From the data set of 321 news articles, 116 were found with the use of thought presentations – a total of 218 instances. The variation of 7 thought reporting verbs was noticed in the data, e.g. think, believe, expect, know, etc. Samples of the news articles’ reported discourse were matched and analysed with their original speeches from their respective interview clips of the original speakers. The result shows that, in most cases, the original speakers explicitly spelt out their thoughts to the press, from which reporters, then, presented those thoughts in the form of speech presentation. This phenomenon can be described as ‘a thought presentation embedded in a speech presentation’, as illustrated in the underlined sentence in an example below.

Example
นายจิรายุกล่าวเสริมว่ากรณีที่กูรูฝ่ายกฎหมายในรัฐบาลแนะนำให้เวลาค่อนคืน...พร้อมฝ่ายคำนุ่งเป็นหน่วยและเห็นว่าต้องยื่นผู้ต้อง
Mr Chirayu added that the advice from the legal consultant of the government would prolong the process … The opposition was worried and thought that (subject omitted) needed to address this issue…

The underlined part above shows a free indirect speech (FIS, hereafter), reported from Mr Chirayu’s utterance. Interestingly, FIS in Thai has no identifying linguistic characteristics, as tenses in the Thai language are not marked by inflectional affix as those in English and subject can be omitted form the sentence (Jiemwongsa, 2017). It causes ambiguity between FIS and narration even where FIS comes immediately after indirect speech (IS). This ambiguity is found in discourse presentation in English as well (Ikeo, 2007), but the case of Thai is more challenging without those linguistic indicators. In addition to the ambiguity between FIS and narration, when reporters use thought presentation embedded in the form of FIS, it becomes more complicated to distinguish whether that thought is from the original speakers (as a thought presented in the form of speech presentation) or the journalists (as an inferred thought reported by the narrator). Readers can interpret the texts in both ways, and it has different effect on their understanding. The findings display this ambiguity in thought and speech presentation in Thai news reports, including its form and usage, which will allow for a better analysis framework for further study of thought and speech presentation in Thai and in a comparative perspective with other languages.

References

Is thinking like saying? The case of Tabasaran

Sofia Yaroshevich
(Russian State University for the Humanities, Moscow)

Keywords: reported speech, indexicals, Nakh-Daghestanian languages

In Tabasaran (Nakh-Daghestanian, Lezgic branch; approximately 130,000 people), speech or thought reports are parallel in that both require embedding by means of the verb pub ‘say’ (either in the form of the matrix verb or as a complementizer):

\[(1) a. \text{Aslan} \quad \text{egzamen} \quad \text{tuvur} \quad *(k'uri)] \quad \text{uzu} \quad \text{e'ir} \quad \text{kap'un} \quad =za\]
\[\text{Aslan test give.FUTII COMPL 1SG shout make.PER 1SG.S} \]
\[I \text{shouted that Aslan would pass the test.}\]

\[b. \text{Aslan} \quad \text{egzamen} \quad \text{tuvur} \quad *(k'uri)] \quad \text{uzu} \quad \text{fikir} \quad \text{kap'un} \quad =za\]
I thought that Aslan would pass the test.

However, the constructions are not similar in some other aspects. When speech act verbs are in the matrix clause, the embedded clause has personal clitic doubling. The inner predicate must be cliticized according to the original speaker’s interpretation (1SG clitic =za means that the subject of the reported clause is coreferent to the subject of the matrix clause). The absence of the clitic causes ill-formedness:

(2) uzu egzamen tuvur *(=za) k’uri] uzu =ra ƙarun =za
say.PER 1SG.S COMPL 1SG too
Saida =ra ƙarunu
Saida too say.PERF
That I would pass the test, I said, and Saida said the same.

For mental act verbs, meanwhile, the opposite is true. Though the inner verb is finite (e.g. it is tense-aspect-mood marked as a predicate of a simple clause), the person clitic must be omitted. Only clauses with a non-cliticized predicate are well-formed:

(3) uzu egzamen tuvur =(*za) k’uri] uzu =ra kmiǯi q’un =za
1SG test give.FUTII 1SG.S COMPL 1SG too even believe.PERF
1SG.S
Even I believed that I would pass the test.

In (3), the mental act verb does not allow person to be expressed in the embedded clause. Instead, it selects the unmarked 3rd person form. Thus, only speech act verbs are able to mark that the original speaker is coreferential to the subject of the inner clause. By contrast, in mental act reports the attitude holder is not marked.

The fact can support the explanation of Tabasaran reporting construction as follows: embedded speech clauses probably contain null items that control clitic doubling, on the one hand, and are logophoric in nature, on the other hand. These are null logophors bearing the person feature (in the spirit of Deal (2018)), silent logophors (indexiphors, in Deal’s terms). For Tabasaran, this could be presented in the following structure:

(4) aslan-di, [OP log [uć, [pro1sg tukandi-z ruşun =za,]] ƙari
Aslan-ERG LD.REFL store-DAT go.PERF 1SG.S
say.AOR
Aslan, said he, went to the store.

Upon the whole, if we have some special items that are used only in speech reports (in contrast to thought reports), it can be concluded that speech reports should be treated as a unique category of constructions in Tabasaran while thought reports are closer to other cases of embedding.

Acknowledgments: This research has been supported by the RSF grant 17-18-01184.

References


WORKSHOP 4

Derivational zero affixes

Convenors: Gianina Iordâchioaia & Chiara Melloni
(University of Stuttgart & University of Verona)

Keywords: derivation, zero affixes, conversion, lexical semantics, linguistic theory

Since the early days of Sanskrit grammars, zero affixes, i.e., phonologically null morphology with syntactic and semantic content, have proved instrumental in describing language and have consistently been employed in this sense in modern structuralist linguistics and their later adherent theories (see Bergenholz & Mugdan 2000, and Dahl & Fábregas 2018 for overviews).

The empirical challenge. If the absence of a significant challenges the traditional conception of the Saussurean sign, the postulation of zero morphemes facilitates a way to maintain the one-to-one-mapping between form and meaning in puzzling cases in which a change in meaning is not accompanied by a visible form. These phenomena are widespread across languages and they can be found in both inflection and derivation. Zero affixes have been postulated for inflectional paradigms (English singular sheep – plural sheep-Ø; I walk-Ø – she walk-s), category change morphology (a table > to table-Ø; to walk > a walk-Ø; clean > to clean-Ø), labile verb alternations (inchoative – causative verb forms: The vase broke – John broke the vase), and other cross-linguistic phenomena. In some languages, like Italian or German, zero affixation seems more limited in scope for reasons that may relate to the morphological typology of the language, language-specific morphophonological constraints and the status of the base (root vs. word), especially for conversion phenomena (Don 2005, Valera 2015). Experimental support for zero affixes has been provided via psycholinguistic studies in language acquisition, as summarized in Dahl & Fábregas (2018), and, at least for conversion from verb to noun in English, also in computational studies such as Kisselew et al. (2016).

The theoretical debate. While inflectional zero affixes seem to be widely accepted by various theoretical frameworks, there is an unsettled debate concerning the legitimacy of derivational zero forms (e.g., Myers 1984, Pesetsky 1995, Plag 1999, Lieber 1992, Lieber 2004, Borer 2013). In general, positing a contentful zero morpheme confronts us with a theoretical quandary in differentiating it from the lack of a morpheme altogether: how do we know where there is a zero affix and where there is nothing? A reasonable condition for positing a zero affix is the existence of an overt affix with a similar function (i.e., its ‘overt analogue’, Sanders 1988). In inflectional paradigms one easily finds such evidence (cf. I walk-Ø – she walk-s), but for derivation it is not always straightforward. For conversion from verbs to nouns in English, for instance, we find pairs such as to climb > the climb-Ø / the climb-ing, but it is debatable whether this zero affix has a coherent enough meaning to be retrieved from all such deverbal nouns: the noun climb-Ø denotes an event or a path, break-Ø may describe an event, a state or a result entity, and cook-Ø unambiguously denotes an agent. Therefore, Ø emerges here as a nominalizer suffix with very vague meaning. Following a strong interpretation of the ‘overt analogue criterion’ (Sanders 1988), it has been argued that a zero suffix would be semantically too diverse to form a lexical entry by comparison to overt suffixes in English (Plag 1999, Lieber 2004). Another reservation about allowing zero affixes in morphological theory comes from the undesired effect of zero proliferation and the indeterminability of their ordering: we may need to distinguish endlessly many zero suffixes and prefixes without being able to establish
sound criteria for their existence or interaction with each other (Anderson 1992, Bergenholz & Mugdan 2000, and Dahl & Fábregas 2018).

Alternative explanations have been put forward to account for these phenomena, among which ‘relisting’ and underspecification are two prominent hypotheses for conversion. Rather than an actual grammatical phenomenon, relisting is a form of coinage triggered by pragmatic needs, where an entry gets listed again in the lexicon with new category and associated meaning (Lieber 1981, 1992, 2004). In underspecification theories, the lexical category of words in conversion pairs gets specified only in a syntactic context (Farrell 2001). However, both relisting and underspecification fail to capture relevant morphosyntactic and phonological constraints which would be explained under a zero derivational approach (Don 1993, 2005, and Darby 2015). Furthermore, and more importantly, the issues of meaning indeterminacy, zero proliferation and ordering do not only concern derivationally linked words; inflectional zero also appears with many more facets than overt inflectional affixes do. Whether we posit several such zero affixes with each individual meaning or a heavily underspecified one, why would this be more legitimate in inflection than in derivation?

In a nutshell, the literature has highlighted various theoretical and empirical advantages and disadvantages in the use of derivational zero affixes. These heavily depend on the foundational principles of each framework and the methodology it employs, beyond the need of a faithful empirical description. For separationist theories of morphology, which keep morphophonology apart from morphosyntax and interpretation (e.g., Distributed Morphology, Lexeme-Morpheme Base Morphology), zero is just a possible spell-out like any other affix. For lexeme/word-based theories of morphology, which concentrate on words as wholes and their paradigmatic relations, zero affixes are nonexistent (e.g., Construction Morphology, Paradigm Function Morphology).

**Research questions.** In this workshop we aim to gather contributions that address derivational zero affixes from the perspective of the following questions:

1. What are the empirical and theoretical advantages and disadvantages of positing derivational zero affixes? How do theories successfully implement them in their systems? Why do other theories find derivational zero undesirable?
2. How do derivational zero affixes differ from the inflectional ones? Are they empirically and/or theoretically less motivated/more difficult to implement than the latter and why?
3. How do derivational zero affixes resemble or differ from overt affixes (their ‘overt analogues’)?
4. How do different theories of morphology deal with empirical phenomena for which it would be tempting to posit a derivational zero affix?
5. To what extent can language-specific properties and typological generalizations explain availability or unavailability of zero affixation across languages? What does the crosslinguistic study of conversion phenomena bring to our understanding of zero derivation, its empirical adequacy and theoretical status?

**Acknowledgments:** Iordăchioaia’s contribution to this workshop has been supported by the German Research Foundation (DFG), via the research grant IO 91/1-1 to the project Zero-derived nouns and deverbal nominalization: An empirically-oriented perspective, at the University of Stuttgart.

**References**


Kisselew, Max, Laura Rimell, Alexis Palmer, and Sebastian Padó (2016), Predicting the direction of derivation in English conversion, in Micha Elsner and Sandra Kübler (eds.), Proceedings of the 14th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology, Berlin, 93–98.


On Romanian zero-derived nominals

Diana Anițescu
(University of Bucharest)

Keywords: nominalizations, zero affix, Romanian, syntax, polysemy

This study investigates the syntax and interpretation of Romanian zero-derived nominals (1), starting from the theory of nominalization put forth by Grimshaw (1990) and later refined by Alexiadou (2001) and Borer (2011), a.o.
(1) a ataca – atac
‘to attack’ – ‘attack’

Grimshaw (1990) distinguishes between English zero-derived nominals and other overtly derived nominalizations (i.e. –ing, -ation, a.o), and argues that zero-derived nominals can never be complex event nominals, but always non-argument structure nominals (i.e. result or simple event nominalizations). Her analysis relies on nominals which do not accept an ‘of’ internal argument in the Genitive (2), showing that the argument structure has not been activated. Therefore, one of the most important tests which diagnose argument structure in nominals is the existence of an obligatory internal argument in the Genitive case, inherited from the transitive verb.

(2) John’s kick *of/to the dog

 Romanian has only one structural Genitive case position (Cornilescu 2001, Iordăchioiaia & Soare 2008, a.o). Thus, the obligatory internal argument of complex event nominals is projected and assigned the Genitive case (see 3a adevărului ‘of the truth’), while result nouns and simple event nouns do not take an obligatory internal argument and the external argument may occupy the structural Genitive case position (see 3b lui Ion –Ion’s):

(3) a. Afirmarea adevărului de către ministru
    State.INF-the truth-the.GEN by minister
    ‘The telling of truth by the Minister’
    b. Afirmația lui Ion a fost neașteptată.
    State.ȚIE-the OBL Ion has been unexpected
    ‘John’s statement was unexpected.’

Furthermore, previous research (Cornilescu et al. 2014, a.o.) has shown that some suffixes specialize for an event reading, while others are specialized for a result reading. This can be seen in Romanian -re-ție doublets with the same verb base, where the infinitive suffix –re gives rise to a complex event noun (3a), whereas its -ție counterpart is a result noun (3b).

Starting from a corpus of 133 zero derived-nominals, with authentic Romanian examples from the online corpus CoRoLa, we show that Romanian zero-derived nominals are the most versatile class, receiving all three interpretations: result (4a), simple event (4b), and complex event (4c) and we investigate the syntactic properties associated with each interpretation.

(4) a. anunțul Președintelui a fost șocant
    announce.o-the President.GEN-the has been shocking
    ‘The President’s announcement was shocking’
    b. Atacul teroriștilor a avut loc ieri.
    attack.o-the terrorists.Gen-the has taken place yesterday
    ‘The terrorists’ attack took place yesterday.’
    c. Munca *(pământului) de către țărani
    work.o-the soil.GEN-the by peasants
    ‘The field cropping by the peasants’

We propose that they are polysemous (Lieber 2004, Melloni 2006, a.o.) because they do not have a phonologically expressed nominalizing affix to develop a specialization. Therefore, they can shift from one reading to another and may be coerced into a complex event interpretation.
To conclude, we argue that the zero affix is responsible for the versatility of Romanian zero-derived nominals which, contrary to previous claims, can appear with obligatory internal arguments and receive a complex event interpretation.

References

### Zero morphology and change-of-state verbs

Pavel Caha, Karen De Clercq & Guido Vanden Wyngaerd
(Masaryk University, Ghent University & KU Leuven)

Keywords: nanosyntax, adjectives, causative verbs, inchoative verbs

3. **Claim.** We present a theory of conversion in terms of phrasal spellout. In this approach, there are no zero morphemes. Instead, the ‘silent’ meaning components are pronounced cumulatively within overt morphemes. We show how this approach works for the case of adjective-verb conversion.

4. **Data.** In English there are two types of verbs based on adjectives. One type is identical to the adjective (1a), the other has the adjective suffixed with *-en* (1b) or *-ify* (1c).

\[(1) \begin{array}{llll}
\text{a. cool} & \text{b. tighten} & \text{c. solidify} \\
\text{narrow} & \text{widen} & \text{prettify} \\
\text{open} & \text{shorten} & \text{simplify} \\
\end{array}\]

All these verbs have both an inchoative (2a,b,c) and a causative (2d,e,f) sense:

\[(2) \begin{array}{llll}
\text{a. The road narrowed} & \text{d. The workers narrowed the road} \\
\text{b. Her stomach tightened} & \text{e. She tightened the lid} \\
\text{c. The mixture solidified} & \text{f. The company solidified its position} \\
\end{array}\]
5. **Analysis.** We take the overt suffixes to realise two abstract heads, INCH and CAUS (cf. Ramchand 2008).

(3) \[ \text{[ CAUS [ INCH ]] } \Leftrightarrow \text{/en, ify/} \]

(3) instantiates phrasal spellout: a complex node with two abstract heads is realised by a single morpheme. Our account extends this idea to the verbs in (1a). We give them the lexical entry (4), which makes them capable of realizing all the meaning components.

(4) \[ \text{[ CAUS [ INCH [ A ] ] ] } \Leftrightarrow \text{/cool, narrow, … /} \]

The ambiguity between the inchoative, causative and adjective in (4) falls out as a consequence of the Superset Principle (Starke 2009), which says that a lexical entry can spell out any subtree which it contains. Both the inchoative sense and the adjective sense correspond to proper parts of (4), and we therefore derive these readings without the need to postulate zeroes. The same reasoning applies to inchoative-causative ambiguity of the suffixes in (3).

4. **Predictions.** An account based on these premises predicts the possibility of a language where only the inchoative reading would be zero-marked, but not the causative. This pattern would arise with lexical entries for roots as in (5), with only the inchoative reading included in the lexical entry of the adjective.

(5) \[ \text{[INCH[A]] } \Leftrightarrow \text{/A₁,A₂,…/} \]

A set of Czech verbs instantiates such a pattern. The table with the relevant data is given in (6). (Some of the verbs require prefixes, for reasons we do not understand.)

(6)

<table>
<thead>
<tr>
<th>ADJ</th>
<th>gloss</th>
<th>INCH</th>
<th>CAUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>tmav-ý ‘dark’</td>
<td>tmav-­l</td>
<td>tmav-i-l</td>
<td></td>
</tr>
<tr>
<td>tvrd-ý ‘hard’</td>
<td>tvrd-­l</td>
<td>tvrd-i-l</td>
<td></td>
</tr>
<tr>
<td>splep-ý ‘blind’</td>
<td>o-slep-­l</td>
<td>o-slep-i-l</td>
<td></td>
</tr>
<tr>
<td>živ-­ý ‘alive’</td>
<td>o-živ-­l</td>
<td>o-živ-i-l</td>
<td></td>
</tr>
</tbody>
</table>

The adjectives are in the first column; they have an agreement marker -ý, which is irrele­vant for the pattern. The INCH column shows that the inchoative verb is identical with the adjective: the past tense (participial) marker -l attaches directly to the root. The causative version, however, has the verbal marker -i between the root and the past tense marker. A second prediction is the nonexistence of languages that are the opposite of Czech, i.e. that would zero-mark the causative but have a dedicated suffix for the inchoative.

**References**


The interaction of compounding and conversion in Construction Morphology

Bożena Cetnarowska
(University of Silesia, Katowice)

Keywords: conversion, deverbal nouns, (para)synthetic compounds, Construction Morphology

This paper considers empirical data from Polish and English which support the postulation of an affixless processes of conversion (instead of derivational zero morphemes).

Semantic diversity of the output of verb-to-noun conversion in English is problematic for scholars advocating zero-derivation since it requires the recognition of several zero affixes with distinct semantic functions (Štekauer 1996, Bauer and Valera 2005), e.g. those deriving names of events (hunt, ride), agents or instruments (cook, brush), results (cut), objects (catch) or locations (dump). Polish data would necessitate the postulation of even more noun-forming zero affixes since affixless deverbal nouns (with various meanings) belong to one of three grammatical genders (Waszakowa 1993, Szymanek 2010).

(1) a. [+masculine]: bieg ‘run, race’ (event), odrzut ‘reject’ (object), rząd ‘government’ (agent)
b. [+feminine]: budowa ‘construction’ (event, location), oprawa ‘binding’ (event, instrument)
c. [+neuter]: łącze ‘link’ (instrument)

As zero affixes are not postulated in Construction Morphology (Booij 2010, Booij and Audring 2018), the Polish and English data indicate the need for multiple construction schemas in this framework to model the verb-to-noun conversion.

The process of verb-to-noun conversion interacts with compounding, as exemplified in (2) for English and (3) for Polish (LV stands for a linking vowel).

(2) oil spill, mouseclick, dog attack, bear hunt, blood test

(3) a. bajk-o-pis (fable+LV+write) ‘fable writer’
b. ręk-o-pis (hand+LV+write) ‘manuscript’
c. list-o-nosz (letter+LV+carry) ‘mail carrier’

The English compounds in (2) are commonly regarded as headed by deverbal nouns, i.e. [N+[V]N]N (Lieber 2009). In contrast, Booij (2013) treats Polish compounds in (3) as exocentric formations [N+LV+N], since their right-hand constituents function as verb stems and do not occur (in isolation) as nouns. However, a parallelism can be perceived between the internal structure of the compounds in (3) and the synthetic or parasynthetic (cf. Melloni and Bisetto 2010) compounds in (4). The latter contain verb+suffix sequences which either occur (4a) or do not occur (4b) as nouns.

(4) a. bajk-o-pis-arz (fable+LV+write+SUFF) ‘fable writer’ (pisarz ‘writer’)
b. ojc-o-bój-ca (father+LV+kill+SUFF) ‘patricide’ (*bójca ‘killer’)

The notion of schema unification (Booij 2010) can be used to account for the co-occurrence of compounding and suffixation in the formation of the Polish nouns in (4) (see Naccarato 2019 for Russian). The unification of the schema for compounding with the schema for verb-to-noun conversion can be employed in the analysis of the nouns in (3).
The advantage of the approach based on schema unification is the prediction that particular conversion schemas can interact differently with compounding schemas. Polish compound nouns represented as $[N+L+V]_N$ are $[+\text{masculine}]$. Moreover, the productivity of selected conversion schemas (e.g. those coining agentive nouns) increases when they are unified with the compounding schema. In contrast, the verb-to-noun conversion schema instantiated by event nouns is rarely unified with the compounding schema. This is demonstrated for derivatives from Polish verbs belonging to selected semantic classes, including motion verbs, verbs of ingesting, and verbs of creation and transformation.

References
Bauer, Laurie and Salvador Valera (2005), Conversion or zero-derivation: An introduction, in L. Bauer and S. Valera (eds.), Approaches to Conversion/Zero-Derivation, Münster: Waxmann, 7–17
Szymanek, Bogdan (2010), A Panorama of Polish Word-Formation. Lublin: Wydawnictwo KUL.

Zero affixation and morphological processing in Italian

Serena Dal Maso & Sabrina Piccinin
(University of Verona)

Keywords: zero affixation, morphological processing, masked priming

Our research is intended to investigate how Italian zero-affixed denominal and deadjectival verbs are processed during visual word recognition compared to derived verbs which display an overt derivational marking, i.e. suffixed and parasyntetic verbs. The three morphological schemas considered, i.e. zero-derivation (it. calmare ‘to calm down’), suffixation (it. legalizzare ‘to legalize’) and parasyntesis (it. ingrassare ‘to put on weight’) are productive in Italian to create denominal and deadjectival verbs, and, from a functional point of view, they behave similarly: i) they change the category of the base; ii) they select the same base category to create verbs; iii) they show similar semantic characteristics (they have mainly ingressive/resultative, locative, and instrumental...
meanings). However, this shared function is realized by means of different formal exponents which imply different degrees of word internal complexity, perceptual salience and iconicity. Specifically, zero-affixed verbs have the highest degree of formal simplicity but, on the other hand, are scarcely iconic, as there is no overt marking of the transcategorization process (except for the inflectional ending with is required by Italian morphology, Dressler 2005, and Thornton 2004). In order to verify how formal simplicity and low iconicity affects the processing of zero-derived verbs in Italian, with respect to more formally complex but also more iconic morphological schemas (suffixation and parasynthesis), we conducted a lexical decision experiment combined with the masked-priming paradigm (Foster & Davis 1984). The experimental design comprised three priming conditions (identity, morphological, and unrelated) as summarized in the table below:

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Prime</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Suffixation</strong></td>
<td>Identity</td>
<td>legale</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>legalizzare ‘to legalize’</td>
</tr>
<tr>
<td></td>
<td>Unrelated</td>
<td>abbaire</td>
</tr>
<tr>
<td><strong>Parasynthesis</strong></td>
<td>Identity</td>
<td>grasso</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>ingrassare ‘to put on weight’</td>
</tr>
<tr>
<td></td>
<td>Unrelated</td>
<td>aderire</td>
</tr>
<tr>
<td><strong>Zero-derivation</strong></td>
<td>Identity</td>
<td>calmo</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>calmare ‘to calm down’</td>
</tr>
<tr>
<td></td>
<td>Unrelated</td>
<td>sognare</td>
</tr>
</tbody>
</table>

Our hypothesis is that the formal simplicity of zero-derived verbs should favor the perception of the morphological relationship between the prime and the target; however, such facilitation could on the other hand be hindered by the low degree of iconicity of such verbs. Preliminary results seem to confirm the hypothesis that Italian speakers are sensitive to both formal simplicity and iconicity when processing morphologically related words since a significant priming effect has been yielded when targets were preceded by zero-derived verbs, as well as by suffixed verbs, but no significant facilitation was observed for parasynthetic verbs. By contrasting zero-derivation against other morphological schemas and by focusing on the role of factors such as formal simplicity, perceptual salience and iconicity on the processing of zero-derived verbs, our analysis takes an alternative view with respect to previous studies which mainly investigated how morphological complexity of zero-derived verbs (one-step vs two-step zero derived verbs, Pliatsikas et al. 2014 and Wheeldon et al. 2019) or directionality of the derivational relationship (Darby & Lahiri 2016) impact on priming effects patterns.

**References**


Zero Morphology, Interface Asymmetry and Computational Efficiency

Anna Maria Di Sciullo
(University of Quebec in Montreal)

Keywords: denominal verbs (English), VN compounds (Italian), null verbs and prepositions, Merge, SEM PHON asymmetry, third factor principles

1. We investigate the hypothesis that zero morphology is a case of interface asymmetry derived by Merge \((a, b) = \{a, b\}\), and enforced by Principles of Efficient Computation (Chomsky et al. 2019). We focus on the derivation of denominal verbs in English and Italian, provide independent evidence for zero morphology, raise questions related to Merge and Conditions on Pronunciation, and draw consequences for linguistic theory.

2. We detail the derivation in (1) for the English (E) denominal verb *to bottle*, and contrast it with its Italian counterpart *inbottigliare* in (2), disregarding inflection.

(1) \([V \quad \text{bottle} <V_{\text{cau}} > [V \quad \text{bottle} <V_{\text{inc}} > [P \quad \text{bottle} <P > [N \quad \text{bottle} ] ]]] \) (E)

(2) \([V \quad \text{in-bottigl-i-a-} \quad [V \quad \text{in-bottigl-i-} \quad [P \quad \text{in-bottigl-i-} \quad [N \quad \text{in-bottigl-i-} ] ]]] \) (I)

The derivation in (1) includes three unpronounced < > heads. The heads and their dependents are associated with features, valued and unvalued \((u)\) eliminated via feature valuing. The noun *bottle* is internally merged to the Specifier of the prepositional (P) projection. The unvalued nominal feature \([uN]\) of the inchoative verb \((V_{\text{inc}})\) is valued by the nominal feature \([N]\) of *bottle* and the unvalued verbal feature \([uV]\) of the causative verb \((V_{\text{cau}})\) is valued by the \([V]\) feature of *bottle*. Principles of computational efficiency (Chomsky 2005, Kayne 2005, Collins 2007), contribute in part to account for the non-pronunciation of heads. Thus, in (1), the noun internally merges to the Specifier of the prepositional constituent, and given Collins (2007) Condition on pronunciation P is not pronounced. In (2), P incorporates to the lower nominal constituent and is pronounced; Vinc is spelled out by -i-, and Vcau is spelled out by -a- (Bisetto 1999, Kayne 2019). Further support for zero affixes in verb formation comes from the fact that the covert verbal affixes in (1) are externalized in English deadjectival verbs, e.g. *formal-i-z-e* as they are also in Italian, e.g. *formal-i-z-a-re*. As expected, zero morphemes are syntactically active. In (3) null causative and inchoative morphology may licence a by-phase and a purpose clause. In (4) null prepositional morphology may license locative modifiers.

(3) The wine is bottled by Masi Agricola to preserve tradition.
(4) In other regions, wine is bottled in aluminum cans.

Independent evidence for zero morphology comes from VN compounds in Italian such as *apribottiglie* (Lit. open bottles) “bottle opener”. We investigate the hypothesis that zero nominal morphology is merged to the extended verbal projection, (Franck 2008, Ferrari-Bridgers 2005). The former can be externalized by -tore (-er) in *apritore di bottiglie* (Lit. opener of bottles), in which case the presence of the preposition *di* (of) follows from unvalued Case feature valuation. We point to fMRI studies (Pliatsikas et al. 2014), indicating that morphological processing by the human brain is based on underlying morphological structure, which may include zero morphology.
3. We raise questions associated with the derivations discussed, including the following: whether exocentric structures might be derived by Pair-Merge \((a, b) = \langle a, b \rangle\), and whether Pair-Merge should be preferred to Rebooting and phrasal Spell-Out (Franco 2015) on explanatory grounds.

References
Franco, Ludovico. 2015. VN compounds in Italian and some other Romance languages, phrasal spell-out and rebooting. Rivista di Grammatica Generativa 37, 81-98.

Succinct Typology of Derivational Zero Affixes
Matthias Gerner
(Shanghai Jiaotong University)

Keywords: host boundness, affix productivity, lexicalism, Warlpiri, Gayo, Russian

Among the criteria that typologists provide to draw a distinction between inflectional and derivational affixes (Aronoff 2005: 160; Haspelmath 2002:71; Whaley 1997:121), only two appear reliable: the productivity of the affix and the boundness of its host. On this view, an inflection is an affix which is productive and whose host is bound, whereas a derivational affix is either unproductive or is attached to a host which is morphologically free.

<table>
<thead>
<tr>
<th></th>
<th>Derivational affix:</th>
<th>host is bound and affix is unproductive;</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Derivational affix:</td>
<td>host is free and affix is unproductive;</td>
</tr>
<tr>
<td>III</td>
<td>Derivational affix:</td>
<td>host is free and affix is productive;</td>
</tr>
<tr>
<td>IV</td>
<td>Inflectional affix:</td>
<td>host is bound and affix is productive.</td>
</tr>
</tbody>
</table>

* Type I: Host is bound and zero affix is unproductive

In Warlpiri, product nominalizations are lexicalized verb roots for a limited number of verbs. As verb roots are bound, they only provide the input of morphological processes (Simpson 2020:160). Examples:

<table>
<thead>
<tr>
<th>Verb root (bound)</th>
<th>Product Nominalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>jarda- ‘sleep’</td>
<td>jarda-Ø ‘a sleep’, ‘a nap’</td>
</tr>
<tr>
<td>kulu- ‘fight’</td>
<td>kulu-Ø ‘a fight’, ‘a combat’</td>
</tr>
<tr>
<td>wajili- ‘chase’</td>
<td>wajili-Ø ‘a chase’</td>
</tr>
</tbody>
</table>

322
Product nominalizations thus involve a derivative zero affix.

- **Type II:** Host is free and zero affix is unproductive

In Gayo (Austronesian, Indonesia), dynamic verbs roots are bound, whereas roots of stative and resultative verbs are morphologically free. Bound verb roots attach several prefixes, in particular the prefix *mu-* , which outputs controlling intransitive predicates, whereas free verb roots are either unmarked or marked by the prefix *mu-* in which case the verb becomes a dynamized intransitive predicate (Eades 2005:72-76, 154-159).

\[
\begin{array}{|c|c|c|}
\hline
\text{Verb root (bound)} & \text{Controlling Intransitive Predicate} & \text{Verb root (free)} \\
\hline
\text{-tuh ‘fall’} & \text{mu-tuh ‘fall’} & \text{Ø-tue ‘old’} \\
\text{-sôt ‘answer’} & \text{mu-sôt ‘answer’} & \text{Ø-pesam ‘warm’} \\
\text{-alak ‘sweat’} & \text{m-alak ‘sweat’} & \text{Ø-beluh ‘gone’} \\
\hline
\end{array}
\]

Not every free verb root may attach *mu-* , for example verbs of emotion, which allow stative/dynamic interpretations, do not use *mu-* (Eades 2005:106-109). We therefore can posit a zero prefix Ø- that encodes stative/resultative meaning but only for verb roots that also take the prefix *mu-*.

- **Type III:** Host is free and zero affix is productive

Almost all Russian verbs have imperfective/perfective forms. The perfective form is encoded by one of 16 selective verb prefixes, while the absence of prefixes expresses the imperfective form. The absence of prefixes is a derivational zero prefix of type II. The purported zero prefix is associated with imperfective meaning, is productive and its host is free. For a small set of 14 nondirectional motion verbs, the verb root is free and actually underspecified for the distinction imperfective/perfective (Tolskaya 2014:8). This suggests that the host is morphologically unbound – independently of the imperfective zero prefix.

Within generative grammar, the Warlpiri zero suffix supports weak lexicalism (‘derivative morphology takes place in the lexicon’, Anderson 1992:70), whereas the Gayo and Russian zero prefixes are morphological processes outside the lexicon.

**References:**
Zero affixes and derivational paradigms in Modern Greek adjectives

Nikos Koutsoukos & Angela Ralli
(F.R.S.-FNRS and Université catholique de Louvain & University of Patras)

Keywords: zero affixes, derivational paradigms, inflection, exocentric, adjectives

Modern Greek is a language which mainly employs overt markers (suffixes) for the transition from one lexical category to another (Ralli 2005, 2013) in contrast to other languages which may prefer categorial shift with no formal marking (Ježek and Ramat 2009, Koutsoukos to appear). For example, in Modern Greek, the adjective *θεατρικός* ‘theatrical’ is derived from the noun stem *θεατρ-* ‘theater’ with the addition of the derivational suffix *-ικ-,* while *-ος* is the inflectional ending. In other languages, the transition from noun to adjective may not always be formally marked, as in the following example from French: “Les décors du film sont (ou font) très théâtre” (cf. Lauwers and Van Goethem 2020). Here, the noun *théâtre* may have an adjectival function without any derivational affix marking this function.

In Modern Greek, only suffixes are category-changing, while prefixes are always category-preserving (Ralli 2005, 2013). Moreover, more than one derivational suffix can be employed for building a given category, the sum of which is assumed to belong to the same derivational paradigm (see van Marle 1985, Bauer 1997, and Booij and Lieber 2004 on a theoretical discussion of derivational paradigms). This is the case of the derivational suffixes shown in (1) which form adjectives out of nominal and verbal stems.

```
(1) a. ksil_N ‘wood’ → ksil-inA (os) ‘wooden’
b. nom_N ‘law’ → nom-iKA (os) ‘legal’
c. ayapi_V ‘love’ → ayapi-tA (os) ‘beloved’
d. katici_V ‘inhabit’ → katici-simA (os) ‘inhabitable’
(The inflectional ending -os appears in parenthesis)
```

However, in Modern Greek, there is an interesting group of formations which has not been thoroughly discussed so far. It refers to several adjectival structures in which there is no overt morphological marker accounting for the transition from the category of noun to that of the adjective. Consider the examples in (2).

```
(2) a. tix_N ‘luck/fortune’ → á-tix(os) ‘unlucky’
b. tolm_N ‘daring’ → pará-tolm(os) ‘reckless’
c. trom_N ‘fear/terror’ → én-trom(os) ‘fearful’
```

As shown in (2), these formations contain a prefix, a noun stem and an inflectional ending. The question is how to account for the categorial and semantic properties of these formations. In this paper, we discuss these formations and we aim at examining whether zero derivational affixes are a valuable theoretical notion.
Following the discussion in Vogel (2005) and Gaeta (2013), we propose that the postulation of zero morphemes is an issue which can be language dependent and that zero derivation relations are to be assumed only when overt exponents are to be found in other similar structures (cf. Sanders 1988). Zero derivational suffixes have been already convincingly argued for Greek exocentric compound formations (Ralli 1988, 2013, Ralli and Andreou 2012). In line with Ralli (2005, 2013) we propose that the derivational formations in (2) are also exocentric, i.e. a zero derivational suffix added assigns their category. In that respect, compounding and derivation can be similar regarding exocentricity. We also propose that the existence of zero derivational affixes are correlated with the parameter of whether a given language has rich derivational paradigms, a property which generally characterizes Modern Greek.

Acknowledgements
Nikos Koutsoukos is indebted to the Belgian Fund for Scientific Research F.R.S.-FNRS for funding the current research.

References
Booij Geert and Rochelle Lieber (2004), On the paradigmatic nature of affixal semantics in English and Dutch, Linguistics 42, 327-357.
Ralli, Angela (2005), Morphology. Athens: Patakis.
Ralli, Angela (2013), Compounding in Modern Greek, Dordrecht: Springer.
When Numbers Surface as Verbs

Costanza Marini & Elisabetta Ježek
(University of Pavia)

Keywords: numbers, verb semantics, conversion, zero-affixation, relisting

In this contribution, we report the preliminary results of a corpus-based investigation of the semantics of an English verb class which – to our knowledge – has never been systematically analysed before, namely those verbs consisting (in part or entirely) of numbers: henceforth denumeral verbs. These verbs can be considered double instances of conversion or transcategorisation (Ježek & Ramat, 2009: 395) since they undergo two categorial shifts without any superficial marking: the first from number to noun – 747 → a 747 – and the second from noun to verb – a 747 → to 747, “to fly on a Boeing 747” (Clark & Clark, 1979: 776).

By querying the EnTenTen15 corpus through the Sketch Engine (Kilgarriff et al., 2014), we identified 47 denumeral verbs, that we provisionally clustered in 10 groups (see Table 1) according to relevant linguistic and sociolinguistic criteria: e.g. the verbs from group II derive from HTTP status codes (1), while those from group III specify amounts of points used by gamers (2).

1. If you have pages that are 404ing, try redirecting them.
2. I didn’t expect him to […] 36 me in one turn.

<table>
<thead>
<tr>
<th>GROUPS</th>
<th>DENUMERAL VERBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>TO 0 SOMETHING, TO 0 SOMETHING OUT, TO 000000 A MOBILE</td>
</tr>
<tr>
<td>II</td>
<td>TO 403, TO 404, TO 500, TO 301 AN URL TO SOMEWHERE</td>
</tr>
<tr>
<td>III</td>
<td>TO 19.5 SOMEBODY, TO 21 SOMEBODY, TO 36 SOMEBODY, TO 50 SOMEBODY, TO 5/5, TO 1 FOR 2 SOMEBODY</td>
</tr>
<tr>
<td>IV</td>
<td>TO 45 SOMETHING, TO 180 SOMETHING, TO 270, TO 360, TO 3/4</td>
</tr>
<tr>
<td>V</td>
<td>TO 212 SOMEBODY, TO 302 SOMEBODY, TO 602 SOMEBODY/SOMETHING, TO 5150 SOMEBODY</td>
</tr>
<tr>
<td>VI</td>
<td>TO 14, TO 10:10, TO 420, TO DEEP 6 SOMETHING</td>
</tr>
<tr>
<td>VII</td>
<td>TO TOP 8 SOMETHING, TO TOP 32 SOMETHING, TO TOP 50 SOMETHING, TO TOP 64 SOMETHING</td>
</tr>
<tr>
<td>VIII</td>
<td>TO CALL OF DUTY 4, TO GRAND THEFT AUTO 4, TO iOS7, TO CHANNEL 9</td>
</tr>
<tr>
<td>IX</td>
<td>TO *69 SOMEBODY e TO 911</td>
</tr>
<tr>
<td>X</td>
<td>TO HIGH 5 SOMEBODY, TO +1 SOMETHING, TO 4X4, TO 69, TO 64, TO 1080 SOMETHING, TO RULE 34 SOMETHING, TO 49, TO 86 SOMEBODY, TO 123, TO 2</td>
</tr>
</tbody>
</table>

Table 1 – Identified denumeral verbs

We also tried to map our denumeral verbs onto Clark & Clark’s periphrases and we noticed that many could not be mapped, like to 420 (3).

3. We are going to meet […] for 420ing.
According to *Lexico.com*, this expression was coined in the 70’s to refer to “the time at which [a group of friends met] to locate an unattended plot of cannabis plants”. Only later it spread through the US becoming synonymous with “smoking pot”. Interpreting similar examples out of their pragmatic context is impossible.

The variety of verb meanings we found appears to be beyond the semantic potential of a zero-affix (Plag, 1999). Therefore, we suggest conceiving *denumeral verbs* as instances of *relisting* (Lieber, 2005): a stance which is not far from Clark and Clark’s (1979) *innovative coinage*, from Wilson and Carston’s (2007) *lexical adjustment*, nor from the different degrees and types of lexicalisation described in Ježek (2016: 7-13).

**References**


---

**Zero suffix in English and Italian deverbal nouns**

Chiara Melloni & Gianina Iordăchioia (University of Verona & University of Stuttgart)

**Keywords:** nominalization, verb classes, argument structure, English, Italian

We argue for a zero suffix in English and Italian deverbal zero-derived nominals (ZNs) (*to climb* > *climb-ØN*), by showing that they are more similar to suffix-based nominalizations (SNs: *examin-ation*) than previously suggested.

Borer (2013:§7) rejects *ØN* in her account of ZNs in English, by highlighting two differences between *ØN* and overt suffixes (cf. overt analogue criterion, Sanders 1988):

i) Unlike SNs, ZNs allegedly cannot instantiate compositional readings with verbal event and argument structure (Grimshaw 1990):

(1) **the importation/*import of goods** from China in order to bypass regulations (Borer 2013:332)

ii) Zero cannot attach to overt verbalizing suffixes like overt nominalizers do:

(2) **crystal(l)-ize** > *crystal(l)-ize-ØN* vs. **crystall-ize-ation**
Borer concludes that ZN-formation follows no derivational rules and crucially differs from SN-formation, as ZNs cannot embed verbal event structure and behave like categorizations of roots in nominal contexts (see also Farrell 2001, Lieber 2004).

We bring evidence for more regularity and similarity between ZNs and SNs in English and Italian than expected in such approaches. Our empirical database is a collection of 1,000 English and 300 Italian ZNs, representing 35% and 25% of the ZN-entries recorded by the OED and DISC. We argue that:

i) A good proportion of English and Italian ZNs derived from particular verb classes (change-of-state verbs, verbs of killing/throwing, etc.) may realize verbal arguments: e.g., (3) (from english-corpora.org and sketchengine.eu).

(3) a. [he] probably witnessed their **murder of his mother** (NOW)
   b. Il mancato **rinovo del contratto da parte di Mediaset** (ItTenTen16)
      *lit. the missed renewal of the contract by Mediaset’s failure to renew the contract*

ii) The incompatibility of ØN with overt verbalizers (-ize, -ify) in English is explained by its selection of native(-ized) bases, while -ize/-ify are Latinate. This is confirmed by Italian, where the native/’popular’ verbalizer -eggiare is often input to ZNs, as in (4), while the ‘erudite’ -ificare and -izzare are less likely to (Thornton 2004; see Portuguese in Rodrigues 2009).

(4) It. passo ‘step’ > **pass-egg-iare** ‘to walk’ > pass-egg-i-o ‘walks’

Our Italian database includes 11 ZNs derived from -eggiare-verbs but only 3 ZNs from -izzare-verbs, of which only utilizzo ‘usage’ is common (smobilizzo, carbonizzo are infrequent technical terms), and no ZNs derived from -ificare-verbs.

Our English database includes 20 ZNs derived from Latinate -ate-verbs (e.g., deviate, incubate), but -ate is not a productive verbalizer, since such verbs are back-formed from -ation-nominals (Plag 1999). Moreover, no -ate-ZN denotes events, as one would expect if -ate were a verbalizing suffix.

For the different verb classes, we will further show that ZNs exhibit semantic uniformity comparable with that of SNs (Kisselew et al 2016; contra Plag 1999, Lieber 2004): namely, event, state, agent, instrument readings. We will closely consider such ZN-SN parallelisms and argue that, even if ZNs frequently realize lexicalized readings, compositional readings are also available to a considerable extent, which cannot be accounted for by a root-categorization analysis, in which they never inherit verbal structure. We argue that zero must be a possible spell-out of a nominalizer, similarly to overt suffixes.

**Acknowledgments:** Iordâchioaia’s work is supported by a research grant awarded by the German Research Foundation (DFG) to the project IO 91/1-1 on Zero-derived nouns and deverbal nominalization: An empirically-oriented perspective, at the University of Stuttgart.

**References:**
ItTenTen16 -- Italian Web 2016 corpus, available at www.sketchengine.eu
NOW -- News on the Web corpus, available at www.english-corpora.org


---

**Semantic Analysis of Verb-Noun Conversions in Princeton WordNet**

Verginica Barbu Mititelu, Svetlozara Leseva & Ivelina Stoyanova

(Research Institute for Artificial Intelligence of the Romanian Academy, Institute for Bulgarian Language & Institute for Bulgarian Language)

Keywords: Princeton WordNet, zero derivation, morphosemantic links, semantic primitives

Our focus is the conversion involved in English verb-noun pairs of words. The perspective adopted is semantic, with two aims: identifying semantic regularities involved in conversion (i.e., meaning relations between the members of a pair) and the semantic conditions in which it occurs (i.e., the semantic classes which the nouns and verbs belong to), in correlation.

The study is based on the conversion pairs identified in the Princeton WordNet (PWN) (Fellbaum 1998). The network entries are the word senses, each represented by a set of synonymous literals (i.e., the synonymous words, with associated sense number, making up a synset), and they are connected through a network of lexical and semantic relations. The PWN team made available a standoff file (Fellbaum et al. 2009) containing 17,740 verb-noun pairs of which 52.8% represent derivatives (i.e., affixes are used to obtain one word from the other, thus they differ both in form and meaning) and 47.2% are conversions (i.e., their form is identical, whereas their meaning is (partially) different). The direction of conversion is not established. The great number of conversions accounts for the fact that in the data set under study we deal with word senses and each pair of verb-noun word senses is represented separately. We should also bear in mind that the synsets in PWN represent fine-grained sense distinctions. As a result, there are many cases when the “same” pair occurs several times. However, this is a false redundancy, as each time, either literal participates with a different meaning. This is an adequate way of capturing the cases when there are different semantic distinctions between the literals in the pair.

We provide qualitative and quantitative analysis of the data: for each pair displaying conversion we identify the semantic primitives (i.e., language-independent semantic classes; Koeva et al. 2016) of the components and consider the different types of semantic class pairs.
In addition, we study the semantic relations associated with these verb-noun pairs. As the preliminary observations show, zero derivation is found pervasively in relations such as Result (dimension (shape or form to required dimensions) - a dimension), Property (silence (keep from expression) - silence), Undergoer (import (bring in from abroad) - import (commodities), Location (chamber (place in a chamber) - a chamber), Vehicle (boat (ride in a boat) - a boat). Other relations are clearly dominated by affix-based derivation, such as Agent (pal (become friends) - a pal), Body-part (finger (examine by touch) - a finger), Material (lacquer (coat with lacquer) - lacquer). There are also relations which are relatively equally distributed between derivation and conversion, such as Event (guarantee (make sure or assume responsibility) - a guarantee) and Instrument (calliper (measure the diameter of something with callipers) - a calliper). An in-depth study of the trends and regularities will enable us to provide a more detailed picture of the distribution of conversion. Conclusions will highlight similarities and differences between the so-called “zero” morpheme form conversion and the affixes involved in derivation from the perspective adopted in this work.

Acknowledgements
The work reported here has been undertaken under the bilateral project Enhancing Multilingual Language Resources with Derivationally Linked Multiword Expressions (2018-2020) between the Institute for Bulgarian Language at the Bulgarian Academy of Sciences and the Romanian Academy Research Institute for Artificial Intelligence.

References

Prohibition against Category Resumption by Cyclic (Zero) Derivation and the Phase Impenetrability Condition

Yoshiki Ogawa
(Tohoku University)

Keywords: zero derivation, category resumption, economy of derivation, Phase Impenetrability Condition, Japanese

In this presentation, assuming that “morphological” derivations involves syntactic operation (Marantz 1997, Harley 2009), that zero derivation involves syntactic merger of zero categorizers (Don 2005), and that grammar allows a phrasal affix (Ogawa 2001, Kishimoto 2006), I will argue that the Phase Impenetrability Condition (PIC) (Chomsky 2001) constrains not only internal merger (syntactic movement) but also external merger. The argument comes from otherwise unexplained restrictions on “word” formation in Japanese.
In Japanese, N, V and A can each change their categories, either by overt derivation, as in (2a,b,d), or by zero derivation, as in (3a,b,d). However, category resumption by cyclic derivation is prohibited:  

(1) *[[[M1]-M2]-M3], where M1=M3 categorically.

Thus, [[[N]-V]-N], [[[N]-A]-N], [[[A]-N]-A], [[[A]-V]-A], [[[V]-N]-V], and [[[V]-A]-V] are all ill-formed, though other combinations are possible. This restriction applies to both overt and zero derivation, ruling out (2c,e) and (3c,f), respectively:

(2)  
- garu\(_v\) selects A; - sii\(_A\) selects V; - sulsi\(_V\) selects N/VN; - kata\(_N\) selects V/VP:  
   a. [[hosi\(_A\)]-gariv\(_V\)]-katan\(_N\)  ‘want\(_A\)-express\(_V\)-way\(_N\)’  
      (=how to express one’s desire)  
   b. [[[urayama\(_A\)]-sii\(_A\)]  ‘envy\(_A\)-feel\(_A\)’ (=envious)  
   c. *[[[hosi\(_A\)]-garav\(_V\)]-sii\(_A\)]  ‘want\(_A\)-express\(_V\)-feel\(_A\)’  
   d. [[[kekkon\(_N\)]-suv\(_V\)]-ru  ‘marriage\(_N\)-do\(_V\)-Nonpast’ (=to marry)  
   e. *[[[kekkon\(_N\)]-siv\(_V\)]-katan\(_N\)]  ‘marriage\(_N\)-do\(_V\)-way\(_N\)’ (=how to marry)

(3)  
zero derivations:  
   a. [[ikiv\(_N\)-kaerv\(_V\)]-(ga)  ‘gov\(_V\)-return\(_V\)-(Nom)’ (=path)  
   b. [[jikov\(_N\)]-(Ru\(_V\)]  ‘accident\(_N\)-\(\varphi_v\)-Nonpast’ (=cause an accident)  
   c. *[[[ikiv\(_N\)-kaerv\(_V\)]-[\(\varphi_v\)-v\(_V\)-u]  ‘gov\(_V\)-return\(_V\)-nominalizer-verb\(_V\)-nonpast’  
      (=go and come)  
   d. [[dai-[[ninki\(_N\)]-\(\varphi_A\)]-da]  ‘large\(_A\)-pop\(_V\)-adj\(_V\)-Cop’ (=is very popular)  
   e. [[ninki\(_N\)]-ga ar-u  ‘pop\(_V\)-be\(_N\)-Nonpast’ (=is popular)  
   f. *[[[dai-[[ninki\(_N\)]-\(\varphi_A\)]-\(\varphi_N\)]-ga  ‘large\(_A\)-pop\(_V\)-adj\(_V\)-nom\(_N\)-nom\(_N\)-be\(_N\)-Nonpast’ (=is very popular)

(3a) contains \(\varphi_N\) because juxtaposition of two antonymous non-nominals behaves as N (cf. Scalise et al. 2009); (3b,c) contain \(\varphi_V\) because the tense morpheme –(r)u selects V; (3d) contains \(\varphi_A\) because the copula must select adjectival predicates.

Interestingly, (1) is suspended when a phase head H (=Comp/Determiner) intervenes between M1 and M3, as in [[[M1]-M2]-H]-M3 or [[[M1]-H]-M2]-M3:  

(4)  
- yasui\(_A\) ‘easy’ selects CP:  
   Omocha-wa  [ap[CP[TP[CP[Top-\(\varphi_A\)-kodomo-ga  [[hosi\(_A\)]-gariv\(_V\)]-[\(\varphi_C\)-yasui\(_A\)]  
      toy-Top  child-Nom  want\(_A\)-\(\varphi_T\)-\(\varphi_C\)-\(\varphi_A\)-easy\(_A\)  
      (=Toys, children are likely to want.)

(5)  
- sa\(_N\) selects AP potentially containing DP:  
   a. [ap[ap[DP-sono-\(\varphi_D\)-\(\varphi_D\)-rasi\(_A\)-\(\varphi_D\)]-sa\(_N\)]  
      that-person-\(\varphi_D\)-typical\(_A\)-nominalizer  
      (=the typicality of each person)  
   b. [ap[ap[DP-otokon\(_N\)]-[\(\varphi_D\)-rasi\(_A\)-sa\(_N\)]  
      man-\(\varphi_D\)-typical\(_A\)-nominalizer’  
      (=manliness)

I propose that (1) is a kind of economy principle that requires derivation to use the smallest number of categorizers to reach its final state (e.g., √root→V→N is more economical than √root→N→V→N), and that (4) and (5) do not violate (1), because the PIC also constrains external merger, where the transferred domain is irrelevant to the computation of (1).
-yasui_A in (4) must select CP because a Nominative subject should be licensed in CP (Miyagawa 2011). Thus, a phasal CP intervenes between -yasui_A and the lowest A in (4). Similarly, -saN in (5b) can select AP that contains DP because it clearly does so in (5a). Thus, a phasal DP intervenes between -saN and the lowest N in (5). In these environments, the PIC prevents cyclic derivation in (4)-(5) from violating (1) because part of it is already transferred when (1) applies.

The acceptability of English examples like manliness might counterexemplify (1). However, -ness can take a phrase containing a functional head, as in (6a) (cf. Allen 1978); hence, we can assume that manliness contains an invisible phase head (D), as in (6b):

(6) a. matter-of-factness, salt-of-the-earthiness, deaf-and-dumbness, out-of-controlness (COCA)
   b. [af[Np[Np[manN]-φ]-li_A]-nessN]

References
Don Jan (2005), On conversion, Relisting and zero-derivation: A comment on Rochelle Lieber: English word formation processes, a paper presented at the Word-Formation Theories workshop, held 25-26 June 2005 at Prešov University.
Miyagawa, Shigeru (2011), Genitive subjects in Altaic and specification of phase, Lingua 121, 1265‒1282.
Scalise, Sergio, Antonio Fábregas and Francesca Forza (2009), Exocentricity in compounding, Gengo Kenkyu 135, 49-84.

The fuzzy boundary between affixation and conversion in Ancient Greek denominal verbs
Matteo Pellegrini & Fabio Montermini

Zero-Derivation by Phase
Marc Richards
(Queen’s University Belfast)
Keywords: zero-derivation, conversion, phases, Transfer, Neg-raising

Recent psycholinguistic studies into the processing and production of zero-derived forms and conversion pairs have added to the long-standing tension between lexical and structural approaches to conversion phenomena (see e.g. Platsikas et al 2014 and Lohmann 2018 for opposing viewpoints). This paper attempts to resolve this tension and to reconcile the two approaches by proposing a syntactic model that can potentially combine the theoretical and empirical advantages of both. Instead of introducing additional structure in the form of a zero morpheme (i.e. a Merge-based approach requiring the postulation of a zero derivational affix), I instead propose and explore a phase-based approach (cf. Chomsky 2001, 2008) in which the categorial distinction between nouns and verbs reduces to basic configurations of categorizers, roots and their complements, and to the timing of Transfer. More specifically, the approach is premised on a configurational conception of categories in which a generalized lexical categorizing head K, characterized as a phase head, combines with the root R and its argument XP according to the two logical possibilities in (1):

(1) a. \{K, R, \{XP\}\} = “[+N]”
    b. \{K, \{R, XP\}\} = “[-N]”

As well as deriving many of the essential differences between nouns and verbs (such as case versus agreement and extraction asymmetries), the above approach implies that the operation Transfer will itself suffice to turn a [-N] structure into a [+N] structure, yielding the effect of conversion, as in (2) (see Chomsky 2013: 44 for the relevant conception of Transfer):

(2) a. \{K, \{R, XP\}\} \rightarrow Transfer \{K, \{R, XP\}\} \rightarrow \{K, \{R, XP\}\}
    b. \{K, R\}

Essentially, word-internal Transfer results in re-categorization, converting between nominal and verbal configurations without the need for an additional zero head. Two main empirical consequences of this approach are considered. Firstly, it follows from (1) that unergative verbs are zero-derived denominals (recapturing the N-to-V incorporation of earlier syntactic approaches based on Hale & Keyser’s seminal work), since every derivation must terminate in a \{K, R\} base-pair. Secondly, evidence in favour of the proposed analysis comes in the form of certain well-known (but thus far less well understood) restrictions on Neg-raising, which now fall into place once we adopt a syntactic approach to the latter (cf. Collins & Postal 2014). Based on the observation that otherwise-expected Neg-raising is blocked just in cases where conversion is indicated, as in the classic contrast in (3), I argue that phase-based zero-derivation predicts precisely this restriction as a consequence of word-internal Transfer (and its corollary, a version of the Phase Impenetrability Condition).

(3) a. I don’t think that John will come to the party.
    \Rightarrow I think that John won’t come to the party.
    I don’t think so.
    b. I don’t hope that John will come to the party.
    \Rightarrow I hope that John won’t come to the party.
    *I don’t hope so.

In (3b), a phase boundary, induced by Transfer, prevents the formation of a cross-clausal negative dependency just when conversion takes place. Such conversion-induced opacity effects would be harder to capture under existing lexical and structural approaches.
Zero-derived nominals in a language with obligatory verbal suffixes: The case of Czech

Magda Ševčíková
(Charles University, Prague)

Keywords: zero suffix, derivation, conversion, aspect, Czech

In Czech, derivational zero suffixes are assumed to occur in nouns that denote an action or a result. These nouns are viewed as derived from verbs because the action meaning is primarily expressed by verbs. As verbs in Czech are obligatorily marked by a thematic suffix conveying grammatical aspect, deverbal zero-suffix nouns belong to the minority of derivations that have a simpler morphemic structure than their base words and violate thus the general account of derivation as an affix-adding process (Iacobini 2000: 866; cf. (1)). In analogy to verb-to-noun conversion in English, zero-suffix action nouns have been referred to as conversion in the literature on Czech (e.g. Dokulil 1968a, 1968b, Bozděchová 2016), though not complying with the canonical type of conversion defined by word-class change and formal identity (Valera 2014, Štekauer 1996).

(1) \( \text{skok-0} \prec \text{skák-a-t} \)
\( \text{jump-0} \) \( \rightarrow \) \( \text{jump-IPFV-INF} \)
\( \text{‘jump’} \) \( \rightarrow \) \( \text{‘to jump’} \)

In the paper, the derivational zero concept is discussed by analysing a corpus-based sample of 300 nouns that lack the suffix and have a direct verbal counterpart (Křen et al. 2015). The sample contains deverbal zero-suffix nouns but also noun-verb pairs with an analogous morphemic structure but a reverse direction of derivation; cf. the verb in (2) the meaning of which is explained by referring to the noun (cf. semantic dependency in Marchand 1964).

(2) \( \text{lék-0} \succ \text{léč-i-t} \)
\( \text{medicine-0} \) \( \rightarrow \) \( \text{medicine-IPFV-INF} \)
\( \text{‘medicine’} \) \( \rightarrow \) \( \text{‘to medicate’} \)

In addition to semantic dependency, which is determined by using a general dictionary in the present study (Filipec et al. 1998), the sample of suffixless nouns and related verbs is analysed for the category of grammatical aspect in order to prove the hypothesis that the direction of derivation

References

Zero-derived nominals in a language with obligatory verbal suffixes: The case of Czech

Magda Ševčíková
(Charles University, Prague)

Keywords: zero suffix, derivation, conversion, aspect, Czech

In Czech, derivational zero suffixes are assumed to occur in nouns that denote an action or a result. These nouns are viewed as derived from verbs because the action meaning is primarily expressed by verbs. As verbs in Czech are obligatorily marked by a thematic suffix conveying grammatical aspect, deverbal zero-suffix nouns belong to the minority of derivations that have a simpler morphemic structure than their base words and violate thus the general account of derivation as an affix-adding process (Iacobini 2000: 866; cf. (1)). In analogy to verb-to-noun conversion in English, zero-suffix action nouns have been referred to as conversion in the literature on Czech (e.g. Dokulil 1968a, 1968b, Bozděchová 2016), though not complying with the canonical type of conversion defined by word-class change and formal identity (Valera 2014, Štekauer 1996).

(1) \( \text{skok-0} \prec \text{skák-a-t} \)
\( \text{jump-0} \) \( \rightarrow \) \( \text{jump-IPFV-INF} \)
\( \text{‘jump’} \) \( \rightarrow \) \( \text{‘to jump’} \)

In the paper, the derivational zero concept is discussed by analysing a corpus-based sample of 300 nouns that lack the suffix and have a direct verbal counterpart (Křen et al. 2015). The sample contains deverbal zero-suffix nouns but also noun-verb pairs with an analogous morphemic structure but a reverse direction of derivation; cf. the verb in (2) the meaning of which is explained by referring to the noun (cf. semantic dependency in Marchand 1964).

(2) \( \text{lék-0} \succ \text{léč-i-t} \)
\( \text{medicine-0} \) \( \rightarrow \) \( \text{medicine-IPFV-INF} \)
\( \text{‘medicine’} \) \( \rightarrow \) \( \text{‘to medicate’} \)

In addition to semantic dependency, which is determined by using a general dictionary in the present study (Filipec et al. 1998), the sample of suffixless nouns and related verbs is analysed for the category of grammatical aspect in order to prove the hypothesis that the direction of derivation
corresponds to the way how the verbs form their aspectual counterparts in Czech (Ševčíková, in prep.). The vast majority of the noun-verb pairs in the sample confirm that deverbal zero-suffix nouns are motivated by verbs that use a suffix to change the aspect whereas denominal verbs prefer prefixation; see (3) vs. (4).

(3) skák-a-t : skoč-i-t
     jump-IPFV-INF       jump-PFV-INF
     ‘to jump.IPFV’      ‘to jump.PFV’

(4) léč-i-t : vy-léč-i-t
     medicine-IPFV-INF   PREF-medicine-IPFV-INF
     ‘to medicate.IPFV’  ‘to medicate.PFV’

Deverbal zero-suffix nouns are compared to other nominalizations with the same roots, particularly to inflectional nominalizations in -ně and to derivational nominalizations with overt suffixes. Nominals of the former type are attested with unmotivated verbs of both aspects (5a,b) as well as with denominal verbs (6a). As for derivational nominals, denominal verbs prefer overt analogues to the zero suffix; cf. -ba in (6b). Differences in lexical meaning, valency features and corpus frequency scores of the nominalizations are pointed out in the paper.

(5) a. skák-a-t > skák-á-ní
     jump-IPFV-INF       jump-IPFV-NMLZ
     ‘to jump.IPFV’      ‘jumping’

 b. skoč-i-t > skoč-e-ní
     jump-PFV-INF       jump-PFV-NMLZ
     ‘to jump.PFV’      ‘jumping’

(6) a. léč-i-t > léč-e-ní
     medicine-IPFV-INF   medicine-IPFV-NMLZ
     ‘to medicate.IPFV’  ‘medication’

 b. léč-i-t > léč-ba
     medicine-IPFV-INF   medicine-NMLZ
     ‘to medicate.IPFV’  ‘medication’

Acknowledgements:
This work was supported by the Grant No. GA19-14534S of the Czech Science Foundation. It has been using language resources developed, stored, and distributed by the LINDAT/CLARIAH-CZ project of the Ministry of Education, Youth and Sports of the Czech Republic (project LM2018101).

References:


Ševčíková, Magda (in prep.), Action nouns vs. nouns as bases for denominal verbs in Czech: A case study on directionality in derivation.


Derivational zero suffixes and their alternatives

Petra Sleeman

<not updated>

The semantics of noun-to-verb zero derivation/conversion in English and in Spanish

Salvador Valera

<not updated>
WORKSHOP 5

Multifunctionality and syncretism in non-finite forms

Pavel Rudnev, Ksenia Shagal & Anna Volkova
(NRU Higher School of Economics, Moscow; University of Helsinki; & NRU Higher School of Economics, Moscow)

Keywords: non-finite forms, multifunctionality, typology, subordination, nominalization

This workshop aims at discussing multifunctionality in the domain of non-finite verb forms. Crosslinguistically, it is common for non-finite forms to occur in a variety of dependent clauses performing different functions. For example, as can be seen from Cristofaro’s (2003) data, in the majority of cases, non-finite forms used for adnominal modification are not specific to this function, but can be found in other subordinate constructions as well (see also Hendery 2012: 171). Van Lier (2009: 206–210) shows that in a genealogically and geographically balanced sample of 50 languages all logically possible combinations of functions are attested, i.e. reference function, adnominal modification, and adverbial modification. For instance, the nominalization in -n- in Kayardild (Tangkic), an example of the most flexible form, can function as a predicate of a relative clause, a complement clause, or an adverbial clause (Evans 1995: 474–476). Forms that combine the function of predicate of a relative clause with that of predicate of a complement clause are especially common (Koptjevskaja-Tamm 1993: 42–44; Shibatani 2009, 2019).

Thus, we believe that these phenomena merit special attention. In particular, the workshop aims to investigate to what extent the internal syntax of non-finite clauses headed by these multifunctional forms is determined by their external syntax as realized in distinct grammatical functions, and vice versa. For example, in Kalmyk (Mongolic), participles show differences in argument marking depending on the type of dependent clause in which they occur: participial relative clauses have a genitive (or in some cases nominative) subject, while complement clauses headed by the same forms require a subject in the accusative. In Muna (Austronesian), nominalizations in ka- behave differently in reason clauses if compared to the argument function. In reason clauses, they take a specialized non-finite negator and combine with a future marker, while in the argument function they take a constituent negator and no TAM markers whatsoever (van den Berg 2013: 144, 211).

The existence of such multifunctional forms and patterns of syncretism implies a certain degree of categorial continuity which poses serious challenges to theories postulating an inventory of discrete lexical categories (see Baker 2011; Kiparsky 2017 for similar observations), especially the most restrictive ones appealing to just two basic lexical categories of, roughly, noun and verb. On the other hand, those formal theories which dispense with lexical categories altogether by viewing categorization as a byproduct of how functional elements and acategorial lexical elements are put together seem to predict a certain amount of variability and continuity (cf. Halle & Marantz 1993, Borer 2005, Caha 2007, among others), but require additional constraints to exclude rare or unattested patterns. In this regard, multifunctional non-finite forms present an ideal testing ground when choosing between lexicalist and non-lexicalist approaches. As regards the forms themselves, their identity is frequently taken — at least in non-lexicalist frameworks — to be the result of underspecification, formalized in terms of Halle’s (1997) Subset Principle or, conversely, the Superset Principle of nanosyntax (Caha 2007; see Lundquist 2008 for an analysis of multifunctional non-finite forms in Swedish along these lines). That way, a language can use the same building blocks for
encoding relations that the grammatical system deems similar or at least similar enough, both synchronically and diachronically (Ramchand 2018).

The questions we want to address at the workshop include the following:

1) What functions are typically combined in one non-finite form? Are there constraints on multifunctionality, and if so, then what are they?
2) What kind of differences can multifunctional forms show depending on the function they perform (e.g. differences in TAM expression or argument encoding)? Is there a connection between the function of such form and the degree of finiteness it demonstrates?
3) What sets multifunctional forms apart in comparison to dedicated (specialized) forms, that is, participles, infinitives, converbs or nominalizations? If not all of the non-finite forms in a language show multifunctionality, which are most likely to be multifunctional? Relatedly, what, if anything, stops the remaining forms from being multifunctional?
4) Can any observations be made regarding the frequency distributions of different functions for specific forms? In other words, if a form can be used in several types of non-finite dependent clauses, do the frequencies tend to be evenly distributed across contexts, or is the distribution typically skewed in favour of one function?
5) Do languages with multifunctional non-finite forms tend to have any particular properties in other domains of grammar?
6) Is multifunctionality of non-finite forms more common in certain geographical areas than in others, and if so, where is it most commonly attested?

We invite contributions from a variety of backgrounds and perspectives, both functional and formal. We particularly welcome submissions on understudied or underdocumented languages provided they offer an explicit and substantial theoretical contribution to the issue at hand. Possible topics for submissions may include (but are not restricted to) the following:

- In-depth case studies of multifunctional non-finite forms in individual languages;
- Typological studies of the distribution of multifunctional non-finite forms;
- Formal and functional approaches to the identity of form and diversity of function;
- Quantitative studies of multifunctionality within individual languages and crosslinguistically.

Acknowledgements: Financial support from the Russian Science Foundation (grant #18-78-10128) is gratefully acknowledged.

References


The loss of Jê nominal verbs in Panará
Bernat Bardagil

Kazym Khanty non-finite forms:
Multifunctionality and variability in the amount of structure

Daria Bikina, Denis Rakhman, Aleksey Starchenko & Svetlana Toldova
(HSE and Harvard University; HSE; HSE; & HSE)

Keywords: syntax, non-finite forms, complementation, relativization, Uralic

In Khanty, most of subordination makes use of non-finite forms (Nikolaeva 1999a: 46), e.g. relative clauses (1), control clauses (2), temporal (3) and reason/purpose clauses (4), sentential complements (5-6).

1. **tuw-e luw isa tut juχ səwər-ti lájm-əl**!
   *bring-IMP.O s/he usually fire wood chop-NFIN.NPSTaxe-POSS.3*
   Bring the axe that he usually chops the firewood with!"
waśaj-en ńañ le-ti wuli-s

Vasya-POSS.2SG bread eat-NFIN.NPST stop-PST[3SG]

‘Vasya stopped eating bread’.

muŋ ropata táj-t-ew-an ɣot-el
we job have-NFIN.NPST-POSS.1PL-LOC house-POSS.3
jem-a tâj-1-ol-ew

good-DAT have-NPST-O-1PL.

‘When we have a job, people in our house live well’.

aŋk-em kor-an punšta-ti pâta ſuţi lot-əs
mother-POSS.1SG stove-LOC close-NFIN.NPST for meat buy-PST[3SG]

‘My mother bought some meat to bake it’.

ma waśaj-en juχɔt-t-al ɣoll’-əs-əm
I Vasya-POSS.2SG come-NFIN.NPST-POSS.3 hear-PST-1SG.SG

‘I heard that Vasya would come’.

luw wuli katəl-ti wer-l ma wo-l-em
s/he deer catch-NFIN.NPST deed-POSS.3 I know-NPST-1SG.SG

‘I know that he is catching a deer’.

Also, mirative forms are derived by attaching a possessive marker to a non-finite form (which is common in Uralic, cf. Jalava 2016):

aʃ-em juz ʃəwɔr-t-al
father-POSS.1SG wood chop-NFIN.NPST-POSS.3

‘It turns out that my father is chopping the wood’.

The syntactic properties of Kazym Khanty non-finite forms vary in different constructions: they may take possessive agreement marker or not; they can occur in independent clauses; they can be postposition complements or noun modifiers, etc. (see 1–7). We claim that Kazym Khanty non-finite constructions fall into five classes, represented in Table 1.

Table 2. Classification of non-finite forms uses in Kazym Khanty

<table>
<thead>
<tr>
<th>Constructions that belong to the class</th>
<th>Agreement placement</th>
<th>Amount of the structure</th>
<th>Other distinguishing syntactic properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Relative clauses</td>
<td>On the head noun</td>
<td>&lt;TP</td>
<td>Interaction with argument alternation (restricted)</td>
</tr>
<tr>
<td>2. Sentential complements of factive predicates with wer ‘deed’ + temporal clauses that have truth presupposition</td>
<td>On the head</td>
<td>TP</td>
<td>Interaction with argument alternation (possible)</td>
</tr>
<tr>
<td>3. Adverbial clauses and sentential complements of non-factive predicates</td>
<td>On the participle</td>
<td>~CP</td>
<td>Optional agreement with null subject</td>
</tr>
<tr>
<td>5. Control clauses</td>
<td>No agreement</td>
<td>~VP</td>
<td>Can only be formed with the NPST form; no overt subject</td>
</tr>
</tbody>
</table>
The proposed classification is based on the agreement properties of the constructions and on the amount of functional structure they involve. The first two classes exhibit possessive agreement on the head but are of different size: relative clauses cannot be modified with TP-level adverbs, while the constructions that belong to the second class can. In uses 3 and 4, the agreement marker attaches to the participle itself, and CP-level adverbs are allowed. Control clauses involve less structure, show no agreement and never allow for an overt subject.

The main question we address in this study is the one about the exact structure of aforementioned clause types and their interrelationships, i.e. which of them constitute synchronically distinct, although homonymous, constructions, and which are the instantiations of a single underlying phenomenon. We claim that different amount of structure involved is the core cause of variation observed.

Acknowledgments
This research has been supported by RFBR, grant № 19-012-00627.

References

The structural underpinnings of the multifunctionality and syncretism in non-finite forms in Irish

Maria Bloch-Trojnar
(The John Paul II Catholic University of Lublin)

Keywords: multifunctionality, syncretism, Irish, Distributed Morphology, AspP layer

The paper aims to present the pervasive multifunctionality and syncretism observed in non-finite forms in Irish through the optics of the Distributed Morphology framework (Halle and Marantz 1993). Traditional grammarians recognize two non-finite verb forms: the verbal noun (VN) and the verbal adjective (VA).

VN can function as infinitives, present participles in periphrastic progressive constructions and nominalizations. Nominals in Irish are ambiguous only between the Simple Event and Result nominal status, thus Irish lacks Complex Event nominals containing the AspP layer in their structure (Alexiadou and Grimshaw 2008: 2). SE-nominals (also found in Light Verb Constructions) are represented as the [nP [vP [Root]]] complex and R-nominals, which behave like ordinary nouns, as [nP [Root]]. Event denoting nominals in Irish can license the internal argument but aspectual modification and external argument licensing are not possible (cf. synthetic compounds in Greek in Alexiadou 2017), which means that the licensing of Argument Structure is independent of an aspectual reading of the event. This ties in with a DM analysis of verb inflection in Irish as lacking an aspect morpheme (Acquaviva 2014). A finite clause with a lexical verb in Irish, contains an AspP located just above vP (corresponding to VoiceP in this analysis) and nested into a Tense Phrase (TP). Asp is the likely locus for preposition-like particles involved in aspectual periphrases, like \( ag \) in the
progressive (McCloskey 1996). What is interpreted as the participle in the progressive or the infinitive in modal constructions is, in fact, a [vP [Root]] complex, which, like nominals, lacks the aspectual projection.

The verbal adjective (VA) discharges the role of the perfective participle. It is used in passive perfective and resultative constructions or as a nominal postmodifier (e.g. pláta briste ‘broken plate’). It is homonymous with the genitive case of the VN, which should be viewed as a positional variant of the present participle.

It transpires that non-finite forms contrast structures where the [vP[root]] complex is dominated by nP or aP (Lundquist 2008). The extensive multifunctionality and syncretism stem from the lack of the AspP layer in their structure.

References

Multifunctionality in dependent clauses in Cariban languages

Natalia Cáceres Arandia

<not updated>

Essential features of Kildin Saami non-finites

Ivan Gamov
(Rice University)

Keywords: non-finites, nominalization, multifunctionality, morphology, dependent-marking

We are focusing on three non-finite forms in Kildin Saami language, which are regularly derived from verbs with affixes -men’, -em and -mušš. These phenomena were barely described in previous works (Kert 1971, Kert 1993, and Rießler 2020). They do not have analogs in other Saami languages. We have determined 1) argument marking strategies for the non-finites as well as 2) their syntactic functions.
The data for this study was obtained during a fieldwork involving 20 native speakers of Kildin Saami. The examples were produced by the consultants themselves and then confirmed by at least two other people. In addition, ungrammatical sentences were discussed with the consultants in a manner of elicitation.

1) For all three forms, the marking strategy appears to be the same: subject in Oblique case and direct object in Nominative case and direct object in Nominative or Genitive- Accusative case (free variation):

(1) oaxxka-n’ koppč’-men’ kuumpar/kuumprat l’ai pir’r’-mus
grandma-com(Emma). ꞉ pick-ptcp1 mushroom.nom.pl/gen(acc).pl be.p.t.3 ꞉
gooder-adj
‘the picking of mushrooms by the grandma was the best’.

(2) munn roamm-guedd-e a3’-an’ pued’d’-em
1 ꞉ joy-vblz.p.t-1 ꞉ father-com(Emma). ꞉ come-ptcp2
‘I rejoiced the coming of the father’.

(3) sonn c’oajed’ kuumpar/kuumprat koppč’-em baal’
3 ꞉ get_loss.3 ꞉ noun phi pick-ptcp2 during
‘he got lost during the picking of mushrooms’.

(4) kül’/kul’ šil’-mušš taal-en l’ai pir’r’-mus
fish.nom. ꞉ noun phi catch-nmlz bear-com(Emma). ꞉ be.p.t.3 ꞉
gooder-adj
‘the catching of fish by the bear was the best’.

Clarification: Kildin Saami has an accusative alignment (Kuruch 1985). Genitive and Accusative cases are homonymic as well as traditional Comitative and Essive cases, which are denoted here as Oblique.

Having examined the data from (Koptjevskaja-Tamm 1993), we cannot find any examples, where subject of both transitive (1) and non-transitive (2) verbs is marked with any oblique case, whereas direct object is marked with either Nominative, Genitive or Accusative. This makes Kildin Saami non-finites uncommon in the typological sense.

2) The syntactic functions of the investigated forms are distributed as follows:

- men’ - noun phrase head (1), adverbial phrase head (5), attribute (6);

(5) sijj eevt-elmexx-t-en’ vijj-men’ c’arr m’il’l’te
3pl front-adv hunt-vblz-3pl.p.t riding-ptcp1 tundra.gen(acc). ꞉ across
they used to hunt, riding across tundra’.

(6) kull-men’ mojnas
hear-ptcp1 tale.nom. ꞉
tale, which was heard.
- em - noun phrase head (2, 3), attribute (7);

(7) kull-em mojnas
hear-ptcp2 tale.gen(acc). ꞉
‘who was hearing the tale’.

In a typological study (Nedjalkov 1990), there are no examples of a form combining 3 functions like -men’. That is why we consider the first multifunctional form idiosyncratic and worth investigation.
We also argue that the latter form -mušš is monofunctional due to its greater degree of nominalization in the sense of (Serdobolskay 2005): it can have case markers (8), while the other two cannot (9).

(8) až-an’ puedd-mušš-xa munn ę’ojv-e
father-com(e-ž()-).G come-nmlz-abe l.G be_bored-1.G.p.T
‘without coming of the father I was bored’.
(9) *až-an’ puedd-men’-xa munn ę’ojv-e
father-com(e-ž()-).G come-ptcp1-abe l.G be_bored-1.G.p.T
‘without coming of the father I was bored’.

I would like to acknowledge Oleg Volkov with gratitude for making this research possible.

References


Nedjalkov, V. P. (1990), Osnovnyje tipy dejeprichastij [The main types of converbs], Tipologija I grammatika [Typology and grammar], Moskva: Nauka, 36 - 59.


Serdobolskay, N. V. (2005), Sintaksicheskij status aktantov zavisimoj nefinitnoj predikacii [The syntactic status of actants of a dependent non-finite predication], PhD dissertation, Moskva: MGU.

Glossing abbreviations

ABE - abessive case
ADJ - attributivization marker
ADV - adverbalization marker
COM(ER) - oblique (comitative-essive) case
GEN(ACC) - genitive-accusative case
NMLZ - nominalization marker -mušš
 NOM - nominative case
NP-T - non-past tense
PTCP1 - past tense
PTCP2 - nominalization marker -men’
PTCP2 - nominalization marker -em
VBLZ - verbalization marker

Gender and number markers in Yukuna (Arawakan): from nominalization to clause-chaining
Magdalena Lemus Serrano
(Université de Lyon 2)

Keywords: <nominalization, finiteness, grammaticalization, multifunctionality, Arawakan>

Yukuna (ycn, Arawakan, Colombia), has a set of gender and number (G/N) suffixes (-ri M, -yo F, -ño PL) used for subject grammatical nominalizations (Shibatani 2019, Kohtjevskaja-Tamm 1993), which is the main strategy to encode relativization. Besides G/N marking agreeing with the referent, these nominalizations are characterized by a subject gap, the presence of finite verbal features (TAM, argument marking) and optional adnominal modifiers (demonstratives, numerals) (1).

(1) Méñö țjó ri=ț-cha=o, [kêlê nô-cha-ri]
where.towards 3SG.M=escape-PST=MID MED.DEM kill-PST-M 2SG.PRO
‘Where did he escape to, that one who killed(M) you?’

This multifunctional set of G/N markers is also found on verbs in other constructions, most importantly, subject focus clauses (2) and same-subject clause-chaining (3). These constructions, although formally very similar to (1), display significant differences that impede a unified synchronic analysis. Despite their many similarities, the nominal features of function (1) are slightly reduced in function (2), and they are almost entirely absent in function (3). Indeed, the construction in (2) has a cleft-clause reading, where the nominalized verb is still compatible with demonstratives, and shows semantically transparent G/N marking, similarly to (1). In contrast, the construction in (3) does not allow the use of demonstratives, and the masculine singular suffix -ri is used almost by default, even with feminine and plural referents.

(2) Nukhá pa'-jé-ri=o amí cháje.
1SG.PRO return-FUT=M=MID mother at.toward
‘I will return(M) to my mother’s place.’
(Alternatively: ‘The one who will return(M) to my mother’s place is me.’)

(3) Ru=yağıi'-cha=o, amí-cha-ri ru=ja'pá-ka tάa kame-rú ŋáni.
3SG.F=look-PST=MID see-PST-M 3SG.F=walk-NMLZ EMPH black-F DIM
‘She, looked, and saw(M), that she, was walking (looking) black(F).’

This synchronic variation reveals that G/N markers used in nominalization/relativization are grammaticalizing into a one clause-chaining marker --ri. However, this process is synchronically still ongoing, as evidenced by a few instances where clause-chaining shows semantically transparent G/N marking, as in (4). In these cases, there is ambiguity between the different functions of G/N markers, and the interpretation is context-dependent.

(4) nekhá li'-cha-ńó kamú'ji ké=ja, jemi'-ché-ńo ri=yukána
3PL.PRO do-PST-PL poor like=LIM hear-PST-PL 3SG.NF=story
‘they were living poorly, and (they) heard his story’
(Alternatively: ?they were the ones who lived poorly, the ones who heard his story’

This study explores the different functions of these G/N markers by taking into account morphosyntactic, semantic, discourse, and frequency-related factors based on a corpus of narratives. The synchronic facts suggest a grammaticalization pathway whereby G/N markers shifted from encoding background information about participants in nominalizations to encoding sequential events.
This functional shift then motivated the semantic bleaching of the G/N markers, which combined with the higher frequency of masculine gender suffix –ri, led to its near adoption as default clause-chaining marker. While formal similarities and historical links between relativizers and temporal adverbial subordinates have been reported in many languages (Heine and Kuteva 2007: 251–253; Epps 2009), the synchronic Yukuna data provides us with new insights into the historical relation between these functional domains with a snapshot of a not-yet-culminated grammaticalization process.

**Abbreviations**

DIM diminutive; FUT future; M masculine; MED.DEM medial demonstrative; MID middle voice; NMLZ nominalizer; PL plural; PRO pronoun; PST past; SG singular.

**References**


---

TAME in multifunctional non-finites: towards a unified account of nominalized forms in Northern Khanty

Nikita Muravyev

<not updated>

**Puzzling patterns in non-finite forms**

Eric Reuland

(Utrecht institute of Linguistics OTS)

Keywords: nominalization, verbalization, infinitive, multifunctionality, underspecification

The Dutch -en affix apparently marks infinitives and nominalizations, the English -ing affix nominalizations, progressives and participles. Such *Multifunctionality* can reflect:

(i) *Underspecification in syntax.* (ii) *Underspecification in PF.*

Claim: What looks like (ii) may instantiate (i).

1. Traditionally, two types of infinitives are distinguished in Dutch, with a *verbalizing* and a *nominalizing* affix, both spelled out as -en:

I. verbal infinitives:

(1) *Cindy zal dat boek lezen.*

‘Cindy will that book read’
II. nominal infinitives

(2) subtype II.i

[Dat boek lezen] is een genot.
Litt: that book read is a pleasure

(3) subtype II.ii

[Het lezen van dat boek] is een genot.
Litt: ‘the read of that book is a pleasure

Type II.i. has a null subject and ‘verbal’ word order. Type II.ii. has a definite article/demonstrative with a null subject or a possessive DP, and a mixed nominal-verbal pattern (a verbal subshell within a nominal shell, Schoorlemmer 2001; Reuland 2011).

Claim: Analyzing the contrast as a distinction between a nominalizing and a verbalizing affix is unwarranted: a. Problematic for Frisian (contra Reuland 2003). b. Standard definitions of ‘nominal’ don’t apply to nominalizations (Baker 2003; Reuland 2011); b. Deriving (3) would require unmotivated high merger of nominal –en followed by lowering.

Alternative:

i. –en is merged to the ∖ and uniformly marks eventuality (neither verbality nor nominality).

ii. Projections of –en (up to VoiceP) have sufficient structure to qualify for T-licensing and case licensing (visible in Frisian).

iii. They can be merged as complements to auxiliaries, see (1), corresponding to their traditional construal as ‘verbal’ infinitives.

iv. They can also be merged as arguments (e.g. the subject in (2)), corresponding to their traditional construal as ‘nominal’ infinitives.

Hence, ‘nominal’ characteristics may be absent as in (2). The cut-off point in ii. accounts for the absence of higher adverbs.

Analysis of (3):

- The head of VoiceP bears ϕ-features (Wurmbrand and Shimamura 2017).
- A determiner or Poss-DP can be merged, and bind the ϕ-features in VoiceP. Hence, these cannot license the object, which then requires a preposition. Thus, the mixed nominal/verbal characteristics follow, without stipulating nominality of –en.

2. Infinitives can be additionally marked by te ‘to’. Te maps an eventuality expression onto an eventuality expression (with interpretive effects), retaining the licensing options contributed by -en. Hence, te-infinitivals qualify as arguments for prepositions, as in adverbial modifiers (Reuland 1979):

(4) Door veel boeken te lezen leerde Cindy Engels.

‘by reading many books Cindy learned English’

3. English –ing (see e.g. Abney 1987, Lundquest 2009). My proposal: –ing contributes the necessary structure for both case licensing (allowing an analysis along the lines of 1. for nominal –ing constructions), and T-licensing (progressive-ing). Bare verbs lack the structure for case licensing. Infinitival to does not contribute such structure. Hence to-infinitives cannot be headed by a preposition, ruling out the English equivalent of (4) but requiring an –ing-form as in (4)’s translation. Using a diagnostic described in Reuland (1979) I show that participial –ing clauses instantiate underspecification at PF.

References
Keywords: non-finite verb forms, syntax, morphology, typology, Turkic languages

This paper demonstrates that the commonly assumed multifunctionality of Turkic non-finite verb (NFV) forms does not hold in any of a range of Turkic languages surveyed. Evidence is also provided of other types of multi-functionality not previously described in the literature.

Descriptive grammars and linguistic studies of Turkic languages typically present two or three categories of NFV: participle (причастие in Russian-language sources), converb (деепричастие), and occasionally infinitive (Şağdan uulu and Batmanov 1933, Дыренкова 1941, Убрятова et al. 1982, Anderson 1998, Anderson and Harrison 1999, Erdal 2004). Forms called participles are said to have both adnominal (1b) and substantive functions (1a).

(1) “Participles” in Kazakh
a. as verbal noun
Маған китапханага барғанunday.
I. DAT library-DAT go-VN be.liked-NPST-3
‘I like going to the library.’

b. as verbal adjective
Маған кітапханаға барған кісі ұнайды.
I. DAT library-DAT go-VADJ person be.liked-NPST-3
‘I like the/a person who went to the library.’

Similarly, the term converb (Johanson 1995) covers the functions of verbal adverbs (head a verb phrase that is adverbially subordinate to another verb phrase, 2b) and infinitives (co-occur with auxiliary verbs, 2a).

(2) “Converbs” in Kyrgyz
a. as infinitive
Досум китепти китепканага тапшыраны турат.
friend-POSS.1SG book-ACC library-DAT turn.in-VADV stand-NPST-3
‘My friend is getting ready to return the book to the library.’

b. as verbal adverb
Досум китепти китепканага тапшыраны кезекте турат.
friend-POSS.1SG book-ACC library-DAT turn.in-VADV line-LOC stand-NPST-3
‘My friend is standing in line in order to return the book to the library.’

Examining corpora and in consultation with native speakers of a representative set of Turkic languages (Kazakh, Kyrgyz, Sakha, Tatar, Turkish, Tuvan), we categorise a near exhaustive list of NFV-forming suffixes in each language (averaging >11 suffixes/language) by function (verbal noun, verbal adjective, infinitive, verbal adverb) based on syntactic and morphological tests. We find that it is common for forms typically labelled as participles and converbs to exhibit only one of the respective two traditionally assumed functions, and all six languages examined have at least one converb-forming and one participle-forming suffix each with only one of the four functions.

Additionally, we identify several affixes that appear to cross the traditional participle/converb dichotomy. Tuvan, -/GIde/ and -/GIj/ each have one function from each of the two traditional categories. Furthermore, a form traditionally referred to as an infinitive in Tatar (historically a verbal noun in the dative case, with cognates in other Turkic languages) may span all four functional categories considered (3). (Other forms traditionally referred to as infinitives tend to be verbal nouns used in nominative or accusative case.)

We additionally discuss forms whose functions are difficult to conclusively identify, and touch on diachronic issues of how forms appear to have gained and lost functions.

(3) “Infinitives” in Tatar
a. as verbal noun
Миңа китапханәгә барырга кирәк иде.
I. DAT library-DAT go-VN necessary be-PST.3
‘I needed to go to the library.’

b. as verbal adjective
Китапханәгә барырга ынолыым житмәде.
library-DAT go-VADV nerve-POSS.1SG be.sufficient-NEG-PST.3
‘I didn’t have the nerve to go to the library.’

c. as verbal adverb
‘Time is needed in order to go to the library.’

References
Дыренкова, Н.П.(1941). Грамматика шорского языка. Москва: Академия наук СССР.
Neglected syntactic functions and non-syntactic functions of applicative morphology

Sara Pacchiarotti & Fernando Zúñiga
(Ghent University & University of Bern)

Keywords: Applicatives, Applicative-like morphology, Focalization, Registration, Semantics, Valence neutrality, Voice

Applicatives appear to be very common cross-linguistically, especially in agglutinative languages with rich verbal morphology (see Peterson 1999, 2007 for a survey of 50 languages with applicative constructions and their features). Most definitions of applicative constructions (e.g. Alsina and McChombo 1993; Bresnan and Moshi 1993; Payne 1997, 2002; Peterson 2007; Mithun 2002; Haspelmuth and Müller-Bardey 2004; Kulikov 2011; Creissels 2016) coincide in claiming that there are at least four fundamental attributes of applicative morphemes: (i) they are verbal derivational processes with syntactic consequences; (ii) they introduce an internal argument (i.e. object argument) to the argument structure of the underived verb root/stem; and (iii) there are multiple typically “peripheral” semantic roles that can be mapped onto the morphosyntactic entity introduced by the applicative (e.g. Beneficiary, Instrument, Possessor, Location, etc.). Several definitions (Payne 1997; Dixon and Aikhenvald 2000; Peterson 2007) also assume that applicative morphemes are optional (i.e. verb roots/stems without an applicative can appear in a construction with an oblique/adjunct, but this peripheral participant can alternatively be expressed as a core argument if the verb root/stem combines with an applicative). These features have become well known and are often assumed to be “canonical” or “prototypical”. However, if one looks across language families, it is unsurprisingly unclear what should be the essential, defining criteria to include a given construction in the category “applicative”, and why.

For instance, in languages where applicative morphology is optional, it can co-occur with the oblique constituent which should have been promoted by applicative derivation (e.g. Taba, Austronesian, Bowden 2001; Indonesian, Austronesian, Chung 1976: 55, Musgrave 2001: 156; Australian Aboriginal languages, Austin 2005; Abaza, Northwest Caucasian, O’Herin 2001). The functions of these constructions are not well understood at the present time, but they could be discourse-related.

By contrast, in many Bantu and Niger-Congo languages (Creissels et al. 2008: 109; Creissels 2010: 30), as well as in many Northwest Caucasian languages (Peter Arkadiev, p.c.) and Iroquoian languages (Mithun 2002), applicative constructions are obligatory. They are the only way to express a given participant in a single clause — such as General Location, where the event described by the verb takes place.

Applicative derivation is not always a valence-increasing operation. For instance, Bantu applicative constructions where the initial direct object loses some object properties and the applied object gains all object properties after applicative derivation, are extremely problematic for any claims related to syntactic valence. Further problems with labeling applicative derivation as a valence-increasing operation are found in Salish redirective applicatives, where the applicative combines with a transitive verb root and the resulting derived stem is still transitive (Kiyosawa and Gerdts 2010).
In some languages and with certain types of verbs, applicative morphology might have purely semantic (e.g. aspectual-related) effects with no syntactic consequences whatsoever for the clause (Maa, Nilotic, Lamoureux 2004; preverberation in Germanic, Slavic, Baltic, Indo-Aryan, Kulikov 2012, Kozhanov 2016; Tabu, Austronesian, Bowden 2001; Bantu, Trithart 1983).

Even though applicative morphology is often characterized as introducing a “semantically peripheral” participant, i.e. a non-Theme (and non-Agent), the (obligatory) use of applicative morphology to bring in core (i.e. non-peripheral) semantic participants such as Themes appears to be common in Australian Aboriginal languages (Austin 2005), Papuan (Onishi 2000: 131), Austronesian (Bowden 2001), Eskimo-Aleut (Mithun 2000: 108), and Mayan (see Lehmann and Verhoeven 2006 who call this construction in Yucatec Maya “extraversion”).

Finally, applicative morphology has been reported to have a focalizing function in Otomí and Mesoamerican languages in general (Hernández-Green 2016), including Eastern Mayan and Mixe languages (Hernández-Green 2016: 357; Mora-Marín 2003), Mojeño Trinitario (Arawak, Rose forthcoming), Bantu (Kimenyi 1980; Trithart 1983) and Atlantic (Voisin-Nougier 2002). This function is particularly noteworthy given that applicative constructions often function as topic-continuity/worthyness devices (Peterson 2007).

In this workshop, we do not take any feature of applicative morphology to be defining, canonical, or prototypical. Instead, our definition of APPLICATIVE morpheme is any derivational morphology occurring on a verb root/stem that has amongst its functions the introduction of a non-Actor semantic argument into a main clause. This non-Actor is usually mapped onto an APPLIED PHRASE. The latter term, coined by Denis Creissels, refers to any morphosyntactic entity introduced and/or semantically/pragmatically manipulated by the applicative without any specifications about its syntactic category and argumenthood status.

The goal of this workshop is to focus entirely on seldom-described functions of applicative morphology that appear to be nevertheless widespread in geographically distant and genealogically unrelated language families. These include:

**Syntax-related:**
- Applicative morphology is the only (or one of) the morphosyntactic means to introduce a General Location (e.g. ‘in the house’, ‘at the sea’, etc.) into a main clause (as in Northwest Caucasian and Bantu).
- Applicative morphology introduces an applied phrase that is syntactically an adjunct/oblique or “registers” the occurrence of the adjunct/oblique in the clause (as in Mayan, Oto-Manguean, Bantu, and Austronesian).
- Applicative morphology is valence-increasing with some verbs but valence-neutral or even valence-decreasing with others (as in Mapudungun and Tswana), i.e., applicative morphology has become fossilized and has lost its original functions.

**Non-syntax-related:**
- Applicative morphology can be used to narrow-focus a syntactic non-core argument like a Location, an Instrument, etc. (as in Bantu, Mayan, Otomanguean and Arawak).
- Applicative morphology does not introduce an applied phrase. Instead, it can add meaning nuances to the meaning of the verb root/stem, such as completeness, repetition, thoroughness, excess, “in vain”, etc. (as in Bantu, Austronesian, and some branches of Indo-European).
- Applicative morphology does not introduce an applied phrase. Instead, it can change the meaning of the whole event described by the construction in which applicative morphology appears (e.g. The man
will jump (over) the rock vs. The man will jump-APPL on top of the rock, as in Maasai, Nilotic; see Lamoureaux 2004).

Other possible topics:
– the semantic and/or pragmatic features/usage contexts of optional applicative constructions (see above) vs. the counterpart construction in which a verb root without an applicative combines with an oblique
– the occurrence of applicative morphology on the verb in subordinate clauses (as in Bantu)
– the occurrence of applicative morphology on the verb in wh-questions (as in Bantu)
– the genesis and evolution of applicative morphology as a way to explain the (dis)appearance of their non-canonical functions

Some contributions within this workshop point to the fact that directionals and other spatial-related morphology can have applicative functions (see also Rose forthcoming, where applicative-like functions are carried out by classifiers). More generally, not only do the contributions below address these less described functions of applicative morphology, but they also point to several other non-described functions which have parallels in other language families.

References
Barðdal, Johanna; Smirnova, Elena; Sommerer, Lotte & Gildea, Spike (eds.). 2015. Diachronic Construction Grammar. Amsterdam: John Benjamins.


Lopis-Sylla, Jeanne. 2010. La langue noon. Dakar: UCAD-IFAN.


Maho, Jouni 2009. NUGL Online. The online version of the New Updated Guthrie List, a referential classification of the Bantu languages.


https://doi.org/10.1515/lingvan-2016-0078


Segerer, Guillaume. 2002. La langue bijogo de Bubaque (Guinée Bissau). Louvain: Peeters.


Wane, Mohamadou Hamine. 2017. La grammaire du noon. LOT Publications.


\[ 2 \times 2 = ? \]

**Applicatives and transitivizers as valence-increasing or valence-maintaining devices in Salish**

Valeria Generalova

(Heinrich Heine University of Dusseldorf)

Keywords: applicatives, valence-increasers, valence-maintainers, semantic and syntactic valence, macroroles.

Salish languages are known for their rich system of valence-increasing suffixes. Namely, almost all of them have at least one suffix called “transitive”, at least one suffix called “causative” and at least two suffixes called “applicative”. In particular, suffixes usually working as valence-increasers sometimes demonstrate the behavior of valence-maintainers (see Kiyosawa 2006, Chapter 5.2 and p. 339). This problem is central to our work.

We find applicative suffixes’ ability to add or not to add a core argument to the construction peculiar. We would also like to investigate the interrelation of these non-prototypical applicative constructions with other valence-increasing constructions in Salish.

In our talk, we abandon the distinction between “relational” and “redirective” applicatives made by Kiyosawa 2006. We focus on constructions that comprise a transitive verb form with some other morphemes from the palette of valence-increasers that can take as direct object a non-core participant of a respective transitive base without further derivation. This approach is inspired by the view of Creissels 2004, p. 3. and helps delimiting causative and ditransitive constructions from applicatives. In this sense, Salish applicative suffixes are clearly different from other Salish valence-increasers as only they can allow a verb to take a recipient as a direct object and a theme as an oblique (consider ex. 29 in D. Gerdts 2010, p. 9).

To examine canonical (valence-increasing) and non-canonical (valence-maintaining) behavior of Salish applicatives, we consider four options. The default option would be to add an argument using an applicative suffix. Three other options are also hypothetically possible: (i) to add an argument
(represented as a non-core structure in a non-derived construction) using a non-applicative suffix, (ii) not to add an argument using an applicative suffix, (iii) not to add an argument using a non-applicative suffix. The third option being out of our interest, for now, the first two are a challenge for the existing analysis of Salish applicatives presented in Kiyosawa and D. Gerdts 2010.

We built our analysis upon the idea of differentiating between S-transitivity (number of direct core arguments a predicate takes) and M-transitivity (number of macroroles), suggested and explained by Van Valin 2005. Although specific for Role and Reference Grammar (RRG), this approach is in line with a well-known distinction between the semantic and the syntactic valence of a predicate. According to Van Valin 2005, p. 64, “from an RRG perspective, the S-transitivity a verb takes is less indicative of its syntactic behavior in simple sentences than its M-transitivity”. Therefore, we find it useful to examine Salish applicatives in respect to both types of transitivity.

Our analysis demonstrates that Salish languages are sensitive to both S-transitivity and M-transitivity. When analyzing a construction, it appears crucial to distinguish between the semantic and the syntactic valence of a given predicate. The outcome of this procedure results in a finer-grained classification of Salish applicatives and better descriptions of each type. It also partly explains the effects of stacking valence-increasing morphemes. In our talk, we present this new scheme in detail.

References

Applicatives and prepositions in Bantu

Hannah Gibson, Lutz Marten, Maarten Mous, & Kristina Riedel
(University of Essex, SOAS University of London, University of Leiden, University of the Free State)

Keywords: Bantu languages, applicative constructions, preposition, micro-variation, information structure

Bantu applicatives are often described as typical examples of the construction, introducing an applied phrase into the clause which would otherwise be more peripheral or not be licensed at all (e.g. Katamba 2003). However, an increasing amount of scholarship has shown that this characterisation is too simple, and that in many cases applicative verbs do not alternate systematically with non-applicative verbs and do not change valency (e.g. Marten 2003, Creissels 2004, Voisin 2006, Cann and
In light of this more detailed understanding of Bantu applicatives, we explore the interaction of applicatives and prepositions. Bantu languages typically have comparatively small and poorly described preposition inventories, which interact with applicatives in different ways. We consider morphologically simple (and presumably old) prepositions such as Bemba na ‘with’ (1), transparently grammaticalized ones, e.g. Mbuun ñgígirá ‘for’ (2), and borrowed ones such as para in Makhuwa or for in Sotho (3).

(1) Mutálé a-léé-ípík(-íl)-a na supuni [Bemba]
Mutale SM1-PROG-cook(-APPL)-FV with 9.spoon
‘Mutale is cooking with a spoon’ (Marten and Kula 2014: 21)

(2) ó-á-mó-dzwillé ñgígirá n-dzim [Mbuun]
SM1-PRS-PROG-OM1-kill.APPL for 9-money
‘He kills her for the money.’ (Bostoen and Mundeke 2011: 192)

(3) a. Ke-rek-etse dikonopo selamba sa-ka [Sotho]
SM1SG-buy-APPL.PERF 10.button 7.jacket 7-POSS.1SG
b. Ke-rek-ile dikonopo for selamba sa-ka
SM1SG-buy-PERF 10.button for 7.jacket 7-POSS.1SG
‘I bought buttons for my jacket’ (Gibson and Riedel 2019)

In the Bemba example (1), the preposition is required while the applicative marking is optional. When both are used together, emphasis is placed on the instrument (the spoon, in this case). In Mbuun, applicative marking and the use of the preposition are alternative coding strategies. Like in Bemba, when both are used together, as in (2), emphasis is placed on the applied object. In Sotho either applicative (3a) or preposition (3b) can be used, but not both, and (as far as we know), there is no distinct information structure associated with either structure.

We show that the interaction of applicative marking and the use of prepositions in Bantu is related to information structure on the one hand (e.g. in Bemba and Mbuun) and to providing alternative syntactic coding strategies on the other hand – as in Mbuun and Sotho, but not in Bemba, where, with instruments at least, the use of the preposition is obligatory. It is interesting to note that with prepositions borrowed through language contact with European languages, in Sotho and Makhuwa, syntactic coding, rather than information structure plays a more important role.

Based on comparative evidence, we illustrate the syntactic, semantic and pragmatic properties of the two constructions and their implications for the typology of applicatives in Bantu and cross-linguistically.

References

362
Unusual applicative constructions in Yaqui

Lilián Guerrero
(IIFL-National Autonomous University of Mexico)

Keywords: Applicative, self-beneficiary, external possessor, non-promotional applied argument

Previous studies on Yaqui (Uto-Aztecan, Mexico) have shown that the morpheme -ria is an optional verbal affix with syntactic consequence: it introduces a new (animate) core argument and, the applied argument adopts object properties (Rude 1996, Guerrero and Van Valin 2004, Guerrero 2007, and Harley et al. 2009). These properties satisfy the expected attributes of applicative constructions (Peterson 2007, Zuñiga and Kittilä 2010). Nevertheless, new data reveal that some applied arguments do not behave as privileged syntactic arguments. This paper examines the properties of these atypical constructions.

In Yaqui, -ria can combine with stative, intransitive, and transitive verbs, but not with ditransitive verbs. When added to states, -ria functions as a causative morpheme since the new argument corresponds to the agent. When added to active intransitives, the applied argument adds a benefactive or malefactive depending on the verb meaning (1a-b). Most commonly, -ria combines with transitive verbs, and the applied argument can only bear the role of beneficiary (1c). In the non-derived version, this participant would be coded as an adjunct marked by the postposition betchi’ibo ‘for’.

(1) a. Empo ye’e-k Peo-ta-betchi’ibo
   2SG.NOM dance-PFV Peter-ACC-for
   ‘You danced for Peter.’
In Yaqui, accusative nominals are marked by -\textit{ta} except when they are plural (1b); some postpositions (including betchi’ibo) demand accusative nominals or oblique pronouns as a complement. Note that in (1c’), the patient and the applied beneficiary are both coded as accusative arguments. In the applicative construction, the beneficiary displaces the patient for syntactic functions. For instance, the passive subject would be ‘the house’ for (1c), but ‘you’ for (1c’).

Furthermore, two atypical applicative constructions have been identified in a larger corpus: in (2a), the agent and the beneficiary are coreferential; in (2b), the applied argument is a possessor. These constructions are unusual for three main reasons. First, there are two direct core arguments which are referentially less distinguishable. Second, the applied possessor can be inanimate. And third, the applied argument cannot function as the passive subject. Whereas (2a’) is completely ruled out, (2b’) corresponds to an impersonal clause.

\begin{align*}
\text{a'.} & \quad \text{\textit{Empo Peo-ta yi’-ria-k}} \\
& \quad \text{2SG.NOM Peter-ACC dance-APPL-PFV} \\
& \quad \text{‘You danced (for) Peter.’}
\end{align*}

\begin{align*}
\text{b.} & \quad \text{\textit{Jesus-Ø muuku-k yoemia-betchi’ibo}} \\
& \quad \text{Jesus-NOM die.SG-PFV people.PL-for} \\
& \quad \text{‘Jesus died for (his) people.’}
\end{align*}

\begin{align*}
\text{b'.} & \quad \text{\textit{Jesus-Ø yoemia muuk-ria-k}} \\
& \quad \text{Jesus-NOM people.PL die.SG-APPL-PFV} \\
& \quad \text{‘Jesus’s people died.’}
\end{align*}

\begin{align*}
\text{c.} & \quad \text{\textit{Kari-ta=ne jinu-k e-betchi’ibo}} \\
& \quad \text{house-ACC=1SG.NOM buy-PFV 2SG.OBL-for} \\
& \quad \text{‘I bought a house for you.’}
\end{align*}

\begin{align*}
\text{c'.} & \quad \text{\textit{Kari-ta=ne enchi jinu-ria-k}} \\
& \quad \text{house-ACC=1SG.NOM 2SG.ACC buy-APPL-PFV} \\
& \quad \text{‘I bought you a house.’}
\end{align*}

Therefore, the applicative morpheme allows a self-beneficiary and a possessor to be coded as a core argument, but it restricts their morphosyntactic privileges. This use of -\textit{ria} represents a puzzle in the Yaqui grammar since it is valence-increasing, but it does not fully promote the applied argument in particular constructions.
Neglected functions of the Bantu applicative in relation to Locations: new insights from Fwe (K402)

Hilde Gunnink
(University of Ghent)

Keywords: applicative, Bantu, information structure, new information focus, contrastive focus

This paper deals with two seldom described discourse functions of the reflexes of Proto-Bantu applicative *-id in the Bantu language Fwe (K402), spoken in Namibia and Zambia (Gunnink 2018): its use to signal narrow completive focus on a Location applied phrase (Trithart 1983), and its use to widen the “orientation” of a locative phrase from involving the object of a transitive verb root to involving also the subject of that transitive verb root (e.g. John saw Mary on the boat implies only Mary is for sure on the boat; John saw Mary-APPL on the boat implies both John and Mary are on the boat) (Trithart 1977, 1983, Hyman et al. 1980, Kimenyi 1995, Rugemalira 2004, Cann & Mabugu 2007, Pacchiarotti 2020). The Fwe data are special in that they reveal so-far unattested variations and interpretations in these two uses. With respect to the focus function, in Fwe the applicative can combine with a cleft construction to express new information focus, as in (1).

kwákátimá ndíβerék(ér)á
φ-kwa-Katima ndíβerek-(er)-á
COP-CL17-Katima S1SG-work-(APPL)-FV <REL>
‘It is in Katima that I work.’ (as an answer to: ‘Where do you work?’)

While the cleft construction alone can be used to express new information focus, contrastive focus can only be expressed by combining the cleft construction with the applicative.

With respect to the orientation function, the applicative can modify the scope of the Location applied phrase not only to include the subject of a transitive verb (as observed elsewhere) but also the speaker who utters a sentence, cf. (2) and (3).

βàkwèsí βàzánà ‘hándyè
βa-kwesi βa-zán-a (ha-N-dyé)
S3:2-PROG S3:2-play-FV CL16-CL9-outside
‘They are playing outside (said by someone who is inside).’

βàkwèsi βàzàniná ńhándjè

S3:2-PROG S3:2-play-APPL-FV CL16-CL9-outside

‘They are playing outside (said by someone who is outside).’

Acknowledgments

I would like to acknowledge funding from FWO for my postdoctoral research project (12P8419N). I am grateful to Sara Pacchiarotti for her valuable input on the analysis of Fwe applicative constructions.

References


Trithart, Mary Lee. 1977. Locatives. In Ernest Rugwa Byarushengo, Alessandro Duranti and Larry M. Hyman (eds), Haya grammatical structure (Southern California occasional papers in linguistics 6), Los Angeles: Linguistics Department, University of Southern California, 89–98.


Non-obligatoriness of an applied phrase in applicative constructions

Katarzyna Janic
(University of Leipzig)
Keywords: beneficiary, obligatoriness, optionality, applicative periphrastic construction, grammaticalization

Applicative constructions (ACs) have a distinct piece of verbal morphology, introducing a non-agent argument (mapped onto an applied phrase) into a main clause (Peterson 2007). In this study, we focus on the non-obligatoriness of this argument in applicative periphrastic constructions (APCs) (Creissels 2010). We observe that this pattern tends to occur in APCs, which developed dedicated benefactive morphology, based on self-benefactive (1b) and benefactive (1c) distinction. Languages, which recognize dedicated benefactive distinction, such as Indo-Iranian, Turkic, Sino-Tibetan, etc., use grammaticalized verbs typically meaning ‘take’ and ‘give’ to signal this distinction, as shown in (1b) and (1c) respectively.

(1) **Punjabi** (Indo-Iranian; Bhatia 1993: 119)
a. ó ne xat likhiaa
he ERG letter.SGM write.PST.SGM
‘He wrote the letter’.
b. ó ne xat likh littaa.
he ERG letter.SGM write take.PST.SGM
‘He wrote the letter (for his own benefit).’
c. ó ne xat likh dittaa.
he ERG letter.SGM write give.PST.SGM
‘He wrote a letter (for somebody).’

We tentatively hypothesize that in APCs, optionality of a beneficiary is not necessarily related to syntax. It may result from the semantics of the verbs ‘take’ and ‘give’ because even if they are grammaticalized into valency operator, they still maintain their lexical content that specifies, which participant benefits from the action described by the main lexical verb. Its presence is thus semantically implied and does not have to be obligatorily expressed.

Yet in other cases, the optionality of a beneficiary can be culturally based, because in some societies referring directly to the first or second person (with a beneficiary role) is avoided for politeness reasons. Meiteilon uses the honorific suffix -bi in a benefactive function to specify that the speaker has more power and greater ability to perform an action than the addressee and that the performed action is not for the benefit of the speaker but for the addressee, (2).

(2) **Meiteilon** (Sino-Tibetan; Betholia 2005: 71)
əy thɔbok-tu tɔw-bi-ge
I work-DET do-BEN-OPT
‘I will do the work (for you).’

Given that non-obligatoriness of a beneficiary is subject to different conditions, this study aims at capturing cross-linguistic variation and possible constraints on the realization of this argument. It is guided by three research: (i) What are defining properties of benefactive ACs, and APCs in particular, across different languages? (ii-iii) How do these properties determine the realization of a beneficiary and why? We will carry out the empirical investigation within Basic Linguistic Theory (Dixon 2012), based on (secondary) data issued from grammars.

Building on our study and observed cross-linguistic tendencies (Peterson 2007), we expect that in most languages the expression of a beneficiary in ACs is obligatory. We also expect that the obligatory-optional line can be drawn on the basis of the semantic role distinction. Our preliminary results, in particular from Tzotzil (Aissen 1983), where a recipient but not beneficiary must be obligatory coded in ACs, go in this direction.
Acknowledgments
This research was supported by the project ‘Grammatical Universals’, funded by the European Research Council under the European Union’s Horizon 2020 research and innovation program (grant agreement no 670985).

References

The Diachrony of Central Kurdish Applicative Markers
Shuan Karim
(The Ohio State University)

Keywords: Kurdish, Applicative, Diachronic, Morphology, Historical

Prevailing theories of the development of applicatives postulate their development from adpositions or serial verb constructions (Pacchiarotti, 2016). Mithun (2002) shows convincing examples of each. Reconstructing either scenario is more difficult in Bantu, because the applicative morphology may date back as far as Proto-Niger Congo (Voeltz (1977) and Trithart (1983)).

Soranî (Central Kurdish) sets itself apart from other Kurdish dialects in that it has developed applicative morphology. Comparison between Soranî and the more archaizing Kurmancî (Northern Kurdish) shows a clear trajectory from PP to applicative marker. This discovery is significant to the diachronic study of applicative markers because enough of the morpho-phonological material has survived to allow this analysis to act as proof of concept for the prepositional source of applicatives.

Only recently has Soranî been described as a language with applicative morphology (Karim and Salehi, 2020). Particles known as Absolute Prepositions (APs) attach to verbs and re-purpose the object agreement morphemes for a participant role. In Kurmancî, when simple prepositions combine with a 3sg.obl pronoun there is an idiosyncratic (cf. bI mIn ‘to me’ and pe ‘to her’ not bI we). Soranî has lost the oblique pronoun we. Therefore, forms like le, pe, and te are no longer separable into their constituent parts. Additionally, 3sg serves as a bridging context for the reanalysis as an applicative because Kurdish verbs agree with their direct object in the past tense which is zero-marked in 3sg (cf. (2) which has no overt morpheme corresponding to 3sg and (1) which is overtly marked). A verb without object agreement (3) was then reanalyzed as having 3sg agreement and an applicative marker (4) instead of a preposition and a pronoun.

(1) wer-im-girt-m
   PV-1SG-hold.PST-3PL
   “I took them.”

(2) wer-im-girt-Ø
PV-1SG-hold.PST-3SG
“I took it.”

(3) nan-im wer l-e grt
bread-1sg PV from-3sg hold.pst
“I took the bread from her.”

(4) nan-im wer le grt-∅
bread-1sg PV APL hold.pst-3sg
“I took the bread from her.”

Because of the shallow time depth, there is a high level of certainty that this analysis reflects the development of the applicatives in Sorani. This is not the case with Niger-Congo, where there is a debate over whether the applicative has its origins in prepositions or serial verbs (Pacchiarotti, 2016). Furthermore, there is an issue reconstructing the original semantic function of Bantu applicatives. One possibility (supported by Hyman (2007)) is that there were originally many more such markers that have merged in the modern languages. Therefore, it is unnecessary to reconstruct a single meaning for Proto-Niger-Congo applicatives. This type of analysis is mirrored by the Kurdish data which shows four APs in the Kurmanji (Thackston, 2002) while Sorani has kept only pe and le but with the semantic range of all four. The full story of Sorani’s development of applicatives shows how as prepositions are recruited into verbal morphology, they semantically merge.

References
Pacchiarotti, Sara (2016), Bantu applicative construction types involving *-id: Form, functions and diachrony, University of Oregon, PHD Dissertation.

Applicatives in Kogi (Chibchan)

Dominique Knuchel
(University of Bern)

Keywords: voice, non-canonical applicatives, location, Kogi, South American languages

369
This talk discusses applicative constructions found in Kogi, a Chibchan language spoken in Northern Colombia. The semantic functions that are targeted by applicativization include comitative, benefaction, malefaction and location. The discussion focuses on constructions formed with the applicative prefix $i$-, whose properties deviate from those of prototypical applicatives in some instances. These properties concern the realization of the promoted argument as direct object (Zúñiga & Kittilä 2019:48) and the optionality of applicativization (e.g. Payne 1997).

The semantic role targeted by $i$- may be, in general terms, be conceived of as a location. Firstly, it promotes a participant that is adversely affected by an event, i.e. a maleficiary, exemplified in (1). Note that the promoted argument is referenced by goal indexes, thus behaving like an indirect object rather than a direct object. The connection between malefactive and locative can be seen in the English translation of (1) and is also attested in other languages (e.g. Völlmin 2010).

(1) $José=hâ$ $hui$ $zing-i-tsin-gwâ$

Josê=ERG house 1PL.G-APPL-destroy.INTR-CAUS

‘José destroyed our house on us.’

With verbs of location or placement, the marker promotes noun phrases that denote the Ground (i.e. the entity onto which an object is placed or is located on). The marker is obligatory whenever a concrete Ground is expressed, as illustrated in (2).

(2) a. $aldimeta$ $sha$

bottle lie

‘The bottle is lying.’

b. $aldimeta$ $mesa=k$ $i-shâ$

bottle table=LOC APPL-loc

‘The bottle is lying on the table.’

The construction in (2b) can be considered atypical in that (i) the applicative marker is obligatory (i.e. there is no non-derived alternative to express (2b)) and (ii) the promoted argument obligatorily carries a locative marker, thus not being coded like an oblique constituent.

In (3), the $i$-applicative similarly promotes a location (‘bird’) where the event occurs. The construction differs from the one in (2b) in that the promoted argument does not carry oblique case marking and constitutes a core argument of the derived clause.

(3) $ma$ $nuba$ $galda$ $i-mun-gwa-ki$

2SG bird feather APPL-grow-CAUS-NEG.IPFV

‘You don’t let the feathers grow on the bird.’ (Preuss 1923:150)

The promoted argument can further refer to the goal (e.g. ‘shout at’) or the source of directed actions, as in (4), where the promoted argument is referenced by patient indexes (indexing direct objects).

(4) $è=hâ$ $shêtsukkwâ$ $na-i-kukâ$

DEM=ERG knife 1SG.P-APPL-take

‘He took the knife (that I was holding) from me.’

Finally, in some instances, the $i$-applicative can be associated with a location in a less straight-forward sense, typically observed with verbs of cutting or tearing as in (5).

(5) a. $zhizhu$ $muan$ $has$ $men-gwâ$
fish middle IDEO cut.INTR-CAUS
‘He cut the fish through the middle.’

b.  zhizhu sánkalda i-men-gwá
fish head APPL-cut.INTR-CAUS
‘He cut off the head of the fish.’

(5b) may be interpreted in the way that only a small part (i.e. a concrete location) is affected, rather than the entire referent. In this case, the promoted noun phrase denotes a part of the referent, rather than a location.

Abbreviations
1 = first person; 2 = second person; APPL = applicative; CAUS = causative; ERG = ergative; G = goal; DEM = demonstrative; IDEO = ideophone; INTR = intransitive; IPFV = imperfective; LOC = locative; NEG = negation; P = patient; PL = plural; SG = singular

References

The functions of applicative morphology in Shiwiar (Chicham, Ecuador)

Martin Kohlberger
(University of Saskatchewan)

Keywords: Applicatives, morphology, Amazonia, Chicham, typology

Applicative morphology is usually defined as valency-increasing morphology that promotes an otherwise peripheral participant into the role of a core argument. In some languages, however, morphemes that satisfy this definition also perform other roles, including non-valency-changing and verbalising functions. This presentation will describe such uses of applicative morphology in Shiwiar, an Amazonian language.

Shiwiar is a Chicham language spoken by about 1,200 people in eastern Ecuador and northern Peru. The analysis in this talk is based on a 30-hour audiovisual documentary corpus of naturalistic speech collected between 2011 and 2016.

Shiwiar verbs are specified for a particular number of arguments (i.e. there are no labile verbs), and any changes in valency are overtly marked by morphology. Among different types of valency-changing morphology, Shiwiar has an applicative morpheme.

In broad terms, the applicative functions as expected: if a peripheral participant is promoted to a core argument, the verb is marked by the applicative. Usually the promoted argument is a
beneficiary/maleficiary or a location. Compare example (1), where the verb is intransitive, to example (2), where the verb is transitive and the experiencer has been promoted to the role of an object.

(1) muukur nahámawaii
muuk-iɣ nahám-a-wa-i
head-1SG.P hurt-IPFV-3.S-DECL
‘My head hurts.’

(2) muukur nahambrútawai
naham-rú-t-a-wa-i
hurt-1SG.O-APPL-IPFV-3.S-DECL.
‘It hurts me.’

However, in naturally-produced speech, not all verbs that appear with an applicative have increased valency. In some cases, the addition of the applicative morpheme simply results in a change in semantics. Compare examples (3) and (4), both of which exhibit intransitive verbs, but in the latter the attention is shifted to a notional (yet unexpressed) location.

(3) táahai
táN-a-ha-i
arrive-IPFV-1SG.S-DECL
‘I’m arriving.’

(4) tandáhai
táN-t-a-ha-i
arrive-APPL-IPFV-1SG.S-DECL.
‘I’m arriving here.’

Furthermore, the applicative morpheme in Shiwiar can also be used as a verbaliser, as in examples (5) and (6).

(5) nasindwai
nasi-t-wa-i
wind-APPL-3.S-DECL
‘It’s windy.’

(6) ñisándwai
ñisaN-t-wa-i
sun-APPL-3.S-DECL
‘It’s sunny.’

These two last uses of the Shiwiar applicative do not correspond to the usual valency-changing function that is typically attributed to applicative morphology. For this reason, the Shiwiar applicative might be described as non-prototypical from a cross-linguistic perspective. Nevertheless, there are indications that many more languages of the northwestern Amazonian region exhibit similar uses of applicative morphology. I will show that the Shiwiar data resembles constructions in neighbouring Kawapanan, Záparoan and Peba-Yaguan languages. In light of the areal similarities, I will also discuss the possible contact-driven historical development of this particular type of applicative morphology in the region.

To conclude, I will suggest that non-valency changing functions of applicative morphology may be more widespread than has been previously acknowledged. I will also argue that exploring naturally-produced speech might help uncover non-prototypical uses of applicative morphology that have so far been neglected in the literature.

The argument-adjunct continuum and the diversity of Circassian applicatives

Yury Lander & Irina Bagirokova

<not updated>
Applicatives and beyond

Marianne Mithun
(University of California, Santa Barbara)

Keywords: applicatives, argument structure, Chumash, nominalization, secundative patterning

Prototypical characteristics of applicative constructions are well known, but they can vary not only in their productivity, but also their syntactic effects and additional functions. Some such differences are illustrated here with examples from Barbareño Chumash, indigenous to California, recorded by J. P. Harrington from speaker Mary Yee.

Basic Barbareño clause structure is predicate initial. Verbs can stand alone as complete sentences. They contain subject pronominal prefixes and object suffixes, but third person singular objects are unmarked. There is no case marking on nouns: all dependents are preceded by a dependent (DEP) proclitic hi=, whether subjects, objects, times, locations, etc., or even clauses.

There are instrumental, locative, and dative/benefactive applicative suffixes. All add an argument to the core, which may be further identified by a determiner phrase. Since third person objects are unmarked in the verb, the only indication of the added argument may be the applicative suffix, as in (1).

(1) 3-INDEF-smear-INS DEP=REM=2-NMLZ-have.poison.oak-LOC.APPL
‘One smears the poison-oaked place with (it).’
JPH 59.657

The instrumental applicative in (1) serves to promote a topical participant, angelica leaves, to core status. The angelica was introduced in the previous sentence: ‘The angelica leaf is a good remedy for poison oak.’. The instrumental and locative applicatives are not used if an instrument or location is considered incidental.

(2) 2-CAUS-through.air-descend DEP=DISPAL=ART=dishrag
‘If you drop a dish-rag on the floor, . . .’

Applied objects, like the ‘lime’ in (3), generally precede any other nominals.

(3) 3-INDEF-CAUS-be.soft-TR-INS DEP= ART=lime DEP=DISTAL=ART=corn
‘They soften the corn with lime.’

The applicative suffixes are derivational. Formations are lexicalized: some have developed specialized meanings (aq-kum-pi quickly-arrive-LOC.APPL ‘blame’), and some no longer have simple root counterparts. But the three applicatives differ in further details.

Dative/benefactive applicatives are pervasive and intimately tied into argument structure. First and second person suffixes show secudative patterning, identifying patients, recipients, and beneficiaries: s-kutiy-it ‘s/he sees me’, s-ikš-it ‘s/he gives it to me’; s-kutiy-in ‘s/he sees you’, s-ikš-in ‘s/he gives it to you’. But third persons show indirective patterning: recipients and beneficiaries require the applicative -us:, s-kuti ‘s/he sees (him/her)’, s-kuti-wun ‘s/he sees them’, but s-ikš-us ‘s/he gives it to (him/her)’, s-ikš-us-wun ‘s/he gives it to them’.

373
Locative applicatives have an additional function not shared by the others: they can serve as nominalizers to form terms for places, like ‘the poison oaked place’ in (1). Furthermore, modification of referents can be accomplished by appositive nominals in the language. This means that locative applicative nominalizers also function much like adjectives or relative clauses.

(4) ‘[The piñones do not grow at Los Piños],
\[k'\ a \ ho=\ l=iy-yux-yuxwowon-pi \ hu=\ l=nip-nipolomol?\]
rather DIS=ART=PL-RDP-be.high-LOC REM=ART=RDP.-mountain
‘but rather in the high mountains.’
JPH 59.666

What appear at first to be simply parallel derivational suffixes have thus developed in divergent directions, one now playing a central role in alignment patterns, and another forming nominal, adjectival, and relative constructions.

The applicative and the expression of location in Rundi (Bantu JD62)

Manoah-Joël Misago, Koen Bostoen & Sara Pacchiarotti
(University of Burundi & Ghent University)

Keywords: Eastern Bantu (Niger-Congo), obligatory applicative constructions, non-syntactic uses of applicative morphology, information structure, diachrony

In this paper we elaborate on the corpus-based PhD research of Misago (2018) to provide evidence for two virtually undescribed non-syntactic functions of the reflex of Proto-Bantu applicative *-id in the Eastern Great Lakes Bantu language Rundi (JD62). First, we show that the applicative can signal narrow focus on a General Location applied phrase (cf. Kimenyi 1980, Trithart 1983), both completive (e.g. in answers to where questions) and replacive (e.g. X and not Y) in the terminology of Dik et al. (1981). This can be seen by comparing (1) and (2). Note that the focus function is possible only when the applicative is not required to syntactically introduce a General Location into a main clause with a given verb root, such as gend ‘walk’ in (1). All examples are from the native Rundi speaker Manoah-Joël Misago. Parentheses indicate a syntactically optional constituent.

(1) arikó aragendra mw’ishamba
    a-rikó       a-ra-Ø-geend-a       (mu i-shamba)
    S3:1-PRS.PROG S3:1-DJ-PRS-walk-FV CL18 CL5-forest
‘He is walking (in the forest).’ (as an answer to: What is he doing?)

(2) (oya), arikó agendera mw’ishamba
    (oya) a-rikó       a-Ø-geend-ir-a       mu i-shamba
    (no) S3:1-PRS.PROG S3:1.CJ-PRS-walk-APPL-FV CL18 CL5-forest
‘He is walking IN THE FOREST.’ (as an answer to: Where is he walking? or to: Is he walking in the mountains?)


(3) umwáan’ayuuzè igitabo mu māazi
The child has thrown the book (in the water).’ (the child is not in the water and throws the book into the water)

‘The child has thrown the book [while he was] in the water.’ (the child is in the water and throws the book somewhere else)

Second, as is common in Bantu and Niger-Congo more generally (Creissels et al. 2008), certain verb roots in Rundi require the applicative to combine with a phrase expressing General Location in a main clause. We present for the first time a list of such Rundi roots and compare it with the list which Rugemalira (1993) compiled for Nyambo (JE21), another Great Lakes Bantu language (Schoenbrun 1994). If overlaps exist between these two closely related languages, the obligatory use of the reflex of PB *-id to introduce General Locations was likely already present in their most recent common ancestor. The reconstruction of this use to a higher node in Eastern Bantu feeds into the hypothesis that cross-linguistic variation in Location-related functions of PB *-id suggests that these are older than other functions (Endemann 1876, Kähler-Meyer 1966, van Eeden 1956, Schadeberg 2003, Hyman 2007) and had more time to develop idiosyncratic behavior.

Abbreviations
APPL: applicative; AUG: augment; CJ: conjoint verb form; CL: noun class; DJ: disjoint verb form; FV: final vowel; PFV: perfective; PROG: progressive;PRS: present; PST: past; REC: recent; S3:1: third person subject of class 1.

References
Applicative Constructions in the Colombian Spanish from the Andes:
Canonical and non-canonical features

Sergio Ibáñez, Alejandra Ortiz & Armando Mora
(Universidad Nacional Autónoma de México, Instituto de Educación Media Superior & Universidad Autónoma Metropolitana)

Keywords: Spanish, applicative, benefactive, indirect object-dative, periphrasis

This work analyzes applicative periphrastic constructions (APC) of the Colombian Spanish from the Andes (1). In this construction DAR ’give’ functions as an auxiliary and as a grammaticalized form that allows the promotion of an adjunct to a core argument of the main predicate (the gerund verb), coded as dative or indirect object (IO). In (1) the promoted argument is a recipient benefactive, which is the most common applied participant in the languages with this type of structure, and this is why Haboud (1998) and Creissels (2010) label the construction as a benefactive applicative periphrasis.

(1) a. Le=da comprando unas=naranja-s a=la=Angelita.
3DAT=GIVE.APPL.PST.2SG buy.GDO ART.INDPL=orange-PL IO=ART=A.
‘(You) buy Angelita some oranges’

So APC are bi-verbal constructions with one verb functioning as the main predicate and the other as a kind of applicative marker (2), and in this sense they differ from the canonical applicative clauses with a morphological affix marker.

(2) (Yoruba-Rowlands, 1969:83) [Creissels 2010: 31]
Our goal is to properly characterize APC in the Colombian Andean Spanish variant and show the presence of some non-canonical and unexpected features. We show, for example, that the construction promotes participants beyond the typical animate recipient benefactive: in (3a) the applied argument is a deputative benefactive (Van Valin & LaPolla 1997). The interesting thing is that this is the most common promoted argument in our data, even it is not a ‘give’ type argument (not a recipient). In (3b), the ‘new’ dative is a non-animated goal; this is already a verbal argument (echar ‘throw’ is a three valency change of place predicate), so it is not being semantically applied (adjunct to argument), but still, it is syntactically promoted from oblique to IO. And in (3c), the applied source remains coded as an oblique; there is not semantic or syntactic promotion. It is a case of a purely pragmatic ‘promotion’, as the participant gain discursive prominence.

(3) a. Mi=abuelita sabe dar=me matando 1PRON.POSS=grandmother know.PRS.IMP.3SG give. APPL=1DAT kill.GDO las=gallina-s ART.PL=chicken-PL
   ‘My grandmother used to kill the chickens instead of me.’

b. Puedes dar=les echando agua can.PRS.IMP.2SG give. APPL=3PLDAT pour.GDO water a=las=flore-s, pobrecit-as
   1O=ART.PL=flower-PL poor-PL
   ‘You can pour water to the flowers, poor flowers.’

c. Juan le=dio quitando las=hoja-s Juan 3DAT=GIVE.APPL.PST.2SG remove.GDO ART.PL=leave-PL del=maíz
   PREP.ART=corn
   ‘Juan took the leaves from the corn.’

The DAR marker in (3a) and (3b) also shows, in addition to its semantic bleaching, a ‘syntactic bleaching’. In those examples, DAR appears in a periphrastic sequence in which another auxiliary verb functions as a hostage of the inflectional features, and DAR is coded as an infinitive form, while the main predicate still appears as a gerund. In this cases, DAR, only functions as a marker that licenses the promotion of the participants coded as datives, these contexts show that DAR in its in way to be fully grammaticalized as an applicative marker.

References
Haboud, Marleen (1998), Quichua y castellano en los Andes Ecuatorianos, los efectos de un contacto prolongado. Quito: Ediciones Abya-Yala.
The Applicative(-like) Functions of Nilotic Directionals

Doris Payne

<not updated>

Canonical and non-canonical applicatives in Waray

Thomas E. Payne & Voltaire Q. Oyzon
(SIL International; University of Oregon, Eugene, & Leyte Normal University, Tacloban City)

Keywords: Austronesian languages, Philippine languages, Waray, applicatives, valence

Waray is the mother tongue and language of wider communication for most inhabitants of Samar, Eastern Leyte, and parts of Biliran islands in the Eastern Visayas region of the Philippines. With over three million speakers, it is the sixth most widely spoken language in the country. In this paper, we document two applicative processes, both of which have canonical and non-canonical applicative functions.

First, we show that Waray verb morphology includes two productive stem-forming affixes, -an and i-, with canonical valence increasing applicative functions (see, e.g., Polinsky 2013, Peterson 2007, Dixon & Aikhenvald 2000). Next, we show that between them, these applicative affixes have at least five non-canonical functions: 1) An applicative affix is required in the basic transitive form of some verbs. 2) The i- prefix always indicates that there is a beneficiary or recipient in the scene being expressed, but without necessarily changing the basic case frame of the verb. 3) The -an applicative may indicate that the patient argument is only partially affected, again without changing the case frame of the verb. 4) There is an “adversative” use of the -an suffix, in which the absolutive argument is someone adversely affected by the situation described in the clause. Finally, 5) there is an “evaluative” usage of the -an suffix, e.g.

(1) Nahúhusáyan hi Pedro ha íya. ‘Pedro considers her beautiful’.

Non-canonical applicative functions similar to 1-4 are attested in the world’s languages (see, e.g., Pacchiarotti, 2017, Peterson, 2007). As far as we are aware, however, the evaluative function is not attested outside of the Philippines. Our conclusion is that the overarching function of the -an applicative is to “deflect” the transitivity of the situation to an alternative central participant. This general function encompasses all the usages of -an, including the evaluative function.

References
Adjunct Incorporation in Soran^{	extsuperscript{V}}

Ali Salehi
(Stonybrook University)

Keywords: Soran^{	extsuperscript{V}} Kurdish, incorporation, adjunct, absolute preposition, valence

This study examines incorporation in Soran^{	extsuperscript{V}} Kurdish (North-Western Iranian) where it appears that the phenomenon affects both internal arguments and adjuncts. Using cross-linguistic evidence from the literature (Mithun 1984), news corpora (Malmasi 2016) and native speaker intuitions, I show that both direct objects and obliques can be incorporated inside a VP in Soran^{	extsuperscript{V}}.

As shown elsewhere, Soran^{	extsuperscript{V}} features two kinds of prepositions: simple (SP) and absolute (AP) (Samvelian 2007; Thackston 2006). Unlike SPs which select independent NPs, APs are proclitics to the verb and their nominal complement is incorporated to the right of the verb. They also contrast with SPs in bearing stress, allowing them to exclusively select clitic pronominals with which they form a single word. Thus, the realization of the preposition (AP or SP) correlates with the clitic or non-clitic (non-a xal) status of the complement. For example, the absolute equivalent of simple ba min will be pe-m ‘to me’ and simple laawan will be absolute le-yan ‘from them’. The fact that absolute forms are obligatorily used with clitic arguments signals the morphological relation between the two. In addition, APs are said to add a specific semantic role interpretation to the adpositional complement (Opengin 2013). For instance, the AP l^e can introduce a participant that is obligatorily interpreted as malefactive. In fact, APs add a set of semantic roles | malefactive, source, addressee, recipient, causee, benefactive, in line with cross-linguistic evidence from other languages such as and Lakotah (de Reuse 1994) and Kinyarwanda (Givon 1984), where PP adjuncts are integrated as adjuncts through applicativization. In Soran^{	extsuperscript{V}}, both a verb’s direct and indirect object can be incorporated in the structure of the verb and past tense verbs get an agreement marker for the object (marked in bold). Compare (1-2) to (3-4).

(1) pirsya^r ek Pe-ka-n
   {\textit{question-indf ipfv-do-3pl}}
   ‘They ask a question.’

(2) pirsya^r ek=yan kird
   {\textit{question-indf=3pl do.pst}}
   ‘They asked a question.’

(3) pirsya^r ek=yan l(e)^eme kird

(4) pirsya^r ek=yan l^e=kird^n
question-indf=3pl sp 1pl do.pst
'They asked us a question.'

question-indf=3pl ap-do.pst-1pl
'They asked us a question.'

Given the syntactic and semantic constraints in Soran^, I argue that Soran^ is peculiar in allowing incorporation by only a certain class of prepositions. While incorporation is usually a valence decreasing operation, it is in this language, akin to applicativization when integrating adjuncts as arguments and triggering an additional agreement marker on the verb to signal the indirect prepositional object (4-6). That these PPs are adjuncts can be seen by the fact that they cannot be coordinated with arguments and relativization of those PPs is grammatical.

(5) mina l-ek-an goranʰ=yan wut bo ^eme
&
child-def-pl song=3pl say.pst for 1pl
'The children sang songs to us.'

(6) mina l-ek-an goranʰ=yan bo=wut-ʰn
&
child-def-pl song=3pl to=say.pst-1pl
'The children sang songs to us.'

References


Applicative Constructions and the Introduction of Attitude Holder in Tibetan

Camille Simon
(LACITO and INALCO, Paris)

Keywords: Tibetan, pragmatics, applicative, imperative, attitude holder

Classical Tibetan and the modern Tibetic languages possess an applicative voice used to introduce the semantic role of a Beneficiary marked in dative case as in (1).

(1) a. Amdo Tibetan (Picture description)

kaʰ ə ti ʰgam fgetsʰon-sək-kə ʰgo ʰje-kəko
big.dog DEM.ERG cupboard-INDF-GEN door open-IPFV.SENS
'This dog opens a cupboard.'

b. kaʰ ə ti ʰgam fgetsʰon-sək-kə ʰgo ʰje-rok â je-kəko
big.dog DEM.ERG person-INDF-DAT cupboard-GEN door open-APPL LIGHTV-IPFV.SENS
‘This dog opens a cupboard for someone.’

However, this valency-related derivation process is usually not recognised as such in the grammars describing Tibetic languages and dialects. The construction is indeed usually depicted as a “polite imperative” (Kretschmar 1986: 72; Schwieger 1989: 39; Tournadre 2009[1998]: 211; Haller 2000: 54, 94; Hoshi 2003: 35-36), a function illustrated in (2).

(2) Classical Tibetan, Milarepa’s biography (14th century)

\texttt{snambu 'di-la smadg.yogs-shig-kyang 'tshem-drogs mdzo-d-cig}

\texttt{serge DEM-DAT lower.garment INDIRECT APPL sew-APPL LIGHT V-IMP}

‘Please, sew also a lower garment in this piece of serge.’

A second, common use of the construction is found in negative utterances like (3).

(3) Lhasa / Common Tibetan (Elicited)

\texttt{pugu di pn-ro' ye-gimindu}

\texttt{child DEM lie.down APPL LIGHT V-IPFV.NEG.SENS}

‘This child doesn’t want to sleep. (Lit. ‘This child doesn’t sleep [even] for me.’)

In a first part, this applicative construction will be illustrated with data from several Tibetic languages, in order to provide evidences of the basic applicative function of the construction. In particular, it will be demonstrated that the construction is attested with diverse TAM markers, and not only in imperative or negative. Then, I will show that in polite commands and negative modal functions, the Beneficiary added to the verb frame is always formally omitted – any argument or adjunct can be formally omitted in the Tibetic languages, and there is no argument indexation on the predicate – but can be reconstructed as corresponding to the speaker herself. Hence, she gets involved into the described event at the pragmatic level, as an attitude holder or \textit{ethical dative} (Bosse et al. 2012). While, in Tibetan, dative-marked NPs can never assume the function of attitude holder, the applicative construction is used to express this pragmatic function more frequently than it is to introduce Beneficiaries.

\textbf{References:}

- Hoshi, Izumi (2003), \textit{A verb dictionary of the modern spoken Tibetan of Lhasa. Tibetan-Japanese}. Tokyo: Research Institute for the Languages and Cultures of Asia and Africa.
Semantic and syntactic functions of western Indonesian applicative morphology

Christina L. Truong & Bradley McDonnell
(University of Hawai‘i at Mānoa)

Keywords: applicative, morphsyntax, Indonesia, Austronesian

Many western Indonesian languages make use of a small number of verbal affixes which can introduce a non-A(gent) semantic argument into the clause, such as a location, instrument, beneficiary, goal, among others. Some constructions formed with these affixes exhibit features consistent with prototypical applicatives (Peterson 2007; Zúñiga & Kittilä 2019), such as the inclusion of the introduced non-A as a core argument, and the “demotion” of an A or P argument to oblique. However, the verbal affixes which form applicative constructions in these languages are consistently polyfunctional. They also form causatives and comparatives, derive verbs from non-verbal bases, change the scope, aspect, or intensity of action, increase affectedness of P arguments, and derive verbs with non-predictable, lexicalized meanings.

One widely-discussed example of this polyfunctionality is found in the Standard Indonesian suffix -i. It can act as a valency-increasing suffix, deriving transitive (or ditransitive) constructions that mark a goal or location, as in (1).

(1) a. 3SG sit  LOC chair  that ‘S/he sat on the chair.’
   b. 3SG AV.sit-APPL  chair  that ‘S/he was sitting on the chair.’ (Slightly modified from Arka et al. 2009)

It also derives transitive verbs from various bases as in kulit (n.) ‘skin’ → kulit-i ‘peel’, basah (adj.) ‘wet’ → basah-i ‘dampen’. Finally, -i in some cases does not increase transitivity but has aspectual effects as in (2).

(2) a. 3SG AV.hit  1SG ‘S/he hit me.’
   b. 3SG AV.hit-APPL  1SG ‘S/he was hitting me.’ (Slightly modified from Arka et al. 2009)

In this paper, we survey the functions of applicative affixes in eleven western Indonesian languages. We pay close attention to the varied syntactic properties and semantic nuances of apparent applicative constructions. For example, some constructions in which a non-A argument is introduced exhibit unexpected syntactic behavior, such as failure to increase valency, lower access to syntactic operations for the introduced argument, and incompatibility with particular voice markers. We further demonstrate that use of apparent applicative affixes for non-prototypical functions is both prevalent in individual languages and broadly distributed across languages of the region. We closely examine these constructions with respect to properties of compatible bases, semantic relationships encoded, degree of
productivity, and distribution of functions across the affix forms available. While there is considerable overlap between functions, forms, and morphosyntactic properties of these constructions across languages, variation among these present a puzzling and complex synchronic and diachronic picture.

Although a few affixes included in the survey have been discussed at length in the literature (e.g., Indonesian -kan and -i (Arka 1993, Arka et al. 2009), and Tukang Besi applicative suffixes (Donohue 2001, Peterson 2007)), on the whole, the polyfunctional nature of these affixes is not widely discussed nor well-accounted for. Western Indonesian applicative constructions pose a number of challenges to a general theory of applicatives due to their non-prototypical functions and large variation in semantic and syntactic behavior.

References

Spatial prefixes as applicatives in Harakmbut

An Van linden
(University of Liège and research fellow KU Leuven)

Keywords: spatial affixes, applicatives, valency changes, Harakmbut, Amazonian languages

This paper, based on first-hand data recorded in the field, focuses on a set of spatial verbal prefixes in the underdescribed language Harakmbut, more specifically the Arakmbut (Amarakaeri) dialect, spoken in the departamentos of Cuzco and Madre de Dios in Peru. Harakmbut is still considered as an unclassified (Amazonian) language (cf. Wise 1999: 307; WALS), although Adelaar (2000) has argued for a genetic link with the Brazilian Katukina family, which may be further linked to Macro-Jê.

The spatial prefixes studied here specify locative or directional circumstances of (participants in) the event denoted by the verb. Examples include ti- (1), which indicates location high up, on~n- (2), which signals the spatial relation of ‘in’, ‘to’ (Tripp 1976: 8) or ‘on’, and ok~k- (3), which expresses ‘separation’ (Tripp 1995: 219). While ti- in (1) is valence-neutral, n- in (2) and ok- in (3) increase the valence of the verb.

(1)  ken on-ti-pok mboerek-ta
    then 3PL-SPAT:UP-pass man-ACC
‘Then they pass the man (who is high up, on a ladder).’ (Pear story)

(2)  

\[ o\text{-}wedn\text{-}ato \quad \tilde{a}n\check{i} \quad \text{bisikleta} \quad o\text{-}n\text{-}kot \]

3SG-lie-MOVE&DO FILLER bicycle 3SG-SPAT:ON-fall

‘He falls (literally: ‘moves and lies down’), eh, he falls on his bike.’ (Pear story)

(3)  

\[ \text{wambo-ta} \quad i\text{-}mba\text{-}uk\text{-}i \quad tare? \quad \text{men-}\text{ok-}mber-e\text{-}ni\check{y}\text{-}ta \]

boy-ACC 1SG-VPL-search-1.IND manioc 3PL>1SG-SEPARATION-steal-REC.PST-REL-ACC

‘I am looking for the boys that stole manioc from me.’ (elicitation)

In (1), \textit{ti-} does not introduce an applied phrase, but specifies the location of the object participant (\textit{mboerek-ta}). In (2), \textit{n-} introduces an applied phrase for the intransitive verb root \textit{kot} ‘fall’ (\textit{bisikleta}), which is zero-marked, as typical for inanimate object participants of transitive verbs. In (3), finally, \textit{ok-} promotes the person from whom the boys stole manioc to a core argument registered on the verb (resulting in a portmanteau person prefix, indexing both agent and applied object participant), and thus turns a transitive verb into a three-place predicate (note that \textit{ok-}~\textit{k-} does not serve as a general malefactive applicative in Arakmbut). As in (2) and (3) the spatial prefixes introduce a non-agent argument into a main clause, they can be analysed as applicative morphemes (cf. Peterson 2007). In (1), however, the function of the spatial prefix is non-syntactic; it characterizes the object participant in terms of location, just like verbal classifiers – also present in the language – characterize object participants (or subject participants in the case of intransitive verbs) in terms of shape or substance.

This paper aims to determine the conditions for valence-neutral versus valence-changing behaviour of this as yet poorly understood category of spatial prefixes, as well as the range of syntactic and non-syntactic functions associated with them. In the process, it will investigate whether constructions like (2) and (3) are obligatory or optional applicative constructions (cf. Mithun 2002), and it will compare spatial prefixes with non-spatial applicatives (e.g. benefactive and comitative applicatives). Finally, attention will be paid to cases in which (combinations of) spatial prefixes seem to have become fossilized, e.g. (4).

(4)  

\[ e\text{-}ma\text{-}ti\text{-}on\text{-}ka \]

INF-VPL-SPAT:UP-SPAT:ON-do

‘hunt’

References


Tripp, Robert (1995), Diccionario amarakaeri-castellano, Yarinacocha: Min. de Educación & SIL.


\[ \]
The neglected syntactic functions of optional applicative morphology in Javanese

Jozina Vander Klok
(University of Oslo)

Keywords: applicatives, syntactic accessibility, morphology, Javanese, Austronesian

Introduction. Javanese (Austronesian; Indonesia) has two applicative suffixes which introduce distinct undergoers as an applied phrase: -i introduces goals/locatives, while -ake/-no has a wider association with benefactives, instrumentals, and themes (e.g., Suhandono 1994; Wedhawati et al. 2006). While their function of introducing the applied phrase as a core argument as in (1a) has been well studied (e.g., Hemmings 2013; Nurhayani 2014), the fact that the applied phrase can also remain syntactically as an adjunct/object with transitive predicates, as in (1b), has never. Drawing on data from elicitation-based fieldwork, this paper investigates this neglected syntactic function of optional applicative morphology in Javanese, focusing on its syntactic consequences.

Proposal. I show that this function has predictable syntactic consequences based on syntactic accessibility (cf. Keenan & Comrie 1977): the direct object (the non-applied phrase) is still eligible for extraction to subject position with applicative morphology, allowing for passivization, relativization, or object questioning (just like the applied phrase). For example, in (1a) where –no introduces a benefactive applied phrase (AP), the AP is a direct object (right-adjacent to the verb and cannot be introduced by a preposition), and can be passivized, (2a). Crucially, in both cases applicative morphology is obligatory as expected by canonical definitions of applicative constructions (e.g. Peterson 2007). Interestingly, if the AP can remain as an adjunct, as introduced by a preposition (1b), then we predict that that base object should still be able to undergo passivization with the applicative morphology. This is borne out, as shown in (2b). I also show that deletion does not show the same symmetry; rather, the theme argument seems to resist deletion (whether an applied phrase or not).

(1) a. Nunung nules*(-no) Duriati surat. ✓ S>V-APPL>NP2>NP1
   Nunung AV.write-APPL Duriati letter
   ‘Nunung wrote a letter for Duriati.’

   b. Nunung nules(-no) surat *(gawe) Duriati. ✓ S>V-APPL>NP1>Prep NP2
   Nunung AV.write-APPL letter use Duriati
   ‘Nunung wrote a letter for Duriati.’

(2) a. Duriati di-tules*(-no) surat *(karo) Nunung. ✓ [APPL PHRASE]
   Duriati PASS-write-APPL letter with Nunung
   ‘Duriati was written a letter by Nunung.’

   b. Surat iku di-tules(-no) Nunung *(gawe) Duriati. ✓ [BASE OBJ]
   letter DEM PASS-write-APPL Nunung use Duriati
   ‘That letter was written by Nunung for Duriati.’

Further consequences. The fact that this syntactic accessibility remains with applicative morphology seems to be redundant, as the base object (NP1) has this availability without applicative morphology as well, (2b). I suggest that this syntactic flexibility is due to the additional aspectual meanings of the applicative, such as indicating pluractionality with –i or implying perfectivity with –ake/-no (e.g., Hemmings 2013). To better understand for which types of predicates are these meaning nuances salient, I am currently compiling results across 175 predicates from Semarang Javanese.


References

“Subjective applicatives” in Mapudungun, Even, and beyond

Fernando Zúñiga
(University of Bern)

Keywords: Applicatives, Even, Mapudungun, passives, voice

A number of Western Austronesian constructions famously conflate functions that are usually kept separate in many other languages, viz. applicativization and promotion to subject. Some valency-changing operations in Mapudungun (unclassified, South America) and Even (Tungusic, Russia) show comparable features but have received much less attention in the literature. Based on the extant descriptive literature of selected languages, the present talk explores these phenomena and discusses their implications for our understanding of grammatical voice. Particular attention is given to our current knowledge of the variation displayed by applicative(-like) constructions, both within and across languages.

In what is arguably the default Mapudungun construction, the verbal suffix -(ñ)ma applicativizes transitive predicates —i.e., it turns a nonagentive extra-thematic participant into a core syntactic argument—, commonly (but not obligatorily) with an adversative connotation (1a-b) (data from field notes, based on Salas 2006: 119-120):

(1)  
a.  
*Kintu-ñma-fi-n*  
search-APPL-3P-1SG.IND mamüll.  
‘I looked for (fire)wood on him/her (e.g., in order to take it from him/her).’

b.  
*Leli-ñma-fi-n*  
look-APPL-3P-1SG.IND ñawe.  
‘I looked at his daughter on him (e.g., in order to somehow harm him by harming her).’

Non-default (ñ)ma-derivation in the language comes in two main guises. First, albeit not discussed in the present talk, some ñma-predicates do not introduce another syntactic argument to the clause, with or without concurrent semantic effects. Second, most relevantly for this talk, with some intransitive predicates (unergative (2a) represents the normal case within this predicate class; unaccusative (2b) represents the special case) and with most atransitive predicates (2c), the marker not only introduces an applied argument (2a-b-c) but also grants it subject status (2b-c); it works as an
applicative-cum-passive, so to speak. (The language does have a passive marker, which would be used with cases like (1a-b). It would be expected in (2b-c) but fails to appear.)

(2) a. *Lef*-ma-fi-n.  
run-APPL-3P-1SG.IND
‘I ran away from him/her.’ (own notes)
b. *Iñche aku-ñma-n*  
1SG arrive.here-ÑMA-1SG.IND *kiñe*  
*küme dangu*.  
‘I received a nice message.’ (Smeets 2008: 303)
c. *Mawüg-ma-n*.  
rain-ÑMA-1SG.IND
‘I got rained on.’ (Salas 2006: 125)

There is a similar phenomenon in Even, where the verbal suffix -w can simply passivize a transitive predicate like *maa-* ‘kill’ with the expected syntactic outcome (3a). Interestingly enough, non-prototypical constructions like (3b-c), where the subject is portrayed as being negatively affected, are also possible, even though the latter is not a semantic argument of the base predicate; it works as a passive-cum-applicative, so to speak (data from Malchukov 1993: 21-23):

(3) a. *Etiken nugde-du maa-w-ra-n.*  
old.man[NOM] bear-DAT kill-W-NFUT-3SG
‘The old man was killed by a bear.’
old.man[NOM] bear-DAT friend-REFL.POSS kill-W-NFUT-3SG
‘The bear killed the old man’s friend.’
c. *Etiken (imanra-du) imana-w-ra-n.*  
old.man[NOM] snow-DAT snow-W-NFUT-3SG
‘The old man is caught by the snowfall.’

References
Malchukov, Andrej. 1993. Adversative constructions in Even in relation to passive and permissive. In:  
Discourse phenomena in typological perspective

Alessandra Barotto & Simone Mattiola
(University of Bologna)

Keywords: linguistic typology, discourse phenomena, information structure, connectives, repetition

It is a well-established idea in functional-typological approaches that grammar is shaped by ‘discourse use’ (cf. Givón 1979, Du Bois 1985, Bybee and Beckner 2010, Diessel 2019). Generally speaking, discourse and grammar tend to be considered as two strictly intertwined domains (cf. Ariel 2009, Mithun 2015 among others, and the notion of emergent grammar in Hopper 1987), in the sense that discourse can be seen as the result of grammar usage in particular contexts, and, vice versa, grammar can be said to reflect discourse (“Usage feeds into the creation of grammar just as much as grammar determines the shape of usage” Bybee 2006: 730). This means that both phenomena that typically occur in oral speech as well as more grammatical phenomena (i.e. morphological and syntactic) should be of great interest for typology. However, typological studies have always focused more on grammatical phenomena, considering them as the natural field of investigation of typology. As a consequence, less attention has been payed to the typological investigation of discourse phenomena (see, as an example, the phenomena investigated in the WALS’s chapters).

With the term ‘discourse phenomena’, we mean linguistic elements and constructions that typically occur in oral speech and help to manage the organization, flow and outcome of communication (cf. Schiffrin 1987, Du Bois 2003). These phenomena are generally considered as the natural field of investigation of pragmatics, semantics, and discourse analysis. They range from connectives and information managing structures (e.g. topic shift strategies or introduction of new reference) to interjections and discourse markers (e.g. reformulation and exemplification markers). As already noted, indeed, systematic typological cross-linguistic investigations on discourse phenomena are relatively rare. Some exceptions are, for instance, Givón ed. (1983) and Lambrecht (1994) respectively on topic continuity and information structure in general; Brown and Levinson (1987) on politeness strategies; Ameka (1992) on interjections, Dingemanse (2012) on ideophones and Lahaussois and Treis (2019) on both interjections and ideophones; Mauri and Sansò (2018) on general extenders and non-exhaustive constructions; and Guérin ed. (2019) on bridging constructions. These preliminary works represent a pivotal starting point for a broader discussion, and, at the same time, they also confirm that typological investigation of this kind of phenomena is viable.

The rarity of systematic typological studies on these issues goes hand in hand with a scarcity of detailed accounts in descriptive grammars, in a circular loop that provides a critical challenge for the cross-linguistic investigation of this kind of phenomena. This empirical problem results in the need for mixed methodologies, providing evidence coming from different types of data, such as corpora, grammars, and behavioral experiments.

This workshop aims at investigating discourse phenomena in typological perspective, that is, adopting the methods and instruments of typology. Our purpose is to create a forum in order to discuss the theoretical relevance of discourse phenomena for typology and, vice-versa, how a typological approach can be useful to provide new insights on the study of these phenomena. At the same time, we also aim at giving an account of discourse phenomena that usually are not described in the literature in typologically interesting languages.
The workshop addresses the following topics:

i) cross-linguistic analyses, synchronically and diachronically oriented, on one or more discourse phenomena;

ii) methodological proposals on how to collect, examine and compare cross-linguistic data on discourse phenomena, as well as methods to extract data on this kind of phenomena from both grammatical descriptions and corpora of specific languages;

iii) language-specific (synchronic and/or diachronic) investigations that follow a typological approach and may raise issues that are relevant for the broad cross-linguistic picture;

iv) studies investigating the correlation between discourse phenomena and typological features;

v) studies addressing patterns of re-analysis or functional extension of grammatical elements towards discourse motivated functions (e.g. politeness, emphasis, etc.).

Possible phenomena to be investigated include:

- Discourse connectives
- Discourse markers
- Interjections and ideophones
- Repetitive patterns
- List constructions
- General extenders
- Intentional vagueness and non-exhaustivity
- Information structure management (e.g. topic shift, introduction of new reference, focus)
- Politeness strategies

References


On the topic-marking function of Left Dislocations and Preposings.
Variation across spoken and written Italian and English

Doriana Cimmino

Left Dislocations (La torta, l’ho mangiata/The cake, I ate it) and Preposings (A Maria, non ho parlato/To Maria, I didn’t speak) are generally regarded as performing an overall topic-marking discourse function (see Benincà et al. 1988, Berretta 2002, Ferrari et al. 2008, for Italian; Reinhardt 1981, Lambrecht 1994, 2001, for English). However, when confronted with real texts, this description fails to account for all the cases, as shown for both spoken and written texts (see, at least, Sornicola 1981, Berruto 1985, Cimmino/Panunzi 2017, for Italian; Prince 1998, Kreyer 2006, Tizón-Couto 2012, for English).

The main aim of this contribution is to challenge the topic-marking as the general discourse function of Left Marked Structures (LMSs), relying on a corpus-based and multilevel investigation. The study focuses both on spoken and written Italian and English texts; more specifically, 400 occurrences of LMSs, comparable within the two languages, have been collected selecting spoken and written texts representing a standard variety of language – namely, spontaneous conversations and daily newspapers articles. The multilevel analysis included LMSs’ syntactic characteristics, information properties, and discourse functions. As for the syntactic level, the presence/absence of resumptive pronouns, the subject’s position and its adjacency to the left marked constituent (Biber et al. 1999) have been described. Subsequently, the Topic/Comment articulation of LMSs has been analyzed following two theoretical frameworks for the segmentation of spoken and written texts: the Language into Act Theory (Cresti 2000) and the Basel Model (Ferrari 2014). Finally, the investigation of LMSs’ discourse functions included the observation of the interaction between their information structure and the topical and logico-semantic progressions (Ferrari/De Cesare 2009) of the texts of occurrence.

The results provide evidence that, in both written and spoken texts, Italian and English LMS’ syntactic properties do not systematically correlate with a topic-marking discourse function; in fact, LMSs do not always exhibit a Topic/Comment articulation, as shown in the following (1) Italian left
dislocation and (2) English preposing (the non-topical left marked constituents are highlighted in italics):

(1) *I mezzi* te li abbiamo fatti capire (familiar conversation, DB-IPIC, Panunzi/Gregori 2012)
(2) *Most discussed* is the future of same-sex marriage in the state (nytimes.com, ICOCP corpus, De Cesare et al. 2016)

Moreover, on a discourse level, LMSs’ functions include not only their interaction with the topical but also with the logico-semantic dimension of texts, and considerable variability between spoken and written texts can be observed.

In the talk, it will be shown that disentangling the syntactic and information features of LMSs not only provides a clearer understanding of the interaction between these levels of analysis but also allows for an accurate description of their discourse functions in spoken and written language. The discussion of Italian and English data is extremely relevant to the current research into cross-linguistic properties of LMSs, as their topic-marking function has recently been questioned for French (Pekarek-Doheler et al. 2015), Burmese (Ozerov 2015) and Hebrew (Ozerov, SLE 2019), but an alternative overall function is still to be identified.

References:
De Cesare Anna-Maria & Davide Garassino / Rocio Agar Marco / Ana Albm / Doriani Cimmino, *Sintassi marcata dell’italiano dell’uso medio in prospettiva contrastiva con il francese, lo spagnolo, il tedesco e l’inglese. Uno studio basato sulla scrittura dei quotidiani online*, Frankfurt am Main, Peter Lang.
Ferrari, Angela, De Cesare, Anna-Maria. 2009. La progressione tematica rivisitata. In Vox Romanica 68, 98-128.


Embedding wh in bilingual Turkish: A corpus-pragmatic study

Annette Herkenrath & Birsel Karakoç
(Adam Mickiewicz University, Poznań & Uppsala University)

Keywords: Turkish linguistics, language contact, bilingualism, subordination, wh, discourse analysis, corpus linguistics

This study looks at the subordination of wh-constructions in spoken discourse data of Turkish-German bilinguals (755,000 words or 213,000 utterances), completing qualitative findings with a quantitative approach. Typologically, Turkish, which marks syntactic subordination by means of nonfinite suffixes on the verb, does not employ wh-words as subordinators; wh-words can remain in situ (Herkenrath 2011). Previous studies on the same data (Herkenrath, Karakoç & Rehbein 2003, Herkenrath & Karakoç 2007, Rehbein, Herkenrath & Karakoç 2009) suggest an innovative bilingual use of wh-words, reminiscent of the contact language German: as subjunctors. In a data-driven take, we developed the following criteria for innovation: (1) finiteness of the embedded wh-construction, (2) position of the embedded wh-construction after the higher clause, (3) position of the wh-word at the beginning of the wh-construction. When occurring together, these phenomena can indicate a reanalysis of the wh-element as a subordinator.

The present study conducts a quantitative analysis, applying corpus-analytical methodology (Schmidt 2010, Herkenrath & Rehbein 2012) to a concordance of 14,000 wh-occurrences. The study takes into account the specific discourse structure of the data: conversations between children and adults in family settings with a number of interlocutors. This situation, which involves frequent turn
changes, makes for short utterances – posing challenges to several aspects of our method. While the finiteness of a *wh*-construction is easy to identify, what poses problems in spoken registers, in which utterance boundaries are subject to interpretation, is the relationship between a *wh*-construction and an adjacent matrix construction – which may or may not be a higher clause (Spronck & Nikita 2019 on syntactic transitions in reported speech). Moreover, in short utterances, the position of the *wh*-word within the *wh*-clause can be ambiguous. Thus, while individual criteria can easily be identified, we need to be able to look at patterns in which we can study all three criteria together. – Given our large empirical basis, the study will be able to present a clearer picture of the status of any innovative uses: where all three features occur together.

This overall situation confronts us with problems related to transitions between discourse and emerging grammar (Hopper 1987, Bybee 2006). *Wh* brings into the conversation elements of non-knowledge (Rehbein 1985) – as topics of potential interest that can be freely expanded on (Herkenrath 2011). With respect to this discourse function, they can be said to have a connective quality, independently of the typological differences. Typology comes into play at the level of smaller units (Mithun 2015), i.e. where the connectivity of *wh* is transferred to the utterance-internal level, when a knowledge deficit is presented as the object of perception or communication, or as attributed to some speaker (Larson 1978: 21ff). In Functional Pragmatics, which conceives of smallest grammatical units as actional units (thereby connecting morphosyntax to discourse, Ehlich & Rehbein 1979), the *wh*-element acquires a new ‘procedural quality’, meaning: a new actional feature.

References
Herkenrath, Annette; Karakoç, Birsel & Rehbein, Jochen (2003), Interrogative elements as subordinators in Turkish: Aspects of Turkish-German bilingual children’s language use, in N. Müller (ed), *In)Vulnerable Domains in Multilingualism*, Amsterdam: Benjamins (Hamburg Studies in Multilingualism 1), 221–270.
Marking beginning or end: Topic-shift conceptions in Ainu, Japanese, and English

Katsunobu Izutsu & Mitsuko Narita Izutsu
(Hokkaido University of Education & Fuji Women’s University)

Keywords: <topic shift; episode shift; end-orientation; beginning-orientation; aspectual marking>

This study presents a cross-linguistic analysis of topic-shift conceptions mainly based on Ainu-Japanese parallel texts (Kindaiichi 1931; WDGKK (ed.) 1984-1989; Kayano 1998) and their English translations. We argue that Ainu is a type of language with an end-oriented conception of topic shift in conversations (or episode shift in narratives), while Japanese and English prefer a beginning-oriented conception. A stretch of discourse usually consists of discourse segments, each of which centers around a topic. Topic shift can involve the “closing” and/or “opening” of such a discourse segment (Goutsos 1996), which are often marked with “digressive” discourse markers or “topic shifter” (Bordería and Arguedas 2009): e.g., English by the way, incidentally, Japanese sate ‘now,’ tokorode ‘by the way,’ etc. As these are also known as “new relevant information markers” (Bordería and Arguedas 2009), they typically signal the opening of a topic or an episode in many languages (e.g., Goutsos 1996; Tsai 2011).

Ainu, an indigenous language of Japan, does not only have discourse markers indicating the opening of a new topic (e.g., sineanta ‘once,’ orowano ‘and then’ etc.) but also has markers specialized for (at least literally) indicating the closing of a current topic or episode (e.g., pakno ne kor ‘(lit.) becoming until (t)here’ eepaki ta ‘(lit.) at its end,’ etc.), as illustrated in (1a). Furthermore, pakno (ka) can be used to mark the end of a narrative discourse or an interactional sequence, for example, when the speaker announces that s/he has finished a story, as in (1b), or declines an offer of some more food or drink, as in (1c). Ainu thus conceptualizes topic shift as the end as well as the beginning of discourse segments.

(1) a. pon menoko pirka suwe p a=e katu an=omonmomo, eepaki ta
   small girl good boil thing I=eat figure I=detail ‘s.end at
tane anakne iyoski utar iyoski nispa hosippa kotom an=esanniyo,
   now TOP be.drunk PL be.drunk men leave. PL like I=infer
pakno ne kor pon menoko a=esikari.... (Kayano 1998: 74)
   until become when small girl I=catch

b. pon menoko
   small girl

(continued)
‘I’ll skip (lit. detail) how I ate what the young girl cooked, and next I found the drunk men were now gone, that’s all, and I (went on and) caught the young girl and....’

b. *pakno (ka) ku=ye.* (Tamura 1996: 507)
until also I=say
‘That’s all about it.’ (lit. ‘I told (the story) to its end.’)

much be FP more eat FP until also now I=be.full
‘There is much of it. Eat more.’ ‘(That’s) enough, I’m full.’

One interesting correlation with this contrast in the topic-shift conceptions is aspectual marking. English and Japanese have more inceptive (start, begin, set about; hazimeru, dasu, kakeru, etc.) than terminative expressions (finish; oeru), while Ainu has a number of terminative expressions (a, wa okere, wa isam, wa an, etc.) with almost none of inceptive. We point out that these aspectual conceptions correlate with how topic shift (or episode shift) is typically conceptualized in each language.

References
Bordería, Salvador Pons and Maria Estellés Arguedas (2009), Expressing digression linguistically: Do digressive markers exist?, *Journal of Pragmatics* 41: 921-936.
Kindaichi, Kyosuke (1931), *Ainu Jojishi Yukara no Kenkyu, nibunsatsu* [A study of Ainu epic yukar, two volumes], Tokyo: Toyobunko.

Polish *że* ‘that’ and its discourse functions

Łukasz Jędrzejowski & Wojciech Guz
(University of Cologne & John Paul II Catholic University of Lublin)

Keywords: spoken Polish, discourse connective, elaboration, reformulation, complementizer

In contemporary standard Polish *że* ‘that’ is best known as a declarative complementizer. In colloquial speech it also serves lesser-known functions (e.g. relativizer and adverbializer). Further, spoken corpus data indicate additional discourse-semantic functions so far overlooked. Specifically, the aim of this paper is to address *że* introducing clauses of explanation/elaboration/reformulation, as exemplified in (1) (elaboration) and (2) (reformulation). Note the corresponding use of *like* in the English translations. The data come from the Spokes corpus of contemporary conversational Polish (http://spokes.clarin-pl.eu/).
While examples (1) and (2) still display some signs of syntactic subordination across the clauses, example (3) illustrates more clearly the departure of *że* from intra-sentential to inter-sentential level: the chunk constituting the elaboration is associated to the preceding material at the level of discourse by commenting on its proposition.

(3) **ludziom bardziej opłacało się wsadzić plastikowe drzewka na przykład na  ektarowym polu | co metr | że to jest oliwne drzewko**

‘It was more profitable for people to stick in plastic trees, for instance on a one-hectare field, every meter. **Like**, this is an olive tree’

We will argue that in such uses *że* is a discourse connective coding relations of explanation/elaboration/reformulation (similar to English *like*, French *genre* (Fleischman and Yaguello 2004) and Russian *tipa* (Kolyaseva 2018)). In contrast to the traditional view whereby a *że*-clause is a complement clause of a matrix predicate, this usage is characterized by lack of associated complement-taking predicates and loose clausal linkage. These *że*-clauses elaborate/comment on the proposition expressed in the preceding clause or a longer stretch of discourse.

To the best of our knowledge, this use of *że* has not been addressed in the literature. This is likely due to the fact that it is limited to colloquial speech only. However, it does not seem to be an innovation emerging contemporarily in speech since signs of the same use are recorded in 16th-century texts. The function thus seems to have survived in colloquial speech but has been banned from contemporary standard language, where *że* is ‘preferred’ as a complementizer. This points to a need for more empirically-grounded investigation of spoken phenomena in discourse studies, also with specific reference to complementizer semantics (cf. Hansen et al. 2016, where Polish *że* is said to be a prototypical complementizer with no semantic content). An intriguing question raised by this study and relevant to cross-linguistic typology is how common it is for complementizer forms to venture beyond sentence-level subordination and serve additional discourse functions; and conversely, how the relations of elaboration/explanation/reformulation are marked cross-linguistically (cf. Guérin’s 2019 recapitulative and summary linkage).

References


Differential indexing as a means of information structure management:

A typological overview

Erika Just
(University of Kiel)

Keywords: indexing, information structure, differential marking, topic, focus

This study provides a typological overview of differential indexing of A and P arguments as a means of marking deviations from information structural default configurations. The typological study based on reference grammars is complemented by corpus based case studies of the languages Ruuli (Atlantic-Congo), Maltese (Afro-Asiatic) and Gutob (Austroasiatic). It has been well known at least since Givón (1976) that indexing is a means of tracking the referent which is highest in topicality, be it the A or the P argument. Differential P indexing has been subject to investigation in both family specific (see e.g. Riedel 2009 on Bantu) as well as typological studies (e.g. Lemmolo 2011). There is a general consensus that more topic-worthy P arguments are often indexed in contrast to non-topic-worthy P arguments. This is not surprising given that, whereas the A argument is usually associated with topic features and a high position on the animacy and definiteness scales, the P argument is generally associated with the opposite, often introducing new information. In Efate, P arguments cannot be indexed on the verb if there is a co-occurring NP, exemplified in (1a). However, P indexing does occur alongside an overt NP whenever P is topicalized and the co-occurring noun phrase in left-dislocation (Thieberger 2006:115, 275), as in (1b).

Efate (Oceanic, Vanuatu, Thieberger 2006: 104, 136)

1a) Ruk=fo wat kineu
   3PL.RS=IRR hit 1SG
   ‘They will hit me.’

1b) Na-fi-asel-wen ni tiawi gakit,
   ART-be:IRR-friend-NMLZ of old.people 1PL.INCL.POSS
   tu=tae pitlak-e-n mes
   1PL.INCL.RS=be.able have-TR-3SG.OBJ today
   ‘And the friendship of the old people, we can still have that today.’

Similar to differential P indexing, differential A indexing, through the omission of an index, might likewise be induced by a shift away from a default information structural configuration, e.g. a shift of the A argument away from high topicality. This has been reported for some languages (cf. Lambrecht & Polinsky 1998, Siewierska 2004:159-162), but the phenomenon has not yet been subject to closer typological scrutiny.

An example comes from Jamsay (Heath 2008: 455-456) where the A argument is generally indexed, except when it is focused. This means that the lack of the index is an indicator of focus status of the referent. There is no indexing on the verbs in (2a) with subject focus, but there is an index in (2b) and (2c), where the subject is not focused.

Jamsay (Dogon, Mali, Heath 2008: 455)

2a) ùrn-ùm kù, bé=ŷ  kö gujôː, bé=ŷ kö kårà,
   child-PL DEF 3PL=FOC it defeather.IPFV 3PL=FOC it cut.IPFV
The children, it is they who defeather them (birds), they cut them up, they prepare them completely.'

2b) ŋē-m  bē=ŷ  ēː-m
    woman-PL 3PL-FOC see.PFV-1SG
    ‘The women, it’s them that I saw.’

2c)  yá:  ũ  ēː-sá-m.
    yesterday 2SG.OBJ see-RES-1SG
    ‘Yesterday, I saw you.’

References

Discourse Functions of Aspectual Clitics in Languages of Southeast Asia
Atsuko Kanda Utsumi

<not updated>

Adverb/postposition ‘after’ to sequential connective ‘then’ to discourse-pragmatic marker in Andi
Timur Maisak
(Institute of Linguistics RAS & HSE University)

Keywords: Nakh-Daghestanian, adverb, postposition, connective, discourse marker

It is often the case that discourse-pragmatic markers find their origin in linguistic elements with locative or temporal semantics. In Andi, a Nakh-Daghestanian language of the Avar-Andic branch (Russia), there exists a set of cognate words with very different functions ranging from locative adverb ‘after, behind’ to emphatic particle. In the present paper, I try to determine which elements should be
distinguished within this set, and which are the diagnostic properties of each member. In particular, I look at the semantics/function of respective words/markers, their class membership (e.g. whether they are adverbs, postpositions, or particles, etc.) and various properties of their morphosyntactic behaviour (such as position in the clause, ability to agree and phonetic weight).

Andi is a left-branching SOV language with postpositions; each postposition determines the case form taken by its nominal complement. One of these postpositions is b-ɛχuɖu, which has both the locative meaning ‘after, behind’ and the temporal meaning ‘after’. This postposition takes a complement in the cont-essive case in -ɛχu. There is gender agreement in Andi, and b-ɛχuɖu includes a prefixal gender slot. (In the citation form b-ɛχuɖu, I use the gender III/IV marker b-; in total, there are 5 genders in most Andi dialects). Agreement at the clause level is with the absolutive NP, e.g. in (1) w-ɔχuɖu agrees with murid-ol ‘disciples’ in the masculine. Like some other postpositions, b-ɛχuɖu can be also used adverbially without a nominal complement.

(1) ḡeγe-ʃ-ɛχu Ṣ-ωχuɖu murid-ol=m lo w-oli-du
   [DEM-OBL(M)-CONT M-after] disciple-PL(ABS)=ADD M-walk-PRF
   ‘and his disciples followed him, lit. went after him.’ (Luke 22:39)

Two more instantiations of (historically) the same item are the sequential connectives rexuɖu and rexu ‘then’, which are both invariant in form (without agreement, but probably with a “petrified” gender V marker r-). The longer one, rexuɖu, can be used sentence-initially, as in (2). The shorter one, rexu, is preferred in clause chaining, being encliticized to the right edge of a non-final clause (3). In the latter use, rexu does not seem to provide any additional semantics, but just reinforces the sequential meaning, i.e. “P=rexu Q” means ‘P, and then Q’, although in principle the order of predicates is enough to express precedence.

(2) rexuɖu isa-di=ɾɔdi rac′in (...) after Jesus-ERG=REP ask.AOR
   {The storm subsided, and all was calm.} ‘Then Jesus asked (...)’ (Luke 8:25)

(3) onšiɬo qabdol-duɖu ḡeγe-l-li Ṣʷaɾʷaɾči=rexu
   then [cut.CAUS-PRF DEM-OBL-GEN belly]=after
   reʃ-ɬo j-ilo-duɖu ḡiri racوارا... [outside-SUP.LAT F-come-PRF red hat]
   {The hunters killed the wolf.} ‘Then they cut its belly, and Little Red Riding Hood came out...’
   (Tales)

Finally, both rexuɖu and rexu are used as purely discourse-pragmatic particles, e.g. emphasizing a request/threat (4) or appealing to some previously obtained shared knowledge.

(4) oh, j-osi-b-ul=rexu!
   PTCL III.PL-PL.stop-IIMP-PL=after
   {The angry wolf says to the piglets:} ‘Just you wait!’ (Tales)

Besides illustrating the various heterosemous items, the paper will present possible semantic links between them. This, while the development from an adverb/postposition ‘behind’ to a sequential marker ‘then’ seems quite straightforward, it is less clear how the sequential developed into a pragmatic particle.

Acknowledgments
This research is part of the project “Andi morphosyntax in a typological perspective” supported by the Russian Foundation for Basic Research (RFBR), grant No. 18-012-00852a.
Small words matter: *Oh* as a marker of discourse transition in English and its equivalents in Spanish

Elena Martínez Caro
(Universidad Complutense de Madrid)

Key words: interjections, discourse markers, discourse transition, English, Spanish

Interjections belong to a broad class of items which have generally received little attention in descriptive grammars. A close examination of their use in large samples of talk in interaction, however, readily reveals that (at least the most common types of) interjections merit further investigation, on account of their high frequency and multifunctionality (Ameka and Wilkins 2006, Cuenca 2000, and Goddard 2014).

Among interjections, *oh* is one of the most highly frequent and multifunctional types, generally considered a discourse marker (Schiffrin 1987). Previous accounts have focused on its epistemic or affective uses, often describing *oh* as expressing the speaker’s surprise as a reaction to an expected situation, or as involving a change of state, presented as immediate (Aijmer 1987, Bolden 2006, Fox Tree and Schrock 1999, and Heritage 1998, 2018).

Within the framework of a project on global discourse coherence, this study examines *oh* from a different perspective, that of discourse transition and topic shift. As such, it attempts to offer insights into the way speakers of English (and Spanish; see below) move from one discourse topic to another and how these transitions are marked. (1) and (2) illustrate this structural meaning of *oh*:

(1) *Oh*, dessert time, are you having any dessert? (BNC Int KS2)
(2) *Oh* and you know another thing that Tim had the audacity to bitch about? (SBC 007)

The findings from the examination of oral samples of three corpora (BNC, LLC and SBC) indicate that *oh* can be used in the context of a topic introduction, resumption or closure in two main discourse situations: narrated stories affecting the speaker personally and in the expression of the speaker’s personal opinions or feelings. The topic shift may imply a more or less evident departure from the main story line, coinciding with the introduction of a completely new discourse topic in the former, or a related or inferred discourse topic in the latter. Further, certain structural and prosodic features noticed enhance this meaning of *oh*.

Note that the claim is not that *oh* alone always marks the topic shift in these uses (cf. *oh incidentally*), nor that further added meanings cannot be found in addition to that of topic transition (e.g. “sudden realisation” or “immediate remembering”). Rather, the study attempts to contribute to the set of markers of topic shift found in the literature and expand the knowledge of *oh* as a marker of discourse organisation and not exclusively as expressing affective and epistemic meanings. A final goal of the study is to explore the equivalent uses in Spanish of transitional *oh* in English, from the examination of several spoken samples from comparable corpora (C-Oral Rom, CORLEC and
ValEsCo). The results show that there is a wide range of equivalent uses, with the interjections *ah* and *ay* being the most common, sometimes in combination with other markers (*ah y; ay Dios mío*), and an omission of the interjection altogether. Some of these results are supported by studies based on audiovisual translations (e.g. Chaume 2004, Cuenca 2006, Matamala 2007).

**Acknowledgements**

I am indebted to the European Comission (Erasmus+) and the Universidad Complutense de Madrid (*Beca del Amo*) for financial funding during my stays in 2019 at the Université Catholique de Louvain (UC Louvain) and University of California at Santa Barbara (UCSB), where the research for this paper was initiated.

**References**


Heritage, John (2018), Turn-initial particles in English: The cases of *oh* and *well*, in J. Heritage, and M.-L. Sorjonen (eds), *In Between Turn and Sequence: Turn-Initial Particles Across Languages*, Amsterdam / Philadelphia: John Benjamins, 149–184.


Reported speech at the intersection of grammar and style: The speech-introducing function of interjections

Tatiana Nikitina & Ekaterina Aplonova

*Bridging linkage in the world’s languages*

*Jesus Olguin Martinez*

(University of California, Santa Barbara)

Keywords: Discourse, complex sentence, typology, subordination, temporality
Bridging linkage is a discourse phenomenon pervasive cross-linguistically. Two types have been traditionally recognized: recapitulating and summary bridging constructions. First, recapitulative constructions involve the repetition of a predicate of the reference clause in the bridging clause (de Vries 2005: 364), as in (1). Guérin & Aiton (2019: 11-12) explain that there are four subtypes of recapitulative constructions involving different instances of repetition, viz modifications, omissions, additions, and substitutions.

Lango (Eastern Sudanic/Nilotic)

(1) té <càmmò dyél > tè ninò.
   3SG.SBJ.and.then-HAB eat-INF goat 3SG.SBJ.and.then-HAB sleep-INF
   ‘He ate the goat and he slept.’

   ì kărë àmë òninò…
   in time REL 3SG.SBJ-sleep-PFV
   ‘After sleeping…’ (Noonan 1992: 253)

Second, summary constructions involve the replacement of the lexical verb of the reference clause by a generic or light verb (de Vries 2005; Guérin & Aiton 2019), as in (2). Guérin & Aiton (2019: 10-20) propose two main subtypes, viz. generic or light verbs accompanied by a deictic element, and demonstrative verbs expressing manner (de Vries 2005; Thompson et al. 2007; Guérin 2015).

SiRoí (Trans-New Guinea/Madang)

(2) piro mbolngɛ ngukina.
   garden in plant.PST
   ‘She planted it in the garden.’

tangamba, nu kinyana.
   doing.thus 3SG.SBJ.sleep.PST
   ‘Having done so, she slept.’ (van Kleef 1988: 151)

This paper aims at contributing to broadening the typology of bridging linkage constructions through examination of a convenience sample of 57 languages. In particular, this work seeks to expand this research by exploring recapitulative and summary bridging constructions not addressed by Guérin & Aiton (2019: 11-12). First, languages may have recapitulative constructions in which reported speech clauses are omitted and only the quotative verb is recapitulated, as in (3). Second, languages may have summary bridging construction composed of a light verb accompanied by a deictic and dependency marker, as in (4).

Barasano (Tucanoan)

(3) to yi, ba-re yà-a-ha. Soe-ya bëê. Ba-to
   that do eat-SBJ be-PRS-3SG burn-PRS.IMP 2SG eat-IMP.2SG
   so-re yi-yu-hu.
   3SG.F-OBJ say-INFER-3.PL
   ‘Doing that. “Here is food. Singe off the outside. Let’s eat” he said to her.’

   o i yi-ro…
   that 3SG.M say-NMLZ
   ‘When he said that…’ (Jones & Jones 1991: 180)

Hup (Nadahup)

(4) yikàn xìn ni-ay-àh, bɔyɔk xìn bìz-ìy.
   over.there 1PL stay-INCH-DECL.tapiri.shelter 1PL make-DYN
‘There we stayed (a while: while we were there), we built a tapiri shelter.

\[ \text{yɨ-nɨh-yó, } \vzık\text{ } \z\text{in}\text{ } \yɨt-\jy,\text{ } \z\text{in}\text{ } \ní-\ay-\ám. } \]

that-be-like-SEQarmadillo 1PL follow-DYN 1PL be-INCH-DECL

‘Having done this, we followed an armadillo: we stayed there.’ (Epps 2008: 531)

Although both strategies are attested in languages from different areas of the world, they are more frequent in South America in the languages of the sample, in particular in East Tucanoan languages (cf. Aikhenvald 2019). Interestingly, other neighboring languages not genetically related to East Tucanoan also have both strategies. It is proposed that the most obvious explanation seems to be language contact. This is because: (1) the languages are spoken in the same region, (2) they are not genetically related, and (3) the probability of chance resemblance is low given the rarity of the strategies.

References


Thinking non-aprioristically of discourse markers, modal particles, and the like

Vladimir Panov

This paper presents a methodological proposal illustrated by case studies. In this talk, I focus on “small elements” which are usually associated with the domain of discourse and are labeled in the literature as discourse particles, modal particles, discourse markers, pragmatic markers, pragmatic particles, etc. I will try to answer the question: what is the best way to typologically account for them?

The organizers of the workshop note that discourse phenomena have been a natural field of investigation of pragmatics, semantics, and discourse analysis, not typology. “Small elements” are no exception: there are hundreds of studies of such elements in particular languages, a few contrastive studies of pairs of languages, and very few large-scale cross-linguistic studies.
Besides language-particular studies, scholars have been concerned with the problem of “distinguishing” between different classes of such elements, e.g. in the influential volumes Degand et al. (2013) or Fedriani & Sansò (2017). Surprisingly, it seems that until now this field of research has been immune to the “comparative concepts revolution” provoked by Haspelmath (2010). The main failure of the majority of works trying to draw lines between, say, modal and discourse particles is their non-distinguishing between language-particular categories (“p-categories”) and cross-linguistic generalizations (comparative concepts). Authors seek to establish cross-linguistic categories of the natural-kind type (Haspelmath 2018) and “diagnose” these categories in particular languages by studying their “symptoms”. As a consequence, previous studies often suffer from shortcomings such as mixing up morphosyntactic and functional properties, and eurocentrism, for instance, uncritically applying the p-category “Modal Particles” of some European languages cross-linguistically.

I argue that, as in more traditional fields of typology like TAM or argument structure, one should rather take a non-aprioristic approach and distinguish between p-categories based upon language-particular criteria, and comparative concepts allowing for cross-linguistic generalizations. In defining comparative concepts, I suggest to follow the ideas of Croft (2001, 2003) and clearly set apart meaning (conceptual space) and form. Typological tendencies manifest themselves in the types of form-meaning pairings and statistical cross-linguistic asymmetries found in such pairings.

Thus, for instance, if one wants to make the German Modal Particle wohl part of typological research, one should treat its meaning (‘~ probability’) and morphosyntactic/phonological properties separately. Given that the morphosyntactic and phonological properties of the German Modal Particles are highly language-specific and cannot be mechanically applied to other languages, one should come up with a comparative concept applicable cross-linguistically, e.g., “sentence particles” defined as “short bound elements taking scope over the whole clause or sentence”. Then one could look at how cross-linguistically typical a form-function pairing such as the German wohl is. At the same time, one should investigate by which morphosyntactic means the same function can be alternatively expressed, and, conversely, which functions are typically expressed by “sentence particles” cross-linguistically.

In my talk, I will elaborate on a number of useful comparative concepts in the field defined above, and will present case studies of how they can be fruitfully applied.

Towards pragmatic construction typology: The case of discourse formulae

Ekaterina Rakhilina & Polina Bychkova
(National Research University Higher School of Economics and The V.V. Vinogradov Russian Language Institute; & National Research University Higher School of Economics)

Keywords: discourse formulae, construction grammar, pragmatic typology, yes/no responses, speech acts

The paper suggests a typological account for a linguistic domain which combines pragmatic meanings and idiomaticity in the sense of Construction Grammar (Fillmore et al. 1988, Goldberg 2003). We consider a particular class of units, discourse formulae, and study them cross-linguistically applying the frame-based approach of Moscow Lexical Typology group (Koptjevskaia et al. 2016). In this approach the semantic oppositions are explored through the combinatorial potential of the lexemes. The methodology was initially developed for the traditional semantic fields such as motion verbs, physical qualities, or names of objects. The present work extends its scope to discourse constructions and investigates typologically a range of formulaic expressions that are used as positive and negative
responses in a dialogue. The discourse constructions of this type are compared in a number of languages including English, Italian, Russian and Slovenian.

The term “discourse formulae” here refers to dialogical remarks such as English *Fair enough*, *Nothing of the sort*, *You bet* or *I couldn’t care less* that are used as verbalized reactions to the utterance of another participant of the conversation:

\[(1) \quad - \text{You wanna grab some coffee?} \]
\[
\text{positive: – You bet! negative: – I’m good.}
\]

Just like in other schematic constructions (cf. *F X A Y let alone B, the X-er the Y-er, or What a X!*), the meaning of these phrases is non-compositional. What distinguishes discourse formulae from standard constructions is that the former constitute an entirely fixed utterance. Their variables are located in a broader context: the preceding utterance of the interlocutor occupies the slot of the discourse formula. Hence, the analysis of discourse formulae is based on the differentiation of stimuli remarks.

The contextual distribution of nearly synonymous formulae highlights the types of stimuli utterances that correspond to a large extent to the category of speech acts, originally elaborated in philosophic literature (cf. e.g. Kissine 2013). In this respect, they could be seen as a linguistic tool for classification of speech acts and detection of typologically relevant pragmatic frames (i.e., typical conversational situations).

An inventory of discourse formulae can presumably be found in any language. They tend to present a system of regular pragmatic classes, such as consent, refusal, confirmation, negation, or surprise. According to our data, languages can exhibit different strategies to organize these classes, revealing pragmatic oppositions within them. What lies behind is a universal system of comparative concepts (Haspelmath 2010) just as it is in the case of grammar or lexicon. We believe that the typological analysis is a key to that system.

**References**


Repetitive Constructions and Stance-Marking: The Case in Korean

Seongha Rhee

<not updated>

**Introducing new referents: A corpus-based cross-linguistic perspective**

405
Keywords: new information, spoken language corpora, presentational, discourse processing

The introduction of new referents into discourse has been assumed to pose challenges to discourse processing, leading to compensatory strategies in morphosyntax (e.g. Chafe 1987; Du Bois 1987, 2003a,b, 2017). Lambrecht (1994: 184–191) postulates the principle of separation of role and reference (PSRR), claiming that speakers use specific clause constructions to establish reference before turning to the expression of event information, for instance presentational or detachment constructions. In the same vein, Du Bois (1987, 2003a,b) identifies objects and intransitive subjects as the preferred entry points for new referents as the preferred entry points for new referents. We test these claims against corpora of spoken free narratives from nine languages (Haig & Schnell, 2020) and Pear Story corpora from four languages. Contrary to Du Bois (1987, 2017), we find that new mentions occur freely in any position except subjects. We thus find that no one specific core argument position is systematically associated with newness over any other position. Speakers hence do not appear to harness specialized morphosyntactic patterns to signal the arrival of new discourse referents to their addressees. Our investigation of Pear Stories likewise show that, contrary to Lambrecht (1994), the most narratively pertinent referents are introduced chiefly via motion predicates, whose selection is motivated largely by semantic and content-related considerations, rather than their use as semantically bleached vehicles for new material. Presentational constructions play only a limited role, and mainly in English.

In sum, our corpus-based findings suggest that morphosyntax and syntactic roles are less sensitive towards the newness of referents than previously claimed; our data suggest that new referents tend instead to be seamlessly integrated alongside event-related content (Schnell et al, forthcoming). Moreover, the cognitive challenges raised by information pressure have little to no effect on morphosyntactic structure. It appears that patterns of referent introduction do not follow from considerations of audience design, an observation that squares well with findings from psycholinguistic work in the area of discourse processing. Those preferred patterns of referent introduction we do find in our data, e.g. a fairly high proportion of new referents in object position and in whatever position comes later in a given sentence, are better accounted for in terms of production-related processing demands, Hence, we outline an account along these lines, drawing on psycholinguistic work by Arnold (2008) or MacDonald (2013), among many others. The rational here is that specific production-constrained patterns can have effects on comprehension via implicit learning, which explains why some specific construction types are “felt” to specialize in referent introduction by listeners. Our research makes a specific contribution to the discussion of reference processing by broadening the empirical basis of production-oriented research across a range of typologically diverse languages beyond English.

In this way, our findings have a much greater potential to outline universal tendencies in reference processing than previous psycholinguistic research as well as discourse-analytical work in the functionalist tradition, both of which draw on fairly confined datasets from a restricted number of structurally similar languages with similar cultural backgrounds.

References
Discourse-pragmatic elements under language contact: The case of Pidgin *like* and Tsova-Tush *k’aco*

Claire Stabile & Bryn Hauk
(Kamehameha Schools & University of Hawai‘i at Mānoa)

Keywords: discourse phenomena, language contact, typology, creoles, minority languages

What can a comparison of borrowed discourse-pragmatic elements tell us about the universality of discourse phenomena? It has been widely observed that discourse-pragmatic elements are among the first items borrowed under language contact (Auer & Maschler 2016; Brody 1987). However, the degree of their similarity in function after borrowing can differ widely depending on the contact scenario: in some cases the discourse-pragmatic elements maintain the functions of the source form quite faithfully (Brody 1987), while in others they retain only restricted functionality (such as English politeness markers in Cypriot Greek; Terkourafi 2011) or develop a life of their own (such as English *like* in Cypriot Greek; Maschler & Dori-Hacohen 2016). These differing trajectories under language contact provide insights into the typology of discourse-pragmatic elements.

We discuss two cases of discourse-pragmatic elements borrowed into typologically and geographically disparate minority languages that have remained in extended contact with the majority language supplying the discourse element. Specifically, we look at English *like* as borrowed into Pidgin (Hawai‘i Creole) and Georgian *k’aco* ‘man (vocative)’ as borrowed into Tsova-Tush (Northeast Caucasian). Despite the differences between the two minority languages, there are considerable similarities in terms of contact scenario and the pragmatic functions of the borrowed items. We use mixed methodologies to show that Pidgin *like* maintains most of its discourse-pragmatic uses as in English, including but not limited to its use as a quotative and discourse marker (Stabile 2019), while *k’aco* has been borrowed into Tsova-Tush chiefly as a quotative, with less evidence of it retaining other discourse-pragmatic functions as in Georgian.

Using the variationist and frequentist methods applied to corpora of Hawai‘i English and Pidgin, we show that in Pidgin, discourse-pragmatic *like* patterns similarly to World Englishes (cf. D’Arcy...
2017), with young men’s higher rates of the discourse marker being a key exception. A matched-guise perception experiment demonstrates that young Hawai‘i Locals accept quotative like as authentically local, a result which adds to the growing body of research on ideologies towards like.

The Georgian discourse marker kac’o has received little prior study. We show that, beyond its original vocative function for addressing adult men, k’aco is used with interlocutors of any gender or generation with discourse-pragmatic functions such as highlighting speaker stance, signaling focus, and introducing reported speech. However, in our corpus-based study of Tsova-Tush, the distribution of k’aco is more limited, used as a specific quotative to mark defiance or surprise. We suggest that other discourse-pragmatic functions may not have been borrowed, in contrast with Pidgin like, due to competition with pre-existing discursive elements and social perceptions of authenticity.

We contribute to the typological theory of discourse phenomena in two ways. First, we highlight the similarities and differences of discourse markers in highly disparate languages, using corpora and experimental evidence. Second, we contribute to the typology of borrowing of discourse features. Although both minority languages in our study remain in extensive contact with the source language, the patterns of borrowing differ in interesting ways, highlighting the need for further study of discourse elements under language contact.

References
Auer, Peter & Yael Maschler (2016), The family of NU and NÅ across the languages of Europe and beyond: Structure, function, and history, in P. Auer and Y. Maschler (eds), (2016), Nu/Nå: A family of discourse markers across the languages of Europe and beyond, Berlin: De Gruyter, 1–47.

Discourse connectives in Beja: From diachrony to typology
Martine Vanhove

In Beja (North-Cushitic, Afroasiatic), discourse connectives, whose origin is still transparent, come from a large variety of word classes and semantic fields. Word classes include nouns (singular and plural), adjectives, finite and nonfinite verbs, negative verb forms, numerals, postpositions, adverbs, interjections, deictics. Semantic fields include dummy nouns (na ‘thing’), time (ho:b, minda ‘time’), manner (bak ‘thus’), size (di:s ‘small’), religion (walla ‘God’), location (ak ‘be there’), emotion (qikw
‘to suffer’, *fad* ‘to love’), distributivity (*=ka* ‘each’), space (*hadi:d* ‘limit’), similarity (*ki:k* ‘equivalent’), cognition (*subbi* ‘reason’). They are grammaticalized to various degrees and target different types of semantic relations between discourse chunks or clauses, such as temporal, concessive and reason clauses, sequentiality, consequence, addition, contrast with a previous utterance or shared knowledge, introduction of a new or surprising idea, restriction, conclusive remark, etc.

Based on the analysis of a corpus of naturalistic oral data of some 40,000 words recorded and annotated by the author, this presentation will first analyse the semantic values of a selection of the most frequent discourse connectives. It will then discuss the advantages of using naturalistic corpora for research in this domain. The second part will discuss the Beja findings both in terms of semantic relations and diachronic sources, in the light of comparative research in this domain, which is mainly restricted to well-known European languages (e.g. Rysová 2017). Beja shows that both semantically and grammatically the origin of discourse connectives are richer than expected and poses a number of challenges for crosslinguistic studies that will be addressed as a final conclusion.

**Reference**
Contact and the architecture of the language faculty

M. Rita Manzini & Greta Mazzaggio
(University of Florence)

Keywords: (micro)variation, change, parametric models, contact, divergence/convergence.

This workshop proposes to investigate the catalyzer role of contact in language variation and change. Contact is seen as a medium to have a better understanding of variation and change and, ultimately, of the architecture of the language faculty.

Linguistic variation and change are interesting to the extent that, no less than language universals, they appear to be not accidental to language, but rather art of its basic design. Specifically, conceptual and inferential components are presumably not subject to variation/change – nor are the computational components (syntactic operations, phonological operations). Rather, variation arises when these components interact (notably at externalization), posing questions of functional optimization to the system. “Why are there so many languages? The reason might be that the problem of externalization can be solved in many different and independent ways”. “Diversity of language results from the fact that the principles do not determine the answers to all questions about language, but leave some questions as open parameters” (Berwick & Chomsky 2015).

In classical Principles and Parameters models, linguistic differences are encoded as different parameter settings, where parameters are all present at the initial state of language acquisition $S_0$ in the form of an extensional finite list of every possible points of variation in human grammars, represented as open binary choices. Though this model has been extremely productive, it does not seem adequate as a really plausible model of the language faculty. Various recent proposals attempt to re-discuss the nature of parameters and their place in the general architecture of grammar, to bring them to bear on current minimalist theorizing. In particular:

- Biberauer et al. (2014), Roberts (2019) propose that parameters are organized along hierarchical schemas, which are identically organized only at general level.
- Longobardi (2017), proposes that parameters are based on a universal set of simple possible parameter schemas (e.g. “Is F, F a feature, grammaticalized?”; “Is F, F a grammaticalized feature, spread on X, X a category?”; “Does a functional category … X have a phonological matrix ?”) which may in principle interact with one another.
- Manzini & Savoia (2011), Manzini (2015) take an even weaker approach, namely that “the parameters interacting with [Externalization] are … the categorical splits” for instance “speaker vs. hearer, 1st/2nd person vs. D” in the realm of nominal properties. Thus even parameters schemas are epiphenomena, the conceptual workspace and the categorical cuts (parameters) that are or are not externalized by the lexicon is all there is. Interactions between categorical cuts take the form “categorial split A is not defined for value 0/1 of … categorial split B”.

Within this general theoretical framework, contact phenomena provide a unique vantage point into models of parametrization, for they are one of the external factors that may act as catalyzers or at least as accelerators for variation and change. By contact, we mean any situation in which two different languages are spoken by the same community, hence in a condition of bilingualism. Specifically, we
are interested in situations of protracted (centuries-long) contact, often historically documented, which allow us to observe internal diversification of the languages involved both on the temporal axis (change) and on the spatial axis (dialectal variation).

By bringing together scholars working on the syntax of varieties spoken in conditions of contact, the workshop will investigate how factors external to the grammar of a given language can influence the shape of the grammar, determining divergence/convergence phenomena with close cognates and with the contact language.

**Range of topics**

Proposed talks focus on the syntax of varieties spoken in conditions of contact. The selected papers

- present a formal analysis of data relating to syntactic variation and change in languages in contact;
- contribute to the analysis of syntactic phenomena involved in variations and change, possibly lying at the interface between syntactic computation and externalization systems.

The papers address one or more of the following core questions:

(i) the scope and limits of morpho-syntactic diversity: what kinds of linguistic properties and relations are subject to variation and what is the admissible extent of this variation, given the constraints imposed by the Faculty of Language?

(ii) the structure of morpho-syntactic diversity: does the analysis of languages in contact provide new evidence for different parametric models (hierarchies, schemas, lexical parametrization)?

(iii) constraints on externally-induced change: is it possible to delimit the possible changes that may happen in a situation of contact? Proposals in this sense are:

a. **Resistance Principle** (Guardiano et al. 2016): “Resetting of parameter $\alpha$ from value X to Y in language A as triggered by interference of language B only takes place if a subset of the strings that contribute to constituting a trigger or value Y of parameter $\alpha$ in language B already exists in language A” (i.e., the resetting of a parameter under the influence of interference data is possible only if the new triggers are similar enough to triggers already unmistakably present in the interfered language.

b. **Given Inertia** (Keenan 2009, “Things stay as they are unless acted upon by an outside force or DECAY”), contact situations provide an “outside force”; then the question arises whether parameter resetting takes place not simply when possible, but rather when needed to fill an independently-created void (we may call this the **Functional Void** principle).

The ultimate goal of the workshop is contributing to the description, analysis and explanation of language diversity, using the “external” impulses acting on it in a situation of language contact as a key to get to the core of the “internal” structures that determine language diversity.

**References**


Peripheral significance: the action at the edge of Kaaps Afrikaans phases

Theresa Biberauer

This paper argues for the importance of structural edge-domains in understanding variation that arises in contact circumstances. Our empirical focus will be on Kaaps, the mother-tongue variety of Afrikaans that has the dual distinction of being both the most contact-influenced and also the oldest spoken Afrikaans variety (Hendriks & Dyers 2016).

Spoken on the Cape Peninsula by a mixed-race community which has always lived in a sociolinguistically complex situation, Kaaps confounds the expectation that a language which has been in long-term intensive contact with i.a. English, Arabic and various Bantu SVO-languages should itself become SVO. Systematic investigation demonstrates the robustness of Kaaps V2 and OV (Biberauer & Pretorius 2018). Macro- and mesoparametric properties, which are known to be acquired early (Tsimpli 2014), thus remain stable in Kaaps, despite intensive contact with systems with different specifications for the relevant dimensions of variation. Strikingly, however, the real interest of Kaaps’ retained OV-V2 profile and of its structural profile more generally arguably only emerges when one looks more closely at micro-properties. More specifically, what we uncover is a range of structures that point to innovations - some of them long-standing, and others more recent - that have taken place at the outermost edge of Kaaps’ CP-, vP- and DP-domains. Here we focus on:

(i) **V2 embedded wh-interrogatives**, the availability of which turns on an innovated CP-peripheral Pol(arity)P. This is also attested in other varieties of spoken Afrikaans that have integrated the additional PolP, which has its origins in a range of emphatic spoken-language structures, strongly tied to the contact situation in which Afrikaans emerged (Roberge 2000); and

(ii) **Differential-Object-Marking vir** (‘for’), involving an extended, DP-periphery which now hosts the erstwhile P-element vir (DOM-marked nominals in Kaaps can be shown to be DPs, just like the a-marked objects in Calabrese or Catalan). The availability of these elements, in turn, has consequences for how West Germanic scrambling manifests in Kaaps, i.e. for how the edge of vP is utilised.

The rise of the vernacular micro-variation described above is ascribed to a minimalist-emergentist hypothesis specifying how “new” elements - both borrowed, and thus (formal) featurally opaque, and newly introduced by means other than upwards-grammaticalisation - are integrated into existing structure. The core proposal is that phase-edges are the “way in” for formal feature([F])-less or formally underspecified elements because such elements must be last-out of their Lexical Arrays.


Essentially, the idea is that Merge combines the regularly [F]-specified elements in a given Lexical Array (corresponding to, say, vP or the higher clausal phasal domain) in accordance with their selection requirements before merging any that either entirely lack or have only a partial [F]-specification (Biberauer 2017). Elements of this kind may gain [F]s over time, as they become more grammaticalised, thereby becoming more integrated at the edge of the relevant domain as a novel micro-structural layer. The presence of this new microstructure, in turn, has knock-on effects for the rest of the grammar, here including more widespread V2, reanalysed D-elements, and a reconfigured scrambling system.

References


Is the placement of subjects to be changed through contact?
The view from Lipovan Romanian and Moldovan Romanian

Adnana Boioc Apintei & Ştefania Costea
(University of Bucharest & ‘Iorgu Iordan-Alexandru Rosetti’ Institute of Linguistics; University of Cambridge)

Keywords: Romanian-Russian linguistic contact, Lipovan Romanian, Moldovan Romanian, verb movement, subject position

In our presentation we aim to discuss the situation of two varieties of Romanian which have been in contact with Russian for the last 200 years, i.e. Lipovan Romanian (spoken in Dobrudja) and Moldovan Romanian (spoken in the Republic of Moldova). We seek to explain, along the lines of Costea’s (2019) proposal for Moldovan Romanian, that the generalized preverbal placement of subjects in pragmatically unmarked contexts is not a direct ‘syntactic borrowing’ from Russian. Of course, Russian presents an EPP feature on TP, generally raising subjects to Spec,TP (Slioussar 2007:40; Gribanova 2017:1081;1084); nevertheless, a more careful examination shows that (standard) Romanian subjects can also raise to Spec,TP, when expressed overtly (Cornilescu 2000:92;95). The difference between (standard) Romanian (1), on one hand, and (standard) Russian (2), Lipovan Romanian (3), and Moldovan Romanian (4), on the other hand, is to be found in the level of verb-movement displayed by these varieties: while (standard) Romanian verbs raise to a high position in the IP-domain, presumably a MOOD-related position à la Cinque (1999) (cf. also Ledgeway&Lombardi 2005:83; Schifano 2018:2), (standard) Russian, Lipovan Romanian, and Moldovan Romanian verbs reach only an ASPECT-related position (Gribanova 2013:94-95; Roberts 2019:359 n.17). Hence, while
(standard) Romanian verbs raise over Spec,TP, (standard) Russian, Lipovan Romanian, and Moldovan Romanian verbs raise below Spec,TP, thus allowing subjects to appear preverbally in pragmatically unmarked contexts.

(1) Te-au  trimis  ei  la  şcoală.
   CL.ACC.2SG=AUX.PERF.3SG send.PPLE they at school
   ‘Your parents sent you to school.’ (standard Romanian)

(2) On  ljubit  evo  druzej.
   he  love.PRES.3SG  his  friends
   ‘He loves his friends.’ (Russian, adapted from Slioussar 2007:146)

(3) Ea  mereu  povesteşte  despre  asta.
   she  always  tell.IND.PRES.3SG  about  this
   ‘She always talks about this.’ (Lipovan Romanian)

(4) Eu  mă  gândeam  la  asta.
   I  CL.ACC.1SG  think.IND.IMPF.1SG  at  that
   ‘I sometimes thought about it.’ (Moldovan Romanian)

References
Costea, Ștefania (2019), When Romanian meets Russian. The effects of contact on Moldovan Daco-Romanian syntax, University of Cambridge: MPhil Thesis.
Ledgeway, Adam, and Lombardi, Alessandra (2005), Verb movement, adverbs and clitic positions in Romance, Probus 17, 79–113.

Trapped in the Parameter: Interrogatives in Contact in the Basque Country

Ricardo Etxepare, Georg Kaiser & Simon Dold

Both Basque and Spanish are wh-fronting languages that typically display wh-phrase-verb adjacency, as the subject inverts in wh-questions (Bosque and Gutierrez-Rexach, 2011: 449; Etxepare and Ortiz de Urbina, 2003: 495ff):

(1) a. Qué compra Pedro?  (2) a. Zer erosten du Pellok?
     what buys Pedro     what buy.NOM.LOC AUX Pello.ERG
     ‘What does Pedro buy?’  ‘What does Pello buy?’
b. *Qué Pedro compra? what Pedro buys
b. *Zer Pellok erosten du? what Pello.ERG BUY.NOM.LOC AUX

Spanish fronting nevertheless allows interpolation of other syntactic material between the wh-phrase and the verb (2) with adjuncts (Torrego, 1984), complex wh-phrases (Goodall, 2004), and (to some extent, subject to inter-speaker variation), with D-linked wh-words too:

(2) A cuál de las chicas tu hermana había visitado ya antes?
‘Which of the girls had your sister already visited?’

None of those possibilities is available in Basque (with the exception of why-questions).

Overall, Basque-Spanish bilinguals rate sentences such as (2) significantly lower than Spanish monolinguals (p<0.01), and 2L1 speakers significantly lower than early sequential learners (p<0.03). The latter also show lower acceptability rates as compared to Spanish monolingual speakers from outside the Basque Country. The results come from a study (Dold, 2018) involving 114 interviews in the region of Gipuzkoa, divided according to the bilingual_monolingual status of the subjects, and their age and degree of exposure to Basque and Spanish. The sociolinguistic distribution of this restriction suggests that it is related to contact with Basque. But how should we model this influence?

Cable (2010) proposes that the syntax of interrogative clauses is mediated by what he calls a Q-particle that on the one hand relates to a left peripheral Force feature, and on the other merges to a phrase containing the wh-word. Association of Q to this phrase may be via Set Merge (as a head) or Pair Merge (as an adjunct) (Chomsky, 2001).

(3) a. [QP Q [XP…wh…]] (Set Merge)
b. [XP Q [XP…wh…]] (Pair Merge)

Basque partial interrogatives are construed by Pair Merge. Spanish ones via Set Merge. This has an effect in the distribution of Q and wh-phrases in the left periphery. Spanish Q+wh-phrases end up as a unit in a high Force projection, which leaves room for topics to be inserted between the wh-word and the verb.


In Basque, Q and the wh-phrase move separately. Wh-words land in the preverbal focus position (Etxepare and Ortiz de Urbina, 2003). Q raises to Force. In such a configuration, there is no room for a topic between the wh-word and the verb.


Once the parameter of Q-association is set on the Basque model, there is no evident way to come back: all sequences compatible with the structure in (3b) are also compatible with (3a), so Basque bilinguals do not reset the parameter.

Modeling syntactic change under contact: the case of Italiot Greek

Cristina Guardiano & Melita Stavrou
(Università di Modena e Reggio Emilia & Aristotle University Thessaloniki)
Keywords: Greek, Southern Italy, syntactic contact, nominal structures, parameter resetting

In this paper, we explore aspects of the syntax of DPs in Italiot Greek, examining data collected in Grecìa Salentina and Bovesìa. We compare Italiot Greek to Modern and Ancient Greek, and we investigate its relationship with Southern Italo-Romance as a potential source of interference.

Nowadays, Greek in Southern Italy is in regression: it is spoken by very few speakers, mostly bilingual. Thus, especially when subtle intuitions are required, speakers are often not able to provide unambiguous judgments. To compensate for this, we gathered data both from native speakers and from a selection of written records, likely to reflect the language spoken by the beginning of the 20th century in the area(s). The comparison between the two sources reveals important differences which allow us to identify steps of changes in progress.

We focus on four subdomains, of which we propose an analysis based on current theories about the structure of the nominal domain in Greek and Romance: adjectives (Stavrou 2012, 2013, Guardiano & Stavrou 2014, 2019a/b), demonstratives (Horrocks & Stavrou 1987, Guardiano 2012, 2014, Guardiano & Michelioudakis 2019, Guardiano & Stavrou 2020), possessives (Guardiano et al. 2016) and genitives (Longobardi & Silvestri 2013, Crisma et al submitted). In at least two of these domains (adjectives, demonstratives) Italiot Greek has undergone important changes which made it superficially diverge from Greek and converge with Romance. Conversely, the two other domains (genitives, possessives) seem to have been immune to change.

It is quite clear that interference with Romance, coupled with the endangered status of the language in the relevant area(s), has played a role in these changes. Yet, the data which we collected lead us to the conclusion that changes crucially happened only in those subdomains where the conditions triggering them were independently available in Greek. In particular, for adjectives and demonstratives, the origin of the observed changes seems to be the loss of a specific structural configuration (the one that generates ‘polydefinite’ DPs), probably due to reasons independent of contact with Romance. As a consequence of this event, the linear patterns originally generated from such a structure either disappeared or were reanalyzed as resulting from other (different) structures, unavailable in Greek but crucially available in the neighboring Romance dialects; significantly, this second circumstance only happened to patterns identical to those found in (Southern Italo-)Romance. No such phenomena are detected in the domains of genitives and possessives, where almost no difference is visible between Italiot Greek and (the rest of) Greek.

These conclusions support the hypothesis (based on Guardiano et al.’s 2016 resistance principle) that syntactic convergence under the influence of interference data cannot just be the byproduct of mere superficial borrowing of word order patterns (or lexical items), but is rather the consequence of the interplay between specific structural conditions, most crucially the availability in the ‘interfered’ grammar of linear patterns compatible with those generated by the ‘interfering’ one, and external pressures, mostly consisting in massive exposure to the linear patterns which generate/favor structural reanalysis.

References
UCLA, May 2, 2014.

At the crossroad of Croatian and Italian Dialects: Subject clitics in Istro-Romanian
Alexandru Nicolae & Adina Dragomirescu

Nominal syntax (and morphology) at the Ligurian/Occitan border
Diego Pescarini

Resilient subject agreement morphosyntax in the Germanic Romance contact area
Cecilia Poletto & Alessandra Tomaselli
A Balkan view on the left periphery: Modal and discourse particles

Anna Roussou
(University of Patras)

Keywords: lexical borrowing, externalization, discourse, modality, particles

In the present paper I consider two phenomena associated with the left periphery of the clause structure in the Balkan languages. The first involves the ‘subjunctive’ (1) and the ‘future’ (2) (modal) particles. These common formations are attributed to geographical proximity, despite the different origins of the languages involved, and are well-discussed in the literature (starting with Joseph 1983, Rivero 1994, a.o.). The second phenomenon involves discourse markers, as in (3). The latter have only recently started to receive attention in the literature (see for example Tchizmerova 2005 for Bulgarian; Hill 2008, Haegeman & Hill 2013 for Romanian; Roussou 2017 for Greek).

(1) {Greek na, Albanian të, South Slavic da, Romanian sâ} + finite verb
(2) {Greek tha, Albanian do, Serbian da} + finite verb
(3) Romanian haide, Greek ande, Bulgarian xajde, …

The modal particles in (1) and (2) correspond to different lexical items in each language. Their common property is the verbal complex they form, which developed out of the loss of the corresponding infinitive. As Asenova (2019) points out, it is this replacement that is due to language contact (and not the loss of the infinitive). Thus the lexical items in (1) and (2) assume similar functions across the languages involved and have a similar syntactic distribution, maintain differences at a micro-parametric level. On the other hand, (3) represents a case of lexical borrowing (possibly from Turkish); the new item is adapted in the recipient languages morpho-syntactically, and is incorporated in the upper left periphery (Speech Act layer) that involves the discourse participants (see also Bağriaçik 2018, Friedman 2019 on evidentials).

So while (1) and (2) involve a ‘surface structure’ effect (Sims & Joseph 2019), (3) is a clear case of lexical borrowing, with morpho-syntactic repercussions. The borrowed items may retain their initial function but adapt to the morphosyntactic properties of the recipient language.

Both patterns in (1)-(2) and (3) are left periphery phenomena. The former is linked to the expression of grammatical TAM properties (i.e. properties associated with tense, aspect or modality), while the latter are associated with discourse properties. It is next argued that despite apparently common externalization patterns, differences do arise, thus arguing for a micro-parametric approach based on the properties of the lexical items and the way they ‘cut’ the conceptual space (Manzini & Savoia 2011).

References


Rivero, Maria Luisa (1984), Clause-structure and V-movement in the Languages of the Balkans, Natural Language and Linguistic Theory 12: 63-120.

Roussou, Anna (2017), Expanding on the Left, Linguistica Brunensia 65(2).


**Contact in Franco-Provençal of Celle di San Vito: possessives**

Leonardo M. Savoia & Benedetta Baldi
(University of Florence)

Keywords: possessives, contact, Franco-Provençal, Southern Italian dialects, kinship terms.

In North Apulian villages of Celle and Facto, a Franco-Provençal dialect - dating to 13th century - is spoken. Although morpho-syntactic Franco-Provençal properties are overall preserved (Manzini and Savoia 2005), the contact with Apulian dialects affected morpho-syntax of possessives of this dialect.

1. In Southern Italian dialects, exemplified by the contact Apulian dialect of Castelluccio, possessives are post-nominal and require the definite article in DPs, (1a), and in predicative contexts, (1c); kinship terms take enclitic 1st/2nd/3rd singular possessives, (1b).

(1) a. l-a kammis-a mijә
    the-FSG shirt-FSG my
    ‘my shirt’

    b. fratә-tә
    brother-my
    ‘my brother’

    c. sә l-i susә-jә
they.are-the-PL  his.M-PL
‘they are his’

In Franco-Provençal varieties like the one of Coazze (Piedmont), possessives precede the noun excluding definite article, (2a,b), and may occur without article in predicative contexts, (2c).

(2) a. m-a ʧəmiz-i
1PSG-FSG shirt
‘my shirt’
m-un / t-uŋ ʧiŋ
1PSG-MSG/ 2PSG-MSG dog
‘my/ your dog’

b. t-uŋ μαι
2PSG-MSG brother
‘your/ our brother’

c. i ʃunt tɔ-i / tje
CIS are 2PSG-PL / 2PSG.F
‘they are yours’

Celle dialect possessives replicate the distribution in Southern Italian dialects, occurring in postnominal position and combining with article, (3a). With kinship terms singular possessors, the 2\textsuperscript{nd} of respect included, retain pre-nominal position excluding the definite article, (3b). (3c) illustrates predicative contexts.

(3) a. l-u / lɔ ʧiŋə mi-ŋə
the-MSG / the-MPL dog my
‘my dog/ dogs’
l-a ʧəmmis-a mi-’a
the-FSG shirt my
‘my shirt’

b. m-a / t-a fiəkə
my-FSG / your-FSG daughter
‘my/ your daughter’
m-uŋ / t-uŋ /vut-uŋ fiəwə
my-MSG / your-MSG / your.RESPECT-MSG son
‘my/ your son’.

b’. l-o ʧiŋə nniŋə noiŋə / voŋə / laurə
the-MP.L nephews our/ your/ their
‘our/ your/ their nephews’

c. i ʃunt l-o mi-ŋə
SC.LPL are the-MPL my-MPL
‘they are mine/ yours’

Contact changed the distribution of possessives, favouring the post-nominal position and the occurrence of the article in DP and predicative contexts. Kinship terms preserve pre-nominal occurrence of singular possessors, with specialized forms; plural possessors are post-nominal.

2. Differently from cartographic approaches, based on movement driven by uninterpretable features, we relate the morpho-syntactic behaviour of possessives to their interpretive content. Specifically,
of/oblique case lexicalize the elementary predicate connecting possessor and possessum represented as part-whole (inclusion \( \subseteq \)) relation in (4):

(4)  \( \text{of/oblique: } \) [possessum] \( \text{in/oblique } \subseteq [\text{possessor}] \)

Postnominal possessives externalize the argumental structure of inclusion by aligning the possessum outside of the possessor phrase. The reverse order, possessor – possessum represents the possessor’s denotational scope. Celle dialect requires the double lexicalization of possessum referential properties, i.e. possessive inflection and article, on a par with Southern Italian dialects. Instead, in the other Franco-Provençal varieties possessives can subsume the deictic/referential properties associated with the left domain of DPs; in Celle this structure characterizes kinship terms, which exclude the article relying on their relational content involving possessor, as in (5).

There are Southern Italian varieties that show an independent lexicalization of the possessive relation by inserting the preposition *di* (Baldi and Savoia forthcoming), borrowed by Cellese in indefinite contexts.

**References**


Circumventing the ‘that-trace’ effect: Different strategies between Germanic and Romance

Alessandra Tomaselli, Andrea Padovan Ermenegildo Bidese

<not updated>
New perspectives on word order flexibility

Savithry Namboodiripad, Natalia Levshina & Alex Kramer
(University of Michigan, MPI for Psycholinguistics, Nijmegen & University of Michigan)

Keywords: constituent order, flexibility, typology, quantitative methods, experimental methods

The bulk of existing research on word and constituent order typology assumes that languages should fit into categories such as VO, OV or lack of dominant order (cf. Greenberg 1963, Dryer 1992). Although there have been attempts to make finer-grained distinctions of word order flexibility in functional linguistics (e.g. flexible languages with and without dominant word order, cf. Mithun 1992, Siewierska & Uhlirova 1998), these studies are predominantly qualitative and compare a small set of languages. Similarly, psycholinguistic methods have been used to investigate the particular factors which influence processing, production, and comprehension of flexible word order within languages (e.g. Kaiser & Trueswell 2004, Miyamoto & Takahashi 2004, Yasunaga et al 2015), but have focused less on cross-linguistic comparisons or explanations for typological variation (though see Bates et al. 1982 and Šimík & Wierzba 2017).

As a wider set of empirical tools is applied to persistent questions in syntactic typology, it is possible to revisit typological distinctions that were previously treated as categorical, and to ask more targeted questions about the sources and dynamics of these distinctions. Recent work using corpus-based entropy measures (Futrell et al. 2015, Levshina 2019) and acceptability judgment experiments (Namboodiripad 2017, 2019) has argued for a gradient approach to word and constituent order flexibility that moves beyond categorical definitions such as “free” or “rigid.” Such approaches aim to describe word order more precisely, avoiding data reduction and a bias towards bimodally distributed patterns of preferred word order (cf. Wälchli 2009). They also enable us to model correlations between word order flexibility and other language-internal and language-external variables, such as case morphology and more general internal word structure (cf. Futrell et al. 2015, Koplenig et al. 2017). Finally, they allow us to answer outstanding questions about how and why word order patterns change over time (cf. Croft 2000), and what the role of flexibility might be in this process (cf. Heine 2008, Friedman 2006).

At the same time, new data and methods raise many pertinent questions about methods and mechanisms. What are the explanatory limits of a given quantitative measure? How do results from experimental and corpus-based studies compare, and what are their respective limitations? How can we take register, modality, and individual variation into account? The functional pressures and processes, such as processing efficiency, grammaticalization, entrenchment, learnability, etc., which are responsible for the emergence of probabilistic constraints (cf. Levshina 2019) are thrust into the foreground. This allows for new opportunities to connect individual-level processes with typological tendencies (Bybee 2010, Christiansen & Chater 2016), requiring conversation and collaboration between researchers who study the same phenomena at different timescales and across typologically distinct languages.

This workshop brings together novel research on variation in word and constituent order, asking about the range of cross-linguistic variation, the sources of this variation, and the methods that might be most appropriate to capture it. The papers represent a wide range of languages and methodological approaches, and each abstract considers what can be learned and what new questions can now be posed by taking this gradient approach. These questions include:
• What are the advantages and disadvantages of using different types of data to describe word order flexibility?
• Where do different methods converge? Where do they diverge?
• What is the appropriate level of lexical and syntactic granularity for measuring flexibility?
• How do communicative and cognitive constraints affect flexibility?
• What changes to theoretical approaches are necessary in order to capture flexibility?
• What is the relationship between flexibility and processing/production pressures?
• What contributes to group-level differences in word order flexibility?
• Are there discernable patterns in how flexibility changes due to language contact?
• How might a more granular approach to flexibility help to explain historical changes in word/constituent order?
• How are constituent order patterns distributed geographically?
• How do the cross-linguistic distribution of word orders with “hard” constraints mirror (or possibly diverge) from “soft” constraints in individual languages? (cf. Bresnan et al. 2001)

In order to tackle the big questions of how and why languages vary and change when it comes to word and constituent order, it is necessary to marshal evidence from distinct languages, methodological approaches, and, crucially, researchers who are looking at diverse aspects of this complex puzzle.

References
Friedman, V. (2006). West Rumelian Turkish in Macedonia and adjacent areas. *Turkic Languages in Contact, 61*, 27
Variable constituent order in Mapudungun: Semantics, information structure, and construals

Rodrigo Becerra
(University of Alberta)

Keywords: constituent order, information structure, semantics, spontaneous speech, Mapudungun.

Constituent order in Mapudungun, an isolate language spoken in southern Chile and Argentina, has been described as either highly flexible (Augusta 1903, and Rivano 1987, 1988) or free but presenting some preferential patterns (Smeets 2008). Using elicitation and acceptability judgments, Rivano (1987) described that transitive clauses can arrange noun phrases in almost any order. Similarly, using elicitation and narratives, Smeets recognized that different word orders are possible, although she proposed a preferential agent-verb-complement order in transitive clauses.

In this study, I identify the existent and the most frequent constituent orders in Mapudungun and analyze the factors that influence the preference for different orders. I focus both on intransitive and transitive clauses, considering clauses with full and partial codification of subject and object NPs. In order to achieve this, I perform a corpus-based study by using conversational data collected from 9 interviews done in Mapudungun, which constitute a corpus of more than 7 hours (55,000 words) altogether. All the participants are native-speakers of the language (including the interviewer: 5 women and 5 men, aged 45-85).

As a result, all SV, VS, OV and VO orders are attested. However, VS and VO are more frequent than V-final pairs in a ratio of 2:1. Firstly, SV orders are used when S is presented as a new discourse topic or a prominent one. On the contrary, VS order is preferred: i) whenever the subject argument codifies a discourse-active referent; or ii) when S is new but generic, nonspecific, or inanimate. In addition, VO is generally preferred, except for OVS cases, in which O is coreferential with the
discourse topic and is presented as the theme (topic) of the clause, and for SOV cases, in which both S and O are presented as salient. Secondly, when all the arguments are expressed as full NPs, VOS and SVO have the highest and similar frequency, OVS appears as the third most frequent order, and SOV as the least frequent one.

I claim that even though constituent order in Mapudungun can be rather flexible, it is by no means unconstrained. First of all, not all logically possible constituent orders were attested in this conversational corpus. No examples of VSO and its specular image, OSV were found. And second of all, constituent order is not unpredictable but constrained by semantics, information structure and the construal selected by the speaker. Moreover, it is set out from a range of options that follow established principles in the language: newsworthiness principle (Mithun 1992), and higher topicality towards the edges. Finally, the use of naturalistic data provided a fundamental basis to assess frequencies and the interaction between word order and semantic-pragmatic factors.

References
Augusta, Félix de (1903), Gramática araucana. Valdivia: Imprenta Central, J. Lampert.
Rivano, Emilio (1988), Morphosyntactic functions in Mapudungu, Revista RLA 26, 57-90.

Word order flexibility across types of argument realizations and argument structures

Laura Becker & Matías Guzmán Naranjo
(University of Erlangen-Nürnberg, Université de Paris, LLF, CNRS)

Keywords: word order, quantitative typology, argument structure, Universal Dependencies

Studies in word order typology have predominantly focused on transitive constructions with lexical subjects and objects (e.g. Dryer 1991, Dryer 1992). Embracing a gradient approach to word order preferences (e.g. Futrell et al. 2015, Guzmán Naranjo and Becker 2018, Levshina 2019), this talk focuses on two additional parameters that allow to better capture word order preferences within and across languages. Building on Levshina (2019), we examine the realization of arguments (pronominal, lexical, clausal), as well the argument structure of the verb (e.g. intransitive, transitive etc.) This allows us to assess to what extent word order flexibility holds across different realizations of arguments and argument structures, and to what extent this interacts with the position and realization of another argument. To do so, we use the Universal Dependencies, version 2.4 (Nivre et al. 2019), including data of 83 languages.

Selected results include the following crosslinguistically robust patterns for subjects: The preference to precede the verb decreases for different realizations from pronominal > lexical > clausal. Also, the verb-preceding tendency is the strongest for clauses that are transitive > intransitive > intransitive + indirect object. Transitive verbs having the strongest overall preference for verb-preceding subjects, this preference is more dependent on the position of the direct object than e.g. the indirect object in other argument structures. Another example to illustrate the need for a more fine-grained distinction of order preferences concerns OV/VO orders. Our results indicate that there is a
relation between the position of the subject and the relative position of the direct object in transitive constructions: The clauses in which lexical direct objects follow the verb have a higher proportion of subjects preceding the verb than those clauses in which the lexical direct object precedes the verb. For pronominal objects, this effect is much weaker. The position of indirect object, lexical or pronominal, on the other hand, does not relate to the position of the subject.

Those types of interactions let us gain a better understanding of the degrees of word order flexibility in individual languages. In addition, including more argument and construction types instead of restricting word order patterns to lexical arguments in transitive clauses shows and confirms stable crosslinguistic trends that can be accounted for by cognitive preferences such as shorter forms (old information) preceding longer forms (new information) (e.g. Hawkins 1994, Hawkins 2014). However, the fact that we see a different positional interaction between subjects and direct objects on the one hand and indirect objects on the other points towards the need for additional explanations along the lines of different semantic properties and discourse functions of S, A, P, and other arguments which can lead to different preferences in their lexical realizations and orders (e.g. Du Bois 1987, Du Bois 2003, Haig and Schnell 2016).

References
Word order variation in Baltic, Slavic, Germanic and Romance

Christian Ebert, Balthasar Bickel & Paul Widmer
(University of Zurich)

Keywords: Indo-European syntax, parallel corpus, areal convergence, phylogenetic signal, variation partitioning

Basic word order is often assumed to bear signals of areal convergence and/or genealogical distance, but the extent to which word order actually varies within and across languages, areas and genealogical groupings is not usually taken into account. In a bottom-up approach, we assess the syntactic variation between 19 languages of four Indo-European branches (Baltic, Slavic, Germanic, Romance) in 979 declarative sentences from a parallel corpus (ParTy corpus, Levshina 2017) and evaluate the impact of areal and phylogenetic relations on the (dis)similarities of these language vectors. Each sentence in the dataset is categorized as either verb-initial, verb-medial or verb-final, and for each language pair the log-transformed adjusted mutual information score is calculated, which ranks languages as similar to each other if the realization of a variable in one language is predictable by the other language.

Fuzzy clustering yields two clusters as the best solution for our data. The first cluster consists of Germanic languages and French (membership coefficients between 0.66 and 0.78). The second cluster contains Baltic, Slavic (except Russian) and Romance languages, with membership coefficients ranging from 0.61 to 0.74. Russian is almost equiprobably assigned to either of the two clusters with membership coefficients close to 0.5 (Figure).

To assess how much of the syntactic variation observed is constrained by geographical and/or phylogenetic proximity, we used variation partitioning (Legendre 2008). Geographical distances are calculated based on the language polygons provided by Ethnologue (Simons and Fennig 2018), with 1000 samples of random points drawn. For phylogenetic distances, we randomly samples 1000 trees from the posterior sample in Chang et al. (2015). After reducing the dimensions so that at least 80% of the variance is retained, the variation partitioning was calculated with a partial redundancy analysis for each predictor while controlling for the other predictor. The variation explained in both models is then partitioned between fractions which are exclusively attributed to phylogenetic and geographical distance and a third, shared fraction of variation that can not be attributed to either of the predictors alone.

A mean 60.17% of variation can be explained in total (mean adjusted $r^2 = 0.6$, standard deviation = 0.03). Phylogeny alone accounts for a median 0.22 fraction of variation (mad = 0.12) of variation, geography for a mean 0.03 (sd = 0.03). The fraction of variance shared by both predictors has a median of 0.35 (mad = 0.12). The fractions of phylogenetic and shared variation sum up to a mean of 0.58 (sd = 0.003) throughout all samples. In 763 out of 1000 samples, the fraction of variation attributed to geography exclusively is greater than zero. Thus, despite the substantial overlap with the phylogenetic signal, the spatial structure offers additional explanation for syntactic variation.

In contrast to established linguistic areas like Standard Average European, our analysis supports only the assumption of areal convergence for the Germanic languages and French, which shows a low level of similarity to other Romance languages. Apart from that, no effect of the geographical proximity of Germanic and neighboring Slavic or Romancelanguages can be detected. Our results suggest that word order variation is mostly affected by geographic proximity when it is reinforced by a phylogenetic component and historical contingencies (e.g. Frankish rule in the case of Germanic and French).
Map of fuzzy clustering coefficients. Language polygons according to Ethnologue (2018). Darker shades of red and blue indicate a high certainty of being assigned to one of the clusters. Lighter areas indicate equiprobable assignment to either of the two clusters. Gray areas indicate missing data.

References

Flexible word order in a contact situation: Investigating syntactic, semantic and pragmatic factors influencing word order in Arabic in a Persian-speaking environment

Dina El Zarka
(University of Graz)

Keywords: Arabic, Persian, word order, language contact, information structure

This preliminary study investigates factors influencing word order in hitherto undescribed Arabic vernaculars spoken in Southern Iran (Persian Gulf). The hypothesis is explored that these dialects
exhibit a beginning contact-induced shift in canonical word order from VO to OV due to the strong influence of Persian, which is the dominant language for most speakers in this region.

Although word order in Arabic vernaculars is fairly variable and strongly dependent on discourse requirements, OV is usually not an option or at least heavily constrained (Dahlgren, 1999; Brustad, 2000). A beginning shift in word order has already been proposed by Matras & Shabibi (2007) for the neighboring variety of Khuzestan. Recent work on this variety (Leitner, forthcoming), however, suggests that VO order is still very robust and that instances of verb-final clauses are limited to a small number of constructions. By contrast, canonical word order in Arabic varieties of eastern Iran and central Asia has been shown to have completely shifted to OV (Ratcliffe, 2005; Seeger, 2013).

Spontaneous speech data from four villages was annotated for syntactic, semantic and pragmatic factors to assess their influence on word order. Due to the small amount of data, we did not apply statistics, but only present a qualitative analysis and frequency counts. The preliminary results, however, reveal a few interesting facts that will be used as the starting point for further quantitative research. Canonical VO order still seems quite stable in South Iranian Arabic, however, with a change in progress in certain constructions, specifically qualitative predications lacking a full lexical verb, such as copular, existential and possessive constructions with a prepositional phrase functioning as a dummy verb. The data also contains a high percentage of clause-final copulae and auxiliaries that are placed after the lexical verb (1). This fact is in line with Haig and Khan’s classification of this phenomenon as an areal feature of Western Asia (Haig & Khan, 2019). It has also been found in other Arabic varieties in close contact with Iranian and Turkic languages, viz. in eastern Anatolia and northern Iraq (Procházka, 2019a, 2019b).

(1) MKF_9
riḥna rih-na n-sawwi-hi čin-na
1PL self-1PL 1PL.IPFV-make-3SG.M be-1PL.IPV
‘we did it all by ourselves.’

Additionally, the data also contains a small number of true OV clauses like in (2).

(2) MFK_151
harīsa wi-geymāt a-sawwi
porridge and-profiteroles 1SG.IPV-make
‘I make porridge and profiteroles.’

Finally, the study investigates the semantic and pragmatic factors that influence word order choice, suggesting a strong effect of information structure in terms of narrow focus and theme-rheme partition and an effect of semantic role, specifically regarding the preferred placement of goals and recipients in post-verbal position (Haig 2014, Haig & Khan 2019). Example (3) shows that narrow focus (indicated by capital letters) can even override the preferred post-verbal position for goal complements.

(3) HJM_26
HINĀK ga’ad
there sit.3SG.M.IPV
‘THERE he settled.’

References
The effect of register on dependency length in two flexible languages

Alex Kramer
(University of Michigan)

Keywords: Corpus linguistics, Japanese, Russian, constituent order, comparative

Recent corpus studies employing dependency corpora have found that SOV-dominant languages such as Japanese and Korean have longer dependency lengths and more rigid word order than might be expected given their typological profiles (Futrell et al. 2015 and Levshina 2019). These languages share features such as argument drop and flexible constituent order, which can result in shorter dependency lengths. However, speakers may be less likely to use these features in more formal registers (Biber 1993 and Wälchli 2009), and indeed, the Japanese data used in these studies came from formal written and spoken sources.

In order to investigate the effect of genre on dependency length, I constructed a corpus of informal spoken Japanese by scraping the captions from the videos of popular YouTubers. Captions were pre-processed in Python and parsed using the StanfordNLP dependency parser (Qi et al. 2018). Replicating the methods from Futrell et al. (2015), I calculated and compared the total combined length of the dependencies (a) in each observed sentence, (b) a random permutation of each sentence, and (c) an “optimized” permutation of each sentence that minimized the distances between heads and dependents.

Comparison of the observed dependencies to the random and optimal baselines revealed overall shorter dependency lengths than found in Futrell et al. (2015), in which more head-final languages like Japanese, Korean, and Turkish showed less minimization than more head-initial languages. In fact, it is likely that the true Japanese dependencies are even shorter: Manual analysis found that non-
canonical orders, many of which reduce dependency length, are common in the data; however, the parser mis-parses main clause verbs in these orders as modifiers, artificially lengthening dependencies. Japanese allows flexible order and argument dropping, both of which may affect dependency lengths; argument drop, which is influenced by register and context, is extremely common in spoken Japanese (Nariyama 2000). To disentangle the effects of these two factors and extend this comparison to another flexible language, I replicated this method with Russian, a language with more restricted argument dropping. Unlike Japanese, Russian dependencies patterned similarly to the results of Futrell et al. (2015). While not a perfect comparison (e.g. Russian is canonically SVO; Japanese is SOV), these results may suggest an important role for argument dropping in dependency length minimization in spoken Japanese.

Diversifying the dependency corpora used in typological research via YouTube captions allowed for the comparison of informal spoken Japanese and Russian, which revealed different patterns of dependency length minimization across languages and registers. Future work aims to improve the parser’s ability to handle non-canonical sentence orders and other aspects of informal speech. In addition, to understand how features such as argument dropping and flexibility contribute to minimizing dependency lengths, this method should be extended to languages that vary systematically with respect to canonical order, head- versus dependent-marking, and other potentially relevant features. As computational models improve, informal sources such as YouTube have the potential to become powerful tools for uncovering typological patterns that are not present in more formal registers and written modalities.

References
How efficient are human languages? Testing trade-offs between word order flexibility, case marking and semantic cues

Natalia Levshina

<not updated>

The Relationship between Word Order Flexibility and Dependency Length Minimization

Zoey Liu
(University of California, Davis)

Keywords: dependency length minimization; word order flexibility; crosslinguistic

Based on the Principle of Dependency Length Minimization (DLM), human languages have the tendency to shorten the distance between syntactic heads and their dependents. Though previous studies have validated the typological preference for DLM across languages (Futrell, 2015), recent experiments have shown that the extent of DLM varies for languages with different structural characteristics (Liu, in press). One potential explanation for the typological variations of DLM conjectured by previous research is word order flexibility, that certain languages have longer dependencies due to their higher degree of word order freedom (Gildea and Temperley, 2010). This study examines whether this conjecture is quantitatively valid. Specifically, I explore the gradient relationship between word order variability and DLM. I ask: do syntactic constructions with more ordering flexibility have a weaker preference for DLM crosslinguistically?

I addressed the question with multilingual corpora for 36 languages from the Universal Dependencies project version 2.5. I used as test cases constructions in which the head verb has a direct object NP and exactly one PP oblique dependent appearing on the same side, with the two dependents adjacent to each other. Here the NP is considered to be more argument-like than the PP.

1. Kobe praised [NP his oldest daughter ] [PP from the stands ].
2. Kobe praised [PP from the stands ] [NP his oldest daughter ].

3. [NP 朝食 で ] [PP パン を ] 食べる
   chōshoku de pan o taberu
   (I) morning in bread -ACC eat

4. [PP パン を ] [NP 朝食 で ] 食べる
   pan o chōshoku de taberu
   (I) bread -ACC morning in eat

`I eat bread for breakfast.`
After preprocessing, the dataset contained 36 languages in six language families: Indo-European (29), Afro-Asiatic (2), Austronesian (1), Niger-Congo (1), Uralic (2) and Japanese (1). The dataset exhibits three ordering patterns: (1) one for languages with NP and PP after the head verb; (2) one for languages with NP and PP before the head verb; (3) one for languages with NP and PP both before and after the head verb.

I approximated ordering flexibility of the NP and the PP within every language with entropy (Levshina, 2019). Estimates for 95% confidence intervals were computed with bootstrapping for 1,000,000 iterations. The extent for DLM was computed as the predictive power of dependency length in the relative order of the NP and the PP. To do this I used logistic regression, while controlling for the effects of two automatically measurable factors: argument status and pronominality, as well as the random effect of the head verb. Estimates for 95% confidence interval of each coefficient were derived with bootstrapping for 10,000 iterations.

Overall, results show a negative correlation between ordering variability and the extent of DLM (Spearman’s ρ = -0.58). In particular, there is on average more ordering flexibility in preverbal domains in contrast to postverbal domains. The effect for dependency length is much weaker in preverbal orderings than that in postverbal orderings.

References

Word-order flexibility in genitive noun phrases:
A corpus-based investigation of contact varieties of Russian

Chiara Naccarato, Anastasia Panova & Natalia Stoynova
(HSE University, Moscow; HSE University, Moscow; & V. V. Vinogradov Russian Language Institute RAS, Institute of Linguistics RAS and HSE University, Moscow)

Keywords: word-order variation, genitive noun phrases, spoken Russian, spoken corpora, language contact

Contact varieties of Russian feature word-order patterns that are not typical of Standard Russian. In this paper, we focus on word-order variation in noun phrases with a genitive modifier. Bilingual speakers of left-branching languages of Russia often use prepositive genitives (GEN+N) in their Russian speech (1). By contrast, in standard monolinguals’ Russian the neutral word order is N+GEN (2).

(1) *Ona vsju žizn’ ne ljubila deduški sestru*

3SG.F all.F.ACC life.ACC NEG love.PST.SG.F
The first hypothesis is that we deal with a simple calquing of the L1 syntactic pattern (3).

(3) \( \lambda \varepsilon \ \text{de-} \chi\varepsilon \ \text{mahandi-}l \ q'alami! \)

\[
\begin{array}{ll}
give.IMP & 1SG-LOC \\
\text{Muhammad-} & \text{pencil}
\end{array}
\]

‘Give me Muhammad’s pencil!’

[From Testelec 1998: 274, Andalal dialect of Avar]

However, in spoken monolinguals’ Russian, especially in dialectal varieties, the pattern GEN+N is also attested. So, the alternative hypothesis is that prepositive genitives in bilinguals’ speech are actually motivated by a weaker pressure of prescriptive norms and the type of texts under consideration (oral informal communication).

To check these two hypotheses, we present a quantitative analysis based on data from spoken corpora of different contact varieties of Russian: Daghestanian (the largest corpus, which was analyzed in more detail), Chuvash, Bashkir, Tungusic (smaller corpora, which differ from the Daghestanian one in terms of text type and sociolinguistics). To compare these varieties with monolinguals’ Russian, we analyzed spoken corpora of Standard Russian and several dialectal varieties (Malinino, Ustyay, Rogovatka, Opochka, Spiridonova Buda).

The results suggest that non-standard word order in contact varieties is probably due to the interaction of both contact and non-contact factors. The comparison of bilinguals’ data with standard monolinguals’ data partly supports the first hypothesis: the rate of prepositive genitives is significantly higher in bilinguals’ data. However, the mechanisms regulating the choice of a specific word order are not directly related to contact. Rather, they are similar to those conditioning differential marking within the noun phrase in non-contact situations: semantic features of the genitive modifier and the head noun (e.g. lexical class, animacy, definiteness), and processing factors (length, givenness). Dialectal data, by contrast, favor the second hypothesis: in some varieties, the rate of prepositive genitives is comparable to that of bilinguals’ Russian.

The paper was prepared within the framework of the HSE University Basic Research Program and funded by the Russian Academic Excellence Project ‘5-100’.

References


Gradience in constituent order within and across languages: Insights from acceptability judgment experiment

Savithry Namboodiripad

<not updated>

What drives word order flexibility? Evidence from sentence production experiments in two Australian Indigenous languages

Rachel Nordlinger, Gabriela Garrido Rodriguez, Sasha Wilmoth & Evan Kidd
(University of Melbourne, University of Melbourne, University of Melbourne & Max Planck Institute for Psycholinguistics and ANU)

Keywords: word order flexibility, sentence production, eye-tracking, Australian languages, animacy
In this paper we discuss flexible word order in Murrinhpatha and Pitjantjatjara, two Indigenous languages of Australia, reporting experimental work examining sentence production and word order using eye-tracking. Recent psycholinguistic studies have revealed significant cross-linguistic differences in sentence processing that are correlated with a language’s word order properties (Sauppe et al. 2013, Norcliffe et al. 2015, Sauppe 2017). This suggests that sentence production studies can provide useful evidence for determining underlying syntactic properties of languages. Moreover, since experimentally controlled conditions can reduce the possibility of discourse-related effects, we can investigate word order flexibility independently of information structure to better understand other factors that may influence word order choice in these radically ‘free’ word order languages (Hale 1983, Heath 1986, Austin & Bresnan 1996).

In our experiments, we asked participants (Murrinhpatha, N = 43; Pitjantjatjara, N = 50) to describe pictures of transitive events while their eye-movements to the characters were recorded with an eye-tracker. There were 48 test items that were manipulated for agent and patient humanness (+/- human) interspersed amongst 93 filler pictures that mostly depicted intransitive events. The resulting picture descriptions were transcribed and coded for word order, and participants’ eye-movements were analysed. Our results suggest no default, underlying word order in these languages: we find all possible orderings of S, O and V in the experimental corpus, and often many from a single speaker. In Murrinhpatha, for example, while SVO was the most common order (45.2%), several other orders were commonly used (SOV: 14.6%, OVS: 8%) and no word order occurred more than 50% of the time. This result is striking given data from similar studies on other languages with word order flexibility (e.g. German), where speakers overwhelmingly use one dominant word order (>75%) and one or two other word orders in very limited circumstances (Sauppe 2017).

A number of factors influenced word order choice. These include: (i) conceptual accessibility: speakers placed human participants in more prominent sentence-initial locations, although this interacted with other features of the language such as the tendency to ellipse human NPs; (ii) bottom-up perceptual cues: participants’ early construal of the event as reflected in their eye-movement patterns was related to resulting word order choice, and (iii) individual speaker preferences. Our analyses of participants’ eye-movements suggest that the highly flexible nature of word order in these languages is reflected at the earliest observable moment – the first 600ms after picture onset, during event apprehension. These early eye-tracking patterns contrast with the findings from other languages with more fixed word orders (Sauppe et al. 2013; Norcliffe et al. 2015; Sauppe 2017), since we see both relational encoding as well as the signature of the resulting word order very early in the eye-tracking record. This study therefore contributes to the limited literature on sentence production and word order flexibility and provides a useful methodology for investigating the ordering patterns employed by speakers of languages where ordering is not syntactically determined.

Acknowledgement: This research was funded by the ARC Centre of Excellence for the Dynamics of Language (Project ID: CE140100041).

References


Constituent order and information structure in an Australian language: Implications for constituent order typology

Eva Schultze-Berndt  
(University of Manchester)

Keywords: constituent order; information structure; topic; discourse annotation; Australian languages

This paper reports on a corpus study of the patterns of clausal constituent order and information structure in Jaminjung/Ngaliwurru, a morphologically ergative, head-marking language of the Mirndi family (Australia). Methodologically, it pioneers a corpus annotation scheme resulting from the combination of the GRAID annotation scheme for referential expression type, animacy status and grammatical role (Haig & Schnell 2014) with basic information structure annotation. Left-edge topics, right-edge topics, narrow and broad focus, and afterthoughts (prosodically detached focal right-edge constituents) are distinguished on the basis of prosodic and discourse-functional criteria. The data consists of 114 texts totalling over 12,000 intonation units. The preliminary analysis concentrates on the instances of transitive clauses with two overt argument phrases (n=121) – the low figure confirming Du Bois’ (1987; 2003: 35–38) “Lexical A constraint”, which also predicts the low frequency of overt transitive subject phrases in the entire corpus (n=531).

Overall, the existence of the six possible constituent orders confirms the assessment of Jaminjung/Ngaliwurru as a language with flexible constituent order, at least if one follows Dryer’s (1997) criterion according to which a dominant order has to be at least twice as frequent as any other. However, the frequencies as shown in (1) reveal clear preferences for S-initial orders and a dispreference for V-initial orders:

(1) SVO (38%) >> SOV (23%) / OVS (24%) >> OSV (5%) / VOS (6%) / VSO (4%)

These (dis)preferences, at first sight, are in line with cross-linguistic tendencies for constituent order in transitive clauses (Dryer 2013), which in turn confirm Greenberg’s (1966 [1963]: 77) first universal predicting a tendency for subject-object order. More recently, this universal has been re-cast as the “agent-first principle”, with support from psycholinguistic studies (Bornkessel-Schlesewsky & Schlesewsky 2015). The preferred constituent orders in Jaminjung/Ngaliwurru also appear to confirm the verb-object constraint as proposed by Baker (2009).

A closer look at the information structure configurations associated with the different constituent orders reveals that the preferred orders are the result of three tendencies:

(i) Clauses with two overt arguments have left-edge topics (51%)

(ii) Topics in two-argument clauses tend to be agents/subjects (80% of the topic-initial clauses in dataset)
(iii) Two arguments on one side of the verb are dispreferred (unless one of them is a topic or an afterthought).

The data do not support a general agent-first principle: if only topicless clauses are taken into account, any preference for clause-initial transitive subjects disappears. Rather, the findings support the importance of information structure for “soft” universal positional constraints at clause level (which may have become generalised across all subjects and objects in languages with a more rigid constituent order). Specifically, the findings support the view that cross-linguistic preferences for subject-initial order and verb-object contiguity result from the tendency for topics to be transitive subjects and for objects to be part of the rhematic part of an utterance.

References
Baker, Mark (2009), Language universals: Abstract but not mythological, Behavioral and Brain Sciences, 32, 448–449.

Rigid vs. free word order in modern Indo-European languages: an information theoretic measure of the relative position of selected syntactic relations in a multilingual, parallel corpus of literary fiction

Luigi Talamo & Annemarie Verkerk
(Universität des Saarlandes)

Keywords: word order, parallel corpus, corpus-based typology, Universal Dependencies, information theoretic measure
As recently pointed out by Levshina (2019), typological databases often offer a simplified, abstract view on structures of world’s languages, reducing the complexity of features to single categorical datapoints. For instance, WALS map no. 86 ‘Order of Noun and Genitive’ classifies English and Dutch, two closely related Germanic languages, as ‘no dominant order’ and ‘Noun-Genitive’ respectively, whereas the ‘Genitive-Noun’ type is widely attested in Dutch as well: mijn moeder’s fiets ‘my mother’s bike’ vs. de fiets van mijn moeder ‘the bike of my mother’. A straightforward solution would be to classify the order of Genitive and Noun in Dutch as in English i.e. as ‘no dominant order’; however, this gives us little information on how the two different orders are actually employed. It has been suggested that such word order variability can be understood through a corpus-based typology, in which traditional comparative concepts are remodeled as token-based comparative concepts. As with other kinds of token-based comparative concepts (visual stimuli, translation contexts used in questionnaires: Dahl 2007, Haspelmath 2018), the functional-semantics content of the corpus tokens should be the same across languages. Accordingly, we employ for our study a parallel corpus featuring translations of 18 literary texts in a balanced sample of seven Indo-European languages (three genera: Germanic, Greek and Italic); each language is represented by at least one original text, whose translation is provided for every language of the corpus. In contrast to previous parallel corpus-based studies (biblical texts: Wälchli and Cysouw 2012, religious magazines: Agić and Vulić 2019, movie subtitles: Levshina 2015, news: Levshina 2019), our study explores the genre of literary fiction, focusing on the intra- and inter-linguistic comparison of — in traditional terms — rigid syntactic constructions: Determiner-Noun and Adposition-Nominal Phrase vs. free syntactic constructions: Adjective-Noun and Relative Clause-Noun. Whereas most studies (Montemurro and Zanette 2011, Cohen Priva and Gleason 2016, Bentz et al. 2017), have privileged n-gram models of entropy, we propose here a vector space model to study word order. Such models are best known in lexical and semantics studies (see Lenci 2018 for an overview), modeling the semantics of lexical items as multidimensional vectors representing distributions (co-occurrence with other lexical items); dealing with just two possible values, our vector space model is bidimensional, taking into account the relative positions of syntactic constructions as annotated in the corpus according to the Universal Dependeny Relations (Marneffe et al. 2014). For instance, the order of the Determiner-Noun is modeled as a vector composed by the probabilities of det-X and X-det where X is a relevant category modified by the Determiner (in order to get a clearer signal, we filter out clear mistakes in syntactic analysis such as verbs that are modified by determiners). In order to quantify the variation of the selected syntactic constructions, we employ entropy measures on our vector space models. Our preliminary findings correlate with general tendencies described in the traditional typological literature, showing however a more complex scenario, in which rigid syntactic constructions such as Adposition-Nominal Phrase display a substantial amount of entropy while Relative Clause-Noun show more rigidity than expected.

References
Recent works on word order flexibility tend to use a gradient approach to provide a more accurate picture of word order patterns in language (Östling 2015, Futrell et al 2015, Hammarström 2016, Ferrer-i Cancho 2017, Levshina, 2019). Most of these studies are based on large-scale annotated corpora such as the Universal Dependencies (Nivre et al 2019). Corpus size and measurements of word entropy are frequently mentioned as a potential source of statistical biases. As an example, “the corpora available vary hugely in their sample size, from 1017 sentences of Irish to 82,451 sentences of Czech. An entropy difference between one language and another might be the result of sample size differences, rather than a real linguistic difference” (Futrell et al 2015: 93-94). Nevertheless, the size of corpora does not seem to play a major role in the literature, c.f., “The fact that entropy measures on subcorpora of 1000 sentences do not diverge greatly from entropy measures on full corpora indicates that this measure is stable with respect to sample size” (Futrell et al 2015: 98). Similar comments are also found for the different measures of entropy but most studies tend to select only one measure of entropy in their analysis.

This study takes the word order of subjects, verbs, and objects as a case study to provide quantitative evidence with regard to the effect of corpus size on different entropy measures. The variance of constituent order entropy is assessed on 34 languages by testing 50 corpus sizes and 8 entropy measurements. The study answers the call of the workshop What are the advantages and disadvantages of using different types of data to describe word order flexibility? and Where do
different methods converge? Where do they diverge? The languages included in this study are extracted from the Universal Dependencies: Arabic, Basque, Bulgarian, Catalan, Chinese, Croatian, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hebrew, Hindi, Indonesian, Italian, Japanese, Korean, Latin, Latvian, Norwegian, Persian, Polish, Portuguese, Romanian, Russian, Serbian, Slovak, Slovenian, Spanish, Swedish, and Ukrainian. The entropy methods included are maximum likelihood, bias-corrected maximum likelihood, Shannon entropy by plug-in of Bayesian estimates of the bin frequencies using the Dirichlet-multinomial pseudocount model with different priors (Jeffrey, Laplace, Schurmann-Grassberger, minimax), the Chao-Shen estimator, and the shrink (James-Stein-type) estimator. The corpus sizes vary from 100 randomly extracted sentences to 5000 sentences. Different methods of random sampling are also compared. The results indicate that different corpus sizes and entropy methods do not have a big impact on the constituent order entropy cross-linguistically. However, some languages are more sensitive than others to the change of experiment settings.

References
Ferrer-i Cancho, Ramon (2017), The placement of the head that maximizes predictability: An information theoretic approach, Glottometrics, 39, 38–71.
Levshina, Natalia (2019), Universal Dependencies in a galaxy far, far away... What makes Yoda’s English truly alien, Proceedings of the Third Workshop on Universal Dependencies, 35–45.
Even a perfunctory browsing through the research literature reveals an incredibly high number of articles tackling the current influence of English on various languages, thus making it a self-evident truth that over the last decades English has acquired the status of a true global language. This could only please those who have dreamt of a universal language, but there are numerous people who have swallowed the idea that the influence of English is unavoidable with a pinch of salt.

The influence of English on various languages has been the topic of numerous studies in all subfields of linguistics and related fields over the last decades: Otheguy and Zentella (2012) discuss Spanish use of personal pronouns and more complex phenomena triggered by Spanish/English bilingualism in New York; Konieczna (2012), Jaworski (2014) focus on word formation and other morphological aspects of Polish, Vakareliyska & Kapatsinski (2014) focus on Bulgarian, Veisbergs (2006) on Latvian and Saugera (2017) on particular aspect of French, whereas Niculescu-Gorpin and Vasileanu (2018) discuss experimentally the processing of luxury Anglicisms in Romanian, to name just a few.

However, the topic is far from being exhausted as the influence is on-going and it is generating significant language change that can be studied from a myriad of perspectives ranging from purely descriptive or normative ones to language acquisition or psycholinguistics, thus discussing the complex and multifaceted influence of English on various languages and developing new insights on the phenomenon.

Such a pervasive influence of a language on other languages has led to a new and unique situation of intensified bi- and multilingualism in the target languages, an aspect that calls for discussion. The phenomenon has deep ramifications for all speakers and it changes the way their mother tongue functions and the way they use their mother tongue.

The workshop aims to bring together researchers whose research focuses on the current transformation many languages undergo due to immediate or mediated contact with English and on the synchronic and/or interdisciplinary study of Anglicisms, lexical or other. We thus encourage any contribution, be it descriptive or multidisciplinary, normative or experimental that discusses the influence of English on other languages and brings new insights into the phenomenon.

We intend to attract researchers focusing on the English influence on any language to be able to provide a sound overview of the current globalising effect of English that triggers a vast array of language changes that can be studied from different perspectives. And this is the underlying reason that has made us launch a call for papers that have at their core only the English influence, and not a particular theoretical approach. The phenomenon under discussing here needs a heterogenous and multidisciplinary approach.

For the purpose of this workshop, Anglicisms are defined as any spelling, phonetic, morpho-syntactic or lexical phenomenon whose origin can be traced back to English and which have not been completely adapted and embraced by the target language.

By considering the current bi- and multilingual situations from most countries due to globalization, the national language policies and local reactions to the English influence, and the lexical, morpho-
syntactic and other changes that are triggered by the influence of English, we are trying to offer some possible explanations for the speakers’ attitudes towards using their mother tongue vs. the English language, for the way in English discourse markers borrowed in other languages behave in the target language, offering a variety of approaches to the phenomenon under discussion. We also focus on experimental studies of, as well as on more descriptive and theoretical approaches to the English influence.

The abstracts that are currently building up the present workshop are centred on the speakers’ perception and processing of the English influence, be it by means of psycholinguistic experiments, or by corpus analysis. The corpora the authors build up their analyses on are made up of real life texts and speech, some of them compiled from the very modern and productive new environment of social interaction and communication: social media (Facebook, Twitter, blogs, etc.), a hybrid communicative setting that is currently systematically studied as it opens up new interactional situations, that most often are online interactions between the participants, but they do not presuppose direct or face-to-face interactions.

As the main objective of the workshop is to offer a complex overview of the influence of the English language, and to provide some possible explanations related to the effects of this influence, we particularly welcome submissions on the English influence on any language provided they offer an explicit and considerable theoretical, applied or experimental contribution to the issue at hand.

The proposed topics are listed below, but we are open to other approaches or perspectives that may be relevant to the study of the English influence and of Anglicisms:
- speakers’ attitudes towards mother tongue vs. the English language;
- experimental approaches to the English influence, to Anglicisms etc.
- bilingualism, multilingualism and the English influence;
- (possible) psycho- and socio-linguistic explanations for language change phenomena triggered by the English influence;
- lexical Anglicisms: diachronic, synchronic, normative and/or descriptive studies;
- English discourse markers borrowed in other languages and their behaviour;
- ‘luxury’ vs. ‘necessary’ Anglicisms;
- the English influence, Anglicisms and corpus analysis;
- national language policies and local reactions to the English influence;
- lexical, morpho-syntactic and other changes due to the influence of English, and their study (any approach);

The workshop *Theoretical, Applied and Experimental Perspectives on the Influence of English Today* proposes an interdisciplinary, round exposé of the different aspects of the English influence on other languages, focusing on sociological, discoursal, and psychological aspects of the phenomenon that has led to new instances of grammaticalization, variation within and across languages, as well as pragmatic shifts.
The role of the cognates in facilitating the English influence on the Romanian morphosyntax

Ana-Maria Barbu
(Institute of Linguistics “Iorgu Iordan – Al. Rosetti”, Romanian Academy)

Keywords: compounds, cognates, English influence, Romanian, morphosyntax

Our analysis addresses two topics less studied, namely the deep English influences recorded at the morphosyntactic level and how cognates (i.e. words in different languages with considerable phonological/orthographic and semantic overlap, e.g. En. comedy, Ro. comedie, Fi. komedia) facilitate these influences. Studies that approach other aspects than lexical influences are considerably less and recent (Renner 2018, Vakareliyska 2014, Bombi 2017, Witalisz 2018). As regarding cognates, a cognate facilitation effect has been highlighted in bilingualism studies (Costa 2000, Assche 2016), but we extend the analysis of cognate facilitation to the integration of foreign structures in a host language.

More precisely, we analyse the influence of English compounds of type X Y (e.g. stand-up comedy, project manager, radio show etc.), which contain at least one cognate, on Romanian. Note that these compounds that are right-headed: *stand-up comedy (the head is underlined) conflict with some typical features of Romanian:

- Compounds are left-headed: *cuvânt cheie (word key, ‘keyword’), *lapte de pasăre (milk of bird, ‘snow eggs’).
- The definite enclitic article stays, both in compounds and free phrases, on the left element: *cuvânt-ul cheie (word-the key), *lapte-le de pasăre (milk-the of bird), except some joined compounds: *scurtmetraj-ul (short-film-the).
- A categorial modifier never precedes the head: *națională economie (‘national economy’), but only economie națională.

The way in which cognates facilitate the English influence on Romanian is illustrated by the stand-up comedy loanword (having 150 occurrences (occ.) in corpus):

a. The Romanian cognate comedie replaces comedy and brings out the right-headed semi-Romanian compound stand-up comedie. Its meaning is “comedie de (tip) stand-up” ‘comedy of the stand-up type’.

b. Stand-up comedy is treated as a tight unit and the replacement of comedy with comedie has been done probably for a better accommodation of the enclitic article on the right element. Compare the 2 occurrences of stand-up comedy-ei ART GEN (article is assigned to the English word) with the 61 occurrences of stand-up comediei (Romanian articulated word).

c. The native left-head order, i.e. comedie stand-up, was restored, so that both variants equally coexist: stand-up comedie (106 occ.) / comedie stand-up (131 occ.).

d. The lexical family: stand-up comedy, stand-up comic, stand-up comedian determined the stand-up element to become a stand-alone word.

The facts in a.-d. illustrate the following English influences:

1. The number of right-headed compounds grows up (see a. and also one-man show, project/proiect manager etc.).
2. A new meaning type appears in Romanian: X Y means “Y de X” (‘Y of X’) (a.)
3. Categorial modifiers precede the head (see a. and also națională arena ‘national arena’, radio show etc.).
4. The enclitic article moves on the right element of the compound even if its elements are not joined (b.).

5. The free order (c.) and the autonomy (d.) of the elements show how English influences on Romanian morphosyntax.

The full study inventories compounds of this type and researches their behaviour in the Romanian corpus Corola (corola.racai.ro) and on the Internet, in the way described for stand-up comedy, in order to determine the extent of this phenomenon.

References


English discourse markers in spoken Romanian: pragmatic borrowings or a code-switching phenomenon?

Valentina Cojocaru

<not updated>

Coming to terms with Anglicisms

Henrik Gottlieb
(University of Copenhagen)

Keywords: Language contact, impact of English, Anglicisms, code-switching, Global Anglicism Database (GLAD)

In order to determine the actual influence of English in today’s so-called globalized world and qualify academic research in the field, we not only need more empirical studies; we also need to sharpen our terminological tools, including our definition of the core notion of ‘Anglicism’.
This presentation takes a critical, yet constructive look at a number of key concepts used in modern Anglicism research – a field experiencing growing attention in academia (GLAD 2020). My aim is to promote a sound scholarly approach to the much-debated question of English (i.e. Anglo-American) impact on languages outside the Anglosphere – an issue often tackled, in the past as even today, in a subjective and pre-scientific fashion (Dunger 1899; Kristiansen 2013).

In an attempt at refining our terminology in this field, I will use a tripartite taxonomy of Anglicisms aimed at encompassing all types of impact from Anglophone sources in any type of speech community worldwide. This taxonomy is based on the following definition of ‘Anglicism’ – including all phenomena resulting from direct or indirect influence from English: “An Anglicism is any individual or systemic language feature adapted or adopted from English – or inspired or boosted by English models – used in intralingual communication in a speech community in which English is not the home language” (Gottlieb 2020: 32).

Apart from ‘Anglicism’, several other concepts commonly used in scholarly discourse in this field will be scrutinized in this presentation. As a case in point, when discussing the global English influence, the standard term ‘language contact’ is a misnomer at two levels:

(1) The unwarranted personification of language may lead to emotional and erroneous statements; the term ‘speech community’ is preferable as it focuses on in vivo usage rather than conjuring up an in vitro phantasm of language as an independent entity.

(2) The seemingly balanced notion of contact is rarely found whenever English is involved. Even in societies in which English is the home language of only a minority, let alone all the countries in which English is a foreign or second language, English may come out on top (Rousseau 1937, de Swaan 2010). The fact remains that globally, the contact between English and all other languages is largely unidirectional; the impact on English is negligible while almost every speech community on earth is open to English influence.

In today’s Europe – as elsewhere in the world – people have embraced English to the extent that English influence is found not only at a lexical and phraseological level (Furiassi et al. 2012); also semantics, pragmatics, and grammar are being impacted by English (Peterson & Beers-Fägersten 2018). Moreover, in those of Europe’s speech communities presently moving from EFL to ESL status, the ‘outlandish’ phenomenon of code-switching into and out of English is now on the rise (Knospe 2014).

In conclusion, not only are many of today’s Anglicisms found ‘underground’, in the shape of semantic loans and morphosyntactic calques; we also have to come to terms with the fact that English influence has ‘left the ground’, as code-switching may redefine the borders between home and second language among Europeans.

References


Borrowing and disappearance of light verbs

Aaditya Kulkarni
(M. S. University of Baroda)

Keywords: borrowing; south Asian languages; light verbs; verb licensing; homophony.

South Asian languages show a tendency of borrowing English words to form Noun-Verb complex predicates (NV CPs), which generally comprise of a nominal and a light verb. An English noun is borrowed directly into Kannada or Hindi-Urdu and can take the host language’s nominal inflections. English verbs however need the support of a light verb which carry the necessary verbal inflections of the host language (as seen in (1) and (2) for Kannada).

1.  *nanu  i:    post-i게    apply-i isSuccess
    I this post-DAT apply-AUX.PST-1.SG
    ’I have applied to this post.’

2.  nanu  i: post-i게    apply ma isSuccess
    I this post-DAT apply do-AUX.PST-1.SG
    ’I have applied to this post.’ (Amritavalli 2017)

Similar observations are made in Bhattacharya and Åfarli (2016), where borrowed English verbs cannot appear without a light verb in Bangla, but borrowed Hindi verbs can; whereas Norwegian can inflect borrowed English verbs directly.

This is explained by Amritavalli by combining the proposals of Kayne (2008) and Bhattacharya and Åfarli (2016). She proposes that a verb, borrowed or otherwise, is licensed by the verbal functional structure of the particular language. As a borrowed English verb appears directly with Norwegian morphology, there must be a verbal functional structure common to both the languages; whereas the inability of English verbs to take the verbal inflection of Indian languages is indicative of a microlevel variance between English and the host languages, which is then resolved by re-verbalizing the borrowed verb by an overt light verb.

However, languages like Marathi (3 and 4) and Bhojpuri (5) can do away with the requirement of the ‘licenser’ light verb while borrowing English verbs, which then directly takes the host language’s verbal inflections.

3.  mi  जा-ला  एका post-with  टॆग ke-1-ə 1.SG 3.SG.M-DAT one post-on tag do-PRF-N
    ‘I tagged him in a post.’

4.  जुम्हा-ला  टॆग-1-ə
2. HON-DAT tag-PRF-be.PRS.N ‘You have been tagged.’

5. eni taim mɯːɖ-wa ko apǝt-ao nəhɪ mɯ:ra
   any time mood-CLF DAT upset-IMP NEG.AUX naïve/innocent
   ‘O innocent one, don’t get upset.’ (lit. Don’t let your mood get upset) (Gangs of Wasseypur 2)

This indicates that Marathi and Bhojpuri exhibit optionality in re-verbalizing (or licensing) some borrowed English verbs. Questions arise as to
   i. why is a particular verb deemed optional for re-verbalizing, and
   ii. are there any (formal) constraints governing their licensing

The present paper looks at the theoretical implicatures of this phenomenon, and argues for a set of constraints on the occurrence of English verbs with Marathi verbal inflections. It argues against purely sociolinguistic explanations presented in the literature (such as Wichmann and Wohlgemuth’s 2007) and shows that the apparent direct licensing of English verbs is governed by certain constraints such as the frequency of usage and phonological shape of the borrowed verbs. The recipient language seems to resist direct licensing of borrowed verbs which can create potential cases of homophony; e.g. Marathi allows direct licensing of ‘to paste’ to form pasteqe, but does not allow licensing of ‘to cut’ for a potentially homophonous verb ‘kɔʈɳeqe’ is already present in Marathi lexicon).

References

Verbal Anglicisms: A Shifting Pattern?

Cristian Moroianu, Anabella Gloria Niculescu-Gorpin & Monica Vasileanu

<not updated>
Transliterated vs. non-transliterated forms of newly imported English loanwords in Modern Greek

Rania Papadopoulou & George J. Xydopoulos
(University of Patras and Université Lumière Lyon 2; & University of Patras)

Keywords: Anglicisms, loanwords, transliteration, Modern Greek, Anglicisation

Nowadays, the influence of English on Modern Greek (MG) is strong due to the English influence in domains such as industry, entertainment, Internet, fashion, sports and lifestyle (Anastasiadi-Symeonidi 1994, and Mackridge 2016). In some European languages, such as Italian or French, that, like English, use the Latin alphabet, English borrowings do not have to be orthographically adapted, as for example the French *les trucs do-it yourself* ‘things to do by yourself’, where the English borrowing *do-it yourself* appears in its original form (Varga et al. 2011). MG has a unique alphabetical system consisting of solely Greek characters and so English loanwords are orthographically adapted (i.e. transliterated), as for example in Armenian: թրիլեր ‘thriller’ (Galstyan 2012). Nevertheless, it is rather common that newly imported English loanwords appear non-transliterated within the MG text, maintaining their original English orthographic form and therefore they are not adapted into MG (Pulcini et al. 2012: 6-7), as in (1) and (2). However, English loanwords that were imported into MG five decades ago are orthographically adapted and hence they appear transliterated, as in (3):

(1) Η *blogger* που πρέπει να ακολουθήσεις
I *blóger* pu prépi na akolouthísis
The blogger that should to follow
‘The blogger that you should follow’

(2) Το *καλοκαιρινό outfit* για τις βόλτες σου
To kalokeríni outif jia tis vóltes su
The summer outfit for the walks yours
‘The summer outfit for your walks’

(3) Τα *χόμπι* που θα σε κάνουν καλύτερο στη δουλειά σου
Ta xóbi pu tha se kánun kalítero sti δuliá su
The hobbies that will you make better in job yours
‘The hobbies that will make you better in your job’

In this work, we use a sample of English loanwords that were imported into MG during the last two decades. To determine the period of import we consulted MG Dictionaries and the Dictionary of European Anglicisms (Görlach 2001). In order to investigate this transliteration tendency in MG, we check the appearance of the transliterated or non-transliterated forms of the newly imported English borrowings in MG corpora (e.g. Goutsos & Fragaki 2015, and Kilgarriff et al. 2014). Furthermore, we examine the appearance of the non-transliterated forms of the English loanwords in the web (e.g. blogs, social media accounts), in comparison to older English loanwords found in MG.

Having studied the orthographic adaptation of the newly imported English loanwords in the MG Dictionaries, the MG corpora and the web, we will argue that the appearance of the non-transliterated forms: depends on the period of import of the English loanword, predominates in “specialized” and English-influenced vocabularies, such as the vocabulary of internet and technology (e.g. *hacker* instead of χάκερ ‘hacker’), sports (e.g. *snowboard* instead of σνόουμπορντ ‘snowboard’) and fashion (*jumpsuit* instead of τζάμσουτ ‘jumpsuit’), and is linked to informal circumstances of everyday communication. Moreover, non-adaptation reflects perceptions of prestige regarding English. It is also
possibly related to the familiarity of MG speakers with the Latin alphabet, also due to Greeklish, the Latinised form of MG writing (Androutsopoulos 2009) by young users. Finally, non-adaptation shows the general influence of English on MG and on Greek society nowadays.

References


Goutsos, Dionysios and Georgia Fragaki (2015), Eisagogi sti Glossologia Somaton Keimenon. [Introduction to Corpus Linguistics], Athens: Kallipos.


Varga, Drazen, Oreskovic Dvorski, Lidija and Sasa Bjelobaba (2011), English loanwords in French and Italian daily newspapers, SRAZ LVI, 71-84.

Children’s preference for English-sounding neologisms: An experimental approach

Gillian Roberts, Eline Zenner & Laura Rosseel
(KU Leuven, KU Leuven & Vrije Universiteit Brussel)

Keywords: developmental sociolinguistics, language variation, contact linguistics, language attitudes, social meaning of language variation

Background: In the past decade, research on the lexical influence of English on Western European languages has undergone a socio-pragmatic turn, prioritizing the social meaning of English influence over its structural integration in the receptor language (Peterson & Beers Fägersten 2018; Zenner,
Rosseel & Calude 2019). The last few years have also seen the growth of the field of developmental sociolinguistics, which investigates the acquisition of socially meaningful linguistic variation (e.g. De Vogelaer & Katerbow 2017). This study integrates a developmental sociolinguistic perspective with socio-pragmatic Anglicism research to explore how children, as potential leaders in the global spread of English (Berns 2007), attach meaning to the use of English lexical items in Dutch.

Research question: Through an experimental design, this study aims to investigate children’s preferences for English- or Dutch-sounding lexical items in order to gain insight into how children acquire and reproduce variation and drive contact-induced language change (Labov 2007). Specifically, we address the question of which factors steer Belgian Dutch preadolescent children’s preference for English- or Dutch-sounding neologisms as names for a series of new objects.

Respondents: 120 monolingual Belgian Dutch children from the same dialect area are included in a sample balanced for gender and age (including 8-9, 10-11 and 12-13-year-olds). Studying the preferences of children from this age range serves as a window on the evolution of children’s social evaluations as they transition from caregiver-oriented models of sociolinguistic variation towards peer-oriented models in adolescence (compare Holmes-Elliott 2016).

Design: The experiment introduces the children to 12 new concepts belonging to three different semantic fields (e.g. home life, IT). Each concept is introduced to the children with two alternative names, viz. artificial neologisms (compare Samara, Smith, Brown & Wonnacott 2017) which can be pronounced in a Dutch or English way and are phonotactically plausible in both languages (e.g. ‘snaster’, Dutch [snɑstәɾ] vs. English [snɑstәɻ]). The children are asked to indicate which of the two names they prefer for the object. The experiment also includes existing word pairs as a benchmark with which to compare children’s evaluations of the neologisms.

The experimental design addresses possible influencing factors on the choice of name for the objects: firstly, the experiment varies on the characteristics of the new object presented (e.g. homely objects vs. IT objects) and on the context in which it is presented (e.g. marketing, journalism). Secondly, the children’s explicit attitudes towards the use of English and their awareness of the alternation between Dutch and English are elicited in post-tests.

Analysis: Multifactorial analyses uncover the complex interplay between lexical preference (dependent variable) and semantic field of the concept, presentation context, age, gender, awareness and reported attitudes (independent variables).

Result: Overall, the results allow us to track the evolution in children’s positioning towards English and Dutch as available lexical resources in their community. Results offer insight into how consistently children associate the use of (new) English words in Dutch with particular usage contexts.

References
De Vogelaer, Gunther & Matthias Katerbow (eds), (2017), Acquiring sociolinguistic variation, John Benjamins: Amsterdam.
Peterson, Elizabeth & Kristy Beers Fägersten (eds), (2018), Linguistic and pragmatic outcomes of contact with English, Journal of Pragmatics, 133 (Special issue).
Samara, Anna, Kenny Smith, Helen Brown & Elizabeth Wonnacott (2017), Acquiring variation in an artificial language: Children and adults are sensitive to socially conditioned linguistic variation, Cognitive Psychology, 94, 85-114.

**Uses and functions of Romanian OK in professional spoken interaction: A corpus analysis**

Cristina Andreea Stan
(Alexandru Ioan Cuza University of Iași)

Key words: English, Romanian, professional environment, corpora analysis, discourse markers

The English influence on Romanian language has been very significant over the last 30 years, with borrowings ranging from nouns to verbs and even discourse markers. Based on two comparable corpora of professional spoken interaction, CIVMP and ITICMC, this paper analyzes the lexeme OK in Romanian, an English borrowing, in an attempt to see whether the Romanian discourse marker has preserved or enriched its uses and functions in this type of discourse as opposed to the ones it has in its original language English, in all genres.

The analysis is focused on two main aspects. First, I briefly describe its morphological and pragmatic features, both with Romanian and English attested examples, i.e. an adverb (She did it ok for her first time public speaking.), a noun (His father gave him the ok to go to the party.), an interjection (ITICMC 60) (Ok, that’s good!), or an adjective (The professor approves this, so it is ok to write!).

Then, the presentation is dedicated to a very attentive analysis of the uses and functions it has in the Romanian professional discourse. Following Schiffrin (1987), I began my inquiry by paying attention to its distribution in discourse. Thus, in the corpus I have analyzed, OK has the following pragmatic functions:

→ agreement:
    "Dragoș. [...] Începem cu tine? / Dragoș, do you start?
    Alex. Da. / Yes.

→ apparent agreement:
    “Ana. […] dacă vorbesc cu madam de la resurse umane care era printre doctoranzi doama [prenume și nume] de la București este mai singură […] singura care e mai… / if I speak with the woman from human resources who was among the PhD students, Mrs. [first name, surname] from Bucharest, she is the only one […] the only one who is…
    Ramona. Ochei. / OK.” (Clarification of statements) (CIVMP 126).

→ phatic confirmation:
    “Cristina. Un referat să mă apuc să-l citesc ce-a făcut / an essay I need to start reading what he has done
    Nela. Să mergem pe ideea de originalitate […] / Let’s be original
Given the ubiquity of English loans in languages around the world, psycholinguists have recently started to investigate people’s processing of Anglicisms in communication (e.g. Niculescu-Gorpin and Vasileanu 2018). We extend this approach by proposing an experimental paradigm for studying speakers’ use of Anglicisms in interaction. Building on theories of communicative alignment (Garrod and Pickering 2004), we investigate whether German speakers can be induced to use Anglicisms by their interlocutor doing so, and whether this effect is specific to individual lexical items.
To explore these questions, we conducted an online experiment with 100 native German speakers recruited via the crowd-sourcing platform Prolific. We designed a modified version of the ‘confederate-scripting’ task (Branigan et al. 2000) in which participants interacted online with a pseudo-confederate (in fact, text scripted by the experimenters) to narrate a story. In each trial, participants read part of the story provided by the pseudo-confederate, which contained either an Anglicism or its German counterpart, and selected a matching picture. They then saw another picture illustrating the next part of the story and described it to their partner. To do so, they completed a forced-choice cloze task in which they filled a sentence gap with either an Anglicism, its German counterpart or an unrelated distractor. The cloze task had two versions: in the ‘lexical overlap’ condition, one of the gap-filling words was identical with the English/German word previously used by the pseudo-confederate; in the ‘no lexical overlap’ condition, the gap-filling words consisted of different lexemes than the ones previously presented.

We analysed our data with a mixed-effects logistic regression model. The results indicate that in all but one condition, speakers’ use of Anglicisms was primed cumulatively, i.e. they produced more and more Anglicisms as the story progressed (p < 0.001). This was true regardless of whether their interlocutor used Anglicisms or German words, which suggests that participants were primed by the English options in the forced-choice cloze task rather than aligning with their interlocutor’s speech. In one condition, however, participants’ responses were affected their partner’s speech (p = 0.035): when the interlocutor used Anglicisms which lexically overlapped with participants’ cloze options, participants produced more Anglicisms right from the beginning of the story, and that proportion did not increase significantly over the course of the experiment. This suggests that participants were primed to produce more Anglicisms when their interlocutor used the same lexical items. We also found a significant main effect of participants’ attitude towards the use of Anglicisms in German: participants with a positive attitude were more likely to produce Anglicisms than those with a negative attitude (p = 0.002).

To address the methodological limitations of the forced-choice cloze task, we will discuss ongoing work with a free-choice version of the experiment. By providing a paradigm for testing speakers’ alignment in the use of Anglicisms, our study forms a starting point for further empirical investigations into the social and communicative factors that motivate speakers’ use of Anglicisms in interaction.

References

Word Formation Patterns in the Age of Global English. The Case of Romanian Lexical Blending

Monica Vasileanu & Anabella-Gloria Niculescu-Gorpin

<not updated>
WORKSHOP 11

Discourse Marker use: from production to comprehension

Liesbeth Degand & Maria Josep Cuenca
(University of Louvain & University of Valencia)

Keywords: Discourse Markers, Discourse comprehension, Processing, Translation studies, Corpus studies

In any kind of human communication, be it written, spoken or signed, formal or informal, computer-mediated or not, Discourse Markers (henceforth DMs) are part of the game. This ubiquitousness informs us of a crucial inherent aspect of human language. Yet, despite an impressive quantity of work starting in the early 1980s, the linguistic description of DMs remains scattered, first and foremost because it is a very heterogeneous linguistic category, fulfilling many different functions in discourse (for a recent overview, see Blühdorn, Foolen & Loureda 2017, Maschler & Schiffrin 2015). To gain deeper insight into this complex linguistic category, we believe more systematic (cross-linguistic) work is needed on the production and especially on the interpretation of DMs in a variety of situational settings, resorting to different methodological approaches (see, e.g., Crible 2017, Fischer 2014). Thus, looking at how we use DMs is a crucial step in finding out why we use them in certain situations and under certain conditions, but also when we do not, i.e. when DMs are left implicit (cf. Asr & Demberg 2012, Taboada 2009, Zufferey 2016). Thus, while the categorization of DMs is certainly not a novel topic (see e.g. Degand, Cornillie & Pietrandrea 2013), we believe more attention should be drawn to the double face of DMs, namely as traces of the speaker’s production difficulties accounting for their “Potentially Disfluent Functions” (Crible 2018: 51), on the one hand, and as signals intentionally used by the speaker to facilitate the addressee’s interpretation of the discourse, on the other hand. To make progress on these questions, we believe there is need to combine more systematically the production perspective with the comprehension perspective, i.e. taking into account how the presence of DMs affects the addressee’s discourse comprehension, thus getting better insight into underlying cognitive and functional principles of human communication. The aim of this panel is to bring together researchers interested in getting a firmer grip on Discourse Markers as a linguistic category by investigating the role that they play in language production and comprehension from an experimental or corpus based perspective. We especially welcome contributions making use of innovative methods to tackle DMs from one of the following research questions:

- Are DMs a trace of the speaker’s production difficulties and/or a signal to facilitate the addressee’s comprehension?
- How does the presence of DMs affect the addressee’s discourse comprehension?
- Under what contextual circumstances do DMs affect language comprehension?
- What is the impact of genre and register on DM production?
- How do DMs spread through sociolinguistic variation?
- How do second language learners understand and use DMs?
- How can translation help uncover cross-linguistic differences in DM use?
References

Hebrew harey: Sometimes old, sometimes new

Mira Ariel

Discourse markers are crucial to human interaction. They help speakers facilitate the addressee’s ability to e.g., assess the discourse and/or argumentative status of the modified content. Indeed, speakers’ imperative to contribute New information (Grice, 1989) motivates a dedicated Accessible information marker (e.g., after all), so addressees are aware that some piece of information is not to be processed as New. Next, specifically marking some part of a complex utterance as a strongly asserted subjective position, also motivates dedicated marking, because this is what the speaker guides the addressee to focus on. Surprisingly, however, the Hebrew discourse marker harey ‘hereby’/‘after all’, originally ‘here is’, serves to either access a piece of information already considered Accessible, or to confer an instantly Accessible and unchallengeable status on a piece of New information. The former harey introduces only auxiliary, hence typically nonprominent assumptions, while the latter introduces an argumentatively prominent speaker’s assertion.

Each harey type is further mobilized for a number of distinct discourse functions. Cataphoric/New harey can be (i) ‘deictic’, (ii) ‘performative’ or (iii) ‘Main clause’, displaying a self-conscious performativity on the part of the speaker, who goes on record as making a declarative statement, thus underscoring her commitment to the assertion. Anaphoric harey may introduce (i) a justification for another, original speaker claim, (ii) a background assumption needed for the coherence of the
discourse, or (iii) a counter-assertion, disputing some discourse-salient stance. All crucially build on the fact that the information is Accessible, and hence “undisputable”.

*Harey’s* multi-functionality between quite opposite functions is less puzzling once we realize that each of the 6 Cataphoric and Anaphoric functions is associated with a unique discourse profile, which I identify based on spoken (87 tokens) and written (268) *hareys* (see Table 1 below). For example, Main clause *harey* is restricted to written Hebrew, while Justification *hareys* are not genre-specific. All Anaphoric *hareys*, but none of the Cataphoric ones, can cooccur with overt connectives. Only the countering *harey* can be discourse-prominent among Anaphoric *hareys*. Different intonation contours, as well as stress/no stress on *harey*, are additional cues in spoken discourse.

It is these different discourse profiles that explain how addressees can benefit from speakers’ use of a single discourse marker which carries so many different, contradicting functions.

<table>
<thead>
<tr>
<th>Function</th>
<th>Accessibility status</th>
<th>Modified constituent</th>
<th>Syntactic position</th>
<th><em>Harey?</em></th>
<th><em>that</em></th>
<th><em>Harey?</em></th>
<th><em>that</em></th>
<th><em>Harey?</em></th>
<th><em>that</em></th>
<th>+Overt Connective?</th>
<th>Genre</th>
<th>Discourse status</th>
<th>Frequency (corpora combined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deictic</td>
<td>New</td>
<td>NP</td>
<td>Initial</td>
<td>NA</td>
<td>NA</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>NA</td>
<td>Written</td>
<td>Prominent</td>
<td>Rare</td>
<td></td>
</tr>
<tr>
<td>Performative</td>
<td>New</td>
<td>Ut</td>
<td>Initial</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Written</td>
<td>Prominent</td>
<td>Rare</td>
<td></td>
</tr>
<tr>
<td>Main clause</td>
<td>New</td>
<td>S</td>
<td>Initial</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Written</td>
<td>Prominent</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Countering</td>
<td>Old</td>
<td>Ut</td>
<td>Initial/medial</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>All</td>
<td>Prominent or non-prominent</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Assumption</td>
<td>Old</td>
<td>Ut</td>
<td>Initial/medial</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>All</td>
<td>Non-prominent</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Justification</td>
<td>Old</td>
<td>Ut</td>
<td>Initial/medial</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>All</td>
<td>Non-prominent</td>
<td>High</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Summary of the different discourse profiles of the 6 *harey* functions

**Is causality processed faster than temporality? An experimental investigation of implicit and explicit relations in French**

Joanna Blochowiak, Cristina Grisot & Liesbeth Degand
(Institut des Sciences Cognitives – Lyon, Université de Genève & Université catholique de Louvain)

Keywords: causal relations, temporal relations, processing data, implicit and explicit marking, French

Causal and temporal relations in discourse are ubiquitous. Their presence can be marked with lexical items such as connectives (*and, then, so or because*) or left implicit. In this paper we seek to compare the cost of inferring the presence of causal vs temporal relations in the absence vs the presence of a
connective indicative of a given relation. For the explicit marking, two types of connectives were tested – one specialized for each relation (donc for causality and puis for temporality) and one underspecified (et in its temporal and causal readings). Two views can be found in the literature concerning the treatment of discourse relations with and without connectives: (i) the **Facilitator Connective hypothesis** – the use of connective facilitates the access to the underlying relation (cf. Sanders & Noordman 2000, Canestrelli et al. 2013) and (ii) the **Causality-by-default hypothesis** – causal links are most salient to our cognition, therefore, they are easier to be processed (Sanders 2005). The Facilitator Connective hypothesis generates the prediction that the use of a connective would result in shorter reading times, at least for the specialized connectives, since they should give direct access to the relation they point to. The **Causality-by-default hypothesis** predicts that causal relations will be treated faster than temporal relations.

To test these hypotheses, we carried out a **self-paced reading experiment**, in which we used a 2x3 within-group design, with two variables: Relation, with two levels (temporal and causal); and Marking, with three levels (conjunction (temporal et and causal et), adverbial (puis and donc) and implicit (temporal implicit and causal implicit)). A total of 47 participants (32 females, age range 19-25) fulfilled the reading task of 32 experimental items.

Our results showed a discrepancy between forward causal and temporal relations regarding the role of connectives. For temporal relations, we observed the facilitating role of the specialised connective puis compared to the implicit marking (cf. Grisot & Blochowiak 2019 for concurring results) but not for the underspecified connective et. For causal relations, in contrast, both the specialised connective donc as well as the underspecified et had a facilitating role in the sentence processing. This seems to suggest that only specific connectives have a facilitating role, while underspecified connectives behave differently in the two types of relations. Our results do not confirm the **Causality-by-default hypothesis** as we found longer reading times with causal implicit relations than with temporal ones on the one hand, and again longer reading times with causal readings of et than with temporal readings of this underspecified connective. We explain these results within the Relevance Theory framework (Sperber & Wilson 1986, Noveck & Reboul 2008), according to which the calculation of pragmatic type of inferences (implicatures or explicatures) by the comprehender is a costly process (see Noveck & Reboul 2008 for an overview). With (implicit) temporal readings comprehenders have to infer only one (temporal) explicature while with the causal readings they have to infer two: the causal one and the temporal one which is necessarily present.

**References**


How egocentric is Discourse Marker use? Evidence from speech production under cognitive load

Zoé Broisson & Liesbeth Degand
(University of Louvain)

Keywords: Discourse markers, speech production, cognitive load, spoken interaction, French

Language has often been described as produced ‘for the hearer’. A case in point is the study of discourse markers (DMs), often presented as linguistic expressions that ‘provide instructions to the hearer’ (Hansen, 2006: 25). But what about their role for the speaker? Psycholinguistic research showing that speakers demonstrate egocentric behavior when their cognitive load increases (e.g. Keysar et al. 1998, Roßnagel 2000, Vogels et al. 2015) suggests that the lack of consideration for the role played by constraints on the speaker leaves us with an incomplete account of language production. To assess whether speakers use DMs with a communicative intention, i.e. purposefully and intentionally, this contribution reports on two experiments investigating the impact of speaker orientation (allocentric vs. egocentric) and cognitive load on the use of DMs by 24 native speakers of French. In experiment 1, participants took part in a collaborative puzzle solving task where their addressee’s perspective differed from their own. This experimental setting was meant to induce allocentric language use. In experiment 2, the egocentric task, they had to recount how to assemble the puzzle in a recording to themselves. Cognitive load was manipulated by the absence or presence of a dual task in experiment 1, and by having participants recall instructions from memory in experiment 2.

We hypothesised (H1) to find a higher density and lower diversity of DMs in the output of the allocentric task than in that of the egocentric task, as well as (H2) a favoured use of DMs geared towards the hearer’s needs (specific DMs of e.g. reformulation, specification, contrast), (H3) a higher density and a lower diversity of DMs in the high cognitive load condition, as well as (H4) a higher proportion of “easy-on-the-speaker” DM uses (highly polyfunctional and/or underspecified DMs).

Overall, 2700 DMs were identified and coded for discourse domain and function (cf. Crible & Degand 2019). In line with H1, we found a significantly higher frequency of DMs in the allocentric task (120.98 ptw) than in the egocentric task (71.99 ptw) (log-likelihood: 131.76, p < .01) and a lower DM diversity (TTR: 3.51 vs. 8.38 in the allo/ego output). In contrast, the cognitive load manipulation did not induce any significant difference in density or diversity of DM use, contra H3. As for the type of DMs used, we found that H2 was only partially confirmed since a slight majority of DM types used in the allocentric task were “easy-on-the-speaker” rather than “specific-to-the-hearer”, while H4 was not confirmed statistically.

Based on these results, we conclude that DM use is triggered in the first place by speaker-hearer interaction, although it is not absent from egocentric speaking situations. In addition, the high number of highly polysemous DMs (“easy-on-the-speaker”) suggests that as defended by MacDonald (2013), the role played by speakers’ processing considerations in shaping language may well have been underestimated. These findings also point to a use of DMs less heavily centered on the hearer’s needs than previously thought (e.g. Hansen, 2006), but rather balanced between those of both speakers and hearers. Yet, the data suggest that higher memory demands do not substantially affect the production of DMs. These findings are taken as evidence against the hypothesis of an egocentric bias in language production, and in favor of a more central role played by speakers in shaping language form.
References

Translating discourse markers: Implicitation and explicitation strategies

Maria-Josep Cuenca
(Universitat de València, Spain)

Keywords: discourse markers, translation strategies, implicitation, explicitation, discourse marker interpretation

Translating a DM is a two-fold task: first, the translator interprets its meaning and then proposes a unit that transfers its relational meaning (both denotation and connotation). This comprehension-production task is affected by factors such as functional equivalence, frequency and conditions of use, as well as by the polyfunctionality and ambiguity that some DMs typically show (Degand 2009). In this presentation, the strategies for translating connective DMs linking at inter-sentential level are illustrated by analyzing a parallel one directional corpus including papers from a bilingual academic journal (Catalan Historical Review) that publishes the same paper in Catalan (source text, ca. 100,000 words) and in English (translated text). Special attention is paid to the factors favouring the omission of a DM or its substitution by a more general one (implicitation strategies) and the addition or specification of a DM (explicitation strategies).

(1) I en aquesta situació, la resposta va ser similar

Under these circumstances, the responses were quite similar

(2) Va restar sense fer-se la part més interessant del projecte, les cobertes. Fins al 1909 l’edifici no seria cobert d’aigües sota la direcció de l’arquitecte Ricardo García Guereta

The most interesting part of the design, the roof, was never built. In fact, the building’s roof was not constructed until 1909 under the supervision of the architect Ricardo García Guereta

Omitting a DM (the conjunction i ‘and’ in (1)) leads to the underspecification of the discourse relation. Adding a DM (in fact in (2)) or substituting it by a more specific marker implies the specification of an instruction, often to improve readability and naturalness in the translated text.

DMs omission has been observed in many contrastive studies (e.g., Bazzanella & Morra 2000, Crible et al. 2019, Cuenca 2008, Dupont & Zufferey 2017, Hoek et al. 2017, Zufferey 2016). DM explicitation has become an outstanding topic, especially in relation with the explicitation hypothesis.
as a translation universal (e.g., Becher 2011a, 2011b; Denturck 2012; Marco 2018; Zufferey & Cartoni 2014).

The research questions are:
(a) Which are the strategies for translating DMs linking at inter-sentential level?
(b) Which factors (including syntactic and semantic/pragmatic) account for the implicitation or explicitation of a DM?
(c) How can translation help uncover cross-linguistic differences in DM use and processing?

The results point to the following tendencies:
(i) Both omission and addition of a DM are significant translation strategies linked to a variety of factors related to both production and interpretation.
(ii) DMs are more frequently omitted or added when they act at the rhetorical/sequential domain (as opposed to the ideational domain, Crible & Degand 2019).
(iii) Contrary to expectations, some DMs with a specific meaning (such as however) are often added.
(iv) Scope, collocations and style have an effect on the implicitation and explicitation of DMs in translation.

Corpus analysis of translated texts is especially relevant to understand why and how DMs are used in discourse production and comprehension. It also contributes to gain deeper insight into how implicit and explicit coherence relations interplay cross-linguistically.

References


Zufferey, Sandrine (2016), Discourse Connectives across Languages: Factors Influencing Their Explicit or Implicit Translation, *Languages in Contrast* 16(2), 264-279.


Focalizers and discourse relations

Eva Hajičová, Jiří Mírovský & Barbora Štěpánková
(Charles University, Prague)

Keywords: focalizers, discourse relations, parallel data, Czech, English

1. The analysis of the so-called focalizers, i.e. particles such as E. *also, only, even*, and their Czech counterparts *také, jenom, dokonce*, based on the data from the English–Czech parallel corpus PCEDT and studied from two aspects, namely (i) their position in the sentence surface word order, and (ii) their semantic scope has demonstrated that the interpretation of the semantic scope of these particles is highly dependent on the previous context and in several respects has an important influence on the interpretation of discourse relations. As a matter of fact, focalizers often act as discourse connectives.

2. In the present contribution we put under scrutiny three research questions, namely (i) in which respects and under which conditions the selected focalizers may be said to serve as discourse connectives, (ii) which particular discourse relations are indicated by the focalizers in question, and (iii) whether and in which way the two languages under investigation, Czech and English, behave in a similar or in a different way.

3. We have based our analysis on the following data resources: (i) for Czech, the Prague Dependency Treebank of Czech (PDT 3.5, Hajič et al. 2018) containing documents of the total of about 50 thousand sentences annotated on the underlying syntactic layer also for information structure (topic–focus articulation, TFA) and containing also annotation of discourse relations (in a slightly modified PDTB style); (ii) for English, the Pennsylvania Discourse TreeBank (PDTB, Prasad et al. 2008); (iii) for a comparison between Czech and English, the English–Czech parallel corpus (PCEDT, Hajič et al. 2012); (iv) the dictionary of Czech connectives (Mírovský et al. 2017, Synková et al. 2019).

4. We have focussed our attention on three focalizers in Czech (taking into account also their synonyms; the figures in brackets indicate their total frequency in the PDT if they occur as a non-compound connective in the “focalizing” function) *dokonce* (71), *jen/jenom/pouze* (29), *také/rovněž/též/zároveň* (347) and their English equivalents *even, only, also/too* both in the PDTB and the PCEDT. As for discourse relations, *dokonce* indicates the relation of gradation or conjunction, *jen/jenom/pouze* serves to express correction, restrictive opposition and gradation, and *také/rovněž/též/zároveň* indicates the relation of conjunction, gradation or synchrony. The total frequency of the types of discourse relations expressed by the above particles functioning as focalizers is as follows: by far, the most frequent relation is that of conjunction (352), followed by gradation (66), restriction (27), correction (2), synchrony (2) and opposition (1).
We have confronted the observations on the types of relations expressed by the three Czech focalizers as listed above with the English data taking into consideration (a) the word order position of the given focalizer in the surface shape of the sentence, (b) the occurrence of the focalizer in the topic (theme), in the contrastive topic or in the focus (rheme) of the sentence, (c) the semantic scope of the focalizer. We have ended up with a more detailed classification of the assigned discourse relations in dependence on these three factors. In the final version of the contribution, we illustrate the above aspects and their interplay by Czech and English examples taken from the mentioned resources.

References
Synková, P., Poláková, L., Mírovský, J. and M. Rysová (2019), CzeDLex 0.6. Data/software, Charles University, Prague, Czech Republic, http://hdl.handle.net/11234/1-3074.
Štěpánková, B. (2014), Aktualizátory ve výstavbě textu, zejména z pohledu aktuálního členění (Focalizers in the structure of text). Charles University, Prague, Czech Republic.
Zikánová, Š. et al. (2015), Discourse and Coherence. From the Sentence Structure to Relations in Text. Charles University, Prague, Czech Republic.
Syntactic parenthesis and prosodic integration: some clues to understand discourse markers production and use

Maria Cristina Lo Baido
(University of Bergamo)

Keywords: Italian, discourse markers, parenthesis, prosody, spoken discourse

We intend to analyse a class of verbal discourse markers in Italian such as penso (‘I think’) addressing the following question: how are such DMs produced and used in spoken discourse? By studying their behaviour on the syntactic, distributional and prosodic level (439 occurrences taken from the KIParla Corpus, Mauri et al. 2019), we will make some considerations on their role in language production. Given that the DMs at issue are typical parentheticals- and parenthesis implies an interruption - we will study to what extent such disruption is prosodically mirrored and how the DMs’ possible prosodic integration can be related to their functions as indexical strategies operating on spoken communication organization (Diewald 2011).

In the sample of 439 DMs, we will specifically study the profile (Praat - Boersma 2001) of 150 DMs in their immediate environment, by spotting the occurrence of silent pauses (interruptions of the sound signal lasting more than 200 milliseconds, Candéa 2000, Crible 2018) on the level of the DMs boundaries. Verbal parentheticals display various behaviours:

1) voi avrete ((0,4s)) e::h non so ((0,9s)) dei foglietti (KIParla, BOD1005)
‘you will find eh I don’t know some instructions’

2) può contare allora presumo su un’ottima conoscenza della lingua inglese (Lip, NA12)
‘So you can rely I presume on an excellent proficiency of English’

The marker non so, in (1) indicates the speaker’s hesitation, which it is reflected by the filled and silent pauses, and the vowel lengthening (Bortfield et al. 2001). Conversely, the similar marker presumo, in (2), facilitates the speaker’s production, through the expression of both the content level (host) and the more functional meaning (Degand and Evers Vermeul 2015: 78) of evidentiality given by the prosodically integrated DM, despite the lack of hierarchization (Haselow 2016). In cases like (2), despite the syntactic parenthesis, the DM production does not (seem to) pose problems for the speaker, as demonstrated by its prosodic integration. Furthermore, such integrated items may arise from complex predicates (such as strong assertive or semifactive predicates, presumably used intentionally), rather than from generic items conveying hesitation. Beside their prosodic integration, the occurrence of DMs in medial position (180/439) contributes to give evidence for their positive role in helping the speaker for the production of elusive meanings with varying scope during the on-line discourse production.

Prosodic integration is a clue indicating the way parentheticals are produced, namely, as items eventually integrated to the speech flow conveying different functional meanings, which contextualize the hosts meaning. When they are integrated and they do not occur with markers of disfluency (as in 1), they are more likely to work as resources. The prosodic integration of parentheticals is not a new issue (see Dehé and Kavalova 2007, Dehé and Wichmann 2010a,b, Simon 2004, inter al.). Therefore, this contribution offers additional evidence for the integration of verbal DMs, by correlating the latter with the status of DMs as a linguistic category (and resource) not exceptional with respect to their hosts, except for the performing of functional meaning(s) on discourse level.
References
Dehé, Nicole and Anne Wichmann (2010a), Sentence-initial I think (that) and I believe (that): Prosodic evidence for use as main clause, comment clause and discourse marker, *Studies in Language* 34:1, 36–74.
Simon, Anne Catherine (2004), *La structuration prosodique du discours en français* (Sciences pour la communication; 74), Bern: Peter Lang.

Principles of Cognitive Processing of Discourse Marking

Óscar Loureda, Inés Recio Fernández & Adriana Cruz
(Heidelberg University)

Key words: discourse markers, procedural meaning, eye-tracking, processing effort

Discourse markers are procedural-meaning expressions: they constrain inferential routes during processing according to their specific morphosyntactic, semantic and pragmatic features (Blakemore 1987, 2002, Portolés 2001[1998]). As such, discourse markers need conceptual representations upon which to deploy their instructional semantics. The interaction of conceptual and procedural semantics is best described in terms of asymmetry and rigidity (Leonetti & Escandell-Vidal 2004; Murillo Ornat 2010; Escandell-Vidal & Leonetti 2011). Asymmetry refers to the property of discourse markers (and procedural-meaning items in general) according to which they act upon conceptual-meaning
expressions, but not the other way around. Rigidity refers to the fact that, by constraining the access to contextual assumptions, discourse markers compel interlocutors to select specific mental assumptions during processing in a way that the instructions encoded by the discourse marker are complied with.

The rigidity and asymmetry postulated theoretically for discourse markers as to conceptual expressions also seems to find empirical support. In this communication, we present experimental evidence gathered in eye-tracking reading studies carried out on Spanish discourse markers and covering different discourse-marking paradigms: focus operators, connectives and reformulation markers. The results of the studies allow us to formulate a series of general principles of cognitive processing of discourse marking (cf. Loureda et al., in press):

1. Discourse markers modify the processing patterns of the utterances in which they occur, in relation to utterances in which no discourse marker is provided. Pattern-modification can be local and affect a discourse segment, and/or global, thus affecting the utterance as a whole.
2. Discourse markers are effort-constraining expressions. Despite adding semantic information to utterances in the form of conventional implicatures, processing marked utterances is not more effortful that processing unmarked utterances.
3. Discourse markers exert an immediate regulatory function that translates in an effort-constraining effect for the segment upon which they operate/their host segment compared to unmarked utterances. This finding is in line with previous experimental findings for other languages (e.g. Cozijn et al. 2011, Canestrelli et al. 2013, Köhne & Demberg 2013, Mak and Sanders 2013, Drenhaus et al. 2014; Xiang & Kuperberg 2015, among others) and with models of incremental processing (Traxler et al. 1997, Noordman and Vonk 1998).

All in all, discourse markers set in motion effort-optimization strategies during reading which are visible during the construction of an initial assumption as well as in assumption-confirmation strategies.

References
Beyond recipient’s lack of understanding: Hebrew ‘ATA LO MEVIN (‘you don’t understand’)

Hilla Polak-Yitzhaki & Yael Maschler
(University of Haifa)

Keywords: Discourse markers, mental verb constructions, grammaticization, stance, embodied conduct

This study explores the Hebrew negative 2nd person present SUBJ-PRED mental verb construction ‘at/a lo mevin/a (‘you don’t understand (MASC/FEM SG)’), arguing that employment of this construction is highly formulaic, rarely denotes lack of understanding, and that its uses are closely tied to its prosodic, morphophonological and syntactic features, to particular actions in which participants engage, and to the bodily-visual conduct of the speaker. Based on a mostly synchronic analysis, we suggest a grammaticization (Hopper 1987) path followed by this construction from an epistemic verb phrase conveying literal lack of understanding to a discourse marker (Maschler 2009) signaling the opening of a telling which seeks recipient alignment towards the speaker’s highly-involved (Tannen 1989) stance.

Employing Interactional Linguistics methodology, and contributing to Pragmatic Typology, we build on earlier studies of cognitive verb constructions in other languages (e.g., Thompson 2002, Polak-Yitzhaki & Maschler 2016, Lindström, Maschler, & Pekarek Doehler 2016, Pekarek Doehler 2019), all of which point to grammaticization of fixed formulas involving mental verbs from specific discourse patterns employed for distinct interactional purposes.

Our data comprise over 20 hours of audio- and video-recorded informal interactions among 812 speakers, manifesting 43 instances of the ‘ATA LO MEVIN construction.

Only 6 tokens are employed literally, claiming the hearer’s lack of understanding. They generally occur with an object complement and display no morphophonological reduction.

Fifteen tokens respond to a recipient’s comment doubting some aspect of the speaker’s immediately preceding telling. These constructions highlight the hearer’s misconception while at the same time intensifying the speaker’s stance towards the telling. Nine are complemented by an object clause; 6 occur with no complement.
Ten tokens, most of them morphophonologically reduced, are deployed to intensify the speaker’s
stance towards some aspect of his/her telling. Six occur with no complement, as in:

1 Adi: ..... ‘ani smexa sheyesh xofesh.
   I happy.F.SG that-EXIST holiday
   .....I’m happy that there’s a holiday [soon].

2 (inhales)

3 ‘ani kol-kax tsxa ‘et ze,
   {------------------smiles------------------}
   I so need.PRS.F.SG ACC this.M.SG
   I need it so badly,

4 ‘at lo mvina.
   you NEG understand.F.SG
   you don’t understand.

5 Shoval: ‘ani gam.
   I also
   me too.

Adi employs the construction (line 4), further intensifying the upgraded expression of her positive
stance towards the upcoming holiday (line 3). Shoval aligns with Adi’s stance (line 5), revealing that
the construction is not used here literally.

Finally, 12 highly morphophonologically reduced tokens (e.g., tlomvina ‘you (FEM SG) don’t
understand’), mark the opening of a telling, seeking recipient alignment towards the speaker’s highly-
involved stance. Most such tokens occur with no complement, fulfil the semantic, structural, and
sequential requirements of a prototypical discourse marker (Maschler 2009: 17), and are accompanied
by the speaker’s leaning forward towards the hearer – embodied behavior unattested for the other
functions.

Embodied conduct as well as prosodic, morphophonological, and syntactic features of the
construction thus correlate with the weakening of its literal sense. This sheds light on the uses speakers
make of this construction, on its possible grammaticization path, and on the interlaced nature of
grammar and interaction.

Acknowledgements: Yael Maschler gratefully acknowledges Israel Science Foundation grants no.
887/12 and 1233/16.

References
Hopper, Paul J. (1987), Emergent grammar, in: J. Aske, N. Beery, L. Michaelis, and H. Filip (eds),
Lindström, Jan, Maschler, Yael and Pekarek Doehler, Simona (eds), (2016), Grammar and Negative
Epistemics in Talk-in-Interaction: Cross Linguistic Studies, Special issue of Journal of
Pragmatics 106.
Maschler, Yael (2009), Metalanguage in Interaction: Hebrew Discourse Markers, Amsterdam: John
Benjamins.


KS

**The discourse marker bueno in spoken and written Mexican and Peninsular Spanish**

Pekka Posio & Malte Rosemeyer  
(University of Helsinki & Albert-Ludwigs-Universität Freiburg; KU Leuven)

Keywords: Discourse marker; Spanish; Dialectal variation; Corpus linguistics; Genre

Previous studies (e.g., Bauhr 1994) on the Spanish discourse marker (DM) *bueno*, literally ‘good’, have identified two main usage domains: (a) discourse organization, such as turn management (1), and (b) reacting to a previous turn either agreeing or disagreeing (2) (data from PRESEEA).

(1)  
*senti pánico... bueno en lo primero que pensé era mi hermano*  
‘I felt panic... BUENO the first thing I thought about was my brother’

(2)  
*–si antes le veía poco ahora no le veo nada*  
‘I didn’t meet him often before, but now I don’t meet him at all’

*–bueno pero no te preocupes*  
‘BUENO but don’t worry’

While previous studies have focused on spoken, dialogical data, less is known of the use of *bueno* in colloquial written texts, where the turn-taking and discourse-organizing functions are marginal. In the present study, we analyze the DM *bueno* in two very different types of data: stratified sociolinguistic interview corpora from Mexico and Spain (PRESEEA 2014) and the WEB/DIALECTS corpus (Davies 2016) consisting of internet forum and blog texts form these two countries. Our study thus specifically addresses two of the research questions proposed for this workshop: (1) are DMs a trace of the speaker’s production difficulties and/or a signal to facilitate the addressee’s comprehension and (2) what is the impact of genre and register on DM production. In addition, we compare the DM uses of *bueno* in Mexican and Peninsular Spanish.

Our analysis of a random sample of 4,000 occurrences of *bueno* as a DM in these data suggests that its main function in the WEB/DIALECTS corpus is related to the introduction of a new perspective or a contrast between the speaker’s stance and the perspective of the others. In (3) *bueno* marks a contrast between the opinion of the speaker and that of ‘others’, while in (4) it marks the speaker’s own judgment as contrary to expectations.

(3)  
*tiran los desperdios ala calle y no les importa bueno pero ami si me importa*  
‘they throw the garbage to the street they don’t care but BUENO I do care’

(4)  
*Pensé bueno, esto no es tan mal.*
‘I thought BUENO it’s not that bad.

The function of signaling a ‘change of perspective’ may explain the frequent use of bueno in the beginning of reported speech sequences (5), where it indicates a change of voice in the narrative.

(5) me dijo bueno juanma como vas a buscar la felicidad hoy
‘S/he said to me BUENO Juanma how are you searching for happiness today’

The frequent use of bueno in both written and oral data suggests that the DM is used in both genres to facilitate the comprehension for the addressee. Our preliminary analysis also reveals dialect-specific differences in the use of bueno. In particular, in the spoken Mexican data the reactive uses of bueno are significantly less frequent than in our Peninsular data, suggesting a greater degree of pragmatization of bueno in Peninsular Spanish. In our presentation we discuss how to account for the medium and dialect specific differences.

References

The Romanian markers altfel and de altfel: Discourse domains and functions in use

Ariadna Ștefănescu & Sorina Postolea
(University of Bucharest & University of Iasi)

Keywords: discourse markers, discourse functions, discourse domains, Romanian markers

This paper focuses on a couple of Romanian discourse markers (DMs) with the same origin and similar form that stand out among the other Romance languages and in comparison to English by their diverging behaviours in use, i.e. altfel [lit. “other kind/way”] and de altfel [lit. “of/by other kind/way”]. Based on the model provided by Crible & Degand (2019) and translation spotting (Cartoni, Zufferey, & Meyer, 2013), our analysis aims to highlight the discourse domains and the functions performed by these two discourse-relational devices.

We start from the assumption that there is an interdependent relationship between the cumulative effects of some discourse resources present in the co-text and the discourse function of a DM (Das & Taboada 2019). Moreover, there is a tight connection between some co-textual resources and the discourse domain to which the DM belongs.

Our paper shows that altfel and de altfel are unequally distributed over the four discourse domains – i.e. ideational, rhetorical, sequential, and interpersonal –, and it details their range of functions. The discourse domains preferred by altfel/ de altfel are the rhetorical and the interpersonal ones. Both altfel and de altfel are multifunctional. However, the main function of altfel
is that of marking Condition in the rhetorical domain (“otherwise”, “sinon”), while de altfel is more flexible and performs several functions across the domains, out of which the most frequent are Addition (moreover/ de plus) and Specification (in fact). Thus, we present the procedural profiles of the two DMs, taking into consideration their domains, functions and the most frequent pragmatic indications they make manifest (such as the implication of an ‘unwanted hypothesis’ in the case of altfel, or the ‘background information first-come to the Speaker’s mind’ in the case of de altfel).

These findings are also highlighted by the translation spotting technique. In a sample of 400 contexts extracted from the DGT and the EuroParl Corpora in Romanian, English, and French, de altfel had no less than 28 different equivalents in English and 32 in French (e.g. En. moreover, by the way, in fact, in contrast; Fr. de plus, au passage, en effet, autrement).

Our study is completed by diachronic discourse variation and frequency analyses which follow the uses of altfel/ de altfel in the course of time. This supplements a previous diachronic study on the older pair altminteri/ de altminteri (Ştefănescu 2019). Altfel has entered Romanian in the 18th century (DA), whereas altminteri was inherited from Latin (like Fr. autrement, Cat. altrement or It. altrimenti) and has attestations in old Romanian texts. The older DM altminteri started to be used in the ideational domain, whereas the derivative de altminteri was mainly used in the rhetorical and interpersonal domains, tendencies which are still true today. Even if altfel tends to replace altminteri in current use due to its shorter form, the discourse behaviour and the procedural meanings of the two pairs are identical. Our diachronic investigations reveal a cline of pragmaticalisation (Hopper & Traugott 2003; Dostie 2004; and others) in which the discourse domain is a key factor.

References


Das, Debopan, Taboada, Maite (2018), Signalling of Coherence Relations in Discourse, Beyond Discourse Markers, Discourse Processes, 55:8, 743-770.


Hopper, Paul J., & Traugott, Elizabeth Closs (2003), Grammaticalization, Cambridge: Cambridge University Press.

Discourse markers and language contact: Evidence from a Berber-Arabic context

Valentina Schiattarella & Valentina Serreli
(University of Naples, “L’Orientale” & University of Bayreuth)

Keywords: discourse markers, language contact, bilingualism, Arabic, Berber

This presentation is about the use of discourse markers (DM) in a context of societal bilingualism. The definition of discourse markers is debated, as there is no consensus about which grammatical classes are included and which functions DMs code in different languages (Schiffrin 1987, Fraser 1990, Matras 1998, Fischer 2006, and Crible 2017). Studies about the use of DMs by bilinguals (Maschler 1994, Matras 2000, and de Rooij 2000) or in language contact situations (Heine 2016, Sankoff et al. 1997) have shown that use and frequency correlate with the speaker’s proficiency in the L2 and his/her degree of contact with L2 native speakers. Moreover, the same DMs can code different functions in the two languages or display different prosodic contours and positions.

In this presentation we analyze the distribution of and the functions coded by the Arabic DMs yaʕni ‘that is’, ṭabʕan ‘of course’ and masalan ‘for example’, in the L1 (Siwi Berber) and the L2 (Arabic) speech of three bilinguals. The three speakers are selected because they were interviewed by both authors, therefore allowing a direct comparison between L1 and L2 speech. The data were collected in the Siwa Oasis (Egypt) between 2011 and 2019 for two independent studies, focusing on Siwi and on Arabic respectively.

The three DMs show a different distribution. In Arabic yaʕni is by far the most used by two speakers, while ṭabʕan is the most used by one speaker; masalan is the least used by the three speakers. In Siwi, however, ṭabʕan is the most frequent, followed by yaʕni and masalan for the three speakers. The main functions coded by the three DMs in both languages are explicative (yaʕni), evidential or conclusive (ṭabʕan), and elaborative (masalan).

L2 (Arabic)

ah ṭabʕan / il-qurʔaan yibʔa fuṣḥa ṭabʕan
yes of_course / the-Quran becomes Classical_Arabic of_course

‘Yes of_course, the Quran is Classical Arabic of_course’

L1 (Siwi)

yʕәṃṃa ә nt әlqaršwá / lәbsis n әṭṭáf ṭabʕan s tini
they_make_it with chickpeas / lebsis of black of_course with dates

‘[The white lebsis], they make it with chickpeas, black lebsis is, of_course, with dates.’

The Siwi DMs displayed in Arabic are limited in number, only used by speakers with low Arabic proficiency, and only in dialogues. Two examples are the particles ya ‘you know’ and na ‘right?’, which in L2 occur in the same position and code the same function coded in L1. This follows patterns typically found in situations of societal bilingualism and is connected to the power imbalance between Arabic and Siwi.

The study shows that the main functions coded by Arabic DMs in bilinguals’ Arabic and Siwi speech are the same, but in Arabic collocations are more frequent and the range of functions is wider. It also shows that bilinguals’ use of L2 markers correlates with the genre (dialogues vs narrations), with the speaker’s contact with L2 native speakers and his/her proficiency and ability to move across L2 registers.
Empirical evidence for the role of connectives on the acceptability of straw man fallacies

Jennifer Schumann & Sandrine Zufferey
(University of Bern)

Keywords: connectives, subjectivity, forewarning effect, straw man fallacy, empirical validation

In this paper, we investigate the role of causal connectives that convey attributive meaning such as since and as for the acceptability of straw man fallacies. This type of fallacious argument operates by misrepresenting the opponent’s position, namely by making it more extreme in order to refute it easily (see Lewiński & Oswald 2013). In two experiments, we assessed the impact of the French connectives puisque (similar to the English since) and comme (similar to as) for the acceptability of fallacious arguments involving a straw man. Based on previous research (Schumann, Zufferey & Oswald 2019), we expected that the presence of attributive connectives would raise participants’ awareness of a potential distortion between the speaker’s original meaning and the reported meaning, compared to implicit relations in which the attributive content has to be inferred. In addition, we expected to find a difference between the two connectives, due to their different degree of subjectivity. While puisque is a highly subjective connective (Zufferey & Cartoni 2012) comme can be used to convey both objective and subjective causal relations (Roze, Danlos & Muller 2012). In previous research, subjective connectives were found to create a forewarning effect that diminished the persuasiveness of argumentative texts compared to texts with objective connectives (Kamalski et al. 2008). We therefore expected a similar effect, leading to a lower acceptability of fallacious arguments conveyed by puisque.
To test the role of connectives for the acceptability of straw man fallacies, we designed an offline task with 40 short dialogues about different societal topics. The first statement of the dialogue contained a standpoint (segment 1) followed by an argument in support of the standpoint (segment 2). This part of the dialogue was identical in all four conditions. The second part of the dialogue contained the manipulated variable. The first segment expressed a possible consequence of the argument given by the first speaker and was kept constant throughout the conditions. The second segment contained the argument which was either a fallacious or non-fallacious reformulation, and was introduced with a connective (puisque or comme) or simply juxtaposed to the previous segment. The participants had to evaluate every dialogue by responding to 4 questions on a 6-point Likert scale with an additional option they could select when they couldn’t give an answer. Questions 1 and 2 targeted specific features of the straw man and questions 3 and 4 aimed at the agreement with both speakers. Our results indicate that people are intuitively able to spot straw man fallacies as they systematically judge them as less acceptable. More importantly for this workshop, our results also show that connectives do indeed play an important role in argumentative discourse as they forewarn people to the presence of a fallacious argument. Crucially however, not all connectives play a similar role, as only puisque created a forewarning effect. We conclude that the notion of subjectivity, crucial in discourse studies, is also highly relevant to account for participants’ reaction to fallacious arguments.

References

Do teenagers understand connectives from the written mode?

Ekaterina Tskhovrebova & Sandrine Zufferey
(University of Bern)

Keywords: connectives, L1 acquisition, written mode, experimental study, French.

Children start producing connectives many years before they fully master them in comprehension tasks (McClure & Geva, 1983). For instance, children below the age of 10 to 12 do not master connectives frequently used in speech. In a set of experiments, Cain & Nash (2011) tested the way 8- and 10-year-old children understand and process frequent connectives such as because, and, but, or after. Their results indicate that children as late as 10-year-old perform significantly less well than adults in a cloze task assessing their use of temporal, causal and adversative connectives. In a set of experiments targeting more specifically the understanding of temporal relations, Pyykkönen and
Järvikivi (2012) demonstrated that children aged 12 still experience difficulties understanding temporal relations when the connective used (i.e. after) implies a reversed order of the related segments with respect to the chronological order of events in the world.

Yet, in addition to these connectives, most European languages also possess a vast repertoire of connectives that are bound to the written mode, such as therefore and nevertheless in English. Even though they are seldom used in speech, in writing many of these connectives are very frequent. Mastering them represents therefore an important step to achieve effective written communication. Moreover, little is known about the acquisition of connectives during teenager years (from the age of 12 to 18). However, these years are particularly interesting to study since it is during this period that teenagers progressively acquire an adult-like ability to understand and produce connectives.

In this presentation, we investigate the ability of teenagers aged from 12 to 18 to correctly use four French connectives bound to the written mode: aussi to convey consequence relations, en outre to convey additive relations, en effet to convey causal relations and toutefois to convey concessive relations (Roze, Danlos & Muller, 2012). These connectives were chosen because of two main reasons. First of all, even adult speakers were found to vary in their ability to judge their uses (Zufferey & Gygax, 2020). The second reason for this choice is that these connectives are well-fitted to assess the role of three different factors that could account for their complexity: they have a variable frequency in the written mode; they encode relations with various degrees of cognitive complexity; they differ in the number of functions that they can convey.

Results from a sentence completion task indicate that teenagers are not yet able to use all four connectives correctly, and that frequency seems to be the best factor to account for these difficulties. In addition, teenagers from high school (aged 16-18) on average perform better than those from middle school (aged 13-15) on the same task, indicating that development is still taking place between these two age groups.

References
Zufferey, Sandrine & Gygax, Pascal. (2020). “Roger broke his tooth. However, he went to the dentist”: Why some readers struggle to evaluate wrong (and right) uses of connectives. Discourse Processes 57(2), 184–200.
Modeling the interplay of rational production and comprehension of ambiguous connectives

Frances Yung, Jana Jungbluth & Vera Demberg
(Universität des Saarlandes, Universität des Saarlandes & Universität des Saarlandes)

Keywords: discourse connectives, RSA model, psycholinguistics, experimental pragmatic, crowdsourcing

Discourse connectives mark the coherence relation between text components. Many English connectives are ambiguous, such as “since” and “as”. Both interpreters and speakers are confronted with choices during the comprehension and production of connectives, because one connective can have multiple meanings (e.g. *reason* and *temporal* senses of “since”) and one coherence relation can be marked by multiple connectives (e.g. “while” and “but” for a contrast relation).

Bayesian pragmatics suggests that the production and comprehension of utterances are the result of the reasoning between speakers and listeners to achieve efficient communication. We model the interplay between the production and comprehension of discourse connectives by a unified framework – the Rational Speech Act (RSA) model (Frank and Goodman 2012).

The RSA model is a Bayesian model of pragmatic reasoning that formalizes how speaker and listener reason about the reasoning of the counterparts during production and comprehension, incorporating production effort as well as meaning salience.

RSA has been applied to model the pragmatics of various linguistics phenomena (e.g. Goodman and Stuhlmüller 2013 and Baumann et al. 2014), including discourse connective processing (Yung et al. 2016, 2017). Yung and Demberg (2018) presented an experimental study to examine the RSA account of connective production. They reported that speakers preferred less ambiguous connectives when it was necessary for effective communication. Evidence, although weaker, was also found that listeners rationally interpreted the speaker’s choice of connectives.

The task in their experiment was however quite different from real life language processing, as producers was shown a list of alternative sentence continuations that the listener could guess, and the listener was shown a list of connectives from which the producer had chosen. This means that participants CAN rationally choose among connectives and CAN rationally interpret connectives, but it falls short in showing whether people do so under more naturalistic settings.

This work’s objective is to test whether the RSA account of connective processing can account for connective production and comprehension in more naturalistic settings.

We focus on the effect of contextual information on connective production.

Our research questions are:

1. Does contextual expectation affect the speaker’s choice of connectives?
2. Does the speaker’s choice follow the prediction of the RSA model?

We propose an experimental design with improved ecological validity, where the ambiguity of connectives is controlled by contextual priming. For example, the connective “while” has a contrast sense (1) and a temporal (synchronicity) sense (2).

(1) This morning, he arrived at the train station by car [while/but] his sister took the bus yesterday. *(contrast)*

(2) This morning, he arrived at the train station by car [while/during that time] the gates were still closed. *(temporal)*
According to RSA, speaker will choose a less ambiguous utterance, e.g. “but” and “during that time” respectively, when they predict that the listener might not interpret the ambiguous version correctly. We expect that such rational choice of connectives is related to the discourse context, i.e. the speaker should choose a less ambiguous connective when the expected discourse meaning and the actual meaning they want to convey are contradictory. For example, the speaker should tend more to use “during that time” instead of “while” in (2) when the preceding context is (3), comparing with the case when the preceding context is (4).

(3) Mark and his sister preferred different means of transportation. (contrast primed)
(4) Mark always arrived way too early. (temporal primed)

We present our findings on the production of the connectives “since”, “as” and “while”, which have distinct alternative meanings, based on a large number of experimental responses collected by crowdsourcing.

Acknowledgements: This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 232722074 – SFB 1102

References