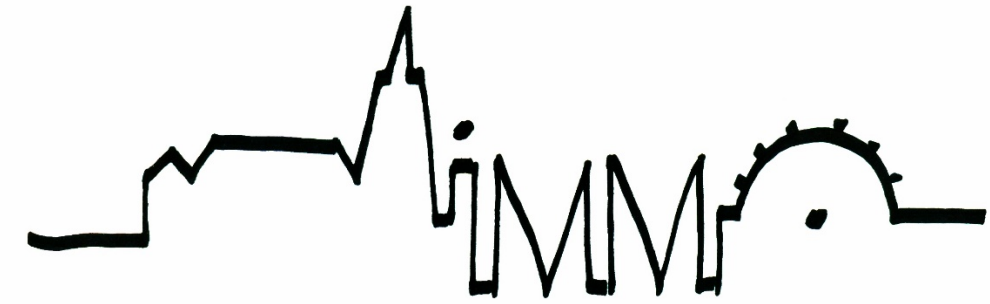
Booklet of abstracts

19th International Morphology Meeting

6 – 8 February 2020, Vienna



Committees

Organizers

Francesco Gardani (University of Zurich), Franz Rainer (WU Vienna)

Conference Manager

Elisabeth Peters (WU Vienna)

Scientific Committee

Evangelia Adamou, Kimi Akita, Nino Amiridze, Giorgio Francesco Arcodia, Peter Arkadiev, Mark Aronoff, Hans Basbøll, Pier Marco Bertinetto, Olivier Bonami, Gilles Boyé, Chiara Celata, Wolfgang U. Dressler, David Erschler, Bernard Fradin, Nanna Fuhrhop, Livio Gaeta, Christina Gagne, Francesco Gardani, Nicola Grandi, Laura Grestenberger, Stefan Th. Gries, Holden Härtl, Claudio Iacobini, Istvan Kenesei, Krista Kerge, Katharina Korecky-Kröll, Johanna Laakso, Lior Laks, Rochelle Lieber, Ana Luis, Andrej Malchukov, Francesca Masini, Veronika Mattes, Felicity Meakins, Marianne Mithun, Peter O. Müller, Vito Pirrelli, Ingo Plag, Frans Plank, Angela Ralli, Dorit Ravid, Elke Ronneberger-Sibold, Thomas Schwaiger, Peter Sells, Lameen Souag, Thomas Stolz, Gregory Stump, Pius ten Hacken, Anna Maria Thornton, Bjoern Wiemer, Zheng Xu

February 2020

Vienna

Table of Contents

To access individual abstracts please use
the links in the program on the following pages.

Program

19th International Morphology Meeting

Vienna, February 6 – 8, 2020

Program

All rooms as well as registration and coffee breaks are located in the Teaching Centre (TC).

Thursday, February 6

8:45 – 9:45	**REGISTRATION AND COFFEE**		
9:45 – 10:00	**OPENING**		**TC 0.01**
10:00 – 11:00	**PLENARY TALK** **Felicity Meakins** Two to tangle: Modelling morphologies in contact		**TC 0.01**
11:00 –11:30	**COFFEE BREAK**		**Ground floor**
	TC 1.01	**TC 1.02**	**TC 3.06**
11:30 –12:00	**Kaarina Pitkänen-Heikilä** Typical foreign word patterns in Finnish special languages and their usage in the general language	**Marcel Schlechtweg and Greville G. Corbett** The acoustics of word-final [z] in English: A comparison of pluralia-tantum and regular-plural nouns	**WORKSHOP** **Dorit Ravid, Elinor Salegh-Haddad, Lior Laks, Sharon Armon Lotem, Elitzur Dattner, Ronit Levie, Orit Ashkenazi, Elisheva Shalmon and Osnat Kandelshine-Waldman** The acquisition of Semitic morphology in Hebrew and Arabic: Developmental cross-modal analyses of corpora
12:00 –12:30	**Jana Hasenäcker, Maria Ktori and Davide Crepaldi** Processing of compound constituents: position-specificity and interpretability	**Ingo Plag, Arne Lohmann, Sonia Ben Hedia and Julia Zimmermann** What is the difference between boys and boys'? The phonetics of plural vs. genitive-plural in English and its implications for morphological theory	
12:30 –13:00	**Roberto de Almeida, Megan Skurnac, Jordan Gallant and Gary Libben** Semantically ambiguous stems and the purpose of morphological processing	**Dominic Schmitz, Ingo Plag and Dinah Baer-Henney** How real are acoustic differences between different types of final /s/ in English? Evidence from pseudowords	
13:00 –14:30	**LUNCH BREAK**		
	TC 1.01	**TC 1.02**	**TC 3.06**
14:30 –15:00	**Yuta Tatsumi** A compositional analysis of deverbal compounds in Japanese	**Olivier Bonami and Matteo Pellegrini** Derivation predicting inflection. The role of families, series, and morphotactics	**WORKSHOP (cont.)** **Dorit Ravid, Elinor Salegh-Haddad, Lior Laks, Sharon Armon Lotem, Elitzur Dattner, Ronit Levie, Orit Ashkenazi, Elisheva Shalmon and Osnat Kandelshine-Waldman** The acquisition of Semitic morphology in Hebrew and Arabic: Developmental cross-modal analyses of corpora
15:00 –15:30	**Dimitra Melissaropoulou** Change in morphological complexity in the language contact perspective: evidence from the Cappadocian Greek derivational domain	**Judit Ács and András Kornai** Learning interpretable patterns for morphological analysis	
15:30 –16:00	**Maia Lomia** The borrowing of markers and semantic differentiation of forms in Georgian and Megrelian	**Matías Guzmán Naranjo** Inflection class and semantic analogies	
16:00 –19:30	**TC 3.06** **WORKSHOP (cont.)** **Dorit Ravid, Elinor Salegh-Haddad, Lior Laks, Sharon Armon Lotem, Elitzur Dattner, Ronit Levie, Orit Ashkenazi, Elisheva Shalmon and Osnat Kandelshine-Waldman** The acquisition of Semitic morphology in Hebrew and Arabic: Developmental cross-modal analyses of corpora	**TC 3.07** **WORKSHOP (end: 18:30)** **M. Silvia Micheli and Giorgio Francesco Arcodia** Word formation in diachrony	

Friday, February 7

<table>
<tr><td>9:00 – 10:00</td><td colspan="2">PLENARY TALK
Lameen Souag
Pattern morphology in contact</td><td>TC 0.01</td></tr>
<tr><td></td><td>TC 1.01</td><td>TC 1.02</td><td>TC 3.06</td></tr>
<tr><td>10:00 –10:30</td><td>Silvio Cruschina
Morphemes in competition: Lexically-conditioned allomorphy, sociolinguistic variation, and change in progress</td><td>Mary O‘ Brien, Sarah Macdonald and Gary Libben
Implicit and explicit awareness of morphemes in L2 German</td><td rowspan="6">WORKSHOP
Marcello Ferro, Claudia Marzi and Vito Pirelli
Implications of psycho-computational modelling for morphological theory</td></tr>
<tr><td>10:30 –11:00</td><td>Martin Maiden
Contamination, segmentation, and abstractive morphology in diachrony</td><td>Shinobu Imai
On the “extended” morphosyntactic and semantic features of Sino-Japanese verbs and native Japanese compound verbs</td></tr>
<tr><td>11:00 –11:30</td><td colspan="2">COFFEE BREAK
Ground floor</td></tr>
<tr><td></td><td>TC 1.01</td><td>TC 1.02</td></tr>
<tr><td>11:30 –12:00</td><td>Jelena Vujić
On N+N nominalizations with the invariable modifying constituent on Serbian</td><td>Arne Lohmann
Acoustic evidence for a category-specific metrical schema? – An extended replication of Sereno & Jongman’s (1995) reading study of noun-verb conversion homophones</td></tr>
<tr><td>12:00 –12:30</td><td>Francesco Gardani, Michele Loporcaro and Alberto Giudici
A borrowed prosodic marker for inherent inflection in Istro-Romanian</td><td>Shani Levy-Shimon, Dorit Ravid and Rachel Schiff
The influence of word frequency and root frequency on the spelling acquisition of an unstable root letter – a case study of the letter Aleph in Hebrew</td></tr>
<tr><td>12:30 –13:00</td><td>Anne Kruijt
Case systems in contact</td><td>Jordan Gallant
Morphological integration in the bilingual lexicon: Evidence from Chinese-English bilinguals</td></tr>
<tr><td>13:00 –14:30</td><td colspan="3">LUNCH BREAK</td></tr>
<tr><td>13:45 – 14:45</td><td colspan="3">GUIDED CAMPUS TOUR (if registered)</td></tr>
<tr><td></td><td>TC 1.01</td><td>TC 3.06</td><td>TC 3.07</td></tr>
<tr><td>14:30 – 15:00</td><td>Alberto Giudici and Chiara Zanini
Plural indefinite quantifier on the Romance-Slavic border</td><td rowspan="8">WORKSHOP (cont.)
Marcello Ferro, Claudia Marzi and Vito Pirelli
Implications of psycho-computational modelling for morphological theory</td><td></td></tr>
<tr><td>15:00 – 15:30</td><td>Alexandra Soares Rodrigues
The impact of Portuguese on the allomorphy of Mirandese prefixation and circumfixation: phonological conditions</td></tr>
<tr><td>15:30 –16:00</td><td>Martin Schäfer and Sven Kotowski
Understanding out-prefixation: merging qualitative and distributional analyses</td><td rowspan="7">WORKSHOP
TC 3.07
M. Silvia Micheli and Giorgio Francesco Arcodia
Word formation in diachrony</td></tr>
<tr><td>16:00 –16:30</td><td>Adam James Ross Tallman
The emergence of empty morphs in Chacobo (Pano) and Araona (Takana)</td></tr>
<tr><td>16:30 – 17:00</td><td>COFFEE BREAK
Ground floor</td></tr>
<tr><td>17:00 – 17:30</td><td>Bogáta Timár and Jeremy Bradley
The expression of epistemic modality in the Volga-Kama Sprachbund</td></tr>
<tr><td>17:30 – 18:00</td><td>Jeremy Bradley, Nikolett F. Gulyás and András Czentnár
Causative and reflexive morphology in the Volga-Kama region: a contact linguistic investigation</td></tr>
<tr><td>18:00 – 18:30</td><td>Márton András Baló
The aftermath of the borrowing of a borrowing pattern in Romani</td></tr>
<tr><td>20:00</td><td colspan="3">CONFERENCE DINNER</td></tr>
</table>

19th International Morphology Meeting

Vienna, February 6 – 8, 2020

Saturday, February 8

9:00 – 10:00	**PLENARY TALK** **Johanna Laakso** Reconstructed and constructed morphology: Can language planning turn back the wheel of time?		**TC 0.01**
	TC 1.01	**TC 1. 02**	
10:00 – 10:30	**Paul Kiparsky** In defense of level-ordering	**Peter Arkadiev** Borrowing non-canonical inverse between Kabardian and Abaza	
10:30 –11:00	**Elisa Mattiello and Wolfgang Dressler** Latinate vs. Non-Latinate English synthetic compounds	**Livio Gaeta** Deconstructing complexity: morphological change and language contact in Walser German	
11:00 –11.30	**Simon David Stein and Ingo Plag** Lexical storage and morphological segmentability effects on the production of English derivatives	**Laila Kjærbæk and Hans Basbøll** The development of derivatives in Danish-speaking children's spontaneous speech	
11:30 –12:00	**COFFEE BREAK**		**Ground floor**
12:00 –13:00	**POSTER SESSION**		**TC 3.06** **TC 3.07**
13:00 –14:30	**LUNCH BREAK**		
	TC 1.01	**TC 1.02**	
14:30 –15:00	**Nikolaos Ntagkas, Despoina Papadopoulou and Anna Anastassiadis-Symeonidis** Late visual morpho-semantic decomposition in Modern Greek derivational morphology	**Lior Laks** Hebrew maCCuC: A tale of a borrowed pattern	
15:00 –15:30	**Jenny Audring** Verbal diminutives between grammar and lexicon	**Thomas Stolz and Nataliya Levkovych** On the contact-induced emergence of grammatical gender	
15:30 –16:00	**Silvia Dal Negro** Borrowing routines in obsolescent languages: borrowing verbs in German minority dialects	**Bernard Fradin** On meaning predictability in derivational domains	
16:00 –16:30	**COFFEE BREAK**		**Ground floor**
	TC 1.01	**TC 1.02**	
16:30 –17:00	**Sven Kotowski** Modelling the semantics of out-prefixed verbs	**Michalis Marinis** On the role of intra-linguistic parameters of change in intensive language contact	
17:00 – 17:30	**Martha Booker Johnson** Reduction of multiple exponence in Kihehe	**Gilles Authier** Verbal morphological complexity in contact: the evidence from Lezgic (East Caucasian)	
17:30	**CLOSING**		**TC 0.01**

Attention: The building will be closed at 18:00.

WORKSHOPS

The acquisition of Semitic morphology in Hebrew and Arabic: Developmental cross-modal analyses of corpora

Convenors: Dorit Ravid[1], Elinor Saiegh Haddad[2] and Lior Laks[2]

Tel Aviv University[1] Bar Ilan University[2], Israel

Thursday | TC 3.06	
11:30 – 12:00	**Dorit Ravid[1]** Introduction: Semitic morphology in developmental perspective
12:00 – 12:30	**Ronit Levie[1] and Elitzur Dattner[1, 2]** Explaining dynamic morphological patterns in linguistic acquisition
12:30 – 13:00	**Osnat Kandelshine-Waldman[1] and Yedael Y. Waldman** The Semitic root from early to late childhood
13:00 – 14:30	**LUNCH**
14:30 – 15:00	**Orit Ashkenazi[1], Ronit Levie[1] and Ma'ayan Shissman Dagan[1]** Morphological sources of causativity in the acquisition of Hebrew verbs
15:00 – 15:30	**Elitzur Dattner[1, 2] and Dorit Ravid[1]** Hebrew pronominal subjects in acquisition
15:30 – 16:00	**Elisheva Shalmon[1] and Elitzur Dattner[1, 2]** Inflected Hebrew prepositions in developmental perspective
16:00 – 16:30	**Elinor Saiegh-Haddad[2]** Introduction: Empirical bases of the Arabic morphology projects
16:30 – 17:00	**Lior Laks[2], Elinor Saiegh-Haddad[2], Ibrahim Hamad[2] & Faten Yousef-Assadi[2]** The distribution of Arabic verbal patterns in text production: Between varieties and modalities
17:00 – 17:30	**Lior Laks[2] & Elinor Saiegh-Haddad[2]** Initiated usage of Arabic case markers in Standard Arabic text production
17:30 – 18:00	**Elinor Saiegh-Haddad[2], Hala Abdelhai[2] & Sharon Armon-Lotem[2]** Plural noun inflection in Palestinian Arabic among children with typical language development (TLD) and with developmental language disorder (DLD)
18:00 – 18:30	**Sharon-Armon-Lotem[2], Rawya Taha[2] & Elinor Saiegh-Haddad[2]** Acquisition of noun and verb morphology in spoken Arabic among children with typical language development (TLD) and children with developmental Language disorder (DLD)
18:30 – 19:00	**Elinor Saiegh-Haddad[2] and Rachel Schiff[2]** Morphological awareness in Arabic diglossia: The role of linguistic distance
19:00 – 19:30	Final words

Word formation in diachrony

Convenors: M. Silvia Micheli (University of Milano-Bicocca) and Giorgio Francesco Arcodia (Ca' Foscari University of Venice)

Thursday | TC 3.07	
16:00 – 16:30	Introduction by the convenors and coffee break
16:30 – 17:00	**Irene Fally and Pamela Goryczka** *Quid manet?* A diachronic approach to the Italian derivational suffixes *-izzare* and *-eggiare*
17:00 – 17:30	**Claudio Iacobini** Romance parasynthetic verbs. A comparative diachronic perspective
17:30 – 18:00	**Pavel Štichauer and Jan Radimský** *Nomina actionis* in the diachrony of Italian: a paradigm-based model of competition
18:00 – 18:30	**Maria Silvia Micheli and Giorgio Frencesco Arcodia** Emerging evaluative meanings in Neoclassical combining forms: two diachronic case studies of *para-* and *semi-*

Friday | TC 3.07	
15:30 – 16:00	**Maria Konoshenko and Christopher R. Green** The development of tonal head markers in Mande compounds: between phonology and grammar
16:00 – 16:30	**Coffee break**
16:30 – 17:00	**Thorsteinn G. Indridason** The diachronic productivity of the derivational suffix *–legur* in Icelandic
17:00 – 17:30	**Martina Werner** Nouns becoming longer? The question of nominal complexity increase in the history of German
17:30 – 18:00	**Fabio Montermini and Matteo Pellegrini** The role of derivational morphology in the shaping of the ancient Greek verbal system
18:00 – 18:30	**Sampsa Holopainen, Juha Kuokkala, Niklas Metsäranta and Ilona Rauhala** Challenges and opportunities of reconstructing Proto-Uralic nominal derivation

Implications of psycho-computational modelling for morphological theory

Convenors: Marcello Ferro, Claudia Marzi and Vito Pirelli

Friday | TC 3.06	
10:00	Introduction by the convenors
10:00 – 10:30	**James Blevins and Petar Milin** Representational agnosticism
10:30 – 11:00	**Fritz Günther and Marco Marelli** CAOSS and transcendence: On role-dependent constituent meanings
11:00 – 11:30	**Coffee break and poster session**
11:30 – 12:00	**Karla Orihuela and Hélène Giraudo** Perceptual salience of affixes
12:00 – 12:30	**Ingo Plag** Morphology? Which units? Which mechanisms?
12:30 – 13:00	**Evelyn Milburn, Mila Vulchanova and Valentin Vulachanov** Collocational frequency and context effects in idiom processing in advanced L2 speakers
13:00 – 14:30	**LUNCH**
14:30 – 15:00	**Elitzur Dattner, Ronit Levie, Orit Ashkenazi and Dorit Ravid** Explaining dynamic morphological patterns in acquisition using network analysis
15:00 – 15:30	**Sabine Arndt-Lappe, Robin Schrecklinger and Fabian Tomaschek** Stratification effects without morphological strata, syllable counting effects without actual counts, and what's in a trigram? – A simulation study of English stress with naïve discriminative learning
15:30 – 16:00	**Christina Manouilidou** Processing morphology: the view from language disorders
16:00 – 16:30	**Francesca Franzon and Chiara Zanini** Functional and semantic properties modulate information in inflectional features
16:30 – 17:00	**Coffee break and poster session**
17:00 – 17:30	**Maria Heitmeier and Harald Baayen** Simulating phonological and semantic impairment of English tense inflection with linear discriminative learning
17:30 – 18:00	**Yu-Ying Chuang, Kaidi Lõo, James P. Blevins and R. Harald Baayen** Estonian case inflection made simple. A case study in word and paradigm morphology with linear discriminative learning
18:00 – 18:30	**Michael Ramscar** Some functions of systems of morphological contrasts in the discriminative
18:30	Concluding remarks

Posters
Marco Marelli, Marco Petilli and Fritz Günther Grounding transparency in vision: image-based distributional models and the perceptual side of compound-word comprehension
Elnaz Shafaei-Bajestan and Harald Baayen Wide learning of the comprehension of morphologically complex words: from audio signal to semantics

Emanuele Casani Morphological errors in typical and atypical reading development. Is it deep dyslexia or morphosyntactic competence? Data by Italian dyslexic children with and without SLI
Romain Brasselet and Davide Crepaldi Minimal parsimonious chunking of written language

POSTER SESSION

Saturday, February 8

TC 3.06 and TC 3.07 | 11:30 – 12:30 h

Nino Amiridze
Borrowing feminine marking in Middle vs. Modern Georgian

Sabine Arndt-Lappe and Melanie J. Bell
Linguistic experience shapes word-formation patterns – evidence from novel formations by native and non-native speakers of English

Hans Basbøll and Laila Kjærbæk
Derivational morphology in Danish: testing a general model of word structure and its relation to lexicon and prosody, and contrasting it with a comprehensive description of Danish word formation

Laura Becker
When verbal markers go nominal: from irrealis to nonspecific and realis to specific

Gilles Boyé
Information theory and morphology: some caveats

Irene Fally
Integration of English verbs in Italian: competing morphological realizations

Eduard Fomin
Adoption of Russian suffixes by the Chuvash language

Francesca Franzon, Alessia Zampieri and Davide Crepaldi
The role of semantics in learning morphological systems. An artificial lexicon experiment

Francesca Franzon, Chiara Zanini and Valentina Pescuma
Form, function, meaning. A study on the distribution of inflectional morphemes in Italian

Adriana Rosalina Galván Torres, Katharina Korecky-Kröll and Michael Schmid
Language contact with English influences learners' production of German comparatives and superlatives: evidence from adult native speakers of Mexican Spanish

Anthony Grant
Bound verbal morphology and borrowing: limits and possibilities

Camiel Hamans
Blends: category at the crossroads of morphology and phonology

Gary Libben, Sarah Macdonald, Learose Pinkham, Mira Reisinger, Michaela Ringseis, Kaitlin Goertz, Katharina Korecky-Kröll and Wolfgang Dressler
A multi-method investigation of morphological processing in German

Fujio Nakamura
An exploration into why the establishment of the negative contraction doesn't was delayed in American English

Maria Rosenberg
Adjectives as a case of word-class changing inflection in the early course of Swedish language development

Thomas Stewart
Morphological construction schemas distribute initial consonant mutation in Scottish Gaelic

Martina Werner
Domain adverbs in German: rivalry or split? A corpus study of the semi-suffixes –technisch and –mäßig

Alina Villalva, Sydelle de Souza and Carina Pinto
Word associations: analyzing activation pathways

Jelena Zivojinovic
The so-called "gerund" in Latin

Plenary Talks

Johanna Laakso (Universität Wien)

Reconstructed and constructed morphology: Can language planning turn back the wheel of time?

The comparative method is based on the interaction of historical phonology – the axiom of the regularity of sound change – and historical lexicology (etymology): the former makes the rules, the latter delivers the material for testing, modifying and enhancing them. This works well with the reconstruction of lexical stems. The reconstruction of grammatical or bound morphs, however, is much more problematic. Affixes play by different rules: they are often affected by system-driven changes ("analogy"), and their structure can be phonotactically more strictly constrained than that of lexical stems. Furthermore, the comparative method in itself has a typologically distorting effect: it is possible to reconstruct invariant proto-forms behind today's variation, but once variation is completely levelled, it becomes irretrievable.

This effect is the cause of the well-known fact (see e.g. Chafe 1959, Korhonen 1974) that reconstructed proto-languages tend to be more regular as concerns morphophonology than the languages used as a starting point for the reconstruction. In other words, reconstructions tend to be closer to what was traditionally called the agglutinative type (for a detailed critique of the term, see Arkadiev forthcoming) – possibly also because reconstructing the substance of affixes is less controversial than reconstructing syntactic structures or patterns. Accordingly, the developments from the reconstructed proto-form to today's language would correspond to the classical typological cycle, or at least part of it: from the agglutinative type towards increasing fusion and/or isolation. However, examples of the opposite, i.e. recreation of clearly segmentable affixal morphology, are not difficult to find.

In morphologically rich languages, language planning and standardization typically involves the regulation of morphology, i.e. taking a stance to variation and morphophonological alternations, or even introducing new affixes or inflection types. Modern Standard Estonian is an example of a highly planned and standardized language, into which some new derivational or inflectional affixes have been implemented by language planning (Raag 1998). The strong tradition of language planning has also provoked criticism. Already Kaplinski (1984) accused Estonian language planners not only of "confusing" the language with "unnatural and artificial" constructs but also of "trying to turn back the wheel of time" by (re)introducing obsolete inflectional or derivational elements or patterns which would have already been replaced by analytic expressions.

In my talk, I will use examples mainly from Estonian and other Finnic languages to demonstrate in what respects and how planned change in morphology can mimic reconstruction and possibly even reverse typological change. The examples will shed light on how different factors – variation, language contact, cultural and political circumstances – conspire in archaizing language planning.

References

Arkadiev, Peter (forthcoming). Morphology in typology: historical retrospect, state of the art and prospects. To appear in Rochelle Lieber et al. (eds.), *The Oxford Encyclopedia of Morphology.* [June 2019 version on academia.edu]

Chafe, Wallace A. 1959. Internal reconstruction in Seneca. *Language* 35:3, 477–495.

Kaplinski, Jaan 1984. Keelekorralduse süvastruktuurist ("Kirjakeele korraldus nüüd ja praegu"). *Keel ja Kirjandus* 27, 456–464.

Korhonen, Mikko. 1974. Oliko suomalais-ugrilainen kantakieli agglutinoiva? Eli mitä kielihistoriallisista rekonstruktioista voidaan lukea ja mitä ei. (English summary: Was Proto-Finno-Ugric an agglutinative language? Or what can and cannot be deduced from linguistic reconstruction.) *Virittäjä* 78: 243–257.

Raag, Virve. 1998. *The Effects of Planned Change on Estonian Morphology* (Studia Uralica Upsaliensia 29). Uppsala: Acta Universitatis Upsaliensis.

Felicity Meakins

Two to tangle: Modelling morphologies in contact

Most cases of language creation occur as languages split and diverge from other languages in processes which are conceptualised in linguistic phylogenies. More rarely, some languages emerge abruptly over the course of one or two generations in language communities as a result of language contact. In this talk, I report on the rapid birth of Gurindji Kriol, which shows a morpho-syntactic and lexical split between Gurindji (Pama-Nyungan) and Kriol (English-based creole). This work is the first investigation of contact-induced change within a single speaker population which uses multiple variants. I will outline an innovative modification of classic population genetics methods to investigating change over time in the Gurindji speech community. This method, which has been developed with Lindell Bromham and Xia Hua, aims to increase our ability to explain language change, with a view to making predictions about how languages will change, in particular the links between morphological simplification and language contact.

Lameen Souag

Pattern morphology in contact

In many Afroasiatic languages, a morphologically complex word is often structured such that the consonants are determined by the input root or stem, while the vowels and word shape are determined by a fixed pattern, largely independent of the input, marking another morpheme. This type of templatic morphology is often called root-and-pattern morphology; since its input is often a stem rather than a root, *pattern morphology* seems a preferable label. While reminiscent of Indo-European ablaut in some respects, and of California Penutian non-vocalic pattern morphology in others, the productive use of transfixational morphemes with fixed vowel segments and word shape appears to be unique to the Afroasiatic phylum (Arcodia 2013). The need to account for this phenomenon without overgenerating has been a spur to the development of morphological theory, from Prosodic Morphology (McCarthy & Prince 1990) to consonantal root rejection (Bat-El 1994; Ussishkin 1999) and beyond. Its behaviour in language contact situations, however, has received much less attention than its system-internal behaviour within specific languages.

Derivational morphology in particular is frequently borrowed in language contact situations, and pattern morphemes are no exception. Arabic has been the principal contributor of such forms in contact situations; recipient languages reported in the literature include Ghomara Berber for the diminutive template *-CCəyCəC-* (Mourigh 2016), Western Neo-Aramaic for the elative template *aCCaC* (Arnold 2007), and Israeli Hebrew for the slang adjectival template *maCCuC* (Bolozky 1999), to name some of the best-described cases.

A general survey of known examples of borrowed templatic morphology, including several previously unidentified ones, reveals a sufficiently large number of examples to identify patterns within the data. The most striking of these patterns is genetic: whereas Afroasiatic languages have in several instances borrowed pattern morphemes from one another, non-Afroasiatic languages have rarely, if ever, borrowed such morphemes productively, even in situations of similarly intense contact. Potential exceptions, such as Arabic broken plurals in Persian (Gardani fc), are not only strikingly marginal within the recipient languages but susceptible to alternative historical analyses. On the widely held assumption that language contact outcomes are primarily determined by social rather than structural factors (Thomason & Kaufman 1988), this result is counterintuitive; Arabic and other Afroasiatic languages have profoundly influenced non-Afroasiatic languages as well as Afroasiatic ones, and the lack of pattern morpheme borrowing in the former cannot be explained in terms of social factors alone.

Morpheme-specific case studies, in particular of the borrowing of the Arabic elative, cast light on the reasons for this difference. Within Semitic, the existence of lookalike cognates often facilitates the reinterpretation of borrowed items as morphologically complex forms built on patterns that can be reapplied. Even beyond Semitic, however, the prior presence of formally similar consonant extraction processes applied in similar contexts appears to give the relevant morphemes a foot in the door. Analysing the factors found to facilitate borrowing across systems provides a new set of data bearing on the question of which synchronic analysis of pattern morphology should be preferred.

References

Arcodia, Giorgio Francesco. 2013. Nonconcatenative morphology in typological perspective. In Giorgio Francesco Arcodia, Federica Da Milano, Gabriele Iannàccaro & Paolo Zubena (eds.), *Tilelli. Studi in onore di Vermondo Brugnatelli*, 1–14. Roma: Caissa Italia.

Arnold, Werner. 2007. Arabic grammatical borrowing in Western Neo-Aramaic. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*, 151–164. Berlin: Mouton de Gruyter.

Bat-El, Outi. 1994. Stem Modification and Cluster Transfer in Modern Hebrew. *Natural Language & Linguistic Theory* 12(4). 571–596.

Bolozky, Shmuel. 1999. *Measuring Productivity in Word Formation: The Case of Israeli Hebrew* (Studies in Semitic Languages and Linguistics 27). Leiden: Brill. https://brill.com/view/title/6571 (9 December, 2019).

Gardani, Francesco. fc. Morphology and contact-induced language change. In Anthony Grant (ed.), *The Oxford handbook of language contact*. Oxford: Oxford University Press.

McCarthy, John J. & Alan Prince. 1990. Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 209–282.

Mourigh, Khalid. 2016. *A Grammar of Ghomara Berber (North-West Morocco)* (Berber Studies 45). Köln: Rüdiger Köppe.

Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.

Ussishkin, Adam. 1999. The Inadequacy of the Consonantal Root: Modern Hebrew Denominal Verbs and Output: Output Correspondence. *Phonology* 16(3). 401–442.

Oral Presentations

Judit Ács and András Kornai

Learning Interpretable Patterns for Morphological Analysis

Abstract

Neural networks are remarkably good at morphological inflection and analysis (Cotterell et al., 2018) but they perform the task in a black box fashion. Obtaining explicit morphemes and morpheme boundaries in machine learning systems remains a major challenge. Even with hard monotonic attention (Aharoni and Goldberg, 2016), machine learning systems remain hard to interpret, especially when compared to traditional rule-based analyzers that explicitly encode inflectional morphology (Koskenniemi, 1983).

Soft Patterns or SoPa (Schwartz, Thomson, and Smith, 2018) is a finite state automaton parameterized by a neural network which has been shown to discover interpretable word patterns for sentiment analysis. We employ this model at a character-level setting and show that that it is capable of discovering morpheme-like patterns. Our results offer the first systematic means of bridging the gap between linguistic intuitions and machine learners.

Sequence-to-sequence neural networks (Sutskever, Vinyals, and Le, 2014), also called *encoder-decoder* models map arbitrary sequences to each other. The encoder creates a continuous representation, typically a vector of a sequence which in turn is used by the decoder to generate an output sequence. The latter is often aided by the so-called *attention mechanism* (Bahdanau, Cho, and Bengio, 2015; Luong, Pham, and Manning, 2015), which allows peaking into the input sequence during decoding. These models were popularized in neural machine translation and have since then become ubiquitous in natural language processing.

We show that SoPa can be used as the encoder of a sequence-to-sequence model and thus it can be applied to morphological analysis. Our setup assigns a sequence of morphological tags to an inflected word such as `V PST` to *walked*.

The patterns learned by the encoder are scored over each character span and the highest scoring span is recovered through the forward algorithm. These scores are used to initialize the LSTM decoder's hidden state as well as the inputs for attention (Luong, Pham, and Manning, 2015) during decoding. We examine these spans for how well they conform to our linguistic intuitions about inflectional morphology.

We experiment with four languages with concatenative morphology (Finnish, Estonian, Hungarian, and Turkish) and we provide detailed analysis of our

Language	LSTM	SoPa low	SoPa high
Estonian	80.9	53.6	75.75
Finnish	82.55	78.95	84.15
Hungarian	93.55	93.5	94.1
Turkish	71.75	72.45	73.55

Table 1: Test accuracy of each model. An output is accurate if it matched the gold sequence.

findings in Hungarian. Although our primary focus is interpretability, our systems are competitive against an LSTM-based encoder-decoder with attention (Hochreiter and Schmidhuber, 1997; Bahdanau, Cho, and Bengio, 2015) (see Table 1).

We experiment with two hyperparameter setting for SoPa, a low-resource one (10 patterns), and a high-resource one (150 patterns) that has roughly that same number of trainable parameters as the LSTM baseline. We train each model on 80,000 training samples, evaluate on 20,000 dev samples and test on 20,000 test samples. Inflected forms of the same lemmas do not overlap between the three splits. Our datasets are sampled from the UniMorph 2.0 dataset (Kirov et al., 2018).

We continue by examining the substrings matched by the learned patterns. We compute the correlation between each pattern and each morphological tag and we find that the substrings matched by patterns highly correlated with tags, resemble morphemes, although morpheme boundaries are not clear. Table 2 lists the top spans for the morphological tags attested in Estonian, Finnish and Hungarian. In the final paper, we will provide full examples for Turkish as well.

Tag	**Estonian**	**Finnish**	**Hungarian**
2	e$, d$, ke$, lda$, a$	tte$, sit$, net$, aa$, tit$	k$, tok$, ok$, tek$
AT+ABL	lt$, t$, lt, elt$, elt	lt, ta$, tä$, lta$, ta	ól$, tól$, től$, ől$, l$
AT+ALL	le$, ele$, e$, ele, le	e$, lle$, ll, le$, le	z$, hoz$, ho, oz, hez$
AT+ESS	l$, el$, el, tel$, il	la$, ll, ill, lla, lä$	ál$, nál$, él$, l$, nél$
COND	ks, ksi, taks, ks$, taksi	isi, si, iso	lná, lné, nk$, nék$, nánk$
FRML	a$, na$, ena$, na, tena	na$, ina$, na, nä$, inä$	én, t$, nt$, ént, ént$
IN+ABL	t$, st, est, st$, est$	sta$, ta$, st, tä$, stä$	ől$, ól$, ból$, ből$, l$
IN+ALL	e$, sse, ss, sse$, se	n$, in$, ihin, iin, hin	ba$, be$, kb, kbe$, áb
IN+ESS	s$, is$, as$, us$, ses$	ssa$, ss, sa$, ssa, ssä$	n$, ban$, an$, ben$, kb
PL	te, de, id, st, ud	si, en, in, isi, ei	k$, kk, ek, ok, kb
TRANS	s$, ks$, ks, tek, iks	si$, ksi, ksi$, si, iksi$	á$, é$, kká$, kké$, ká$
N	te, t$, a$, st, ^k	a$, n$, ta$, si, ti	l$, t$, ól$, rt$, ől$
V	ta, ks, taks, sid, it	tan, tais, tte, taa, tak	k$, d$, nk$, a$

Table 2: Top scoring patterns for tags attested in all three Uralic languages. Patterns correllating with open classes, namely nouns and verbs, do not exhibit systematic patterns, which further confirms that SoPa indeed learns morpheme-like patterns when possible.

Pattern distributions corresponding to open and closed classes exhibit different statistical cues such as higher variance and a higher number of patterns. We will examine these in the final paper.

References

Aharoni, Roee and Yoav Goldberg (2016). "Morphological inflection generation with hard monotonic attention". In: *arXiv preprint arXiv:1611.01487*.

Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio (2015). "Neural machine translation by jointly learning to align and translate". In: *International Conference on Learning Representations (ICLR 2015)*. arXiv: 1409.0473 [cs.CL].

Cotterell, Ryan et al. (2018). "The CoNLL–SIGMORPHON 2018 Shared Task: Universal Morphological Reinflection". In: *Proceedings of the CoNLL–SIGMORPHON 2018 Shared Task: Universal Morphological Reinflection*. Brussels, Belgium: Association for Computational Linguistics.

Hochreiter, Sepp and Jürgen Schmidhuber (Nov. 1997). "Long Short-Term Memory". In: *Neural Computation* 9.8, pp. 1735–1780.

Kirov, Christo et al. (May 2018). "UniMorph 2.0: Universal Morphology". In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Ed. by Nicoletta Calzolari (Conference chair) et al. Miyazaki, Japan: European Language Resources Association (ELRA). ISBN: 979-10-95546-00-9.

Koskenniemi, Kimmo (1983). "Two-level model for morphological analysis". In: *Proceedings of IJCAI-83*, pp. 683–685.

Luong, Thang, Hieu Pham, and Christopher D. Manning (2015). "Effective Approaches to Attention-based Neural Machine Translation". In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Lisbon, Portugal: Association for Computational Linguistics, pp. 1412–1421. DOI: 10.18653/v1/D15-1166. URL: http://www.aclweb.org/anthology/D15-1166.

Schwartz, Roy, Sam Thomson, and Noah A. Smith (2018). "SoPa: Bridging CNNs, RNNs, and Weighted Finite-State Machines". In: *Proc. 56th ACL Annual Meeting*. Melbourne, Australia, pp. 295–305.

Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le (2014). "Sequence to Sequence Learning with Neural Networks". In: *Proc. NIPS*. Montreal, CA, pp. 3104–3112. URL: http://arxiv.org/abs/1409.3215.

Peter Arkadiev

BORROWING NON-CANONICAL INVERSE BETWEEN KABARDIAN AND ABAZA

In this paper I shall discuss a hitherto unreported case of pattern-borrowing (Matras & Sakel 2007) of a typologically non-trivial morphological pattern between two distantly related polysynthetic languages of the Northwest Caucasian family, Kabardian and Abaza. Kabardian is spoken by more than 500 thousand speakers in the Russian republics of Karachay-Cherkessia and Kabardino-Balkaria, while Abaza is spoken by ca. 35 thousand people in several compact districts in Karachay-Cherkessia. The two languages, which are typologically similar but mutually unintelligible, are in a state of intense contact: most speakers of Abaza also know and often use Kabardian; it is common for Abaza men to marry Kabardian women, who then have to learn Abaza. Abaza has numerous lexical and some morphological borrowings, as well as morphosemantic calques from Kabardian.

Both languages are morphologically ergative and head-marking, expressing up to four participants by pronominal prefixes in the verb, in the following order: Absolutive – (Indirect objects) – (Ergative) – Stem. In Abaza, all pronominal prefixes are overt, while in Kabardian 3rd person absolutive and some 3rd person indirect object prefixes are null. Bivalent verbs come in two partly semantically motivated types: transitive verbs take an ergative subject and an absolutive object (1a,b), while intransitive verbs take an absolutive subject and an indirect object (2a,b). Ditransitive verbs take an absolutive theme and an indirect object recipient (3a,b). (Kabardian examples from Kumakhov ed. 2006; Abaza examples elicited.)

(1) a. ***sə-b-****ew-h* Kabardian
1SG.ABS-2SG.ERG-PRS-carry
'You (sg) are carrying me.'

b. ***wə-l-****bá-ṭ* Abaza
2SG.M.ABS-3SG.F.ERG-see(AOR)-DCL
'She saw you (man).'

(2) a. ***s-j-****e-ẑ-a-ŝ* Kabardian
1SG.ABS-3SG.IO-DAT-wait-PST-DCL
'I waited for him/her.'

b. ***hə-j-****pšə́-ṭ* Abaza
1PL.ABS-3SG.M.IO-look(AOR)-DCL
'We looked at him.'

(3) a. ***∅-f-e-s-****t-a-ŝ* Kabardian
3.ABS-2PL.IO-DAT-1SG.ERG-give-PST-DCL
'I gave it to you (pl).'

b. ***j-ŝə́-s-****t-ṭ* Abaza
3SG.N.ABS-2PL.IO-1SG.ERG-give(AOR)-DCL
'I gave it to you (pl.)'

All Northwest Caucasian languages have complex marking of spatial relations in the verb, including two deictic prefixes, cislocative (Kabardian *q̇(V)-*, Abaza *ʕ(a)-*) 'hither' and translocative (Kabardian *n(V)-*, Abaza *n(a)-*) 'thither'. Kabardian has integrated these prefixes into its person paradigms of polyvalent verbs with indirect objects in a fashion reminiscent of inverse marking. The cislocative, occupying the slot immediately following the absolutive, appears when the indirect object is higher than the ergative or absolutive subject on the person hierarchy (4a,b), and the translocative is used in some 1>2 combinations (4c).

(4) a. ***∅-qə-****w-jə-t-a-ŝ* Kabardian
3,ABS-**CISL**-2SG.IO-3SG.ERG-give-PST-DCL
's/he gave it to you'

b. *∅-**qə**-zə-b-ew-t*
3.ABS-**CISL**-1SG.IO-2SG.ERG-PRS-give
'you (sg) give it to me'

c. *sə-**n**-w-ew-ẑe*
1SG.ABS-**TRAL**-2SG.IO-PRS-wait
'I am waiting for you.'

Basically the same pattern is found in the fieldwork data from Abaza collected in 2017–2019 in the village Inzhich-Chukun (5a,b).

(5) a. *j-**ʕa**-wə́-l-t-ṭ* Abaza
3SG.N.ABS-**CISL**-2SG.M.IO-3SG.F.ERG-give(AOR)-DCL
'She gave it to you (man).'

b. *h-**na**-w-pšə́-ṭ*
1PL.ABS-**TRAL**-2SG.M.IO-look(AOR)-DCL
'We looked at you (man).'

That the quasi-inverse uses of the Abaza deictic prefixes is a case of pattern replication from Kabardian is supported by the following considerations: 1) The use of the cislocative prefix in person paradigms is a feature Kabardian shares with its close relative West Circassian, which has never been in contact with Abaza. By contrast, nothing similar is found in Abkhaz, the close relative of Abaza spoken on the other side of the Caucasian range. 2) The quasi-inverse use of the cislocative is recorded in all grammars of Kabardian, while for Abaza this is a not yet stabilized and fully recognized pattern, conspicuously ignored by existing sources (e.g. Genko 1955, Tabulova 1976, Lomtatidze et. al. 1989, O'Herin 2002). 3) In their original spatial meanings the Abaza deictic prefixes occupy the slot closer to the root (6a,b), which indicates that their quasi-inverse uses shown in (5a,b) mirror both the distribution and position of their Kabardian models.

(6) a. *j-rə-z-**ʕə́**-ʕ-ga-ṭ* Abaza
3SG.N.ABS-3PL.IO-BEN-**CISL**-1PL.ERG-carry(AOR)-DCL
'We brought it to them.' (textual example)

b. *čə-r-zə-**na**-hə-r-χa-rnəs*
RFL.ABS-3PL.IO-BEN-**TRAL**-1PL.ERG-CAUS-turn-PURP
'for us to turn ourselves towards them' (textual example)

References:

Genko, Anatolij N. 1955. *Abazinskij jazyk. Grammatičeskij očerk narečija tapanta* [*The Abaza Language. A Grammatical Sketch of the Tapanta Dialect*]. Moscow: Izdatel'stvo Akademii Nauk SSSR.

O'Herin, Brian. 2002. *Case and Agreement in Abaza*. Arlington: SIL International.

Kumakhov, Mukhadin A. (ed.). 2006. *Kabardino-čerkesskij jazyk* [*The Kabardian Language*], Vols. 1-2. Nal'čik: Èl'-Fa.

Lomtatidze, Ketevan, Rauf N. Klychev & B. George Hewitt. 1989. Abaza. In: B. George Hewitt (ed.), *The Indigenous Languages of the Caucasus. Vol. 2. The North West Caucasian Languages*. Delmar, N.Y.: Caravan, 91–154.

Matras, Yaron & Jeannette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31.4, 829–865.

Tabulova, Nurya T. 1976. *Grammatika abazinskogo jazyka. Fonetika i morfologija* [*A Grammar of Abaza: Phonetics and Morphology*]. Čerkessk: Karačaevo-Čerkesskoe otdelenie Stavropol′skogo knižnogo izdatel′stva.

Jenny Audring

Verbal diminutives between grammar and lexicon (oral presentation)

Germanic has classes of complex verbs known as "verbal diminutives" or "frequentatives", because they typically denote repetitive activities of low intensity. The two derivational suffixes associated with this pattern are *-el* and *-er* (and their variants). Examples are:

Dutch
snuffelen 'sniffle', *grommelen* 'grumble', *stotteren* 'stutter', *piekeren* 'worry, brood'
German
krabbeln 'scratch', *baumeln* 'dangle', *räuspern* 'clear one's throat', *zwitschern* 'twitter'
Englisch
sparkle, crumble, simmer, bicker
Norwegian
fomle 'fumble', *somle* 'dither', *sitre* 'shiver', *flagre* 'flutter'

These verbs, which are often overlooked or marginalized in the linguistic literature, present interesting challenges. They occur in substantial numbers and show evident morphological structure, yet the pattern is virtually unproductive in the standard languages. Despite the native suffix, non-lexical roots abound. Moreover, many verbs are derivationally ambiguous due to the existence of the homophonous nominal suffixes *-el* (for diminutives and instrument nouns) and *-er* (for agentives and instruments) and productive N-to-V conversion.

In this talk, I discuss verbal diminutives as a representative of what is often considered marginal morphology. Yet, such patterns offer valuable insights for morphological theory, in particular with regard to the relation between morphology and phonology, and between the grammar and the lexicon. As to the first, I show that phonological similarity can go hand in hand with semantic alignment effects with the morphological schema: words can behave like verbal diminutives even though their suffix is derivationally nominal (Weidhaas & Schmid 2015, Audring et al. 2017). Similar "Gestalt" effects (for which see also Köpcke & Panther 2016) can be seen in other domains, e.g. in gender assignment, where phonological 'impostors' can participate in morphological assignment systems. As to the second issue, I argue that the lack of lexical bases and the absence of synchronic productivity make verbal diminutives unsuitable for modelling in terms of word-formation rules: the existing complex forms must all be listed. Yet, the pattern clearly belongs to the system of derivational morphology, and hence to grammar. I interpret both observations as evidence in favour of a constructional approach (Booij 2010, Jackendoff & Audring 2019) involving declarative schemas that can have a generative, but also a *motivating* function. In particular, I explore the notions of *partial motivation* and *multiple motivation*, for which the verbal diminutives provide an instructive illustration.

Audring, Jenny, Geert Booij & Ray Jackendoff. 2017. Menscheln, kibbelen, sparkle: Verbal diminutives between grammar and lexicon. In Bert Le Bruyn & Sander Lestrade (eds.), *Linguistics in the Netherlands 2017*. Amsterdam: Benjamins.

Booij, Geert. 2010. *Construction Morphology*. Oxford: OUP.

Jackendoff, Ray & Jenny Audring. 2019. Relational Morphology in the Parallel Architecture. In Audring, Jenny & Masini, Francesca (eds.), *The Oxford Handbook of Morphological Theory*. Oxford: OUP.

Köpcke, Klaus-Michael & Klaus Uwe Panther. 2016. Analytische und gestalthafte Nomina auf -er im Deutschen vor dem Hintergrund konstruktionsgrammatischer Überlegungen. In Andreas Bittner & Constanze Spieß (eds.), *Formen und Funktionen. Morphosemantik und grammatische Konstruktion.*, 85–101. Berlin: De Gruyter.

Weidhaas, Thomas & Hans-Jörg Schmid. 2015. Diminutive verbs in German: semantic analysis and theoretical implications. *Morphology* 25(2). 183–228.

Gilles Authier

Verbal morphological complexity in contact: evidence from Lezgic (East Caucasian)

Different types of language contact can lead to simplification or complexification. While the first type is easy to identify as leading to the minimal morphologies found in pidgins and creoles, it proves more difficult to relate the state of affairs in morphologically complex languages to contact, or to explain why genetically related languages may become simpler or more complex in similar contact environments, especially in the absence of historical documents and sources.

Our paper will attempt to draw the general lines of morphological evolution of verbs in Lezgic, a deep and clearly articulated branch of East Caucasian (EC, aka Nakh-Daghestanian) comprising a dozen languages which have all undergone prolonged contact with unrelated (Turkic, Iranian) or related languages. Thanks to external comparison (i.e. the availability of at least five other branches for EC, against which internally reconstructed proto-Lezgic morphology can be checked), the set of inherited verbal inflectional segments is well established and borrowed formatives can be identified.

The main categories involved in the morphological complexity of verbs in East Caucasian are 1) gender/number agreement with S/P arguments, 2) synthetic, root-adjacent aspect marking, 3) layered deictic preverbs, 4) tense and modality formatives, 5) synthetic negation. All these categories can be reconstructed for proto-Lezgic, and have only rarely been lost altogether in the daughter languages. Additionally, 6) person indexing developed fully and independently in two distantly related Lezgic languages, Udi and Tabasaran, as a structural copy of Azeri Turkic person agreement.

We will assess the level of complexity found in each Lezgic language's verbal morphology in correlation with the amount and types of language contact they are known to have undergone. This branch has been chosen for the variety of contact situations and the various levels of verbal complexity it offers, from quite low in Lezgian and Agul to high in Kryz, Archi, and Rutul, to very high in Budukh, Tsakhur and Tabasaran. The complexity of expression of these categories will be evaluated along the parameters of a) number of subcategories, b) inflectional exponence, c) lexical exponence (bipartite stems), d) allomorphy, e) transparency/syncretism, f) suppletion, g) fusion.

Our main finding is that the level of complexity does not seem to have risen or fallen significantly from reconstructible proto-Lezgic to the average level found in the contemporary languages. In general, most of the verbal affix systems found in Lezgic languages continue systems already present in their reconstructed parent language, with the exception of personal indexes, which all derive from inherited free pronouns, and just one instance of affix borrowing[i]. Some inherited categories like gender agreement have dwindled – thereby leading to further complexification as in Tabasaran – or disappeared entirely as in Udi, Agul and Lezgian, but a few new categories like person indexation or non-indicative modalities have emerged in individual languages, and those most probably as copying of Turkic verbal categories : Kryz and Budukh have developed a debitive mood calqued on the same category found in Azeri; in Rutul and Tsakhur as in other, non-Lezgic languages of Daghestan, the development of optative morphemes from univerbation of an auxiliary 'say' on imperative forms is a parallel innovation probably triggered by contact with Turkic.

Our second finding is that in the majority of Lezgic languages, complexity parameters and domains tend to compensate for each other, to the extent that complexity in one part of the verbal morphology of a language seems to correlate with simplicity in another. For instance, Tabasaran has reduced its gender system to a human / non-human opposition, but allomorphy and syncretisms have increased the complexity of verbal agreement in terms of transparency; On the other hand, Tabasaran innovated a dichotomy between the older class of derivative preverbed verbs, which lost most of the original stem aspect marking, and a new class of 'strong' verbs, without semantic preverbs, which use inflectional preverbs to mark subcategories of the perfective (aorist and perfect). This process cannot be historically connected to the similar grammaticalization of deictic element into aspect-marking 'bounders' in Indo-European or Kartvelian. Also in Tabasaran, innovative person marking has drifted

further than attested in the source language (Azeri), where it indexes only S/A (subject) arguments, to index also P arguments, non-core arguments (dative, potential subjects) and even non-arguments (genitive modifiers). Along the same lines, Tsakhur has developed fusional morphology at the juncture of gender and other prefixes, including Turkic-type vowel harmony and other sandhi phenomena, but its TAM system, being suffixal and strictly concatenative, remains perfectly transparent.

Structural complexity may be downplayed in many languages, including conservative ones like Kryz or Archi, by reducing the number of synthetic verbs in favour of light verb constructions: the famously large verbal paradigm of Archi (1.5 million forms, according to Chumakina and Corbett[ii]) is compensated for by a high degree of regularity in its processes, no fusion, and the low number of synthetic verbs overall. In Archi, a definitely low-contact situation accounts for the preservation of inherited complexity. But a formerly high level of contact with Lak (another branch of EC) accounts for the parallel grammaticalisation of 'say' as a highly productive light verb and more generally the development, in Archi, of a large class of verbs much simpler to conjugate than older synthetic verbs.

Such light verb constructions, often overlooked in studies of verbal complexity, are a major source of morphological change in all EC languages that have been in prolonged contact with non-EC or EC but distantly related languages. In the case of Lezgic languages, the number of synthetic verbal lexemes compared to analytic (verb compound) formations has certainly decreased, but periphrastic renewal of the verb stock did not always lead to real simplification. It is sometimes the case that complex verbs can be a source of greater complexity when their analytic structure is blurred and the inflectional morphology becomes trapped inside bipartite stems: this typologically rare development has taken place in about half of the Lezgic languages, with various morphophonemic consequences. Although trapped inflectional morphology generally tends to migrate to a more external position, an intermediate stage is represented by languages like Kryz and Ikhrek Rutul, which show multiple exponence of gender/number morphology in the present and future tense due to incomplete univerbation of the tense-forming copulas[iii], thereby violating the 'inflection-outside-derivation principle' (Haspelmath 1993[iv]).

In Budukh, bipartite stems are purely morphological, having fused to the level of morphomes and represent much higher complexity, to the point of a new Semitic-like morphological type, best described in terms of root-and-pattern. Without significant increase of inflectional synthesis, both Budukh and Tsakhur represent the highest levels of opacity: verb stems have become hard or impossible to segment. Socio-linguistically, both languages have been in regular contact with lowland languages (mainly Azeri, but also Georgian in the case of Tsakhur, and Tat in the case of Budukh) during winter pastures, but remained closed, endogamous and tight-knit communities.

In other languages, verbal locutions have gone the whole way to compounds and univerbation, and have usually brought about simplification, for instance in Lezgian, where negation is still prefixal in some non-finite forms but has otherwise migrated to suffixal position after being attached to tense-forming copulas or, in the prohibitive mood, to the conative auxiliary 'to do'. Such regularisation processes happened mainly in higher contact situations, where more contact between adults is documented, with correlate imperfect acquisition, and prevented the preservation of existing complexity. The case of Lezgian is not surprising: being a large language (400 m speakers) spread over a wide area including lowlands and in contact with Azeri (Turkic) and Tat (Iranian), adult second-language acquisition, exogamy and multilingualism have probably been more widespread than in smaller, more isolated Lezgic language communities.

---

[i] in Udi, the negator *nu-* seems to be a Kartvelian or Iranian loan.

[ii] The unique challenge of the Archi paradigm. In Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society: Special Session on Languages of the Caucasus (2013)

[iii] Kryz *b-a-b-q-re-b* HPL-PV-HPL-keep-PRS-HPL 'They are kept.'

[iv] The Diachronic Externalization of inflection. Linguistics 31 (1993), 279-309.

Márton András Baló

The aftermath of the borrowing of a borrowing pattern in Romani

This paper concerns the current state of the dichotomy between the inflection of the inherited and borrowed nouns in Romani. In a broader context, I would also like to use the example to illustrate the idea that language change is not a one-way street in the strict sense, and different paths of morphological and lexical borrowing may interact in a way that inherited morphological patterns are not lost; on the contrary, they may gain more significance. The observations are mostly based on data from Northern Vlax, a major dialect group of Romani, collected by the author during their recent fieldwork in Hungary.

There are two distinct inflectional patterns in the nominal morphology of Romani, which split the lexicon into two easily distinguishable components, morphologically separating inherited nouns from borrowed nouns (cf. e.g. Boretzky 1989, Bakker 1997 or Matras 2002).[1] One of the patterns is applied to inherited lexical elements, as well as early loans from Persian, Armenian and Greek, while the other pattern is used with lexical items borrowed from Greek and other (Romanian, Serbian, Hungarian etc.) contact languages at a later stage in the history of the language. The basic difference between the two patterns is in the oblique form of masculine nouns: the oblique singular form of the inherited pattern is *-es-* (nominative singular *čhāvo* 'boy' > oblique singular *čhāves-*), while the oblique form of the pattern applied to borrowed words is most commonly *-os-* (see for example Romungro, a Southern Central dialect: nominative singular *vodro* 'bed' > oblique singular *vodros-*, from Serbian *одар*), but other forms, such as *-is-* (*kočiši* 'coachman' > *kočišis-*, from Hungarian *kocsis*, as noted by Vekerdi 1981) and *-us-* also occur (cf. Elšík 2000). The oblique form serves as the stem for additional case markers.

Bakker 1997 also argues that the borrowing pattern used in Romani is borrowed from Greek: both languages add an unstressed vowel or syllable to the stem of consonant-final nouns, and the nominative endings used for this purpose in Romani (*-os*, *-is* and *-us*) are borrowed from Greek. For many Romani dialects, the final *-s* is often lost then in the nominative but it is always retained in the oblique. This borrowing pattern and several Greek lexical elements are borrowed simultaneously by Romani during the lengthy sojourn of the Romani people in Byzantium in the first half of the second millennium. It comes as no surprise then that Greek loans are inflected according to the inherited pattern, and the borrowed pattern only begins to affect subsequent loans, apparently without exception: all nouns borrowed after the Greek period have been integrated into the new class based on the borrowed pattern, rather than the class that has thus come to be restricted to inherited nouns only.

This is reflected in descriptions of various Romani dialects (see Lovari for example: Hutterer & Mészáros 1967, Cech & Heinschink 1999), with minor differences, and even borrowed masculine nouns with a stem-final *-i* are reported to take the oblique in *-os-* (Cech & Heinschink 1999: 22). This is, however, contradicted by the newly collected data: borrowed nouns with a stem-final *-i* invariably take the oblique in *-es-* (nom. sing. *kopáči* 'tree trunk' > obl. sing. *kopáčes-*, from Romanian *copac*).

Elšík 2000 already reports that there is some interplay between the two classes: in Roman, another Southern Central dialect of Romani, the oblique singular form of *grofo* 'count', a parade example of a loanword from Hungarian, can be either *grofes-* or *grofos-* (Elšík 2000: 23). Our fresh data not only corroborate this, but also allow us to draw further conclusions. The attested oblique forms of several nouns borrowed from Serbian, Hungarian and Romanian vary between the inherited and the borrowed patterns: nom. sing. *čokano* 'hammer' > obl. sing. *čokanes-/čokanos-* (from Romanian *ciocan*); nom. sing. *duhano* 'tobacco' > obl. sing. *duhanes-/duhanos-* (from Serbian *дуван*); nom. sing. *mobilo* 'mobile phone' > obl. sing. m*obiles-/mobilos-* (from Hungarian *mobil*). Data from Sinti, a Northern dialect of Romani, provide even stronger evidence, aș there are no lexical items, borrowed or inherited, that would take the *-os-* form: nom. sing. *švigefatro* 'father-in-law' > obl. sing. *švigefatres-* (from German *Schwiegervater*); nom. sing. *frento* 'stranger' > obl. sing. *frentes-*

1 In the English language literature focussing on Romani, the terms "thematic" or "oikoclitic" and "athematic" or "xenoclitic" are used to refer to the inherited and borrowed components, respectively.

(from German fremd); nom. sing. *hamro* 'hammer' > obl. sing. *hamres-* (from German *Hammer*). The evidence strongly points to some form of levelling between the two patterns, with the inherited pattern "reconquering" parts of the lexicon from the borrowed pattern, and the weakening of the borrowed pattern. As the data were collected in Hungary, one might speculate that psycholinguistic factors may interfere in the form of the extent to which a native speaker feels that a certain word is borrowed or not. This is, however, very difficult to measure, and the fact that there is hesitation concerning newly coined words such as *mobilo* does not suggest that older loans from Romanian, Serbian or German, whose origins have already become obscured, would feel more Romani than newer loans; the possibility cannot be excluded, though. While variation in the case of Northern Vlax Romani may also be related to frequency (words with higher token frequency tend to vary less, while words with lower token frequency are more inclined to vary), the complete lack of the borrowed pattern in the Sinti data seems to confirm the erosion of the strict inherited-borrowed dichotomy of Romani.

References

Bakker, Peter. 1997. Athematic morphology in Romani: The borrowing of a borrowing pattern. In Matras, Yaron, Peter Bakker & Hristo Kyuchukov (eds.), *The Typology and Dialectology of Romani*, 1–21. Amsterdam: John Benjamins.

Boretzky, Norbert. 1989. Zum Interferenzverhalten des Romani (Verbreitete und ungewöhnliche Phänomene). *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42. 357–374.

Cech, Petra & Moses F. Heinschink (1999). *Basisgrammatik. Arbeitsbericht 1a des Projekts Kodifizierung der Romanes-Variante der österreichischen Lovara (hrsgg. von Dieter W. Halwachs)*. Vienna: Verein Romano Centro.

Elšík, Viktor. 2000. Romani nominal paradigms: their structure, diversity, and development. In Elšík, Viktor & Yaron Matras (eds.), *Grammatical relations in Romani: The noun phrase*, 9–30. Amsterdam: John Benjamins.

Hutterer, Miklós & György Mészáros. 1967. *A lovári cigány dialektus leíró nyelvtana.* Budapest: Magyar Nyelvtudományi Társaság.

Matras, Yaron. 2002. *Romani: A linguistic introduction.* Cambridge: Cambridge University Press.

Vekerdi, József. 1981. *A magyar cigány nyelvjárás nyelvtana. (Tanulmányok a cigány gyerekek oktatásával-nevelésével foglalkozó munkacsoport vizsgálatából IV.)* Pécs: Pécsi Tanárképző Főiskola.

Derivation predicting inflection
The role of families, series, and morphotactics

Recent research has focussed on the internal predictability of inflection systems (e.g. Ackerman et al., 2009; Stump & Finkel, 2013), i.e., to what extent does partial knowledge of a lexeme's paradigm allow one to predict its inflectional behavior. Guzman (2019) highlights that inflectional behavior is also predictable from inherent properties of the lexeme, such as stem phonology and lexical semantics. In this presentation we will elaborate on this idea and explore how word formation helps predict inflection.

The literature documents situations where the process deriving it determines a lexeme's inflection (Bonami & Boyé, 2006) or where a lexeme inherits the inflection of its base (Stump 2001). Our goal here is to explore the interplay between these two possibilities through quantitative analysis of large lexica. We focus on data from LatInfLexi (Pellegrini & Passarotti, 2018), an inflected lexicon of 3348 Latin verbs, acquired semi-automatically from Lemlat 3.0, a Latin morphological analyzer (Passarotti et al., 2017). The present method extends straightforwardly to other data.

We first quantify the contribution of a lexeme's derivational history to the predictability of its inflection. To this effect, we compare the conditional entropy of paradigm cell *A* knowing paradigm cell *B* to the conditional entropy of paradigm cell *A* knowing paradigm cell *B*, as well as the lexeme's ultimate derivational base (if any) and the nature of the derivational process. All computations were conducted using Qumin (Beniamine, 2018). Table 1 shows that derivational information reduces unpredictability drastically. Hence speakers can predict inflection much more accurately if they know a lexeme's place in the derivation network.

$H(A\|B$, ancestor, process)	0.08
$H(A\|B)$	0.28

Table 1: Conditional entropy of cell A given cell B, averaged over all pairs of cells.

We then explore the respective role of derivational families and derivational processes in predicting inflectional behavior. Table 2 indicates the conditional entropy of inflection class given a variety of derivational predictors. Knowledge of the family is an overall excellent predictor. Verbs that belong to the same family usually display the same inflection, even if highly irregular: witness the suppletive alternation between the present and perfect stems of FERO 'bring' (*fer-* vs. *tul-*) and CON-FERO 'bring together' (*confer-* vs. *contul-*). There are exceptions though, e.g. CON-DO 'put together' is 3rd conjugation, whereas DO 'give' is 1st conjugation.

Conversely, prefixes and suffixes are rather poor predictors, both independently and in combination; but joint knowledge of the three pieces of information leads to almost perfect predictability.

H(class)	1.950	H(class|family,prefix)	0.074
H(class|family)	0.160	H(class|family,suffix)	0.026
H(class|prefix)	1.890	H(class|prefix,suffix)	1.775
H(class|suffix)	1.837	H(class|family,prefix,suffix)	0.002

Table 2: Conditional entropy of inflection class given various combinations of predictors. Dataset of 2747 verbs, excluding families and affixes with a unique type.

These results hide an asymmetry that becomes apparent when focusing on either prefixed or suffixed verbs. Among these, knowledge of the prefix provides little to no information, whereas knowledge of the suffix leads to perfect predictability, as shown in Table 3. Indeed, verbs that are formed by means of the same suffix always belong to the same conjugation in Latin (e.g., all verbs with the inchoative suffix *-sco* are 3rd conjugation), while verbs that are formed by means of the same prefix can belong to virtually any conjugation (compare 1st conjugation AD-NATO 'swim toward' and 2nd conjugation AD-MOVEO 'move toward'). This contrast is not apparent in Table 2, because suffixed verbs make up only 8% of the dataset.

Among prefixed verbs		*Among suffixed verbs*	
H(class)	1.95	H(class)	0.91
H(class|prefix)	1.88	H(class|suffix)	0

Table 3: Conditional entropy of inflection class given affix identity

This suggests that inflection class membership is mostly determined by the last morph in the stem, i.e., the morph adjacent to inflectional affixes. In the talk we will discuss the theoretical significance of this result, in connection with debates on (a-)morphousness and the split morphology hypothesis.

References

Ackerman, Blevins, and Malouf. 2009. Parts and wholes: implicative patterns in inflectional paradigms. Blevins & Blevins, *Analogy in Grammar*, OUP.

Beniamine. 2018. *Typologie quantitative des systèmes de classes flexionnelles*. Ph.D. dissertation, Paris Diderot.

Bonami and Boyé. 2006. Deriving inflectional irregularity. In *Proceedings of the HPSG 2006*.

Guzman Naranjo. 2019. *Analogical classification in formal grammar*. Language Science Press.

Passarotti, Budassi, Litta, and Ruffolo. 2017. The Lemlat 3.0 Package for Morphological Analysis of Latin. In *Proceedings of NoDaLiDa 2017*.

Pellegrini and Passarotti. 2018. LatInfLexi: an Inflected Lexicon of Latin Verbs. In *Proceedings of CLiC-it 2018*.

Stump and Finkel. 2013. *Morphological Typology: From Word to Paradigm*. CUP.

Jeremy Bradley, Nikolett F. Gulyás and András Czentnár

Causative and reflexive morphology in the Volga-Kama Region: A contact linguistic investigation

The genealogically diverse Uralic and Turkic languages of the Volga and Ural Regions of European Russia are on the basis of their striking structural convergence frequently subsumed as members of the *Volga-Kama Sprachbund* (e.g. Helimski 2003: 15). Contact patterns between these languages have historically been primarily studied on the level of the lexicon (e.g. Wichmann 1903, Räsänen 1923), morphophonology (e.g. Saarinen 1997), and code copies (e.g. Hesselbäck 2005). Morphosyntax has been comparatively understudied from a contact linguistic perspective.

The contribution at hand represents an attempt to study contact linguistically motivated convergence and divergence within the Volga-Kama Region as regards the usage of valence-changing derivational suffixes. Excuding Russian, all languages of the region (Turkic: Tatar, Bashkir, Chuvash; Uralic: Mari, Udmurt, and peripherally Mordvin, Komi) make extensive use of both valence-increasing (i.e. transitivizing, causative, etc.) and valence-decreasing (i.e. reflexive, passive, etc.) derivational suffixes. In all languages there is a strong preference for valence-increasing operations, but Russian with its strong preference for valence-decreasing operations (cf. Nichols et al. 2004) might be strengthening the position of valence-decreasing suffixes in the region.

While isolated borrowings of valence-changing suffixes can be observed, most valence-changing suffixes used in the languages of the region are not borrowed; their cognate forms are widely used throughout their respective language families. In spite of the seemingly coincidental superficial typological overlap between the languages of the region in this domain, the minutiae of valence-changing operations show striking convergence and divergence between languages and language varieties of the region. Questions investigated (consulting reference materials, corpora, and native speakers) on the basis of observed structural differences in the region include:

i) Can more than one causative suffix be attached to a verbal stem in a variety (1)?
ii) Does a variety allow for double accusative marking in a causative construction (2)?
iii) Does a variety allow for quasi-causatives (3)?
iv) Does a variety allow for impersonal (subjectless) passives that allow for a direct object marked with the accusative case (4)?
v) Does the variety allow for a canonical passive construction with an overt Agent (5)?
vi) Does a variety allow for a valence-decreasing suffix to be used to indicate the non-volitional nature of an activity (6)?

Our talk will examine these questions within the languages of the region, and in Uralic and Turkic languages from outside the region as a frame of reference, focussing on convergence across genealogical boundaries and divergence within genealogical groupings. As the microvariational study of these facets can yield additional data for the study of the diachrony of language contacts in the region, which to date remains heavily based upon the microvariational study of the spread of lexical items, special attention will be given to closely related Uralic language varieties that have been subject to different contact situations (Meadow Mari ↔ Hill Mari, Moksha Mordvin ↔ Erzya Mordvin, Komi-Zyrian ↔ Komi-Permyak). We will examine how differences between these varieties correlate to structures in their respective contact languages.

Examples (provided/checked by native speaker consultants)

(1) Tudo vaty-ž-lan vüd-ym yry-kt-ykt-en.
3SG wife-PX3SG-DAT water-ACC heat-CAUS-CAUS-3SG.PST2
'He made his wife heat water.' (Meadow Mari)

(2) Maša Saša-jez kńiga-jez lydʑy-t-iz.
Masha Sasha-ACC book-ACC read-CAUS-PST.3SG
'Masha made Sasha read the book.' (Udmurt)

(3) Mone ösky-t-e.
I.ACC vomit-CAUS-PRS.3SG
'I feel sick.' (Udmurt)

(4) Ala-kunam ožno suas jylmy-m=at tunem-alt-yn.
sometime earlier Tatar language-ACC=and study-REFL-PST2.3SG
'Sometime in the past, Tatar was studied as well.' (Meadow Mari)

(5) ʔTyn-eš pur-ty-maš [...] akušerka dene yšt-alt-yn.
faith-LAT enter-CAUS-NMLZ [...] midwife with do-REFL-PST2.3SG
'The baptism was done by a midwife.' (Meadow Mari)

(6) Menam onmöśśi-ś-öma.
I.GEN fall_asleep-REFL-PTCP.PST
'I fell asleep (unintentionally).' (Komi-Permyak)

References

Helimski, Eugene 2003. Areal groupings (Sprachbünde) within and across the borders of the Uralic language family: A survey, *Nyelvtudományi Közlemények* 100: 156–167.

Hesselbäck, André 2005. *Tatar and Chuvash Code-copies in Mari*. Studia Uralica Upsaliensia 35. Uppsala: Uppsala Universitet.

Nichols, Johanna & A. Peterson, David & Barnes, Jonathan 2004. Transitivizing and detransitivizing languages. *Linguistic Typology* 8. 149-211.

Räsänen, Martti 1923. *Die tatarischen Lehnwörter im Tscheremissischen*, Mémoires de la Société Finno-Ougrienne L. Helsinki: Suomalais-Ugrilainen Seura.

Saarinen, Sirkka 1997. Language contacts in the Volga region – Loan suffixes and calques in Mari and Udmurt. In Heinrich Ramisch & Kenneth Wynne (eds), *Language in time and space – Studies in honour of Wolfgang Viereck on the occasion of his 60th birthday.* Stuttgart: Franz Steiner Verlag. 388–396

Wichmann, Yrjö 1903. *Die tschuwassischen Lehnwörter in den permischen Sprachen*. Mémoires de la Société Finno-Ougrienne XXI. Helsinki: Suomalais-Ugrilainen Seura.

Morphemes in competition: Lexically-conditioned allomorphy, sociolinguistic variation, and change in progress

1. Introduction: Plurals in central Sicily. In central Sicilian dialects the plural feature value [PL] in nouns can be realized by three allomorphs:

(i) the morph *-i*, which is the most widespread plural ending across the island and realizes the plural for both masculine and feminine nouns (e.g. *carusu* 'boy' & *carusa* 'girl'→ *carusi* 'boys/girls');
(ii) the morph *-a*, etymologically connected to the plural of the Latin second declension neuter nouns (e.g. *jitu* 'finger' → *jita* 'fingers') (see Sornicola 2010);
(iii) the morph *-ura* (< Lat. *-ORA*), which is characteristic of the outcomes of the Latin neuter nouns from the third declension (e.g. Lat. *corpus* 'body' → *corpora* 'bodies'), together with those nouns from the fourth declension which form the plural by analogy with the same model (*jocu* 'game' → *jocura* 'games', Rohlfs 1968; see also Retaro 2013):

As described by Retaro (2013), the distribution of these three morphs in central Sicilian dialects is rather complex: several lexemes allow for two or even three options, yielding an emblematic case of overabundance (in the sense of Thornton 2011, 2013, 2019); only in very few cases is the allomorph selection semantically or lexically motivated. In addition, the high degree of inter- and intra-speaker variation makes it difficult to connect the alternations to precise extra-linguistic factors.

2. Aims of the paper. In this study, I investigate the allomorphy in the nominal plural formation in the central Sicilian dialect of Mussomeli, Caltanissetta, focusing on the *-ura* plurals. I show that the complex situation that characterizes this dialect is the result of an intricate interplay between lexical allomorphy, sociolinguistic variation, and on-going language change. Numerous factors seem to govern the realization of nominal plural, but none is phonological or morphological in nature. The three allomorphs realize the same morphosyntactic value, but are in fact in competition only with respect to a limited number of lexemes, indicating that this kind of allomorphy is to a remarkably large extent a matter of the lexicon. However, even with these lexemes, the plural allomorphs are neither in complementary distribution nor in free variation, calling for the need to resort to sociolinguistic variables such as the age of the speakers. The sociolinguistic variation, in turn, reflects a change in progress which is leading to a partial resolution of the allomorphic competition, resulting in the loss of the *-ura* allomorph.

3. Methodology. To investigate the distribution of the three plural allomorphs, I administered a questionnaire to 34 native speakers divided into three age groups:

GROUPS	AGE RANGE	NO. OF SPEAKERS
1	14–30	15
2	31–60	15
3	61–93	15

The questionnaire consisted of 45 lexemes for which the plural in *-ura* was attested (Retaro 2013). In order to obtain reliable results, production data were elicited first: speakers were asked to provide the plural form of the relevant lexemes (only one answer was possible). However, since speakers may be unaware of some variation in their speech, their passive competence was also tested by eliciting grammaticality judgments on the alternative plurals on a scale from 0 to 3.

4. Results. The results are summarized in Figure 1:

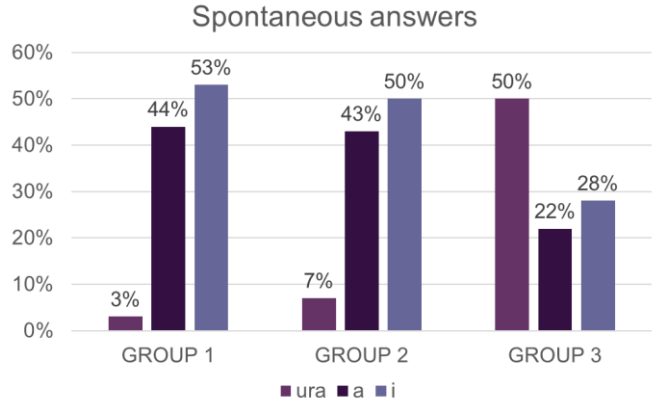


Figure 1: Average % scores for plural form across age groups

a) Only speakers from group 3 spontaneously produced plurals in *-ura* (50%), while the rate is much lower for group 1 (3%) and group 2 (7%). Speakers from the first two groups only show a limited and varying passive competence of *-ura* plurals in the grammaticality judgment task.
b) The *-i* plurals seem to prevail in terms of productivity, especially in groups 1 and 2. Even when they are not provided as first answer, they are accepted as good alternatives by all speakers in the grammaticality judgement task.
c) The *-a* plurals are still much alive in all groups, although the alternative *-i* plurals (with the same lexemes) were actually preferred by most speakers. Only with a few lexemes did *-a* plurals consistently obtain higher scores (e.g. *ligna* 'wood', *ossa* 'bones'; Sornicola 2010).
d) These results confirm Retaro's (2013) observations about inter- and intra-speaker variation, making it difficult to draw clear-cut generalizations. Indeed, in several cases speakers provided or accepted more than one allomorph for the same lexeme (e.g. *furnu* 'oven, bakery' → *furni*, *furna*, *furnira*).

5. Conclusions. Sociolinguistic factors must be taken into account in order to explain the lack of a complementary distribution of the competing allomorphs and the high degree of variation: the frequency of the third variant (i.e. the *-ura* allomorph) differs within the same community according to age. These differences can be used as diagnostics of a change in progress that is leading to the gradual disappearance of the *-ura* allomorph. The tendency towards the generalization of the *-i* allomorph is evident, presumably due to the (partial) proximity of this allomorph to Italian plurals. The *-a* plurals are however still frequent and vital, especially with specific lexemes (Sornicola 2010). The lack of a one-way reassignment of the plural further shows that we are not dealing with an independent (neuter) gender or a distinct inflectional class (see Loporcaro & Pedrazzoli 2016 for a case of systematic reassignment).

Selected References:

Retaro, V. 2013. La morfologia del plurale nelle varietà della Sicilia centrale. *Bollettino del Centro di studi filologici e linguistici siciliani* 24.

Sornicola, R. 2010. I dialetti meridionali e la sorte del neutro: alcune riflessioni su una varietà siciliana. In M. Iliescu et al. (eds), *Actes du XXVe Congrès International de Linguistique et Philologie Romane*. De Gruyter.

Thornton, A. 2019. Overabundance: A Canonical Typology. In F. Rainer et al. (eds), *Competition in Inflection and Word-Formation*. Springer.

Silvia Dal Negro

Borrowing routines in obsolescent languages: borrowing verbs in German minority dialects

This proposal focuses on loan verb integration and in particular on borrowing routines from Italian (the donor language) into an obsolescenting Alemannic dialect of North-Western Italy (the recipient language). By using the term "borrowing routines" (derived from Heath 1984) we want to emphasize the productivity of ongoing accommodation processes and the interplay between individual occurrences and community norms as can be inferred by analyzing natural speech data (Poplack 2017).

A conspicuous number of German varieties are spoken in Northern Italy where they have been in contact with Standard Italian and/ or Italo-Romance dialects for a time span ranging from one to eight-nine centuries. Besides this temporal dissimilarity, German varieties in Italy differ on many other levels: status, official recognition, geographical continuity, access to Standard German. Despite such differences all German varieties in Italy share an asymmetric contact condition with Italo-Romance – notably with Italian – which causes a wide range of unidirectional contact phenomena. Differences in sociolinguistic structure, however, might be responsible for qualitative differences in loanword accommodation strategies (besides quantitative differences in terms of borrowing and code-mixing in speech), a hypothesis that is pursued in this project.

In the dialect under investigation, belonging to the Walser German dialect group settled south of the Alps in the 13th century, verbs of Italo-Romance origins are integrated by means of an indirect insertion strategy (Wohlgemuth 2009), making use of the well-known Middle High German morpheme *-ier* (itself a borrowing from French). Noteworthy no exaptation process seems to have occurred here (as described in Gardani 2016) and this suffix only derives allogenous (Italo-Romance) verbs. On the other hand, some variation can be observed in loan verbs as regards inflection: differently from Standard German, *-ier* verbs are attracted into more productive verb classes, presenting past participle *k-* prefix (as in *kschtudért* vs. *studiert* 'studied') or even *–u* infinitive ending (as in *schtudéru* vs. *studieren*), on the basis of the regular and productive verb inflectional class in *–u* (< MHG *–ôn*). Finally, in analogy with past participles, *–ért* is extended to loan adjectives ending in *–ato* in Italian, such as *fortunérts* 'lucky-NT' derived from Italian *fortunato*.

Working on a corpus of contemporary Walser German, all occurrences of loan verbs have been extracted and accounted for absolute frequency, diffusion within the corpus and inflection. This makes it possible to observe that this borrowing routine is still productively available to bilingual speakers as they interact in the minority language and provide a ready-made scheme for *ad hoc* adoptions from the majority language (most instances found in the corpus are in fact *hapax*), thus allowing speakers to keep a fluent speech despite the fact that the inherited vocabulary is now strongly reduced for most of them.

As said, not all German varieties in Italy, and not even all Walser dialects, present the same loan verb accommodating strategy. In order to use the *–ier* suffix productively, a minority German dialect needs to have been in contact with German varieties in which this accommodation strategy was common and productive at that time and to have been able to transfer it into its own variety. As contact with Italo-Romance started to become more and more intimate, the accommodation strategy turned into a productive borrowing routine, but if this contact with other German varieties was not available in the first place other accommodation strategies are were to be developed instead. A comparison with other dialects, also based on contemporary conversational data, will focus on such different strategies and routines.

REFERENCES

Gardani F. (2016), "Allogenous exaptation", in Norde M./ Van de Velde F. (eds.), *Exaptation and Language Change*, Amsterdam, Benjamins: 228-261.

Heath J. (1984),. "Language Contact and Language Change", in *Annual Review of Anthropology*13: 367–384.

Poplack S. (2017), *Borrowing. Loanwords in the Speech Community and in the Grammar*, Oxford, OUP.

Wohlgemuth J. (2009), *A Typology of Verbal Borrowings*, Berlin/ New York, de Gruyter.

Semantically ambiguous stems and the purpose of morphological processing

Words such as *barking*. and *perching* possess a type of ambiguity that has the potential to reveal fundamental characteristics of morphological processing. Both words seem to be formed by the addition of the English suffix *-ing* to the stem forms of the verbs *bark* and *perch*, respectively. Yet, those verb forms can also be noun forms. The word *bark* can refer both to the sound that a dog makes and it can also refer to the (unrelated meaning of) the outside of a tree. The word *perch* can refer to what a bird does when it alights on a branch and it can also be the name of a freshwater fish.

The critical characteristic of these words is that their unsuffixed forms (e.g., *bark*, *perch*) are ambiguous, but their suffixed forms (e.g., *barking*, *perching*) are not ambiguous. Thus, they provide a means by which we can ask the question of whether affixation can suppress ambiguity.

Previous research with stimuli such as these has shown that it cannot. Libben and de Almeida (2002) reported priming results in which unaffixed ambiguous words prime semantic associates of both their noun and verb forms (e.g., *bark* primed both *dog* and *tree*). For suffixed forms, however, they found positive priming of the semantic associates of the verb forms, but negative priming for semantic associates of the noun forms (e.g., *barking* primed *dog* but inhibited *tree*).

These results suggest that the ambiguity of stems can 'shine through' even when one reading is ruled out by affixation. It is important to note, however, that the Libben and de Almeida (2002) study was restricted to single word processing in a primed lexical decision task. It remains unclear whether comparable effects would also be obtained during sentence processing, which better reflects natural language use. We report on an eye-tracking experiment that was designed to assess whether the 'ambiguous stem effect' is evident in the context of sentence reading.

Participants were 114 native speakers of English. Each participant read each of 64 ambiguous suffixed stems (e.g., barking, perching) in one of four sentence contexts

Table 1. The design of the stimulus sentences (bolding is used here for emphasis only and was not used in the presentation to participants).

Condition	**Example sentence**
Anomalous - verb associate completion	He heard loud **barking** during the **dog** on Saturday
Anomalous - noun associate completion	He heard loud **barking** during the **tree** on Saturday
Anomalous – non-associate completion	He heard loud **barking** during the **fail** on Saturday
Unanomalous completion	He heard loud **barking** during the **night** on Saturday

The dependent variables that were analyzed were reading duration for both the ambiguous suffixed stem (e.g., *barking*) and the completions (e.g., *dog, tree, fail, night)* as well as regressions into the regions of the ambiguous suffixed stem and the completions.

We found results consistent with Libben & de Almeida's (2002). As expected, all anomalous conditions (*dog/bark/fail*) yielded significantly longer gaze durations than the unanomalous condition (*night*). More importantly for our purposes were the differences between anomalous conditions: longer gaze durations for the unrelated condition (*barking- fail*) than for the target

consistent with the whole-word prime (*barking-dog*). Facilitatory effects for the unbiased meaning of the stem(tree) did not reach significance, although numerically faster than the unrelated condition. The lack of difference be- tween the two targets related to the stem (*dog, tree*) suggests that both meanings of the ambiguous stems were accessed in the processing of the sentences and, indeed, that suffixation cannot prevent activation of both stem meanings.

We interpret these results to support the conclusion that the purpose of morphological processing is to enable, rather than suppress opportunities for meaning interpretation. In the case of ambiguous suffixed stems, this leads to activation of both meanings of an ambiguous stem.

Reference

Libben, G. & de Almeida, R. G. (2002). Is there a morphological parser? In S. Bendjaballah, W. U. Dressler, O. E. Pfeiffer & M. D. Voeikova (eds.), Morphology 2000, pp. 213-225. Amsterdam: John Benjamins.

Bernard Fradin

On meaning predictability in derivational domains (oral)

1. The present study is rooted in a very specific derivational domain, namely the set of lexemes derived from nouns denoting animals in French. From a list of about 320 animal names extracted from the website https://www.bestioles.ca/animaux/ and dictionaries, 72 nouns were selected on the basis of their occurring in morphologically derived lexemes: *cheval* 'horse' / *chevalin* 'horsey; equine'. About 250 lexemes derived in this way have been found. These lexemes can be classified into two groups in function of the nature of the relationship they instantiate. In the first, a natural relationship is involved, such as: generic relation *ours* 'bear' / *ourson* 'bear cub', *poulain* 'foal' / *pouliner* 'to foal'; living place *termite* 'termite' / *termitière* 'termite mound'; typical feature *tigre* 'tiger' / *tigré* 'striped'. In the second, the relationship is external inasmuch as it involves an interaction with men or other animals. Breeding *lapin* 'rabbit' / *lapinier* 'rabbit breeder', taming *faucon* 'hawk' / *fauconnier* 'falconer', or hunting *héron* 'heron' / *héronner* 'to hunt heron' are examples of this interaction. To what extent is the meaning of these derived lexemes predictable? This is the main issue addressed in this study.

2. It must be straightaway noted that the same exponent can form derived lexemes with different meanings *porc* 'pig' / *porch-erie* 'pigsty' vs. *singe* 'monkey' / *sing-erie* 'monkeying around'. Moreover, most exponents occurring in the lexemes under discussion have other uses and may express additional contents *sorcier* 'sorcerer' / *sorcellerie* 'witchcraft' (affix polysemy problem). Nevertheless, it is possible to predict the meaning of many of the lexemes in question, if we base this prediction on two (or more) sources of information.

(1) If N denotes an animal, then N-eau denotes its child: *lion* / *lionceau* 'lion cub'.

(2) a. If N denotes a wild animal, N-ière denotes the place that it lives in: *renard* / *renardière* 'fox's earth'.

b. If N denotes a domestic animal, N-erie denotes the place where it is raised: *vison* / *visonnerie* 'mink farm'.

Three pieces of information are involved in (3).

(3) If N denotes an animal with which a typical behavior is stereotypically associated, N-erie denotes this behavior *singe* 'monkey' / *singerie* 'instance of monkeying'

In this account, the suffix is not correlated with any particular meaning since the meaning of the suffixed lexeme is selected on the basis of (at least) two properties that jointly come from the base and the affix. This makes the affix polysemy problem dissolve insofar as the affix qua morpheme disappears. The affix only exists as a morph and entailments (1)-(3) are a way to state how morphosemantic contents are correlated with morphs (Crysmann & Bonami 2015: 4). The fact that meaning predictability is improved when several sources of information are taken into account reminds us of what Bonami and Beniamine (2016) have shown for the cell filling problem in inflection. The proposal made here is not completely new: something similar has been sketched by Schulte (2014: 316) and more factual versions appear in dictionaries (cf. s.v. *–erie* in *TLFi*). It also opens new leads to describe how derivational phenomena are embodied in languages.

3. A derivational domain is a set of morphological derivational series related by some formal and semantic links, here base = N and base content = 'animal'. A morphological derivational series is constituted of lexemes instantiating the same derivational pattern *lion* 'lion' / *lionne* 'lioness', *hérisson* 'hedgehog' / *hérissonne* 'female hedgehog', etc. A derivational pattern (abbreviated as CATEGORY÷MEANING÷EXPONENT) is marginal if it yields (very) few lexeme types for the derivational domain in question and central otherwise. Here some central patterns

for the investigated derivational domain: (4) N÷CHILD÷EAU *chevreau* 'kid', N÷CHILD÷ON *ourson*; N÷FEMALE÷ESSE *ânesse* 'she-donkey', N÷FEMALE÷Ṽ/V *hérisonne*; N÷LOCATION÷IÈRE *termitière*, N÷LOCATION÷ERIE *visonnerie*; A÷SIMIL÷É *tigré*; V÷GIVE_BIRTH÷CONV-V *agneler* 'to lamb'. Marginal ones: (5) N÷TAX÷AGE *moutonnage* 'tax on sheep'; N÷CONTENT÷ÉE *ânée* 'burden that a donkey can carry'. In a complementary way, expressing the sound/meaning correlations at the appropriate level, as in (1)-(3), helps us to put to light what the dominant vs. non-dominant meanings are for a given morphological marker, conceived of as a pair <output.category, exponent>. For each derivational marker, it is possible to list the various semantic contents it is correlated with. A semantic content instantiated by many lexemic types is said to be dominant, while one instantiated by few ones will be non-dominant. The semantic content involves three types of information: bse.category_bse.semantic.type_derived.lxm.meaning. These descriptive tools allow us to show that derivational marker <N, ière> displays many semantic contents, none of which is dominant: (6) N_object/subst_container: *cartouchière* 'cartridge belt', N_animal_dwelling: *termitière*, N_body.part_protection: *jambière* 'greave'. On the contrary, <N, eur>, <N, ée> have a dominant interpretation. It can be shown that derivational markers fall into three distributional types for what regards the (non-)dominant criterion.

Bonami, Olivier & Sacha Beniamine. 2016. Joint predictiveness in inflectional paradigm. Word Structure 9.156-82.

Crysmann, Berthold & Olivier Bonami. 2015. Variable morphotactics in Information-based Morphology. Journal of Linguistics 51.1-64.

Schulte, Marion. 2014. Accounting for affix polysemy with semantic maps—a diachronic study of **-age** suffixation in English. In *Proceeding of the Décembrettes 8th. International conference on morphology* (eds) S. Augendre, G. Couasnon-Torlois, D. Lebon, C. Michard, G. Boyé & F. Montermini. Toulouse: ERSS, CNRS & Université Toulouse-Jean Jaurès.

Livio Gaeta

Deconstructing complexity: Morphological change and language contact in Walser German

Recently, a big debate has started focusing on the concept 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 therefore to systematically 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 – insofar as every speaker in these communities is at least bilingual – 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 – which however is also common to other Walser varieties not found in island condition (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. *ʃri:ba* 'I/we write', *ʃribʃt* / *ʃri: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 *ʃri: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*: *lɒ:* / *la:v* 'I / we let', etc. In Greschòneytitsch 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'gròssò lougò ésch 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'husgspònnto wollschtrangna 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". In fact, since the morphological changes at stake increase systematicity and add redundancy, they also facilitate the speakers' learning and processing. In sum, we cannot but conclude again with Nichols (2009: 112) that "the whole question of whether they increase or reduce complexity needs more chewing by more linguists", while the issue of isolation needs to be taken seriously into account as a main factor influencing complexity (cf. Baechler 2016 on the German Walser community of Issime).

References

Baechler, R. et al. (eds.) (2016), *Complexity, Isolation, and Variation*. Berlin.

Baechler, R. (2016), Inflectional complexity of nouns, adjectives and articles in closely related (non-)isolated varieties. In Baechler, R. et al. (eds.), 15-39.

Baerman, M. et al. (eds.) (2015), *Understanding and Measuring Morphological Complexity*. Oxford.

Dal Negro, S. (2004), *The Decay of a Language: The Case of a German Dialect in the Italian Alps*. Bern.

Meakins, F. (2019), Birth of a contact language did not favor simplification. *Language* 95.2: 294-332.

Karlsson, F. et al. (2008), Introduction. In Miestamo, M. et al. (eds.), vii-xiv.

Miestamo, M. et al. (eds.) 2008. *Language complexity: typology, contact, change.* Amsterdam.

Nichols, J. (2009), Linguistic complexity: A comprehensive definition and survey. In Sampson, G. et al. (eds.), 110-25.

Sampson, G. et al. (eds.) (2009), *Language complexity as an evolving variable*. Oxford.

Jordan Gallant

Morphological Integration in the bilingual lexicon:

Evidence from Chinese-English bilinguals

The bilingual lexicon is characterized by extensive cross-linguistic structural interconnectivity evidenced by cross-linguistic activation of lexical entries (Libben, Goral & Libben, 2017). Although connectivity on the basis of morphology has been posited in the morphological integration hypothesis (Libben, Goral, & Baayen, 2017) little empirical evidence has emerged in support of it. From a developmental perspective, bilinguals show greater success acquiring morphologically complex words with corresponding compound structure across their two languages (Cheng, Wang & Perfetti, 2011). Similarly, cross-linguistic constituent priming has also been reported across languages with widely differing morphological systems (Libben, Goral & Baayen, 2017). However, behavioral studies have yet to detect effects of cross-linguistic morphological correspondence at the whole-word level.

This study combines morphological structure overlap with another factor known to amplify cross-linguistic activation: L1-specific phonological variation (LaGrou, Hartsuiker & Duyck, 2012). I report on the investigation of both these factors in the processing of English by Chinese-English bilinguals living in Canada. I investigated whether cross-linguistic lexical co-activation in the recognition and production of English words is influenced by (a) the pressence of L1-specific phonological features in stimulus presentation, and (b) by similarities in morphological properties of the stimulus. Factor (a) was operationalized by presenting stimuli orally in either Canadian English or in Chinese-accented Canadian English. Factor (b) was operationalized by selecting 28 English compounds with the same constituents and composition as their Chinese equivalents (e.g. moonlight, 月光, 月= moon + 光 = light). Participants performed an auditory stimulus typing task (Libben, Curtiss & Weber, 2014), whereby stimuli were presented auditorily and participants typed the words on a keyboard. Word typing onset times and inter-keystroke intervals were the dependent variables.

Preliminary results (Chinese-English bilinguals N=10, native English speakers N=45) indicate significant differences in typing latencies at the morpheme boundary for compounds with corresponding compound structure when the compound word was presented in a Chinese accent. In this condition, morpheme boundary latencies were significantly greater ($p=0.0371$) than when spoken by a native English speaker. Spurious activation of the second compound constituent (e.g. 光 in MOONLIGHT) slowed down the typed production of the compound at the morpheme boundary. These findings support the morphological integration hypothesis, suggesting that the Chinese-English bilingual lexicon shows connectivity at the level of morphological compound structure and that phonological variation consistent with the L1 of the listener magnifies cross-linguistic co-activation.

References

Cheng, C., Wang, M., & Perfetti, C. A. (2011). Acquisition of compound words in Chinese English bilingual children: Decomposition and cross-language activation. *Applied Psycholinguistics, 32,* 583-600.

Lagrou, E., Hartsuiker, R. J., Duyck. W. (2012). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingual Language Cognition*, 16, 508–517.

Libben, G., Curtiss, K., & Weber, S. (2014). Psychocentricity and participant profiles: Implications for lexical processing among multilinguals. *Frontiers of Psychology*, 4, DOI: 10.3389/fpsyg.2014.00557

Libben, G., Goral, M., & Baayen, R. H. (2017). Morphological integration and the bilingual lexicon. In Libben, M., Goral, M., & Libben, G. (2017). *Bilingualism: A Framework for Understanding the Mental Lexicon* (p. 311-340). Amsterdam ; Philadelphia : John Benjamins Publishing Company.

Libben, M., Goral, M., & Libben, G. (2017). The dynamic lexicon: complex words in bilingual minds. In Libben, M., Goral, M. & Libben, G. (Eds.) *Bilingualism: A Framework for Understanding the Mental Lexicon*. John Benjamins (pp.1-8).

Francesco Gardani, Michele Loporcaro and Alberto Giudici

A borrowed prosodic marker for inherent inflection in Istro-Romanian

(category: ORAL)

"[C]omparative constructions are generally highly diffusible" (Haig 2001: 206) and various cases of borrowing from different language families have been reported for languages as diverse as, e.g., (R(ecipient)L(anguage)/S(ource)L(anguage) are given in this order) Baram/Nepali (Dhakal 2014: 177), Lithuanian dialects/Bielorussian (Wiemer 2009: 353–358), Pipil/Spanish (Campbell 1987: 254f.). However, the case to be discussed here seems to be cross-linguistically almost unique in the borrowing research landscape, as it involves the borrowing of a prosodic pattern to realize morphosemantic values pertaining to the (inherent) inflection of the adjective: comparative vs superlative.

Our case study concerns comparative/superlative formation in Istro-Romanian (henceforth IR), a highly endangered Eastern Romance language, spoken in northeastern Istria (Croatia), under strong pressure from Croatian, with which it has stayed in century-long and intimate contact. Drawing from the existing literature (see Kovačec 1971: 108) and first-hand data (from fieldwork in Istria in 2017 as well as with diaspora speakers in New York City in 2018), we will show that nowadays IR displays an innovative pattern of comparative/superlative formation arisen from diachronic processes involving borrowing and reanalysis. (Daco-)Romanian, the closest variety of the RL, forms the comparative by adding the adverb *mai* 'more' before the adjective (e.g., *mare* 'big' -> *mai mare* 'bigger'), while superlative formation involves an articoloid (*cel mai mare* 'the biggest'). In Croatian, the SL, when the derived adjective stem used for comparative (e.g., *jȁsnij* 'clearer') is prefixed to form the superlative, tone/stress shift (Jachnow 2001: 494) occurs (e.g., *NÂI-jasnij* 'clearest'). IR has created an innovative pattern along the following path: a) it has kept the inherited periphrastic comparative with *mai*; b) a superficial similarity ('lookalike') between Romanian *mai* and Croatian *nâi* led the speakers to align the Romance construction with the Slavic one; c) construction alignment made the speaker reanalyze stress retraction on Croatian *nâi* as the only contrastive element; d) the contrast between comparative and superlative is now realized by purely prosodical means involving adjectives in the positive degree (unlike in Croatian): IR *mai MÅRE* 'bigger' vs. *MÁI måre* 'the biggest'.

To the best of our knowledge, this kind of prosodic pat-borrowing to convey the comparative/superlative contrast hardly finds any matches cross-linguistically. While there are reported cases of contact-induced change in comparative/superlative in the context of Romance-Slavic contact, most (such as those discussed in Breu 1996) involve mat-borrowing

or non-prosodic pat-borrowing. In the literature, the closest match to the IR facts is the prosodic calque on Bulgarian which occurs in comparative formation in Bulgarian Djudezmo (Romance; see Andreeva et al. 2017). Here, the Spanish comparative (*más* + adjective) is reshaped to adopt the stress pattern of the contact language: Djudezmo *MÁS fuerte* 'stronger' (vs. Spanish *más FUERte*), corresponding to Bulgarian *по-силен* (*PO-silen*) 'stronger'. While, however, in Bulgarian Djudezmo the similarity is only superficial/phonetic, in IR the prosodic device has acquired a grammatical function (signalling the comparative vs superlative contrast), which makes the case we discuss a virtually unique instance of contact-induced morphological change.

References

Andreeva, Bistra, Snezhina Dimitrova, Christoph Gabriel, Anna Gazdik and Jonas Grünke. 2017. Intonation and convergence: evidence from Bulgarian Judeo-Spanish. In Ekaterina Tarpomanova and Krasimira Aleksova (eds.), *Надмощие и приспособяване. Международна научна конференция на Факултета по славянски филологии. Софийски университет "Св. Климент Охридски"* [DOMINATION AND ADAPTATION. International Conference of the Faculty of Slavic Studies, Volume II. LINGUISTICS, Sofia University "St. Kliment Ohridski"], 169–177. Sofia: University "St. Kliment Ohridski".

Breu, Walter. 1996. Überlegungen zu einer Klassifizierung des grammatischen Wandels im Sprachkontakt (am Beispiel slavischer Kontaktfälle), *STUF* 49(1). 21–38

Campbell, Lyle. 1987. Syntactic change in Pipil. *International Journal of American Linguistics* 53(3). 253–280.

Dhakal, Dubi Nanda. 2014. Contact-induced changes in Baram. In Gwendolyn Hyslop, Linda Konnerth, Stephen Morey & Priyankoo Sarmah (eds.), *North East Indian linguistics 6,* 167–190. Canberra: Asia-Pacific Linguistics.

Haig, Geoffrey. 2001. Linguistic diffusion in present-day East Anatolia: From top to bottom. In Alexandra Y. Aikhenvald & Dixon, Robert M. W (eds.), *Areal diffusion and genetic inheritance*: *Problems in comparative linguistics*, 195–224. Oxford: Oxford University Press.

Jachnow, Helmut. 2001. Komparationskategorie und Graduierung im Kroatischen/Serbischen. In Helmut Jachnow, Boris Norman & Adam Suprun (eds.), *Quantität und Graduierung als kognitiv-semantische Kategorien*, 483–511. Wiesbaden: Harrassowitz.

Kovačec, August. 1971. *Descrierea istroromânei actuale*. Bucureşti: Editura Academiei Republicii Socialiste România.

Wiemer, Björn. 2009. Zu entlehnten Präfixen und anderen Morphosyntaktischen Slavismen in litauischen Insel- und Grenzmundarten. In Lenka Scholze and Björn Wiemer (eds.), *Von Zuständen, Dynamik und Veränderung bei Pygmäen und Giganten. Festschrift für Walter Breu zu seinem 60. Geburtstag*, 347–390. Bochum: Brockmeyer.

Alberto Giudici and Chiara Zanini

Plural indefinite quantifier on the Romance-Slavic border

1. INTRODUCTION

The development of the indefinite marker for singular nouns from the numeral 'one' is well attested in a worldwide range of languages from Germanic to Amerindian (Givón 1981). It is present in Romance too where some branches have maintained also the plural forms from Latin UNOS/-AS: Catalan uns/unes, Portuguese uns/umas, Spanish unos/unas and Old French (uns/unes) (Ledgeway 2011: 409).
In Italo-Romance, uni/une are pronouns and they never occur in attributive position (Loporcaro 2018: 75). Intriguingly enough, this is instead attested in Istriot, a Romance dialect spoken in Istria, nowadays Croatia. For centuries Istriot has been in contact with Slavic varieties such as Croatian and Čakavian dialects (Tekavčić 1976), that have a quantifier originated from the plural form of the numeral 'one' (Croatian jedni/jedne/jedna; Leko 2009: 25) meaning either 'a pair of' or 'some' and usually occurring with pluralia tantum nouns (Kalsbeek 1998: 175). In this light, Istriot uni ociai 'a pair of spectacles' is probably calqued on Croatian jedne naočale.

2. GOALS AND METHODS

At least to our knowledge, Istriot uni/une have never been investigated in the literature. For the first time, we conducted a survey to observe how and to what extent contact with Croatian affects the use of uni/une in the Istriot dialect.
Fifteen informants (F=8; age: mean=63.7, range=24-88; education: mean=10.7, range=5-21) participated in the study. We selected 24 concrete plural nouns, subdivided in 3 groups: (i) pluralia tantum (PT; ociai 'spectacles'); (ii) plural dominant (PD; calsini 'socks'); (iii) count plural (CP; capoti 'coats'). Two pictures were created for each noun: in one picture the object representing the target noun was depicted once in the case of PT and twice in the case of PD and CP; in the second picture, the same object was replicated four times. In addition, each noun was included in three different phrases: definite plural article +N (le braghe longhe 'the long trousers'); 'a pair of' +N (un per de braghe longhe 'a pair of long trousers'); uni/une +N (une braghe longhe).
The first task aimed to elicit uni/une in the semi-spontaneous speech. Informants were invited to look at each picture and to describe it. In a second task, participants were asked to rate the phrases on a five-point Likert scale after looking at the picture. Both tasks were administered orally. Data were analysed by means of generalised linear mixed models (Baayen, Davidson & Bates 2008).

3. RESULTS AND CONCLUSION

Participants spontaneously produced uni/une with PT (67/240) and PD (56/240), preferably in response to pictures illustrating one pair of objects (87/360) w.r.t. those illustrating more pairs (36/360), but not with CP (0/240). Phrases uni/une +N were rated lower overall (mean=3.701) than the other two types of phrase (means=4.93 and 4.816; all Ps< .05). Yet, they were accepted more in response to pictures illustrating one pair of objects (mean=4.569) than those illustrating more pairs (mean=4.396; p<.001), and with PT and PD (means=4.586 and 4.525) than CP (mean=4.337; all Ps<.05).
Informants still prefer to use the definite plural article or the equivalent of 'a pair of', but the use of uni/une is nevertheless raising, especially where it seems to be particularly relevant on morpho-semantic bases. The overall idea is that the use of uni/une will probably spread from the PT and PD to the CP in the future, since the influence of Croatian on Istriot is growing stronger.

References

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of memory and language, 59(4), 390-412.

Givón, T. (1981). On the development of the numeral 'one' as an indefinite marker. Folia Linguistica Historica, 15 (Historica vol. 2, 1), 35-54.

Kalsbeek, J. (1998). The Čakavian Dialect of Orbanići near Žminj in Istria. Studies in Slavic and general linguistics 25. Amsterdam - Atlanta: Rodopi.

Ledgeway, A. (2011). Syntactic and morphosyntactic typology and change, in M. Maiden, J. C. Smith, A. Ledgeway (edd.), The Cambridge history of the Romance languages, Vol. 1, Cambridge: Cambridge University Press, 382-471.

Leko, N. (2009). The Syntax of Numerals in Bosnian (Lincom Studies in Slavic Linguistics 33), Munich: Lincom Europa.

Loporcaro, M. (2018), Gender from Latin to Romance. History, Geography, Typology, Oxford: Oxford University Press.

Tekavčić, P. (1976), Per un Atlante Linguistico Istriano (con speciale riguardo ai dialetti istroromanzi). Studia Romanica et Anglica Zagrabiensia 41-42, 227-240.

Matías Guzmán Naranjo

Inflection class and semantic analogies

Background Inflection classes are usually considered semantically empty, meaning that the semantics of a cell in a paradigm is independent of inflection class (Carstairs-McCarthy, 1998). For example, it is assumed that *работу* ('work' ACC.SG) is semantically equivalent to *место* ('place' ACC.SG), modulo lexical meaning. Using semantic analogies I show that this is not completely true.

Methodology Mikolov et al. (2013) showed that it is possible to build semantic analogies of the form A:B::C:D (the meaning of A is to the meaning of B, as the meaning of C is to the meaning of D) using semantic vectors. For example, if we know the analogy *man:king* and the meaning of *woman*, we can predict the meaning of *queen*. We can extend this to inflection and calculate semantic analogies like ACC.SG:GEN.PL and test whether these generalize across inflection classes.

I use the semantic vectors provided by Kutuzov and Kuzmenko (2017) for Russian, trained with Word2Vec on a CoNLL17 corpus. These vectors provide 100 semantic dimensions for each form. The dataset contains semantic information for 67691 inflected nouns. Using the Zaliznyak (1967) Russian dictionary I extracted the inflection class of all forms, giving a total of 430 inflection classes.

The idea is to learn a semantic analogy within one inflection class, and compare how well it works on items of that inflection class vs. items in other inflection classes. I built the analogies using a mixed effect model: `cell-1` $\sim$ `cell-2 + (1 + cell-2 | dimension)`. This predicts the semantics of `cell-1` from the semantics of `cell-2`. Because semantic vectors do not distinguish syncretisms, I tested the models on groups of inflection classes which shared the same syncretic forms. For example, the analogies for *работа* ('work') and *место* ('place'), belong to two different inflection classes and contain the same syncretisms.

(1) a. (DAT.SG) *работ*ой - *работ*ы (GEN.SG, NOM/ACC.PL)
b. (DAT.SG) *мест*ом - *мест*а (GEN.SG, NOM/ACC.PL)

In the dataset the number of forms for each inflection class is not homogeneously distributed. To adress this I equalized the number of members of the inflection classes to a maximum of 100, dropping inflection classes with fewer members (different frequency thresholds produced very similar results). This is important because otherwise the models would perform much better for inflection classes with many members.

- trained a semantic analogy for each inflection class
- calculated its performance (as R2 of predictionvs vs observations) on the same inflection class using 10-fold cross-validation
- calculated its performance against each other inflection classes in that group and computed the difference between the within class R2 and cross-class R2

Results In all cases the models achieved very good (> 0.5) R2 values for the within class predictions, which shows that building semantic analogies in this manner is possible. As seen in Figure 1, the cross-validated R2 value of the model within an inflection class was almost always significantly higher than the R2 value of the model predicting a different inflection class, with a mean difference of 0.04 for the different inflection class groups. Additionally, the results strongly indicate that inflection classes which share more exponents were closer semantically. These results show that semantic analogies are in fact sensitive to inflection class, and to the exponent analogies within inflection classes.

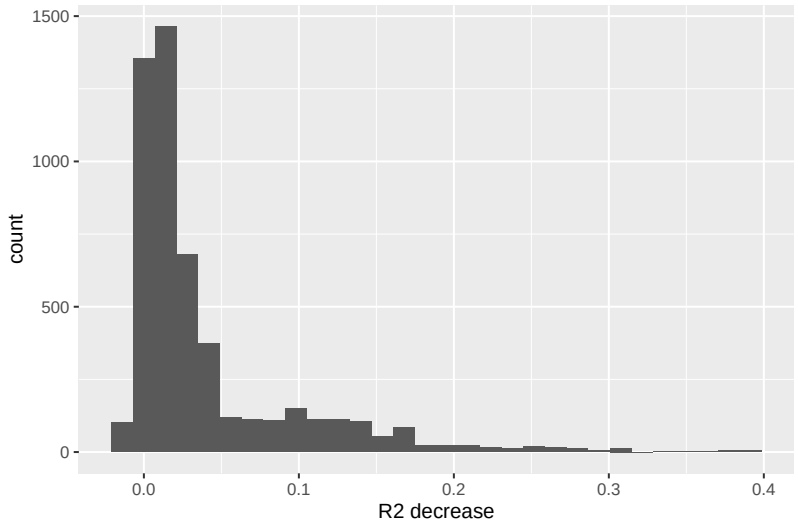


Figure 1: R2 difference between intra-class and cross-class predictions

References

Zaliznyak, Andrey Anatolyevich (1967). *Russkoe Imennoe Slovoizmenenie*. Moscow: Nauka.

Carstairs-McCarthy, Andrew (1998). "How Lexical Semantics Constrains Inflectional Allomorphy". In: *Yearbook of Morphology 1997*. Ed. by Geert E Booij and Jaap van Marle. Springer, pp. 1–24.

Mikolov, Tomas et al. (2013). "Distributed Representations of Words and Phrases and Their Compositionality". In: *Advances in Neural Information Processing Systems*, pp. 3111–3119.

Kutuzov, Andrey and Elizaveta Kuzmenko (2017). "WebVectors: A Toolkit for Building Web Interfaces for Vector Semantic Models". In: *AIST 2016*. Cham, pp. 155–161.

Jana Hasenäcker, Maria Ktori and Davide Crepaldi

Processing of compound constituents: position-specificity and interpretability

[Category of submission: Oral or poster]

It is quite well established by now that the constituent morphemes within complex words (e.g., *play+ground*) are identified during visual word processing, and likewise the presence of real morphemes slows down pseudoword rejection (i.e., the so-called *morpheme-interference effect*; see Amenta & Crepaldi, 2012, for a review). Previous studies suggest that the identification of constituents in compound words is, to a certain degree, position-independent (Crepaldi, Rastle, Davis, & Lupker, 2013), while affix identification is position-bound (Crepaldi, Rastle, & Davis, 2010). This reflects the fact that constituents can occur in any position within compounds across the language (e.g., *boathouse - houseboat*), while, for example, suffixes are constrained to occur after a stem. In practice, however, some stems occur more often in the first position of a compound (e.g., *doorstep, doorstop, doorknob, doorbell* vs. *trapdoor*), while others occur more often in the second position (e.g., *background, playground, underground* vs. *groundwork*), and some stems virtually never appear in a given position (e.g., *chinbone*, *chinrest*, but no compound with *chin* in final position), mimicking the positional constraints that affixes have as a class. The question therefore is: Is the cognitive system sensitive to this specific positional information of stems or does it just make a general distinction between position-independent stems and position-bound affixes?

We constructed a set of noun-noun pseudocompounds in Italian by combining constituents that occur only in the initial position (e.g., *carta*, engl. paper) or only in the final position (e.g., *nave*, ship) of compounds with stems that never occur in compounds (e.g., *pace*, peace). In these combinations constituents could appear either in their typical position (e.g., *cartapace, pacenave*) or in their atypical position (e.g., *pacecarta, navecarta*). Forty-four Italian native speakers participated in a lexical decision task, during which the pseudocompound stimuli were presented among real compound words and morphologically simple words and pseudowords. We hypothesized that if constituent identification is position-independent, we should not see differences in the rejection times between typical and atypical position. If, however, the language system is sensitive to the positional distribution of constituents, rejection times should be slowed down more when constituents are in their typical position as compared to their atypical position, showing, thus, position-specific morpheme interference. In addition, after completion of the lexical decision task, participants were asked to rate the interpretability of each of the pseudocompounds they read before.

Response times of the lexical decision task were analyzed using linear mixed-effects models. We did not find an indication for position-specific morpheme interference: pseudocompounds with constituents in

their typical position were rejected equally fast as pseudocompounds with constituents in their atypical position and this was the same for both initial-only and final-only constituents. However, we did find a strong effect of interpretability: pseudocompounds with higher interpretability ratings were rejected more slowly. In a post-hoc analysis, we looked at the effect of constituent position on the interpretability ratings. This revealed that pseudocompounds with constituents in their typical position had higher interpretability ratings than pseudocompounds with constituents in their atypical position.

From these findings we conclude that a stem's positional distribution throughout language is taken into account as an additional source of information in compound processing (cf. Libben, 2014), but the effect of position-specificity enters through the backdoor of interpretability, at least when the visual presentation is not time-constrained. By comparing these data to a related study from our lab, we also discuss how time-constrained visual presentation can elicit effects of position-specificity that seem to be primarily based on processing at the orthographic level. Taken together these findings allow interesting insights into the dynamics of learned probabilities in orthographic and semantic processing of morphologically complex words.

Amenta, S., & Crepaldi, D. (2012). Morphological Processing as We Know It: An Analytical Review of Morphological Effects in Visual Word Identification. *Frontiers in Psychology, 3*. doi:10.3389/fpsyg.2012.00232

Crepaldi, D., Rastle, K., & Davis, C. J. (2010). Morphemes in their place: Evidence for position-specific identification of suffixes. *Memory & Cognition, 38(3),* 312–321. doi:10.3758/mc.38.3.312

Crepaldi, D., Rastle, K., Davis, C. J., & Lupker, S. J. (2013). Seeing stems everywhere: Position-independent identification of stem morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 39(2),* 510–525. doi:10.1037/a0029713

Libben, G. (2014). The nature of compounds: A psychocentric perspective. *Cognitive Neuropsychology, 31(1-2),* 8–25. doi:10.1080/02643294.2013.874994

Modern Japanese is known to have four lexical strata: i) native Japanese, ii) Sino-Japanese (words of Chinese origin), iii) foreign, and iv) mimetics (Kageyama and Saito 2016). This study deals with members of the first two strata and argues the following: 1) native Japanese multi-verb compounds share morphological, syntactic and semantic properties with Sino-Japanese verbs, and 2) this commonality is due to the fact that they both have "extended" semantic structure which cannot be expressed by native simplex verbs.

Kageyama (2010) points out that native verbs and Sino-Japanese verbs differ in that Sino-Japanese verbs may occur as the head of post-syntactic compounds and subject compounds whereas native verbs do not. However, upon closer look, we do see native compound verbs occurring in such constructions. A native compound verb serves as the head of the post syntactic compound in (1c) and as the head of the subject compound in (2c), compared to the Sino Japanese verbs of (a) and the native simplex verbs of (b).

(1) *Post-syntactic Compound*

a. *[kosyo **koonyuu]** no sai wa, ...*
secondhand_book purchase GEN time TOP
'when (you) purchase a secondhand book, …' (Kageyama 2010: 1)

b. **[huruhon **kau]** sai wa, ...*
secondhand_book buy.PRS time TOP
(intended meaning) 'when (you) buy a secondhand book, …' (Kageyama 2010: 1)

c. *[kosyo **kai-tori]** no sai wa, ...*
secondhand_book buy.INF-take.INF GEN time TOP
'when (you) purchase a secondhand book, …'

(2) *Subject Compound*

a. *[Mootsaruto **sakkyoku]** no kookyookyoku*
Mozart compose GEN symphony
'symphony composed by Mozart' (Kageyama 2010: 2)

b. **[Mootsaruto **tukuri]** no kookyookyoku*
Mozart make.INF GEN symphony
(intended meaning) 'symphony made by Mozart' (Kageyama 2010: 2)

c. *[hahaoya **te-dukuri]** no bentoo*
mother hand-make.INF GEN lunchbox
'mother's homemade lunchbox '

This distinction is systematic in that the same pattern can be observed in other constructions, such as the suffixation of *-tyuu* ‘during’, *-sumi* ‘completed’ and *-suru* ‘do’.

In addition to the aforementioned morphosyntactic distinctions, there are also systematic differences in meaning between Sino-Japanese and native complex verbs, on the one hand, and native simplex verbs, on the other. Levin and Rappaport Hovav (2006) suggest that there is a complementarity between the manner and result components of meaning lexicalized in verbs. If this is correct, one single verb cannot simultaneously express both manner and result; moreover, if both manner and result are to appear in a single clause, whichever is not lexicalized by the verb must be

expressed by an ancillary element. But there is another possibility: extension of the verb to accommodate both components. For example, the Japanese compound verb in (3) and the Sino-Japanese verb in (4) express both manner and result.

(3) *hamigakiko o tyuubu kara* ***sibori-dasu***
toothpaste ACC tube ABL squeeze.INF-get_out.PRS
'squeeze toothpaste out of the tube'

(4) *gekai ga syuyoo o* ***tekisyutu****-suru*
surgeon NOM tumor ACC extract-do.PRS
'a surgeon removes a tumor'

As in (3), in which two native simplex verbs are fused to form the compound *sibori-dasu*, two Sino-Japanese morphemes are combined to form the verb in (4): *teki* 'pick' and *syutu* 'get out'. In both (3) and (4), the first element expresses manner and the second element expresses result.

Another type of verb compound is relevant to the discussion of the “extended” nature of compound verbs. In (5), native verbs *iku* ‘go’ and *kuru* ‘come’ form a compound, but unlike ordinary compound verbs with the structure [V.INF-V.PRS], the second verb always takes infinitive form and is followed by the light verb *suru* ‘do’.

(5) kaidoo *o* ***[iki-ki]****-suru* *hitobito*
highway ACC go.INF-come.INF-do.PRS people
'people coming and going along the highway'

A nearly identical meaning can be expressed by the Sino-Japanese verb *oorai*, which is composed of two Sino-Japanese morphemes, *oo* ‘go’ and *rai* ‘come’.

(6) *kaidoo o* ***oorai****-suru* *hitobito*
highway ACC come_and_go-do.PRS people
'people come and going along the highway'

Interestingly, compounds of the structure exhibited in (5) are not acceptable without the verb *suru*.

(7) a. **kaidoo o* ***[iki-kuru]*** *hitobito*
highway ACC go.INF-come.PRS people
(intended meaning) 'people coming and going along the highway'

The [V.INF-V.PRS] form is “extended” in that it can include both manner and result, as mentioned above. However, [V.INF-V.INF]-*suru* and Sino-Japanese verbs represent a further extended meaning, which Makino and Tsutsui (1994) call “an inexhaustive listing of actions or states”, which cannot be expressed by the [V.INF-V.PRS] structure.

This study concludes that the similarity between Sino-Japanese verbs and native compound verbs can be explained by their "extended" nature in morphosyntax and semantics and that there are two degrees of "extension" among native compound verbs, [V.INF-V.PRS] and [V.INF-V.INF]-*suru*.

References

Kageyama, T (1993) *Bunpoo to Gokeisei* [Grammar and Word Formation]. Hituzi Syobo.

Kageyama, T and M. Saito (2016) Vocabulary Strata and Word Formation Processes. In T. Kageyama and H. Kishimoto (eds.), *Handbook of Japanese Lexicon and Word Formation*. Mouton de Gruyter.

Levin, B. and M. Rappaport Hovav (2006) Constraints on the Complexity of Verb Meaning and VP Structure. In H.-M. Gaertner, R. Eckardt, R. Musan, and B. Stiebels (eds.) *Between 40 and 60 Puzzles for Krifka*.

Makino, S. and M. Tsutsui (1994) *A Dictionary of Basic Japanese Grammar*. The Japan Times Publication.

Martha Booker Johnson

Reduction of Multiple Exponence in Kihehe
Oral or Poster

What is the interplay of factors that determines the reduction of multiple exponence and resulting affix order? There is a typological generalization that inner affixes are more commonly lost than outer affixes in reduction of multiple exponence, with "optimal affix order" (e.g. contextual inflection outside inherent inflection) as one motivation (Haspelmath, 1993; Mithun, 2000; Harris and Faarlund, 2006; Harris, 2017). However, I show based on fieldwork data collected in Mufindi District, Tanzania, that Kihehe speakers are instead innovating a construction that deletes the outer affix. I argue that the loss of the outer affix can be explained by language-internal analogy, which suggests that the existing system is the more important factor in language development (Harris, 2008). While deletion of the outer affix is initially surprising, it provides an opportunity to explore how structural patterns in Kihehe interact with semantic and other factors in reduction of multiple exponence.

Before presenting the relevant construction, I lay out some basic information about Kihehe verbs. In simple tenses, Kihehe affix order (1) is consistent with prototypical Bantu verbal affix order (Meeussen, 1967): subject marker (SM), TMA prefix, object marker, stem.

(1) va-ka-ndi-lafile igolo
3PL.SM-PST-1SG.OM-look.at.PST yesterday
'they looked at me yesterday'

In auxiliary constructions, Kihehe requires a SM on the auxiliary, in addition to the SM on the lexical verb (2).

(2) *(tu)-va tw-i-geenda
1PL.SM-be 1PL.SM-PRS.PROG-walk
'we will be walking'

Like auxiliaries, the construction that is the focus of this paper also allows double-SM (3-a).

(3) a. (tu)-ke tu-kalite
1PL.SM-TMA 1PL.SM-dry.PST
'we dried'
b. *tu-ke kalite

However, *ke* constructions differ from auxiliary constructions in several ways. First, the four consultants I worked with optionally omitted the SM on *ke* (3-a), in contrast to (2). The optionality of the first SM is not documented in older grammars on Kihehe (Velten, 1899; Dempwolff, 1912; Priebusch, 1935), although the double-SM form is documented. This suggests that the double-SM form is older and the optionality of the outer SM is a recent innovation. It also suggests that *ke* is being reanalyzed as a prefix to the lexical verb. The non-optionality of the second SM is surprising given the typological generalization that the inner affix is more likely to delete (Harris and Faarlund, 2006; Harris, 2017). Also, deletion of the inner SM would have brought (3-a) in line with the Kihehe template for simple tenses.

Ke constructions also differ from auxiliaries in terms of coordination. In coordinated VPs, the auxiliary SM-VA is optional in the second VP and when absent, the SM-VA at the

left edge scopes over both lexical verbs (4).

(4) tu-va tu-negite kono (tu-va) tu-telike, pe
1PL.SM-be 1PL.SM-carry.PST where 1PL.SM-be 1PL.SM-cook.PST, when
tw-i-heega
1PL.SM-PRS.PROG-leave
'we will have carried and (we will have) cooked, when we leave'

In contrast, if *ke* is omitted from the second VP, the *ke* at the left edge does not scope over both verbs (5-b).

(5) a. **ke** tu-negite kono **ke** tu-telike
TMA 1PL.SM-carry.PST where TMA 1PL.SM-cook.PST
'we carried and we cooked' [simultaneous action]
b. %**ke** tu-negite kono tu-telike
TMA 1PL.SM-carry.PST where 1PL.SM-cook.PST
'we carried and we have cooked' [carrying happened first, then cooking]
c. tu-negite kono **ke** tu-telike
1PL.SM-carry.PST where TMA 1PL.SM-cook.PST
'we have carried and we cooked' [cooking happened first, then carrying]

In (5-b), *tu-telike* (without *ke*) is a tense form on its own and the verbs were interpreted as being in two different tenses. The contrasting coordination facts for auxiliaries (4) and *ke* (5) show that *ke* does not behave like an auxiliary. In my paper, I detail additional characteristics of auxiliaries, affixes, and *ke* with regards to separability, gapping, and tense marking and conclude that *ke* has behavior that is between that of affixes and auxiliaries. I then argue that speakers are (optionally) omitting the outer SM based on analogy with verbal modifiers (relatives, adverbials, etc.) that bear a phonological resemblance to *ke*.

For Kihehe, the univerbation of *ke* with the lexical verb and the loss of the outer SM signal variation in the affix template that could develop into a permanent change in affix order. In addition to posing problems for syntactically-driven approaches to morpheme order (Baker, 1985), the loss of the outer SM is unexpected based on simple tense affix order in Kihehe, canonical Bantu affix order, and typological tendencies in reduction of multiple exponence (Harris and Faarlund, 2006; Harris, 2017). Ultimately, the Kihehe facts suggest that language-specific characteristics can bring about typologically unusual forms of reduction of multiple exponence.

Paul Kiparsky

In Defense of Level-Ordering

The LEVEL-ORDERING HYPOTHESIS says that phonology and morphology are interleaved at two levels, the stem and the word, which are identified by their proprietrary phonologies and by the constituency and ordering of affixes. I address the criticisms of level-ordering most frequently voiced by morphologists working on English (e.g. Raffelsiefen:2000, Bauer 2013, Saarinen & Hay 2014, Lieber 2019), show that they underestimate the scope and nature of the theory's claims, and present new evidence in support of it.

One source of apparent counterevidence to level-ordering is the commonly used shortcut diagnostic of "stress-neutrality" as the criterion for identifying word-level affixes. For example, *-ée*, *-ésque*, and *-étte*, though stress-changing, are word suffixes by every morphological and phonological criterion; they have an underlying inherent stress, which is *not* a stem-level characteristic. The proper phonological stem-level criteria, as the Lexical Phonology literature makes clear, are adherence to the quantitive-sensitive English stress rule and to syllable-structure constraints such as those on hiatus, syllabic sonorants, coda clusters, and superheavy syllables. These criteria cleanly distinguish 82 derivational stem suffixes and 48 word suffixes in English, with the predicted morphological correlates.

A more serious problem is the conflation of level-ordering with the stratification of vocabulary into native and non-native items (reflected in the call to our IMM19 meeting) . These stratifications are actually independent, even in English. The stem-level *-al* of *signal* and the word-level *-age* of *signage*, and the above-mentioned stressed suffixes, are all non-native. Stem-level *-th* and word-level *-ness* are both native, and so are *-ful*$_A$ (unstressed and reducible to [-fəl]) and word-level nominal *-ful*$_N$ (with secondary stress). English native inflections are level-ordered, e.g. *slept* vs. *leaped*, *sat* vs. *flitted*, *mice* vs. *rats*, *stimuli* vs. *caucuses*, as is compounding, irrespective of the members' native or borrowed origin, e.g. *handcuff* vs. *handcart* (Allen 1978: 129). Moreover, native English suffixes may select non-native stems and vice versa, and native and non-native derivational affixes can be combined in either order (Bauer, Lieber & Plag 2013, Lieber 2010: 189). Even the weak correlation that has led to this conflation in work on English morphology is entirely missing in Dakota, Arabic and in fact in almost all of the 40-odd languages for which we have stratal analyses so far; in these the entire level-ordered morphology is native.

CHAMELEON AFFIXES that make stems from stems, and words from words (Giegerich 1999, Ch. 2) are often cited as problem for level-ordering. I argue that the systematic morphology/phonology covariation revealed by their dual affiliation actually *supports* the theory. The prediction is that where the morphology is consistently word-level, so is the phonology, and where the morphology is consistently stem-level, so is the phonology. This prediction is largely correct, down to the level of individual lexical items.

Haspelmath & Sims (2010: 227) maintain that productivity is sufficient to account for English level-ordering effects. To test their conjecture, I extracted the nouns in the online OED that end in the commonest stem and word suffixes, and counted how many of them occur first in the 20th century, following one of their measures of productivity (Haspelmath & Sims 2010: 130).

(1)

Stem suffixes ("integrated")				*Word suffixes ("neutral")*			
	Total	1900-			Total	1900-	
-ic, A.	7397	1750	24%	-age, N.	1229	150	12%
-ite, N.	3088	610	20%	-dom, N.	326	35	10%
-ity, N.	2994	418	14%	-less, A.	1793	111	6%
-al, A.	7204	874	12%	-ish, A.	1347	85	6%
-ive, A.	2151	176	8%	-hood, N.	314	15	5%
-ance, N.	880	62	7%	-ness, N.	4290	230	5%
-ous, A.	5376	172	3%	-like, A.	403	20	5%

These data show that the commonest stem-level derivational suffixes are on the whole *more* productive than the commonest word-level ("neutral") derivational suffixes, in absolute as well as proportional terms. This undercuts H&S's productivity-based explanation for level ordering effects.

I conclude that level-ordering achieves better empirical coverage than any other single overall generalization about affix ordering so far proposed, and more importantly, that it makes cross-linguistically valid predictions about the relation between morphology and phonology. Stratal organization is therefore essential to the understanding of word structure.

The development of derivatives in Danish-speaking children's spontaneous speech

Children coin new words from an early age as a supplement to the conventional words they have already acquired. These new words are often compounds and derivatives, as e.g. the innovative noun compound *drengeunge* [boy kid] coined from the nouns *dreng* [boy] and *unge* [animal young/kid] by the Danish girl Anne at age 2;2 (Kjærbæk & Basbøll 2017), and the innovative verb derivative *perleficere* [beadsify] coined from the noun *perle* [bead] and the derivational suffix *-ficere* (which turns the noun into a verb) by the girl Karoline at age 11 (used about the act of pulling beads on a cord (personal communication)).

Earlier studies of English-speaking children have shown that 2-year olds use nouns as verbs for actions where they do not yet know the established term for the relevant action (e.g. *I broomed her* [hit her with a broom] and *Daddy's rugging* [vacuuming]), and 3-year olds produce some derivational affixes like agentive and instrumental *-er* (e.g. *I'm a good cooker!* [cook] and *That's a climber* [ladder], Clark 1993). This is something that – to our knowledge – has never been studied for children speaking Danish.

The present study is a first attempt to describe and characterize the development of morphological structures in monolingual Danish-speaking children's early spontaneous speech focusing on derivation. Three central issues for the understanding of the acquisition of derivational morphology are addressed: i) the distributional properties of derivational morphology in the Danish lexicon; ii) the mechanisms of learning from input; and iii) the developmental patterns and trajectories in a corpus of spontaneous parent-child-interactions of four Danish-speaking children from the first occurrence up to the ages of 2;5, 2;5, 2;11 and 3;11, respectively – with brief outlooks into the development up to the age of 6;0.

First, we give a brief introduction to the Danish derivational system including a discussion of important classification issues. Second, we present an operationalization allowing us to make, thirdly, a corpus analysis of the development of conventional as well as innovative derivatives in child speech (CS) and child directed speech (CDS).

Derivatives (derived words) contain at least one prefix (e.g. *be-*, *u-*, as in *betale* [pay] and *uven* [enemy, lit. unfriend]), or one derivational suffix (e.g. *-agtig*, *-lig*, *-skab*, as in *nøjagtig* [exact], *venlig* [friendly], *klogskab* [wisdom]). Derivation is productive in Danish both with regard to conventional and more innovative forms.

Derivation must be distinguished from *inflection* on the one hand and *compounding* on the other. There are at least three cases where it is rather difficult to draw the border line between derivation and inflection: (i) certain infinite verb forms, viz. infinitive and participles (e.g. *studerende* [studying, student] with the present particple ending *-ende*); (ii) adverbs like *ind* [in(to)], *inde* [within], where the morpheme *-e* in the latter example represents static (as against dynamic: the zero form); (iii) the ending of ordinal numbers (e.g. *-ende* in *syvende* [seventh], from *syv* [seven]) (Diderichsen (1946: 20-21). Concerning the relation between compounding and derivation (both considered cases of word formation in traditional terminology), it is a problem that there are many "mixed forms", e.g.

containing both (i) derivational suffix(es) and/or prefix, and (ii) at least two roots; (i) points towards derivation and (ii) towards compounding. This necessitates a detailed grammatical analysis in layers (depending on the theoretical framework).

The results of the present study show that derivative suffixes are far more frequent than prefixes, both in the Danish lexicon, CDS and CS. The most frequent type of derivation in CS is conversion, i.e. zero-derivation (e.g. *sut* [suck] (verb stem) → *sut*-Ø [dummy] (noun); *kridt* [chalk] (noun) → *kridt*-Ø [chalk] (verb stem)). It can be impossible, though, to determine the word class of semantically related and homophonous word pairs – a well-known methodological problem within early language acquisition research (e.g. Ambridge & Lieven 2011). This will be discussed in the presentation. The majority of derivational patterns is acquired after the age of 3;11.

References

Ambridge & Lieven (2011). *Child Language Acquisition*. Cambridge: Cambridge University Press.

Basbøll (2005). *The Phonology of Danish*. Oxford: Oxford University Press.

Clark (1993). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.

Clark (2014). Acquisition of derivational morphology. In Lieber & Štekauer (eds.) *The Oxford Handbook of Derivational Morphology*, 424-439.

Clark (2015). Lexical meaning. In Bavin & Naigles (eds.) *The Cambridge Handbook of Child Language*. Cambridge: Cambridge University Press, 351-368.

Diderichsen (1946). *Elementær Dansk Grammatik*. Copenhagen: Gyldendal.

Kjærbæk & Basbøll (2017). Compound nouns in Danish child language. In Dressler, Ketrez & Kilani Schoch (eds.). *Nominal Compound acquisition*. John Benjamins, 39-62.

Sven Kotowski

Modeling the semantics of *out*-prefixed verbs

Like many derivational processes, English *out*-prefixation gives rise to various semantic and syntactic categories (Bauer et al. 2013: ch.16), e.g.:

(1) a. **locative nouns** outhouse, outstation, outmate
b. **locative verbs** útgas, útsource, út-migrate
c. **comparative-scalar verbs** outrún, outdóllar, outstúbborn

Accounting for the relationship between such different interpretations poses a classic problem for derivational semantic analysis. Are different constructions sufficiently alike for assuming a common semantic core or a basic reading from which others can be derived? And if so, what forms the locus of semantic similarity: affixes themselves, morphological processes, or constructions (Plag et al. 2018; Rainer 2014). Along these lines, this paper investigates the locative and comparative verbs in (1b-c).

Comparative *out*- has generated only few studies (e.g., Kotowski 2019; Talmy 2000; Tolskaya 2014), locative *out*-verbs have been outright neglected in the literature, and both their semantics and the possible relation between the two constructions remain ill-understood. This paper presents the first systematic juxtaposition of comparative and locative *out*-verbs. It builds on a large set of attested data from various corpora (~700 types, 1,000 tokens) and shows that the fundamental differences between the two prefix uses clearly outweigh their commonalities. The PATH=SCALE-metaphor (e.g. Tolskaya 2014) is shown to be insufficient as an explanation for the synchronic relationship between these two prefix uses.

The semantic analysis of the data leads to the following generalizations: First, locative *out*- is largely restricted to verbal bases and occurs both transitively and intransitively, showing some form of causative-inchoative alternation that mirrors the respective FIGURE-GROUND-constellation. In contrast, comparative *out*- is clearly applicative, giving rise to transitive structures only. The two senses appear to be systematically differentiated by primary stress assignment on either the base in comparative or the prefix in locative cases (indicated by the acutes in (1)). Comparative cases are by far more productive and more promiscuous regarding possible base classes. Explaining polysemy resolution on the basis of a highly underspecified affix-meaning and affix-base-interaction appears unfeasible for *out*-verbs: despite being rare, all locative base verb classes are also attested in comparative constructions, with ~20% of concrete locative base verbs also attested with comparative *out*- in my database. At the same time, both constructions allow for interpretational variation of different kinds.

I propose distinct lexeme-formation rules (Bonami & Crysmann 2016) for different senses of *out*- modeled in frame semantics (Löbner 2014; Plag et al. 2018). Comparative *out*- is analyzed as a subevent-adding operation on a morphological base that reflects its (weak) causative-resultative nature, while additional constraints govern interpretational flexibility. The analysis also circumvents the putative problem of selecting the same thematic role twice by accounting for the prefix's sense-specific contribution of correlating two events (Talmy 2000). The proposed frames reflect that the equation PATH=SCALE for locative vs. comparative verbs and the reduction of *out*-'s contribution to the mere transgression of a threshold falls short of the data on both empirical and structural grounds. Construction-specific semantics are of paramount importance in the modeling proposed: comparative *out*- always includes three, while locative *out*- comprises only two sub-events. Irrespective of their undoubted diachronic

relation, both the data and the analysis proposed in this paper speak in favor of two distinct morphological processes rather than one polysemous affix.

REFERENCES

Bauer, L., R. Lieber & I. Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Bonami, O. and Crysmann, B. 2016. Morphology in constraint-based lexicalist approaches to grammar. In A. Hippisley, A. & G.T. Stump (eds.), *The Cambridge handbook of morphology*, p. 609-656. Cambridge University Press, Cambridge.

Kotowski, S. 2019. The semantics of English *out*-prefixation. Submitted to *English Language and Linguistics*.

Löbner, S. 2014. Evidence for frames from human language. In T. Gamerschlag, D. Gerland, R. Osswald & W. Petersen (eds.), *Frames and concept types*, 23–67. Springer.

Plag, I., M. Andreou & L. Kawaletz. 2018. A frame-semantic approach to polysemy in affixation. In O. Bonami, G. Boyé, G. Dal, H. Giraudo & F. Namer (eds.), *The lexeme in descriptive and theoretical morphology*, 546–568. Berlin: Language Science Press.

Rainer, F. 2014. Polysemy in derivation. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Derivational Morphology*, 338-353. New York: Oxford University Press.

Talmy, L. 2000. *Toward a Cognitive Semantics. Vol. 2: Typology and Process in Concept Structuring*. Cambridge, MA: MIT Press.

Tolskaya, I. 2014. Out- and over- prefixation: a scalar approach. Paper at *Workshop on Aspect and Argument Structure of Adverbs and Prepositions* (WAASAP 2); Tromsø/Norway June 12-13.

Anne Kruijt

Abstract: 19th Morphological Meeting (Vienna 2020)
Category: Either oral presentation or poster

Case systems in contact

The regions of Tre Venezie in Northern Italy form a multilingual border region which is home to a large number of linguistic varieties and languages. The Italo-Romance varieties Standard Italian, Venetian, Lombard, and Ladin are spoken side-by-side with the Germanic varieties Standard German, South Tyrolean, Cimbrian and Mòcheno. This abundance of linguistic varieties on a relatively small area and the resulting long-term language contact situation, make the area particularly interesting for research into microvariation and language contact phenomena.

This research focuses on the morphological case systems of the personal pronouns and definite articles of the languages in the area. In general it is correct to state that Northern Italo-Romance varieties have no or almost no case morphology left. In the pronoun systems there has been a general drift towards extending the accusative form across the entire paradigm, and the definite articles do not have morphological case at all. However, Salvi (2016:159) shows that in the Gardanese variant of Ladin, there is a difference between accusative second person singular pronouns (1a) and dative second person singular pronouns (1b).

(1) a. *l a'niəl ros iə pra* ***te***
the lamb brown is near you.SG.**ACC**
'The brown lamb is near you.'
b. *kæʃˊ gwant tə feʒ i a* ***ti***
this dress to-you= make =I to you.SG.**DAT**
'This dress I am making for you.'

(Salvi 2016:159)

This is very uncommon for Northern Italian varieties, but it is not unheard of in other Rhaeto-Romance varieties in contact with Germanic languages, like in Switzerland (Seiler 2004), indicating this might either be a contact-induced innovation or a contact-motivated retention of an older system.

Most of the Germanic varieties in the area have a three-way case system: nominative, accusative, and dative (in most, if not all, varieties the genitive is no longer used productively). However, how this case system is structured is highly variable. For some varieties there is a split in the case syncretisms of the masculine and feminine gender, while others are shifting towards the same two-way split throughout the entire system. There is also the emergence of other type of case marking phenomena. Prepositional Dative Marking (PDM) (Seiler 2004) means that the dative case is necessarily preceded by a preposition. This pattern is obligatorily found in the Germanic minority language Mòcheno (2a), though not in Cimbrian (2b), another Germanic minority language of the region (Rabanus 2018). The PDM pattern is highly similar to the Italo-Romance way of marking a dative (2c), namely with a preposition *a*, which is obligatorily present before nouns and (full) pronouns.

(2) a. *I gib-en nèt doin kòrtn* ***in im*** *ma* ***in si***
(*Mòcheno*, Rabanus 2018)
b. *I gibe net di lettar* ***imen*** *ma* ***irn***
(*Cimbrian*, Rabanus 2018)
c. *Non ghe dagho la lettera* ***a lu*** *ma* ***a ela***
(*Trentino dialect*, VinKo 2019)
'I do not give the letter **to him**, but **to her**.'

The research works with oral data collected through online crowd-sourcing from a large sample of speakers and localities. Apart from general speaker background information, the online platform also registers which municipality participants are from and this allows for the creation of a comprehensive geographical overview of the area. The resulting areal perspective provides valuable insights into language contact areas, variation and similarity across linguistic varieties and the spread of features.

The aim of the presentation is to present an overview of the language contact situation and the effects it has on the morphological marking of case in the region. It will also serve to illustrate the use of online tools for oral data collection in a multilingual region, with a very high degree of variation.

REFERENCES

Rabanus, Stefan. 2018. Varietà alloglotte – tedesco, Versione 1 (23.07.2018, 17:00) in: Thomas Krefeld & Roland Bauer 2018. *Lo spazio comunicativo dell'Italia e delle varietà italiane. Korpus im Text*. Versione 32. url: http://www.kit.gwi.uni-muenchen.de/?p=13187&v=1.

Salvi, Giampaolo. 2016. Ladin. In Ledgeway, Adam and Martin Maiden (eds.). 2016. *The Oxford Guide to the Romance Languages.* Oxford: University Press: 154-168.

Seiler, Guido. 2004. *Präpositionale Dativmarkierung in Oberdeutschen*. Franz Steiner Verlag: Wiesbaden.

Varieties in Contact (VinKo), University of Verona and University of Trentino, accessed on 18-03-2019, <https://www.dipsco.unitn.it/vinko/index.php?lang=en>.

Lior Laks

Hebrew maCCuC: A tale of a borrowed pattern

This study examines formation in the *maCCuC* pattern. Some Hebrew adjectives have *maCCuC* doublets. The adjectives *maxrid* (1a) and *maxrud* (2b), both denote 'awful', share the stem consonants *x-r-d*, but are formed in different patterns. A similar case is presented in (2) for the adjectives *misken* and *maskun* 'poor'.

(1) a. lavašti jins **maxrid**
'I wore an awful pair of jeans'
(https://bike.co.il/?p=2239)
b. hi xorešet al oto jins **maxrud**
'she wears the same awful jeans'
(http://tmi.maariv.co.il/style/Article-609396)

(2) a. eyze **misken**, kol šniya tokfim oto
'what a poor (guy), he is attacked every second'
(https://trwomen.com/id/%D7%91%D7%95%D7%90%D7%95-%D7%A0%D7%A9%D7-%D7%A7-%D7%A1%D7%95%D7-%D7%A8-%D7%94%D7%99%D7%9D-%D7%9E%D7%A9%D7-%D7%A7-%D7%99%D7%A9%D7%A8%D7%90%D7%9C%D7%99-%D7%91%D7%A2%D7%91%D7%A8%D7%99%D7%AA-%D7%90%D7%95%D7%9C%D7%93%D7%A1%D7%A7%D7%95%D7%9C-p1fUlh98k8.html)

b. eyze **maskun**, kol paam ani yocet alexa
'what a poor (guy), I lash out at you every time'
(https://www.inn.co.il/Forum/Forum.aspx/t851240)

Both words in each pair share the same meaning and are used in similar contexts Not all speakers accept *maCCuC* forms like the ones in (1b) and (2b) (Bolozky 1999,2010), yet web searches reveal that they are productive. In contrast, there are adjectives that do not have *maCCuC* counterparts, e.g. *metunaf* – **matnuf* 'filthy'. I show that *maCCuC* formation (and lack thereof) can be predicted based on the interaction of semantic and morpho-phonological criteria. The talk will focus on the morph-phonological aspect. I argue that cases of *maCCuC* formation can be better explained under word-based approaches without separate reference to the consonantal root.

maCCuC hosts two types of adjectives:
(i) Borrowed adjectives from Arabic, e.g. *majnun* 'crazy', *mastul* 'high (on drugs')[1]
(ii) Forms based on existing Hebrew adjectives, e.g. *magʔil-magʔul* 'disgusting'

Bolozky (1999) shows that *maCCuC* forms started occurring in Hebrew as a result of direct borrowing from Arabic (i), and with time, speakers started using this pattern for doublets of existing Hebrew adjectives. In contrast to cases like (ii), there are adjectives without *maCCuC* counterparts, e.g. *maksim-*maksum* 'charming', *satum-*mastum* 'blockheaded'. Why is it so? Semantically, *maCCuC* adjectives must have negative meaning. The borrowed adjectives (i) are mostly negative, and *maCCuC* is associated with this meaning. However, there are negative adjectives without *maCCuC* counterparts. These gaps are explained by a morpho-phonological factor. Adjectives with *maCCuC* counterparts have medial consonant clusters. *maCCuC* formation is faithful to the base, as it involves only vowel(s) changes and preserves the syllabic structure (3a). This brings about structural transparency between the

[1] The borrowed adjectives undergo phonological adaptation, which is not discussed here. For example, the Arabic pattern has a long vowel (*maCCu:C*), which does not exist in Modern Hebrew.

forms. *maCCuC* formation based on adjectives without medial clusters involves modification of the syllabic structure and therefore it is blocked (3b).

(3) a. ma**xr**id - ma**xr**ud 'awful'
b.me**tn**af - *ma**tn**uf 'filthy'

Such gaps are better explained under word-based approaches (Aronoff 1976, 2007). It provides support to the theory of stem modification (Steriade 1988, McCarthy & Prince 1990, Bat-El 1994,2017, Ussishkin 1999,2005) rather than the extraction of a consonantal root. Root extraction could be performed on any base, regardless of its structure. Interestingly, adjectives without *maCCuC* counterparts have counterparts in patterns like *CeCeC*/*CaCeC*, e.g. *satum-setem* 'thickheaded', *metunaf-tanef* 'filthy'. Again, such formation does not modify the syllabic structure of the base.

While most studies on language contact examine borrowed words and borrowed grammatical elements like affixes, there are less studies that examine borrowed prosodic patterns. This case study sheds further light on the nature of non-concatenative morphology and the status of the consonantal root, in addition to the conditions of the usage of this borrowed pattern.

References

Aronoff, M. 1976. *Word Formatiom in Generative Grammar*. Cambridge, Mass: MIT Press.

Aronoff, M.. 2007. “In the beginning was the word”. *Language* 83, 803-830.

Bat-El, O. 1994. “Stem modification and cluster transfer in Modern Hebrew”. *NLLT* 12, 572-596.

Bat-El, O. 2017. "Word-based items-and processes (WoBIP): Evidence from Hebrew morphology". In C. Bowern, L. Horn, and R. Zanuttini (eds), *On Looking into Words (and beyond)*, 115-135. Berlin: Language Sciences Press.

Bolozky, S. 1999. *Measuring Productivity in Word Formation: The Case of Israeli Hebrew*. Leiden: Brill.

Bolozky, S. 2000. "Stress placement as a morphological and semantic marker in Israeli Hebrew". *Hebrew Studies* 41, 53-82.

Steriade, D. 1988. “Reduplication and transfer in Sanskrit and elsewhere”. *Phonology* 5:1, 73-155.

Ussishkin, A. 1999. “The inadequacy of the consonantal root: Modern Hebrew Denominal Verbs and Output-output Correspondence”. Phonology 16, 401-442.

Ussishkin, A. 2005. “A fixed prosodic theory of nonconcatenative templatic morphology”. *Natural Language and Linguistic Theory* 23, 169-218.

McCarthy, J. & A. Prince. 1990. “Foot and word in Prosodic Morphology: the Arabic broken plural”. *Natural Language & Linguistic Theory* 8, 209-283.

Shani Levy-Shimon – shanilevshim@gmail.com, Bar Ilan University & David Yellin College, Israel

Dorit Ravid - doritr@tauex.tau.ac.il, Tel Aviv University, Israel

Rachel Schiff – rschiff@biu.ac.il, Bar Ilan University, Israel

The influence of word frequency and root frequency on the spelling acquisition of an unstable root letter - A Case Study of the Letter *He* in Hebrew

Spelling words that comprise phonologically and orthographically irregular root letters poses a major challenge to writers of morphologically rich languages. The current study examines the influence of root frequency and morphological family size on the written representation of vowel letters. Specifically, we studied the letter *He (ה')* in the Hebrew verb system that is manifested both as a homophonic consonant (the pharyngeal spirant *h*) and a vowel (mostly standing for *a* and *e)*, and hence exhibits irregular representation. The study is based on the idea that inconsistent behavior at the phonological-morphological-orthographic juncture that straddles speech, grammar and written representation poses especially difficult challenges in acquisition. The letter *he* is particularly suitable for this investigation as it needs special emphasis in speech while often interchanging with *y* (the letter י) in writing. Participants were 133 school-age children (in grades 3, 5, 8, and 11) and college students, who completed a verb dictation test. All verbs contained the letter *He (ה')*, while they varied in frequency and family size. Results show that the acquisition of the letter *He (ה')* develops gradually until complete attainment is demonstrated by the 11th grade. In addition, the consonantal representation of the letter is acquired more rapidly than the vowel representation. Word frequency appears to influence the letter's written representation throughout development, particularly when the letter functions as a vowel. Family size impacts the phonological representation of the letter in all age groups, while its effect on the consonantal representation remains only up to 8th grade. The results highlight the significant role of orthographically and phonologically irregular root letters on the development of writing skills. This study emphasizes the importance of focusing on the intersection of phonology, morphology and orthography in language instruction as an aid to uncovering regularity in order to improve and ease spelling acquisition.

Arne Lohmann

Acoustic evidence for a category-specific metrical schema? – An extended replication of Sereno & Jongman's (1995) reading study of noun-verb conversion homophones

The present paper describes an extended replication of a widely cited reading study that reports an effect of grammatical category on the acoustic correlates of stress in English noun-verb conversion homophones (Sereno & Jongman, 1995). More specifically, Sereno & Jongman (1995) compare noun and verb pronunciations of English disyllabic, non-stress-shifting conversion pairs, e.g., *answer* (v) vs. *answer* (n) or *design* (v) vs. *design* (n). The main result reported is that the noun tokens exhibit a tendency toward trochaic and the verb tokens toward iambic pronunciation, reflecting the general difference in position of stress between nouns and verbs (e.g., Davis & Kelly, 1997). This means, for example, that speakers still pronounced *design* with iambic stress across the two categories, but the noun exhibited a relative shift of duration, intensity and fundamental frequency toward the first syllable, and the verb toward the second syllable. Sereno & Jongman (1995) presented the words in isolation, this way controlling for the difference in prosodic position between the two categories in discourse (e.g., Sorensen et al., 1978). Thus, the effect indicates a direct influence of grammatical category on the phonetic modulation of stress, which could be explained via the existence of a category-specific metrical schema.

However, a closer look at the results obtained by Sereno & Jongman (1995) indicates that the empirical evidence for the effect of grammatical category is rather slim, as the acoustic differences between nouns and verbs do not consistently reach statistical significance. The original study was based on a very small sample size (n=160), as only five speakers were recorded. The non-significant results may therefore reflect a lack of statistical power in the original experiment.

The current replication investigates whether the trends found in the original study can be corroborated when increasing power through eliciting a considerably larger number of target word productions (n=2,560). To that end the production study was replicated with 40 speakers, each of whom produced twice the number of target words. A power analysis employing Monte Carlo simulations based on the original data indicates that the replication study has more than sufficient power to find effects of the size observed in the original data. An acoustic analysis of the newly acquired recordings that tested shifts between the two syllables on the acoustic parameters of duration, fundamental frequency, intensity and spectral balance does not indicate robust, statistically significant effects of grammatical category on the acoustic realization of stress. Given that the study was considerably overpowered, the current replication suggests that the category effect either does not exist or is infinitesimal in size.

The null-result is compatible with models of speech production that assume the phonetic realization of stress to be solely impacted by a metrical template that is part of the lexical entry and which is the same for both the noun and verb homophone tested.

References

Davis, S. M., & Kelly, M. H. (1997). Knowledge of the English Noun–Verb Stress Difference by Native and Nonnative Speakers. *Journal of Memory and Language*, *36*(3), 445–460.

Sereno, J. A., & Jongman, A. (1995). Acoustic Correlates of Grammatical Class. *Language and Speech*, *38*(1), 57–76.

Sorensen, J. M., Cooper, W. E., & Paccia, J. M. (1978). Speech timing of grammatical categories. *Cognition*, *6*(2), 135–153.

Maia Lomia

The Borrowing of Markers and Semantic Differentiation of Forms in Georgian and Megrelian

Georgian and Megrelian are related Kartvelian (resp. South Caucasian) languages. Georgian as a state language is the language of education and religious rituals, whereas Megrelian is a non-written language which equals a dialect in its sociolinguistic function. The mutual influence of related languages, including the borrowing of markers, is a natural process. In such cases, as a rule, the literary language is the influencer and the non-written language is the borrower. The given abstract focuses on a reverse process, namely, the borrowing of a marker by Georgian from Megrelian and the semantic change of the word-form, which results in the semantic differentiation of forms in Megrelian and Georgian.

In Megrelian, there are particles -me and -ge which are added to certain nominal lexemes in the nominative case and express the semantics of oath. In the Georgian linguistic literature, they are termed as “particles of oath”. In fact, these particles are the markers of the semantics of oath. The particles -me and -ge are used for different persons: -me is used when the speaker (first person) makes an oath, whereas -ge is used when the speaker forces the second person to make an oath. Georgian borrowed the particle -me from Megrelian to form a concrete lexeme. Therefore, we analyze examples with Megrelian particle -me, its Georgian correlates, the reasons for borrowing and the issues of semantic differentiation of forms.

Examples of the usage of particle -me in Megrelian and its Georgian correlates:

(1) Megr. ghoront-i-me/God-NOM-PTC = “I swear by the grace of God”.

(2) Megr. bzhash mard-i-me/grace of the Sun-NOM-PTC = “I swear by the grace of the Sun”.

In Megrelian, the particle -me is used when mentioning (swearing by) the soul of deceased parents:

(3) Megr. didash shur-i-me/Mother’s soul-NOM-PTC = “I swear by the soul of my mother”.

(4) Megr. mumash shur-i-me/father’s soul-NOM-PTC = “I swear by the soul of my Father”.

In Megrelian, including all the above-mentioned examples, apart from the semantics of oath, the particle -me expresses great respect and veneration to the object of the oath. Hence, the entire phrase is of positive connotation.

In the middle of the 19th century, a lexical unit rusetume (<ruset-i-me/Russia-NOM-PTC) appeared in Georgian. It was formed by the particle -me, which was borrowed from Megrelian, and literally meant: “I swear by Russia”. This term denoted the young people who had studied in Russia, did not care about the welfare of their country and had a nihilistic attitude to everything national. This term appeared in Georgian as an opposition to the term “tergdaleulebi” (literally, those who had drunk the waters of the Terek river). Unlike the young people who were called “rusetume”, the “tergdaleulebi” used their Russian education in support of national ideas. Thus, Georgian borrowed the Megrelian particle -me, but, unlike Megrelian, formed a lexical unit of negative connotation. As a result of semantic differentiation, a pair of antonymous terms“tergdaleuli” and “rusetume” appeared in Georgian.

The paper has been prepared within the scientific project – „The Category of Negation in the Kartvelian Languages” (#FR17_388), implemented with the financial support of Georgian Shota Rustaveli National Science Foundation.

Contamination, segmentation, and abstractive morphology in diachrony

This study considers some examples of diachronic 'contamination' (Paul 1890=1970:161) in Romance verb-forms, such that the exponent of one lexical meaning undergoes formal modification under the attraction of the exponent of some other, closely associated, meaning. The problem is that the 'contamination' occurs in the 'wrong place' within the apparent morphological structure of the word-form.

The word-forms of Romance verbs generally lend themselves to segmentation into a leftmost portion which is the exponent of lexical meaning, followed by inflexional matter to the right. Surprisingly, the contaminatory effects at issue here are not manifest in those lexical roots, but rather in the following inflexional endings and especially in the marking of inflexion class. That is to say that the 'contamination' occurs in a part of the relevant word-forms shared with a large number of other lexemes and which has no inherent relation to the lexical meaning of those verbs. The implication is that transparent segmentability of a lexical root 'morpheme' is deceptive, and that the primary exponent of the lexical meaning is the entire-word-form.

Romance languages have several major verb inflexion classes involving a complex, mutually-implicational, paradigmatic array of thematic vowel, alternant forms of the thematic vowel, stress alternation, and often distinctive fusional combination with tense, person, and number endings. For most Romance languages, neologisms are systematically assigned to just one highly productive inflexion class, the 'first conjugation'. The phenomena discussed in this study are notable by virtue of being neologisms which acquire the full array of inflexional characteristics of otherwise closed, wholly unproductive, inflexion classes, namely the 'second' or 'third' conjugations.

The main class of examples to be discussed involves the verb 'snow', which in nearly all Romance languages is a neologism historically derived from a noun *ˈneve 'snow'. Predictably, this new formation was almost everywhere assigned to the first conjugation. However, both in Francoprovençal and in Ladin dialects it sometimes acquires second or third conjugation morphology. Significantly, whether it shows second or third conjugation morphology depends directly on whether the inherited verb 'rain' locally shows second or third conjugation morphology (cf. Jaberg 1906:68n7). The following Francoprovençal data from *AIS* point 131 (Brusson) show how 'snow' follows the local third conjugation morphology of 'rain', and differs inflexionally from first conjugation verbs (see Duraffour 1932:77 or Ahlborn 1946:45 for Francoprovençal examples where these verbs both follow the second conjugation):

'rain' - 'snow': third conjugation						first conjugation		
INF		3SG. PRS.IND		PST.PRT		INF	3SG. PRS.IND	PST.PRT
						'dig'	'wash'	'thunder'
'pjovri	- 'nɛvri	'pjoːti	- nɛːt	pju'voj	- nu'oj	ga'taːr	'laːvat	tru'naː

The semantic closeness between 'rain' and 'snow' has allowed the former morphologically to 'seduce' the latter. I shall also review the tendency, already detectable in early Romance languages, but repeated in later centuries, for verbs of broadly 'modal' meaning to be corraled counteretymologically into the otherwise closed Romance second conjugation.

In my conclusion I will propose that such facts are inherently irreconcilable with a morpheme-based and 'constructivist' approach to morphological structure, but strongly consistent with an 'abstractivist' perspective (see e.g. Blevins 2016), which puts words and their place within paradigms (or 'items' and 'patterns': Blevins, Ackerman, and Malouf 2019:277) at the centre of morphological analysis. To understand how lexical 'contamination' apparently occurs in the 'wrong place' within words, we have to say that each of the *entire*

word-forms of the paradigm of a lexeme serves as an exponent of that meaning, not just a lexical root, however clearly segmentable it may appear.

Ahlborn, G. (1946). *Le Patois de Ruffieu-en-Valromey (Ain)*. Gothenburg: Wettergren & Kerbers.
AIS = Jaberg, K. & Jud, J. (1928-40). *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen: Ringier.
Blevins, J. (2016). *Word and Paradigm Morphology*. Oxford: Oxford University Press.
Blevins, J. Ackerman, F., & Malouf, R. (2019). 'Word and paradigm morphology'. In Audring, J. & Masini, F. (eds). *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press, 264-84.
Duraffour, A. (1932). *Description morphologique avec notes syntaxiques du parler franco-provençal de Vaux (Ain) en 1919-1931*. Grenoble: Institut phonétique de Grenoble.
Jaberg, K. (1906). *Ueber die assoziativen Erscheinungen in der Verbalflexion einer südostfranzösischen Dialektgruppe. Eine prinzipielle Untersuchung*. Aarau: Sauerländer.
Paul, H. ([1890] 1970). *Principles of the History of Language*. College Park, Maryland: McGrath.

On the role of intra-linguistic parameters of change in intensive language contact

Michalis Marinis
University of Patras
marinis@upatras.gr

The reduction of Griko nominal paradigms has been exclusively attributed to contact with the Romance varieties (Melissaropoulou 2017). However, a closer look at other Modern Greek dialects reveals that similar patterns of evolution occur in systems such as the Tsakonian dialect, that has evolved in a long-term isolation (Liosis 2009). Therefore, a question which arises about the real causes triggering this change is whether it is due to inter-linguistic or intra-linguistic factors.

My study of the diachronic change of Griko's inflection will be based on: (a) the 60 hours of oral data of the Laboratory of Modern Greek Dialects of the University of Patras, (b) the existing literature (among others, Melissaropoulou 2017, Gemma & Lambroyorgu 2001, Karanastasis 1997, Rohlfs 1950: 69-81), and (c) older documents presented in Minas [1993] (2003) and Manolessou & Pantelidis 2018.

Following Ralli's (2000, 2005) analysis of inflection classes in Standard Modern Greek, I show that today Griko has only five inflection classes, contrary to what has been proposed in the literature where the number of inflection classes is proposed to be higher than ten. I argue that compared to the picture presented in older sources, Griko's inflectional morphology has been shrunk. This change could be easily attributed to language contact with Romance (see Melissaropoulou 2017). However, more careful study to inter-dialectal data from Modern Greek shows that phenomena appearing in Griko are also present in other systems, some of which are in a long-term isolation.

In this presentation, I will show that the reduction of inflectional classes, while greatly facilitated by linguistic contact, is triggered by analogy either between different paradigms (inter-paradigmatic levelling) or between different cells of the same paradigm (intra-paradigmatic levelling). Moreover, I will also show that some intra-linguistic features of the system, such as its tendency to prefer the less-marked open syllabic structures, have played a major role in the procedure of paradigmatic restructuring.

Selected references:

Gemma, I., & Lambroyorgu, G. (2001). *Grammatica del dialetto greco di Sternatia*.

Καραναστάσης, Α. (1997). *Γραμματική των Ελληνικών Ιδιωμάτων της Κάτω Ιταλίας*. Αθήνα: Ακαδημία Αθηνών.

Liosis, N. (2009). Counterfactuality in the Tsakonian dialect: a contribution to the history of 'ήθελα'and 'ήμουν'. *Modern Greek Dialects and Linguistics Theory, 4 (1),* 111-123.

Manolessou, I. & N. Pantelidis. (2018). Speaking Greek in Southern Italy: historical, geographical and documentary aspects. Key-talk at *Greeks&Aliens - It's all Greek to me: Hellenization and Greek language survival in Calabria*. Athens: Scuola Archeologica Italiana di Atene.

Marinis, M. (2017). Language contact and paradigmatic uniformity in Greko dialect. Talk at *13th International Conference in Greek Linguistics*. London: University of Westminster.

Melissaropoulou, D. (2017). On the role of language contact in the reorganization of grammar: A case study on two Modern Greek contact-induced varieties. *Poznan Studies in Contemporary Linguistics, 53 (3)*, 449-485.

Μηνάς, Κ. (2003) [1993]. *Η γλώσσα των Δημοσιευμένων Μεσαιωνικών ελληνικών εγγράφων της Κάτω Ιταλίας και της Σικελίας*. Αθήνα: ΙΛΝΕ.

Ralli, A. (2000). “A Feature-based Analysis of Greek Nominal Inflection”. Glossolojia 11-12. 201-228.
Ράλλη, Α. (2005). *Μορφολογία*. Αθήνα: Πατάκης
Rohlfs, G. (1950). *Grammatica storica dei dialetti italogreci (Calabria, Salento)*. München: H. Beck.
Thomason, S. G. (2010). Contact explanations in linguistics. *The handbook of language contact*, 31-47.

Elisa Mattiello and Wolfgang Dressler

Latinate vs. non-Latinate English synthetic compounds

The main innovative aim of this contribution, beyond what is dealt with in the Call for Papers for this conference and in the literature, is that the question of which features of morphology can be borrowed or extracted from borrowed words and word forms can be extended to what happens later with these borrowings. Latinate synthetic compounds (SCs) could have neither been borrowed nor extracted from borrowed material, but have been newly developed within English. Still they developed differently from non-Latinate synthetic compounds, as we can show.

Despite massive borrowing of Latinate bases and affixes and their integration into Middle English morphology (cf. Dalton-Puffer 1996), Latinate suffixes (but not prefixes, cf. re+build, sub+heading, super+man) still avoid combining with non-Latinate bases. Counterexamples of suffixes are rare: e.g., odd+ity, the nonce-word between-ity; flirt-ation.

However, henceforth, no such restrictions have been observed for compounding. This study wants to fill this lacuna by dealing with restrictions on, and features of, Latinate synthetic compounds ending in -er, which non-Latinate SCs do not share.

First, formal families of -er SCs have been formed much later from Latinate than from non-Latinate verbs: e.g., the formal families of shoe+mak+er [1381], wine+mak+er [1382], house+hold+er [a1382], man+kill+er [a1500], heart+break+er [1674] all precede the Latinate formal families of office+manag(e)+er [1866], company+manag(e)+er [1869], money+manag(e)+er [1874], gas+produc(e)+er [1841], oil+produc(e)+er [1859], play+produc(e)+er [1891], service+provid(e)+er [1954], etc. Thus, Latinate SCs developed much later than productive Latinate affixation.

Second, type and token frequencies of SCs from non-Latinate verbs largely exceed the frequencies of Latinate SCs: e.g., in COCA, film(-)maker has 3,177 tokens (5.62 pmw), shareholder 2,356 (4.17 pmw), wind(-)breaker 518 (0.91 pmw), and pain(-)killer 456 (0.83 pmw), whereas the occurrences of the Latinate SCs resoucemanager (16/0.03), service-provider (9/0.01), and oil-producer (1/0.00) are far more limited. This confirms that Latinate SCs represent a recent phenomenon that is still at its birth or of modern expansion.

Third and more importantly, SCs from non-Latinate verbs have developed several semantic subfamilies related to the polysemy of the base verbs, while Latinate SCs, due to being mainly restricted to specialised sectors (esp. economics), are not subdivided into semantic subfamilies. For instance, within the non-Latinate X-breaker family, seven different subfamilies have been developed: from 'violate' (lawbreaker), to metaphoric 'open' (groundbreaker), from 'cause rupture' (bonebreaker) and 'interrupt' (prison breaker), also metaphorically (heart-breaker, ice-breaker), to 'solve, change' (tiebreaker). In Mattiello & Dressler (2019) we argue that, if formal families are subdivided into semantic subfamilies, the latter are the basis for expansion and productivity. The productivity of the X-breaker family, for instance, is demonstrated by the profitability of this pattern for the formation of new words, also occasionalisms (e.g. barrierbreaker, 3 occ. in COCA) and hapax legomena (brain-breaker).

A fourth distinction between Latinate and non-Latinate SCs is that Latinate families, such as X-producer, X-manager, or X-provider, primarily support a derivation analysis (i.e. agent/instrument suffixation from the phrases produce X, manage X, provide X), while non-Latinate SCs support our hypothesis of being derived either from phrases (e.g. heartbreaker ← break hearts), or from compounding after derivation (e.g. gamebreaker ← breaker of the game), or, even more frequently, from both (e.g. met. timekiller ← kill time / killer of time), whereas, in the literature (as discussed in the Journal of Word Formation 1,1, 2017), there has been a heated debate whether the derivation or the compounding analysis is correct. Our claim is that English -er SCs have an ambiguous nature, which can be resolved by assuming, as argued in Mattiello & Dressler (2019), dualism and superposition of suffixing and compounding, similar to Albert Einstein's and Erwin Schrödinger's

assumption of dualism and superposition of waves and particles (photons) for light. What our analysis also provides, beyond the analogy to quantum physics (cf. Libben 2017), is the weighting of the impact of both compounding and derivation in SC families and subfamilies in a probabilistic way. In this weighting, Latinate and non-Latinate SCs differ as well.

As a general and innovative goal, this study aims at demonstrating that Latinate SCs have developed differently from non-Latinate ones. Hence, rather than representing a case of borrowed morphology, they represent a case of only partial and late integration of Latinate verb bases into the non-Latinate patterns of forming -er synthetic compounds.

References

Dalton-Puffer, Christiane (1996) The French Influence on Middle English Morphology: A Corpus-based Study of Derivation. Berlin/New York: Mouton de Gruyter.

Libben, Gary (2017) The quantum metaphor and the organization of words in the mind. Journal of Cultural Cognitive Science 1, 49-55.

Mattiello, Elisa & Wolfgang U. Dressler (2019) Dualism and superposition in the analysis of English synthetic compounds ending in -er.

Dimitra Melissaropoulou

Change in morphological complexity in the language contact perspective: evidence from the Cappadocian Greek derivational domain

This paper aims to offer further insights on the notion of formal 'autonomous' morphological complexity (Anderson 2015), in the light of the evidence provided by language contact, a parameter which is thought to be interrelated with change in the complexity of linguistic systems (among others Trudgill 1999, 2011; Nichols 1992), emphasizing the derivational domain. To this end, we draw our data from Cappadocian, an Asia Minor Greek variety spoken for great many centuries in a situation of regressive bilingualism due to intense contact with the agglutinative Turkish language (among others Dawkins 1916; Karatsareas 2011; Janse forthcoming). The following morphological phenomena are put under scrutiny as follows:

a. Loss of prefixation

While Greek has always displayed a wealth of prefixes and prefixoids (Ralli 2005), in Southern Cappadocia, prefixation as a productive derivational process is almost extinct (Melissaropoulou 2016). This phenomenon, which changes radically the overall affix ordering picture of Cappadocian paving the way towards a purely suffixing system, has important implications on the complexity of the Cappadocian morphological system since established research in the field suggests that suffixes are unmarked and provide a processing benefit in comparison with prefixes (Greenberg 1966; Cutler, Hawkins & Gilligan 1985; Hawkins & Cutler 1988).

b. Elimination of rivalry among competing derivational suffixes

While Greek in general displays a variety of deverbal suffixes expressing the action or the result of an action denoted by the verbal base, in Southern Cappadocian the only productive suffix in use is *-ma* (Kesisoglou 1951). E.g.:

(1)	ðin(o) 'to give'	ðini-ma	'giving'	instead of ðosimo
	vriz(o) 'to insult'	vrizi-ma	'insult'	instead of vrisimo
	feɣn(o) 'to go off/away'	feɣni-ma	'going off/away'	instead of fevʝo

What can be seen is that the suffix *-ma* is highly productive, in that it assumes the role of the formation of nouns that express the result of an action denoted by the corresponding verbal base, to the complete exclusion of all other competing suffixes (Melissaropoulou 2016), which were subject to different selectional restrictions and whose selection was not interpretable in all cases (cf. *Principle of Economy* by Kusters 2003).

c. Elimination of multiple stem allomorphs realized in derivational affixation

In South Cappadocian a strong tendency is observed towards the abolishment of the more idiosyncratic, less systematic allomorphic stems or even suppletive forms and the expansion in all cases of one systematic allomorphic pattern in ~Xi (X = the verbal stem used in the present indicative). E.g.:

(2)	xan(o) 'to lose'	xani-ma	'loss'	instead of xasi-mo
	troo 'to eat'	troi-ma	'eating'	along with fai-ma
	pen(o) 'to go'	peni-ma	'going'	instead of pije-mos

Expanding Cruschina, Maiden & Smith's (2013) view, we assert that a stem displaying multiple allomorphs (lacking independent motivation) is seen as more complex than one exhibiting only

one stem form, or one systematic allomorphic pattern (cf. 'one-meaning-one-form' principle, Kusters 2003).

A thorough analysis of the occurring phenomena reveals the following: the first phenomenon is seen as a change of system complexity that is strictly dependent on the model language and is in need of compensation by other levels of linguistic analysis (the resulting lacuna is compensated for by native means (effort in the lexical/pragmatic domain) and by lexical borrowing. Moreover, it goes hand in hand with other changes (e.g. word order). On the other hand, the other two phenomena, are seen as instances of exponence complexity (although system complexity might also be argued in the case of b.) complexity, following Anderson's (2015) categorization and they are not due to a direct influence by the model language (Melissaropoulou in print). Crucially, these phenomena result in change in the morphological complexity that is not in need of compensation by other levels of linguistic analysis (no complexity tradeoffs).

Admittedly, these phenomena could also easily be associated with language decay (see among others Dal Negro 2004; Dressler 2011) and be interpreted as -necessary but not sufficient- indicators of language decay. Nevertheless, they cannot be considered as purely dysfunctional (Dressler 2011) resulting in irreparable loss. Corroborative dialectal evidence shows that the occurring simplification phenomena are more adequately accounted for as part of the re-arrangements the system is subject in situations of intense contact (regularity, economy and loss of redundancy, through minimization of rivalry among elements and categories with similar function cf. Trudgill 2011) and as part of an on-going contact induced typological shift (cf. Melissaropoulou in print).

Thus, we propose that in contact settings phenomena that at one stage might correctly be accounted for as loss of complexity in another one might be interpreted as indicators of shift and later -but not obligatorily- of decay or death. In this perspective, the amount, the speed and the structural consequences of the occurring phenomena could serve as diagnostics / predictors towards one over the other possible scenarios of language change (Campbell & Muntzel 1989). On this basis, we align with the view that they should be treated under a unified account (Dorian 1989; Romain 1989; Dal Negro 2004) seen as possible routes or manifestations of contact-induced language change.

The above discussion is intended as a contribution to the discussion of change in morphological complexity in contact settings and its relation to other possible paths of contact-induced language change.

References

Anderson, Stephen R. 2015. Dimensions of Morphological Complexity. In M. Baerman, D. Brown & G. Corbett (eds.), *Understanding and measuring morphological complexity*, 11-28. Oxford: Oxford University Press.

Campbell, L. & M. C. Muntzel. 1989. The structural consequences of language death. In Nancy Dorian (ed.), Investi-gating obsolescence: studies in language contraction and death, 181-196. Cambridge: CUP.

Cruschina, S. Maiden M. & J. C. Smith. 2013. *The boundaries of pure morphology. Diachronic and Synchronic Perspectives*. Oxford: Oxford University Press.

Dal Negro, S. 2004. Language contact and dying languages. *Revue française de linguistique appliquée*, 2, 9, 47-58.

Dawkins, R. 1916. *Modern Greek in Asia Minor*. *A study of the dialect of Silly, Cappadocia and Pharasa*. Cambridge.

Dorian, N. C. (ed.). 1989. *Investigating Obsolescence: Studies in Language Contraction and Death*. Cambridge: Cambridge University Press.

Dressler, W.U. 2011. Early indicators of language decay. In E. Miola & P. Ramat (eds.), Language Contact and Language Decay. Pavia: IUSS Press, pp. 89-108

Greenberg, Joseph H. 1966. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Greenberg, J. H. (ed.), *Universals of Language*. Cambridge/ Massachusets/London, England: MIT Press, 73-113.

Hawkins, J. A., & A. Cutler. 1988. Psycholinguistic factors in morphological asymmtery. In J. A. Hawkins (ed.), *Explaining language universals*. Oxford, England: Basil Blackwell, 280–317.

Hawkins, J. A., & G. Gilligan. 1988. Prefixing and suffixing univerals in relation to basic word order. *Lingua* 74, 219–259.

Janse, Mark forthcoming. Cappadocian. In Tzitzilis, C. (ed.), *Η ελληνική γλώσσα και οι διάλεκτοί της* [Greek language and its dialects]. Thessaloniki: Institouto Neoellinikon Spoudon (Manolis Triantafyllides Foundation).

Karatsareas, P. 2011. A study of Cappadocian Greek nominal morphology from a diachronic and dialectological perspective. PhD Dissertation. University of Cambridge.

Kusters, C. W. 2003. Linguistic Complexity: The Influence of Social Change on Verbal Inflection, Utrecht: LOT.

Melissaropoulou, D. 2016. Variation in word formation in the light of the language contact factor: the case of Cappadocian Greek. *Journal of Language Sciences*, 55, 55-67. doi:10.1016/j.langsci.2016.02.005

Melissaropoulou, D. in print. Accounting for morphological complexity vs. simplification in situations of language contact: Evidence from Cappadocian Greek. In Ch. Tzitzilis (ed.), *Language Contact in the Balkans and Asia Minor*. Thessaloniki: Institute of Modern Greek Studies.

Nichols, Johanna 1992. *Linguistic Diversity in Space and Time*. Chicago, London: The University of Chicago Press.

Ralli, A. 2005. *Μορφολογία.* Athens: Patakis.

Romaine, S. 1989. Pidgins, creoles, immigrant and dying languages. In Dorian, Nancy C. ed. Investigating Obsolescence: Studies in Language Contraction and Death, 369-83. Cambridge: Cambridge University Press.

Trudgill, P. 2011. *Sociolinguistic Typology*. *Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.

Late visual morpho-semantic decomposition in Modern Greek derivational morphology

Either oral or poster

A large body of morphological processing research refers to the role of semantic transparency in late visual word recognition. In the absence of semantic transparency effects, it has been claimed that all morphologically complex words are decomposed, irrespective of meaning (Smolka et al., 2009). Conversely, semantic transparency effects indicate late decomposition for semantically transparent but not for semantically opaque complex words (Rastle & Merkx, 2011). Furthermore, graded effects of semantic transparency have been reported (Feldman et al., 2004; Xu & Taft, 2015). Interestingly, semantic transparency effects in late morphological processing seem to correlate with cross-linguistic differences (Smolka et al., 2009), since they are typically detected in morphologically impoverished, concatenative languages (e.g., English) but not in morphologically rich, non-concatenative languages (e.g., Hebrew).

The present study examines the parameter of semantic transparency in derivational morphology of Modern Greek (MG), an Indo-European (IE) language that is described as both morphologically rich and concatenative (Ralli, 2005). Specifically, we deal with denominal suffixed verbs and adjectives, aiming to investigate whether semantic transparency affects late morphological processing and –if this is the case– whether semantic transparency effects are "all-or-none" or graded. To test these hypotheses, we conducted a visual lexical decision experiment, employing the overt priming paradigm with 250 ms prime-target stimulus of asynchrony (SOA), which is considered to tap into late intra-modal processing.

Forty eight university students participated in the experiment. Ninety morphologically related prime-target pairs, each comprising a (pseudo-)derived prime and a (putative) noun-base target, were equally distributed in three semantic transparency conditions (transparent, e.g., *ψηφίζω* /psi'fizo/ 'to vote' – *ψήφος* /'psifos/ 'vote'; semitransparent, e.g., *στολίζω* /sto'lizo/ 'to decorate' – *στολή* /sto'li/ 'costume'; opaque, e.g., *θυμίζω* /θi'mizo/ 'to remind' – *θυμός* /θi'mos/ 'anger'), according to mean rating scores from three semantic relatedness pretests. Critical primes and targets were matched across conditions for several extraneous variables (e.g., length, lemma and word frequency, syllable and bigram frequency, prime-target formal similarity, orthographic neighborhood size, uniqueness point, morphological family size, grammatical category, inflectional and derivational suffixes). Critical stimuli along with 90 control primes (matched to critical primes but morphologically, orthographically and semantically unrelated to targets) were divided into two counterbalancing lists in a Latin square design, so that each participant was exposed to all priming conditions but saw each target only once. Moreover, each list included 180 word-nonword pairs and 90 filler word-word pairs, such that the overall prime-target relatedness proportion was reduced to 25% and target responses could not be predicted from prime-target formal similarity, derivational suffixes of primes, or total word-nonword proportion (50%).

Statistical analyses revealed no priming for the opaque condition. In contrast, significant priming was found for the transparent and semitransparent conditions (Figure 1), with the amount of priming being equal between them (Figure 2). Results suggest that late morphological processing in MG is semantically constrained, as is the case with other morphologically rich, concatenative IE languages (e.g., Polish, Serbian; but not German [Smolka et al., 2009]). Importantly, the observed priming

effects posit whole-word processing for opaque (pseudo-)derivatives on the one hand and morpho-semantic decomposition for transparent and semintransparent derivatives on the other hand, without further gradation between transparent and semitransparent derivatives. Namely, morpho-semantic decomposition seems to be driven by a predictable part of meaning (Corbin, 1987/1991; Lieber, 2004), which brings together transparent and semitransparent derivatives. Therefore, the present findings favor a dichotomous rather than graded account of semantic transparency in late visual morphological processing.

References

Corbin, D. (1987/1991). *Morphologie dérivationnelle et structuration du lexique* (2 Vols.). Tübingen/Villeneuve d'Ascq: Max Niemeyer Verlag/Presses Universitaires de Lille.

Feldman, L. B., Soltano, E. G., Pastizzo, M. J., & Francis, S. E. (2004). What do graded effects of semantic transparency reveal about morphological processing? *Brain and Language, 90,* 17-30.

Lieber, R. (2004). *Morphology and lexical semantics*. Cambridge: Cambridge University Press.

Ralli, A. (2005). *Morfologia* [Morphology]. Athens: Patakis.

Rastle, K., & Merkx, M. (2011). Semantic constraints on morphological processing. In G. Gaskell & P. Zwitserlood (Eds.), *Lexical representation: A multidisciplinary approach* (pp. 13-32). Berlin: De Gruyter Mouton.

Smolka, E., Komlósi, S., & Rösler, F. (2009). When semantics means less than morphology: The processing of German prefixed verbs. *Language and Cognitive Processes, 24,* 337-375.

Xu, J., & Taft, M. (2015). The effects of semantic transparency and base frequency on the recognition of English complex words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 904-910.

Figures

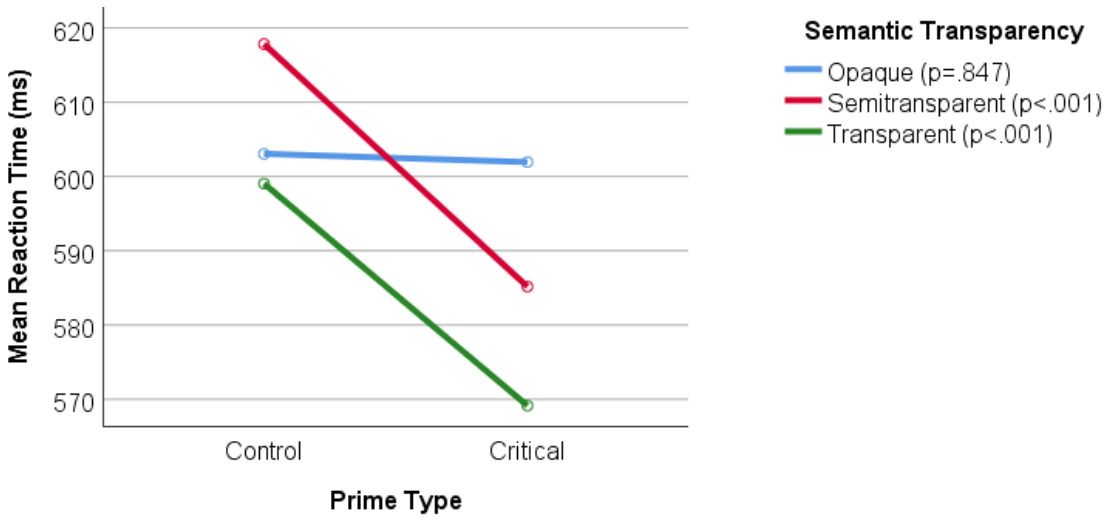
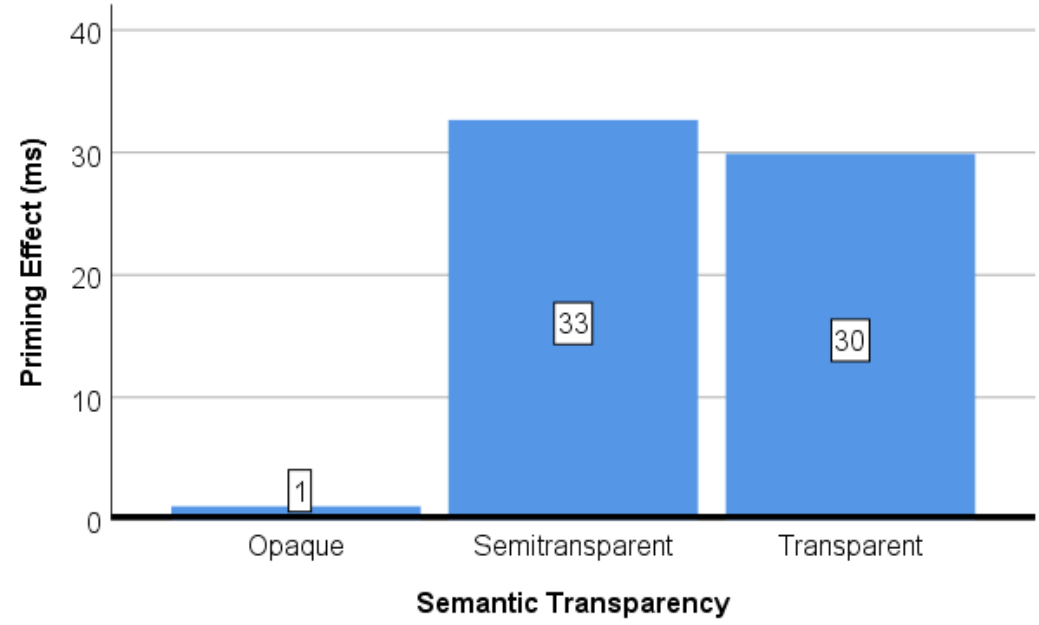


Figure 1: Differences in mean reaction times between control and critical conditions



Transparent = Semitransparent (p=1)
Opaque < Semitransparent (p<.001)
Opaque < Transparent (p=.001)

Figure 2: Differences in priming among semantic transparency conditions

Mary O'Brien, Sarah Macdonald and Gary Libben

Implicit and explicit awareness of morphemes in L2 German

The morphological system of German is highly productive (Clahsen, Sonnenstuhl & Blevins, 2003), and native speakers are able to create new words through compounding and derivation. Native speakers activate morphological information automatically (e.g., Smolka & Libben, 2017), and researchers have concluded that they decompose complex and compound words into their component parts (e.g., Clahsen et al., 2003). Relatively little research has looked at the extent to which second language (L2) learners of German rely on morphological information in their processing of L2 words, but studies investigating L2 morphological processing have generally concluded that L2 learners may rely more on lexical storage than on morphological parsing (e.g., Silva & Clahsen, 2008). In addition, these studies have demonstrated first language (L1, e.g., Alonso & Villegas, 2016) and proficiency effects (e.g., Coughlin & Tremblay, 2015). In the current study we aim to determine the extent to which native German speakers and L2 learners of German process the constituents of morphologically complex and compound words. We also investigate the effects of L1, L2 proficiency, and word type in L2 morphological processing.

Participants were 14 German native speakers tested in Germany and 28 adult L2 learners with intermediate to advanced proficiency in German, an equal number of whom were French and English native speakers. They completed a typing task with progressive demasking. That is, they saw a word that was slowly revealed and then typed it. Response times from the typing task provide information about where participants place morpheme boundaries in that they tend to slow down at boundaries. The 80 target items belonged to four categories, as in (1).

(1) Categories Example Translation
stem.stem.stem Roll.schuh.bahn roller skating rink
stem.suffix Vater.schaft fatherhood
prefix.stem.suffix Haupt.film.lein main short film
prefix.stem Nach.saison post season

Twenty simplex words (e.g., Hornisse ‘hornet’) served as a baseline. An additional task required participants to explicitly place morphological boundaries.

The results of the study point to similarities in the patterns of processing for native speakers and L2 learners. Both native and non-native participants demonstrated evidence of explicit and implicit morphological decomposition. Participants in all of the groups showed near-ceiling performance in the explicit segmentation task. In addition, typing times slowed at morphological boundaries and were affected by the complexity of words, such that simplex words took the least amount of time and tri-constituent compounds took the longest. We interpret these data patterns as supporting a view in which the processing of German words is morphologically informed and perhaps morphologically driven.

We discuss these results with respect to current models of L1 and L2 lexical processing and with respect to the methodological issues associated with comparing the performance of native speakers and L2 learners. Results indicate that lexical familiarity is a more valid measure than lexical frequency for L2 speakers and that the typing paradigm offers an effective measure of online lexical processing that enables both quantitative and qualitative comparisons of native speaker and L2 learner performance.

References

Alonso, J. G., & Villegas, J. (2016). English compound and non-compound processing in bilingual and multilingual speakers: Effects of dominance and sequential multilingualism. Second Language Research, 32(4), 503-535.

Coughlin, C., & Tremblay, A. (2015). Morphological decomposition in native and non-native French speakers. Bilingualism: Language and Cognition 18 (3), 524–542.

Clahsen, H., Sonnenstuhl, I., & Blevins, J. (2003). Derivational morphology in the German mental lexicon: A dual-mechanism account. In H. Baayen & R. Schreuder (Eds.), Morphological structure in language processing (pp. 125–155). Berlin: Mouton de Gruyter.

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. Bilingualism: Language and Cognition 11 (2), 245–260.

Smolka E., & Libben, G. (2017). “Can you wash off the hogwash?” – Semantic transparency of first and second constituents in the processing of German compounds. Language, Cognition and Neuroscience, 32(4), 514-531.

Kaarina Pitkänen-Heikilä

Typical foreign word patterns in Finnish special languages and their usage in the general language

Term formation has been studied in Finnish particularly with regard to the scholarly terminologies of the 19th century. At that time, linguistic purism influenced term formation and loan words were rarely used, but subsequently, special languages began to also use foreign elements in term formation.

Term formation is part of the onomasiological dimension of word formation but the word formation methods of special languages have not typically been described in Finnish word formation research (cf. Pitkänen-Heikkilä 2016). Typical term formation patterns in special languages, however, can also have an influence on word formation in the general language: for example, foreign prefixes have also increased in compounds in the general language as demonstrated by Tyysteri (2015). In the 19th century, new terms were mainly created by forming loan translations (Pitkänen 2013: 72–72), but in the 20th century, terms were borrowed and merely assimilated to Finnish structural patterns. Finnish word formation has adopted international influences particularly through special languages.

Common word patterns in special languages are, for example, neoclassical compounds, which have been studied, for instance, in English and German (e.g. Bauer 1983, Lüdeling et al. 2001). Closely related to neoclassical compounds are derivatives formed with prefixes. Finnish has traditionally only suffixal derivation and prefixes are foreign elements in Finnish words. In the 19th century, the creator of Finnish botanical vocabulary, Elias Lönnrot, translated foreign prefixes and produced repeating compound parts (e.g. *ob-* > *vasto-*, *gamo-* > *yhdis-*) that reflected botanical concept systems (Pitkänen-Heikkilä 2013: 77–79).

In my presentation I describe patterns of complex terms in scientific special languages in Finnish that include a foreign element, and examine their existence in the corpuses of the general language. The examined word patterns (e.g. *-anssi, -aalinen, -iivi, -iikka*; *anti-, de-, hyper-, mono-*) were obtained from the Helsinki term bank of arts and sciences (HTB, tieteentermipankki.fi) that includes approximately 52,000 head words in Finnish. From the general language corpuses, I used the newspaper and periodical corpus of the National Library of Finland between 1820 and 1930. My material shows that foreign word patterns – affixes or compound parts – are mainly borrowed as part of a complex loan word (e.g. *-anssi*; *performance* > *performanssi*), are understood as simplexes and have patterns that are not used in words with domestic roots. However, there are also some suffixes that are examples of indirect borrowing (e.g. *-al* > *-aalinen*) and some prefixes that can also be seen as examples of direct borrowing (e.g. *hyper-*) (see Seifart 2015: 511–513).

References

Bauer, Laurie 1983. *English word formation.* Cambridge University Press.

Lüdeling, Anke, Tanja Schmid, Sawwas Kiokpasoglou 2001. Neoclassical word formation in German. *Yearbook of Morphology*, 253–283.

Pitkänen-Heikkilä, Kaarina 2013. Term formation in a special language: How do words specify scientific concepts? Hacken ten, P. & Thomas, C. (eds.) *The Semantics of Word Formation and Lexicalization*. Edinburgh: Edinburgh University Press, 66–82.

– 2016. Finnish. – Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen, Franz Rainer (eds.), *Word-Formation. An International Handbook of the Languages of Europe*. Vol 5. 3209–3228. Mouton de Gruyter.

Seifart, Frank 2015. Direct and indirect affix borrowing. – *Language* 91(3), 511–532.
Tyysteri, Laura 2015. *Aamiaiskahvilasta ötökkätarjontaan. Suomen kirjoitetun yleiskielen morfosyntaktisten yhdyssanarakenteiden produktiivisuus.* Annales universitatis turkuensis 408. Turku.

What is the difference between *boys* and *boys'*? The phonetics of plural vs. genitive-plural in English and its implications for morphological theory

Recent research on the acoustic properties of morphologically complex words has shown unexpected effects of morphology on phonetic realization. For instance, it has been demonstrated that the phonetic properties of final /s/ and /z/ (henceforth: 'S') in English may differ systematically by morphological status (e.g. Zimmermann 2016, Plag et al. 2017, Seyfarth et al. 2017, Tomaschek et al. 2019). Such findings are unexpected since standard feed-forward theories of morpho-phonology (e.g. Lexical Phonology) and of speech production (e.g. Levelt et al. 1999) do not have a mechanism that would allow morphology to influence articulation, or that would model such behavior.

In this paper we test several hypotheses concerning the potential durational contrast between plurals and genitive-plurals (as in *boys* vs. *boys'*), which is a contrast about which existing studies are silent or inconclusive. Standardly, it is assumed that the two forms do not show systematic phonological or phonetic differences (cf., for example, Zwicky 1975, Bauer et al. 2013: 145, Palmer et al. 2002: 1595).

We report the results of an experiment in which 82 participants read aloud sentences which contained twelve different word pairs (plural vs. genitive-plural forms, such as *boys* vs. *boys'*) in very similar contexts. 462 plural tokens and 417 genitive-plural tokens were phonetically annotated, and the duration of S as well as the duration of the whole word were analyzed using mixed effects regression models with pertinent co-variates (e.g. speech rate, voicing, lemma frequency). The results show that plural S is significantly shorter than genitive-plural S, with a mean difference of 7 to 8 ms (as predicted by different regression models). The duration effect is, however, not restricted to the final S, but extends over the whole word, with (monosyllabic) plural nouns being 14 ms shorter than genitive-singular nouns.

The statistical analysis reveals that word-form frequency is predictive of word duration and S duration, with word-forms of higher frequency showing shorter durations (see Figure 1 for word duration). Since genitive-plural forms are generally of lower frequency than their corresponding plural forms (see Figure 2), and since the two forms systematically covary in word-duration (see Figure 3), the general durational difference between plural and genitive-plural can be interpreted as a word-form frequency effect. This is in line with recent studies that have found that the frequency of inflected word-forms is predictive of word durations (Caselli et al. 2016, Lõo et al. 2018).

The word-form frequency effect on both word duration and S duration can be naturally accounted for in word-and-paradigm models of morphology (e.g. Blevins 2016), in which individual word-forms may have representations in a network of morphologically related forms (see Tomaschek et al. 2019, Baayen et al. 2019 for a modern interpretation of such paradigmatic effects in a dynamic system).

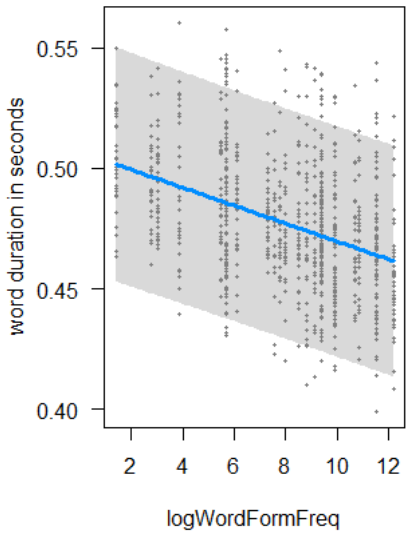


Figure 1: Word duration by word-form frequency (logged)

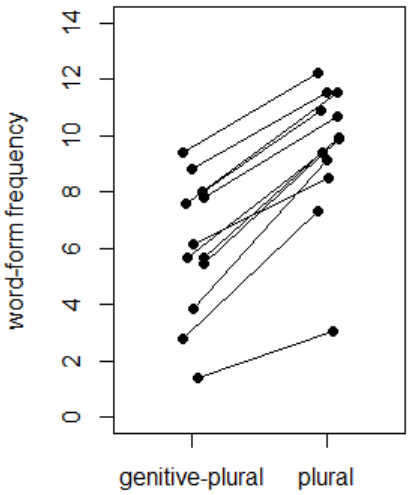


Figure 2: Word-form frequency (logged) by morphological category. Pairs of dots represent the plural and genitive-plural forms of one lexeme

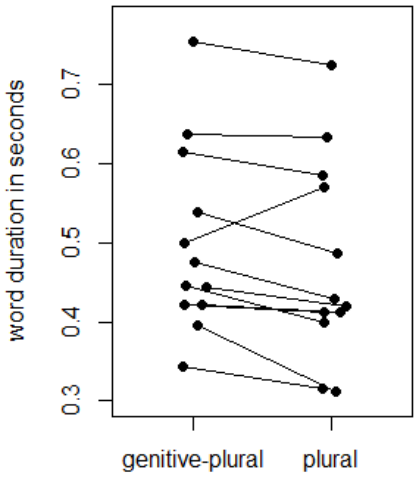


Figure 3: Mean word duration by morphological category. Pairs of dots represent the plural and genitive-plural forms of one lexeme

References

Baayen, R. H., Chuang, Y., Shafaei-Bajestan, E. & Blevins, J.P. (2019). The Discriminative Lexicon: A Unified Computational Model for the Lexicon and Lexical Processing in Comprehension and Production Grounded Not in (De)Composition but in Linear Discriminative Learning. *Complexity*, 2019(1), 1-39.

Bauer, L., Lieber, R. & Plag, I. (2013). *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford: Oxford University Press.

Caselli, N. K., Caselli, M. K. & Cohen-Goldberg, A. M. (2016). Inflected words in production: Evidence for a morphologically rich lexicon. *Quarterly Journal of Experimental Psychology*, 69(3), 432-454.

Levelt, W. J., Roelofs, A. & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–38.

Lõo, K., Järvikivi, J., Tomaschek, F., Tucker, B. V. & Baayen, R. H. (2018). Production of Estonian case-inflected nouns shows whole-word frequency and paradigmatic effects. *Morphology*, 28(1), 71–97.

Palmer, F., Huddleston, R. & Pullum, G. K. (2002). Inflectional morphology and related matters. In R. D. Huddleston & G. K. Pullum, eds., *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press, pp. 1565–1619.

Plag, I., Homann, J. & Kunter, G. (2017). Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics*, 53(1), 181–216.

Seyfarth, S., Garellek, M., Gillingham, G., Ackerman, F. & Malouf, R. (2018). Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience*, 33(1), 32–49.

Tomaschek, F., Plag, I., Ernestus, M. & Baayen, R.H. (2019). Modeling the duration of word-final S in English with Naive Discriminative Learning. To appear in *Journal of Linguistics*.

Zimmermann, J. (2016). Morphological status and acoustic realization: Findings from New Zealand English. In C. Carignan & M. D. Tyler, eds., *Proceedings of the 16th Australasian International Conference on Speech Science and Technology*. Sydney: University of Western Sydney, pp 6-9.

Zwicky, A. M. (1975). Settling on an underlying form: The English inflectional endings. In D. Cohen & J. R. Wirth, eds., *Testing linguistic hypotheses*. New York: Wiley, pp. 129–185.

Understanding *out*-prefixation: merging qualitative and distributional analyses (oral or poster)

We show interpretational patterns behind English *out*-prefixation to rely on property dimensions provided by base verbs within the limits of their verb-class. Applying distributional analyses, we further demonstrate that distributional semantics measures reflect these distinctions and allow us to isolate relevant verb-class and lemma-related characteristics.

Most authors argue that scalar-comparative *out*-verbs, as in *to outrun someone*, rely on underlying sameness of the events that Subject- and Object-argument participate in (Baker 2019; Talmy 2000). This alleged constraint is reminiscent of what the semantics literature calls incommensurability in comparative constructions with two distinct, overt adjectives (Doetjes 2010; Kennedy 1999). For *out*-prefixed derivatives, this sameness condition is taken to explain the oddness of examples such as (1):

(1) ??I outplayed the singer. [In the sense that I played better than the singer sang; marked as ungrammatical by Talmy 2000: 260]

As shown by Kotowski (2019) for *out*-verbs, comparing two events along some dimension of a property scale is possible even in cases in which the arguments do not participate in the same event types. In (2a-b), for example, FLYING- and RUNNING- and RAPPING- and BEAT-EMISSION-events, respectively, are compared regarding the dimension SPEED. Arguably, event-similarity for these items exists on a higher level, in the domains of LOCOMOTION in (2a) and SOUND-EMISSION in (2b):

(2) a. “I wasn't going to run,” Mr. Paxton said later after the game. “I figured I'm not going to **outrun an eagle**, so we might as well just see what happens.” (forbes.com)
b. The flow is a slow drawl to fit with the song and I like that you're not **outrapping the beat**. (iWeb)

Analyzing such attested examples, this study provides evidence for the existence of higher-level generalizations that facilitate comparison with *out*-verbs. An iWeb-corpus study was conducted (Davies 2018), using *out*-verb types with all possible bases from five different VerbNet classes (Kipper et al. 2008) and 100 tokens each of ten *out*-lemmas from these classes. Looking at concrete contextual information in attested examples, we show that all verb-classes show clear-cut preferences regarding the dimensions they base comparisons on. Divergent behavior of individual lemmas can be explained by lemma-specific characteristics as well as by cross-listing in different VerbNet-classes. For example, *out*-verbs with bases from the PERFORMANCE-class show an overall bias for the QUALITY-dimension, while its member OUTSING shows a preference for LOUDNESS. This discrepancy is reflected in the cross-listing of *to sing* as a SOUND-EMISSION-verb.

Using distributional semantics measures extracted from ukWaC (Ferraresi et al. 2008), we provide evidence that similarity-based hierarchical clustering of *out*-bases within their VerbNet classes is intricately connected to the interpretational preferences of *out*-derivatives. For example, the distribution of property dimensions across verbs within the EXIST-class is mirrored in their clustering. In addition, interpretational outliers like *outrap* are mirrored in the distinct positioning of its base in the PERFORMANCE-cluster. Comparing base-to-base to derivative-to-derivative similarity, we find that *out*-prefixation makes resulting derivations less similar based on two contrary processes. One source are strongly lexicalized derivations, e.g. *outlive* and *outstay* with the latter’s large idiomatic component via the ‘*outstay* POSS *welcome*’ construction. This explanation fails for other productive derivations, e.g. *outrun* and *outfly*.

These are semantically transparent in that they predominantly exploit the SPEED-dimension. Since resulting derivations rely on one concrete interpretation of the base verb, they are overall more specific, limiting the number of overlapping contexts.

Our findings underline the importance of semantics in understanding derivational patterns, in particular semantic narrowing of base interpretations, and show how distributional semantics can support and refine qualitative analyses.

REFERENCES

Baker, J. 2019. Split intransitivity in English. *English Language and Linguistics* 23(3). 557-589.

Davies, M. 2018-. The 14 Billion Word iWeb Corpus. Available online at https://corpus.byu.edu/iWeb/.

Doetjes, J.S. 2010. Incommensurability. In M. Aloni, H. Bastiaanse, T. de Jager, K. Schultz (eds.). *Logic, Language and Meaning, 17th Amsterdam Colloquium, Revised Selected Papers*, 254-263. Berlin: Springer.

Ferraresi, A., E. Zanchetta, M. Baroni, and S. Bernardini. 2008. Introducing and evaluating ukWaC, a very large web-derived corpus of English. In *Proceedings of the WAC4 Workshop at LREC 2008*, Marrakech. ELRA.

Kennedy, C. 1999. *Projecting the adjective. The syntax and semantics of gradability and comparison*. Garland: New York.

Kipper, K., A. Korhonen, N. Ryant & M. Palmer. 2008. A large-scale classification of English verbs. *Language Resources and Evaluation* 42(1). 21–40.

Kotowski, S. 2019. The semantics of English *out*-prefixation. Submitted to *English Language and Linguistics*.

Talmy, L. 2000. *Toward a Cognitive Semantics. Vol. 2: Typology and Process in Concept Structuring*. Cambridge, MA: MIT Press.

The acoustics of word-final [z] in English: A comparison of pluralia-tantum and regular-plural nouns

The interest in the interface between morphology and phonetics has grown considerably over the last years. One example in this context is the acoustic realisation of different types of word-final *s* in English. Schwarzlose & Bradlow's (2001) and Seyfarth et al.'s (2018) analyses revealed the surprising result that affixal *s* is longer than non-affixal *s*. Plag et al. (2017), however, found the opposite pattern. The present paper aims at expanding this line of research by comparing the acoustic realisation of the final *s* in English regular-plural nouns (= RPN; e.g., *browsers*) and pluralia-tantum nouns (= PTN; e.g., *trousers*).

In a reading study, native speakers of British English will read test sentences containing one of 18 test items each.[1] The 18 test items represent nine pairs of RPN and PTN (see Table 1).

Table 1. Test items

RPN	**PTN**
yearnings	*earnings*
pods	*odds*
browsers	*trousers*
screens	*jeans*
freezers	*tweezers*
beers	*shears*
gongs	*tongs*
fires	*pliers*
toggles	*goggles*

For each PTN a comparable RPN was selected. That is, in each pair, both nouns end in the fricative [z], have the same rhyme in the final syllable (hence the same segment before the target fricative [z]), the same stress pattern and length (in number of syllables), are inanimate, embedded in the same sentence type and position, and followed by the same segments.

Further, the mean frequencies of the two groups were matched (RPN: 4.2 (SD: 4.7); PTN: 4.2 (SD: 5.4)). Moreover, we ensured that the PTN do not have an additional singular form and that, in the case of the RPN, the singular form was always more frequent than the respective plural form.

All speakers will read both members of a pair (to avoid possible speaker-related effects) and the two conditions will be counterbalanced. The influence of the independent variable TYPE OF NOUN (RPN vs. PTN) on the dependent variable ABSOLUTE DURATION [Z] will be investigated by taking the random variables SUBJECT and ITEM as well as further aspects, such as speech rate, into consideration. With the study described above we intend to test the following hypotheses:

> **Hypothesis 1**: (the null-hypothesis) There is no difference in the duration of [z] between the two conditions. A reason to expect this outcome is that the PTN behave like the RPN in controlling plural agreement.

[1] The data will be collected in the fall of 2019 and will be analysed right after that.

Hypothesis 2: The [z] in RPN is longer than in PTN. One reason for this is the complexity of the former. While RPN consist of a stem and a suffix, the singular forms of these nouns only have a stem. PTN, in turn, do not have a singular form at all, that is, they do not have a non-complex counterpart. An alternative reason to expect this potential outcome refers to the issue of segmentability (see, e.g., Hay 2007). Complex words of a high degree of segmentability, for instance, complex words whose frequency is lower than that of their stem, tend to have longer affixes than complex words of a low degree of segmentability. In our case, while the RPN are less frequent in the plural (i.e., in the complex form) than in the singular (i.e., as a stem), the PTN are more frequent in the plural (as they have a frequency of zero in the singular). Hence, the higher degree of segmentability could be another explanation for a longer [z] in the RPN than in the PTN.

Our results will be interpreted against the background of the interface between morphology and phonetics and the role of these domains in models of speech production. We are deliberately "pre-registering" our study with this abstract (Nosek et al. 2017).

References

Hay, Jennifer. 2007. The phonetics of 'un'. In Judith Munat (Ed.), Lexical creativity, texts and contexts (Studies in Functional and Structural Linguistics 58), 39−57. Amsterdam: John Benjamins.

Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven & David T. Mellor. 2017. The Preregistration Revolution. OSF Preprints. June 16. doi:10.1073/pnas.1708274114.

Plag, Ingo, Julia Homann & Gero Kunter. 2017. Homophony and morphology: The acoustics of word-final S in English. Journal of Linguistics 53. 181−216.

Schwarzlose, Rebecca & Ann R. Bradlow. 2001. What happens to segment durations at the end of a word? The Journal of the Acoustical Society of America 109. 2292.

Seyfarth, Scott, Marc Garellek, Gwendolyn Gillingham, Farrell Ackerman & Robert Malouf. 2018. Acoustic differences in morphologically-distinct homophones. Language, Cognition and Neuroscience 33(1). 32−49.

Dominic Schmitz, Ingo Plag and Dinah Baer-Henney

How real are acoustic differences between different types of final /s/ in English? Evidence from pseudowords

Recent research suggests that homophonous morphemes show systematic differences in their phonetic realization (e.g. Seyfarth et al. 2017, Plag et al. 2017). Such findings contradict basic assumptions of standard feed-forward theories of morphology-phonology interaction (e.g. Kiparsky 1982) in which morphological information is only available at the lexical level. All further phonological processes occur at the post-lexical level, which has no access to morphological information.

A good test case for this distinction is English, which has a number of bound {s} morphemes; plural, genitive, genitive plural, 3rd person singular, the clitics of *is* and *has*, and the pronoun *us*. Previous research on this question found durational differences in the realization of these types of {s}; however, there is no agreement on the nature of these differences. Experimental studies, for example Walsh & Parker (1983) and Seyfarth et al. (2017) for NAE, found non-morphemic realizations to be shorter than plural and 3rd person singular /s/. In contrast, corpus studies on NZE (Zimmermann 2016) and NAE (Plag et al. 2017, Tomaschek et al. 2019) find results of the opposite direction for unvoiced realizations: the duration of /s/ is longest in non-morphemic contexts, somewhat shorter with suffixes, and shortest in clitics. As the aforementioned experimental studies show several flaws, e.g. no use of proper statistical methods (Walsh & Parker 1983) or a lack of differentiation between voiced and unvoiced variants of {s} (Seyfarth et al. 2017), there is need for carefully controlled experimental data to shed more light on the realization of morphemic and non-morphemic {s}.

Previous studies have suffered from the potentially confounding effects of the lexical and contextual properties of the items under investigation, e.g. potential storage effects (e.g. Caselli et al. 2016). To address this concern, the present study uses pseudowords to study the phonetic properties of different types of {s}. We tested whether there are durational differences between non-morphemic, plural, and the *is*-clitic /s/. A production study with forty native speakers of Southern British English was carried out, adopting Berko-Gleason's (1958) pseudoword paradigm. Speakers produced almost 1500 pertinent forms in a sentence production task with carefully controlled stimuli.

Linear mixed effects regression analyses show two main results. First, significant differences in duration between the different types of /s/ are only found in targets followed by a pause. In this environment, non-morphemic /s/ is longest, plural /s/ is shorter, and the *is*-clitic /s/ is shortest. This pattern is the same as that of previous corpus studies, and differs from the previous experimental results.

The results can be interpreted as follows. Differences in duration are subtle and seem to be only strong enough to be clearly observable in environments where final segments are lengthened (as before a pause). Where the differences are observable, pseudowords behave like real words in conversational speech, that is as shown in the corpus studies mentioned above. This means that pseudowords are subject to the same paradigmatic and contextual effects that have been discerned by Tomaschek et al. (2019) for real words.

The present study is the first study to show differences in duration of types of /s/ by utilizing pseudowords. By this, we can show that durational differences of types of /s/ appear to be of a robust morphological nature rather than a by-product of confounding effects of storage.

Hence, morphological information must be accessible in later stages of speech production, calling for a revision of standard feed-forward theories of morphology-phonology interaction.

References

BERKO-GLEASON, J. 1958. The Child's Learning of English Morphology, *WORD*, 14:2-3, 150-177. doi: 10.1080/00437956.1958.11659661

CASELLI, N. K., M. K. CASELLI, & A. M. COHEN-GOLDBERG. 2016. Inflected words in production: Evidence for a morphologically rich lexicon. *Quarterly Journal of Experimental Psychology* 69 (3), 432-454.

DRAGER, K. 2011. Sociophonetic variation and the lemma. *Journal of Phonetics* 39 (4), 694-707.

GAHL, S. 2008. Time and thyme are not homophones: The effect of frequency on word durations in spontaneous speech. *Language* 84 (3), 474-496.

KIPARSKY, P. 1982. Lexical morphology and phonology. In In-Seak Yang (ed.). *Linguistics in the morning calm: Selected papers from SICOL*, 3-91. Seoul: Hanshin.

LEVELT, W. J. M., A. ROELOFS, & A. S. MEYER. 1999. A theory of lexical access in speech production. *Behavioral and Brain Science* 22 (1), 1-75.

PLAG, I., J. HOMANN & G. KUNTER. 2017. Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics* 53 (1), 181–216.

SEYFARTH, S., M. GARALLEK, G. GILLINGHAM, F. ACKERMANN, & R. MALOUF. 2017. Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience*, 1-18.

TOMASCHEK, F., I. PLAG, R. H. BAAYEN & M. ERNESTUS. 2019. Phonetic effects of morphology and context: Modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics*, 1–39. doi:10.1017/S0022226719000203

WALSH, T., & F. PARKER. 1983. The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics*. 11 (2), 201-206.

ZIMMERMANN, J. 2016. Morphological status and acoustic realization: Findings from NZE. In Carignan, Christopher and Michael D. Tyler (eds.), *Proceedings of the Sixteenth Australasian International Conference on Speech Science and Technology (SST-2016), Parramatta, Australia, 6–9 December 2016*. Canberra: ASSTA, 201-204.

Alexandra Soares Rodrigues

The impact of Portuguese on the allomorphy of Mirandese prefixation and circumfixation: phonological conditions

Mirandese is an Iberian-Romance language belonging to the Asturian-Leonese branch, spoken in the Northeast of Portugal, specifically in the territory known as Tierra de Miranda. Mirandese has about 10000 to 15000 speakers, according to the *Ethnologue* 2019. All the speakers are bilingual, since they also speak Portuguese. The language was subject to social ostracism until the 1990s, when a group of educated native speakers promoted it in collaboration with scholars. Although Mirandese has been a statutory language of provincial identity (Law no. 7/99) since 1999, it is an endangered language. Nowadays, it is used more specifically by educated speakers who aim to improve Mirandese as a language capable of conveying erudite culture. Some published translations, such as a number of books from the *Bible*, parts of the *Odyssey* and the complete epic work by Luís Vaz de Camões from 1572, *Os Lusíadas*, as well as some periodical texts in local newspapers, are examples of this purpose.

The raising of Mirandese's erudition and elitism makes it permeable to Portuguese, the other language of Mirandese bilingual speakers. Permeability occurs in lexical borrowings, as well as in morphological structures, specifically in word formation. The aim of this work is a) to describe the prefixation and circumfixation processes of Mirandese word formation; b) to analyse the permeability of Mirandese to Portuguese in word formation, specifically with regard to these affixation processes; and c) to identify the phonological conditions that model this permeability. To date, there are only two studies on Mirandese word formation (Bautista 2013 and Meirinho 2016). However, Bautista (2013) follows a traditional approach that identifies morphemes by their phonological realization, leaving aside the abstract identity of affixes that takes account of their functionality in mental patterns and allow for allomorphy. Meirinho (2016) studies the influence of translation on the lexical innovation of Mirandese. Although this study departs from Bautista's approach, the systematisation of the derivational processes is not complete, nor is it immersed in allomorphic phenomena and it does not present a study on its conditions. Thus, the study we present here is seminal.

The methodology we have employed is the following: since there are no corpora of Mirandese, we analysed traditional texts such as legends, songs, etc., and erudite texts resulting from translations. We also analysed the only two dictionaries available (Pires 2004 and Ferreira, A., Ferreira, J.P.). The results we obtained are as follows:

The Latin prefix IN-, bearing a semantic [illative] feature, which results from a grammaticalization of the preposition IN (Diez 1874: 394), had evolved in Mirandese into the prefix and the first constituent of a circumfix as AN- /ɐ̃/ (*antrançar* 'to braid'; *ampalhar* 'to stuff'; *ambelhecer* 'to grow old' ↔ *belho* 'old') and in Portuguese as EN- /ẽ/ (*entrançar* 'to braid'; *empalhar* 'to stuff'; *envelhecer* 'to grow old'↔*velho* 'old'). (Both in Portuguese and in Mirandese *-m-* vs. *-n-* at the end of the prefixes do not correspond to a phonological distinction.) In Portuguese, EN- displays allomorphy (*e-*/*i-* if the base starts with a nasal consonant (*emagrecer* 'to slim down'; *emigrar* 'to pack'↔ *mala* 'luggage'); *im-* for Latinate formations (*inseminar* 'to inseminate'); *en-* /ẽ/ for the remaining situations (Pereira 2007; 2016)). In Mirandese, the morpheme AN- is displayed autochthonously as *an-* /ɐ̃/ (*amalar* 'to pack'↔ *mala* 'luggage'; *amalhadar* 'to put the cattle in the shelter' ↔ *malhada* 'shelter for the cattle'). When a lexeme is borrowed from Portuguese, there is an oscillation between two forms: one that evidences a morphological analysis of the word (*anarbar* 'to annoy' ↔ *nierbo / nerbo* 'nerve'; *anmagracer* 'to slim down') and another one that manifests borrowing a whole word (Mirandese: *einerbar* 'to annoy' / Portuguese: *enervar* 'to annoy'; Mirandese: *eimagrecer* 'to slim down' / Portuguese: *emagrecer* 'to slim down'). In the first situation, the morphological identity of the prefix/first element of the circumfix is preserved and it suffers a phonological adaptation to Mirandese (/ɐ̃/). In the second situation, there is no identification of the prefix in Mirandese. There are only phonological changes that adapt the Portuguese word as a whole to the phonology of Mirandese, regardless of the identity of the prefix/first element of the circumfix. The phonology of Mirandese does not admit front vowels at the beginning of the word. This phonological condition is observable in borrowed non-prefixed words such as *eiquipaije* /ɐjkiˈpajʒ/ 'equipment' and *einigma* /ɐjˈnigmɐ/ 'enigma', borrowed from Portuguese words with an initial front vowel (*equipagem* /ekiˈpaʒɐ̃j̃/ and *enigma* /eˈnigmɐ/ and /iˈnigmɐ/).

Another interesting prefix is the result of the Latin negative IN- (Diez 1874: 404), which had developed in Mirandese into the prefix AN- (/ɐ̃/) and in Portuguese into IN- /ĩ/. In Mirandese, it displays allomorphy (*an-* /ɐ̃/ if the base starts with a [- sonorant] consonant (*anjusto* 'unfair'↔ *justo* /ˈʒuʃtu/ 'fair'); *ein-* /ɐjn/ if the base starts with a vowel (*einútele* 'useless'↔ *útil* 'useful'); *ei-* /ɐj/ if the base starts with a [+ sonorant] consonant

(*eimortal* ‘unmortal’ ↔ *mortal* ‘mortal’). In Portuguese, the allomorphy of IN- goes as follows: /ĩ/ if the base starts with a [- sonorant] consonant (*impossível* ‘impossible’; *injusto* ‘unfair’); /i/ if the base starts with a [+ sonorant] consonant (*imortal* ‘inmortal’, *ilegal* ‘illegal’, *irreal* ‘unreal’); /in/ if the base starts with a vowel (*inacessível* ‘inaccessible’). In Mirandese, the prefix’s allomorphy is preserved in the borrowed words (*einoumano* ‘unhuman’ / Portuguese *inumano*; *eilegalizar* ‘to outlaw’ / Portuguese *ilegalizar*; *ampenetrable* ‘impenetrable’ / Portuguese *impenetrável*). This fact may be interpreted as the prevalence of the morphological analysis of the word into its base and the prefix (*oumano* ‘human’; *legalizar* ‘to legalise’; *penetrable* ‘penetrable’), due to the semantic transparency of the prefix.

The Latin prefix DES- had developed in Mirandese as Ç-/Z-, manifesting allomorphy: *ç-* /s/ if the base starts with an unvoiced consonant (*çcherumar-se* ‘to weaken’; *z-* /z/ if the base starts with a voiced consonant (*znudo* ‘naked’) or a vowel (*zamburrar* ‘to untangle’). In Portuguese, it has developed into *des-* (*deshonra* ‘dishonour’). In erudite texts in Mirandese, the form *des-* also appears besides the above-mentioned autochthonous allomorphs (*dezalmado vs. zalmado* ‘heartless (literally: ‘soulless’; *dezamparado vs. zamparo* ‘forlorn’).

The conclusions of the study are: a) if the prefix/first element of a circumfix of the borrowed word has a phonological structure in Portuguese that Mirandese does not admit, phonological conditions remain above direct borrowing; b) if both semantic and phonological transparencies of the Mirandese prefix are high, the autochthonous allomorphs prevail; d) if the phonological transparency of the Mirandese prefix is lower than the correspondent in Portuguese, the Portuguese form will be introduced into Mirandese, as long as it obeys the phonological conditions.

References

Bautista. A. (2013). *El mirandés. Contexto y procesos de formación de palabras*. Tesis doctoral presentada a la Universidad Complutense de Madrid (ms.).

Diez F. (1874). *Grammaire des langues romanes*. Paris: A. Franck, 3rd ed.

Ferreira, A. & J. P. Ferreira. *Dicionário mirandés-portugués.* Unpublished.

Pereira, R. 2007. *Formação de Verbos em Português: Afixação Heterocategorial*. Müchen: Lincom Europa.

Pereira, R. 2016 Formação de Verbos. In, G. Rio-Torto; A. S. Rodrigues; R. Pereira; I. Pereira & S. Ribeiro, *Gramática Derivacional do Português* 297-355. Coimbra: Imprensa da Universidade de Coimbra.

Pires, M. ([2004] 2019[2]. *Pequeno vocabulário mirandês-português*. Miranda do Douro. Câmara Municipal de Miranda do Douro.

Simon David Stein and Ingo Plag

ABSTRACT – 19TH INTERNATIONAL MORPHOLOGY MEETING

Lexical storage and morphological segmentability effects on the production of English derivatives

Submission category 3: Either oral or poster

Keywords: morpho-phonetics, morphological segmentability, frequency, lexical storage, lexical retrieval, speech production, prosodic word

In how far is fine phonetic detail in the pronunciation of morphologically complex words affected by morphological segmentability? This question has received much attention in morpho-phonetics and psycholinguistics in recent years, as possible answers to this question have important implications for morphological processing, speech production, and lexical storage.

Morphological segmentability has standardly been operationalized by the lexical frequency of the target word relative to the lexical frequency of its base word (e.g. Hay 2001, 2007). For the relation of segmentability and morphological processing, these frequency measures are interpreted as follows: If acoustic detail varies with target word frequency only, it seems likely that both morphologically simplex and complex words are stored morphologically unanalyzed in the mental lexicon. If acoustic detail also varies with base frequency, however, morphemes must be stored separately and be compositionally assembled during production. Finally, if acoustic detail varies with relative frequency, both types of storage may exist, and which one is processed faster will depend on the degree of segmentability.

Previous studies found phonetic detail to vary with these three frequency measures (e.g. Caselli et al. 2015, Hay 2003, 2007, Plag and Ben Hedia 2018), but some of them also yielded null results for some affixes. For example, Hay (2003, 2007) found segmentability effects for the affixes *un-* and *-ly*, while Plag and Ben Hedia (2018) found effects for *dis-* and *un-*, but not for negative *in-*, locative *in-* and *-ly*. It is not clear to this day why the three frequency measures sometimes show phonetic effects and sometimes do not. One possibility is that frequency effects might surface differently for affixes that are more or less strongly prosodically integrated (Raffelsiefen 1999, 2007).

This study investigates potential effects of frequency and prosodic structure with the English derivational affixes *-ness*, *-less*, *pre-*, *-wise*, *-ize*, *-ation*, *dis-*, *un-*, and *in-*. Using corpus data from the AudioBNC (Coleman et al. 2012) and multiple linear regression, we first explore whether word frequency, base frequency, or relative frequency (segmentability) affect acoustic duration. We find that affixes differ regarding their sensitivity to the different frequency measures. For example, the duration of *-ation*-suffixed bases is affected by all three frequency measures, while the duration of *-less*-suffixed bases is not affected by any of the frequency measures. Second, additional analyses indicate that this variation does not pattern according to affix-specific prosodic structure, and that predictions for duration based on Raffelsiefen's (1999, 2007) prosodic account are not supported.

Our finding that duration is influenced by segmentability for at least some affixes implies that the morphology-phonology-phonetics interface is not adequately captured in models which traditionally do not allow for post-lexical access of morphological information, like stratal and feed-forward models (e.g. Kiparsky 1982, Levelt et al. 1999, Bermúdez-Otero 2018). However, only some affixes are sensitive to segmentability effects, and this is not dependent on prosodic structure. This suggests that neither segmentability nor prosody are satisfactory explanations for the influence of morphology on phonetic realization.

References

Bermúdez-Otero, Ricardo. 2018. Stratal Phonology. In S. J. Hannahs & Anna Bosch (eds.), *Routledge handbook of phonological theory*, 100–143. London: Routledge.

Caselli, Naomi K., Michael K. Caselli & Ariel M. Cohen-Goldberg. 2015. Inflected words in production: Evidence for a morphologically rich lexicon. *The Quarterly Journal of Experimental Psychology* 69.3: 432–454.

Coleman, John, Ladan Baghai-Ravary, John Pybus & Sergio Grau. 2012. *Audio BNC: The audio edition of the Spoken British National Corpus*. Phonetics Laboratory, University of Oxford. http://www.phon.ox.ac.uk/AudioBNC.

Hay, Jennifer. 2001. Lexical frequency in morphology: Is everything relative? *Linguistics* 39.6: 1041–1070.

Hay, Jennifer. 2003. *Causes and consequences of word structure*. New York, London: Routledge.

Hay, Jennifer. 2007. The phonetics of *un*. In Judith Munat (ed.), *Lexical creativity, texts and contexts*, 39–57. Amsterdam, Philadelphia: John Benjamins.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In In-Seok Yang (ed.), *Linguistics in the morning calm: Selected papers from SICOL*, 3–91. Seoul: Hanshin.

Levelt, William J. M., Ardi Roelofs & Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22.1: 1–38.

Plag, Ingo & Sonia Ben Hedia. 2018. The phonetics of newly derived words: Testing the effect of morphological segmentability on affix duration. In Sabine Arndt-Lappe, Angelika Braun, Claudine Moulin & Esme Winter-Froemel (eds.), *Expanding the lexicon: Linguistic innovation, morphological productivity, and ludicity*. Berlin, New York: de Gruyter Mouton.

Raffelsiefen, Renate. 1999. Diagnostics for prosodic words revisited: The case of historically prefixed words in English. In Tracy A. Hall & Ursula Kleinhenz (eds.), *Studies of the phonological word*. 133–201. Amsterdam, Philadelphia: Benjamins.

Raffelsiefen, Renate. 2007. Morphological word structure in English and Swedish: The evidence from prosody. In Geert Booij, Luca Ducceschi, Bernard Fradin, Ernesto Guevara, Angela Ralli & Sergio Scalise (eds.), *Online proceedings of the Fifth Mediterranean Morphology Meeting (MMM5)*, Fréjus, 15–18 September 2005, 209–268.

Thomas Stolz and Nataliya Levkovych

On the contact-induced emergence of grammatical gender

There is as yet no comprehensive account of the behavior of grammar gender (GG) under the conditions of language contact. A particularly intriguing question is whether an originally GG-less language can acquire GG via language contact. Field (2002: 192) counts the genesis of GG among the cross-linguistic rarities whereas Matras (2009: 174) argues that "[g]ender may also be introduced into a language along with borrowed forms." Gardani (2012: 77) even considers GG to be relatively easy to copy. GG as defined by Corbett (1991) manifests itself in agreement of noun (= controller) and attribute (= target). In Thomason's (2001: 71) borrowing scale the copying of agreement patterns is possible only in prolonged and intensive language contact situations. Aikhenvald (2000: 388) assumes that "[b]orrowing of an agreement system is extremely rare" but mentions two cases herself (Ayacucho Quechua and Ilokano). Stolz (2012: 94–104) discusses (potential) evidence of examples of GG-less languages copying GG from their partner in a given contact situation. The data are often inconclusive since the illustration of the phenomenon is restricted to isolated words or NPs so that it is not always possible to rule out codeswitching.

What is needed is a cross-linguistic inventory of all those cases which (presumably) attest to the genesis of GG in contact situations involving a donor language with GG and a replica language without GG. Especially telling are instances like (1)–(4). In these sentences, feminine GG is marked overtly on the target (bold): adjectives in (1)–(3) and definite article in (4). The controller (underlined: pronoun in (1), nouns in (2)–(4)) is neither a copy nor marked for gender. For Karaim (Csató 2001: 18), Yucatec Maya, Tetun Dili (Hajek 2006: 170–171), and (Correntinean) Guaraní alike, these feminine forms contrast with (originally) masculine forms to yield a binary paradigm.

(1) Karaim (Éva Csató p.c.) [donor: Russian]
Ol *e-d'i* *inteligentn-**a**.*
3SG COP-PAST.3SG intelligent-**F**
'She was intelligent.'

(2) Yucatec Maya (Chamoreau 2012: 84) [donor: Spanish]
*Bek'ech-it-**a*** *u* *y-íits'in.*
thin-DIM-**F** A3SG POS-younger_sister
'His/Her younger sister is slim.'

(3) Tetun Dili (Bible Tetun Dili, Hahuu / Jénesis 12:14) [donor: Portuguese]
ema *sira* *haree* *feto* *nee* *bonit-**a*** *lahalimar.*
person PL see woman DEM.PROX pretty-**F** really
'[…] the people saw this really pretty woman'.

(4) (Correntinean) Guaraní (Cerno 2010: 26) [donor: Spanish]
Ani *na* *re-'u* *mamíta* ***la*** *so'o!*
NEG.IMP PAR 2-eat mammy **DEF.F** meat
'Don't eat the meat, mammy!'

In the talk further evidence from e.g. (Lekeitio) Basque, (Erzya) Mordvin, Chamorro, etc. is analyzed. Three questions are paid special attention:

(a) Is GG-agreement possible without the participation of lexical copies in a given construction (i.e. do copied gender-markers attach to autochthonous stems)?
(b) Is GG-agreement copied only if massive borrowing of nouns and adjectives occurs?
(c) Does GG-agreement start with human nouns?

The talk marks the point of departure of a project which investigates the fate of GG in language contact situations. It contributes not only to the research program dedicated to GG but also to language contact studies in general.

Abbreviations: A = a-set, COP = copula, DEF = definite (article), DEM = demonstrative, DIM = diminutive, F = feminine, IMP = imperative, NEG = negative, PAR = modal particle, PL = plural, POS = possessive, PROX = proximal, SG = singular

References:

Aikhenvald, Alexandra Y. 2000. *Classifiers. A typology of noun categorization devices.* Oxford: Oxford University Press.

Bible – Tetun Dili Bible. *Genesis e o Novo Testamento em língua tétum Dili* [Genesis and the New testament in Tetum Dili]. Wycliffe Bible translations. https://bible.cloud/ebooks/pdf/TDTUBB/TDTUBB.pdf (accessed 11 June 2018).

Cerno, Leonardo. 2010. Spanish Articles in Correntinean Guaraní. A Comparison with Paraguayan Guaraní. *STUF – Language Typology and Universals* 63 (1), 20–38.

Chamoreau, Claudine. 2012. Spanish diminutive markers *-ito / -ita* in Mesoamerican languages. A challenge for acceptance of gender distinction. In: Vanhove, Martine et al. (eds.), *Morphologies in contact.* Berlin: Akademie, 71–90.

Corbett, Greville G. 1991. *Gender*. Cambridge, New York: Cambridge University Press.

Csató, Éva. 2001. Karaim. In: Stolz, Thomas (ed.), *Minor languages of Europe. A series of lectures at the University of Bremen, April–July 2000*. Bochum: Brockmeyer, 1–24.

Field, Fredric W. 2002. *Linguistic Borrowing in bilingual contexts*. Amsterdam, Philadelphia: John Benjamins.

Gardani, Francesco. 2012. Plural across infection and derivation, fusion and agglutination. In: Johanson, Lars & Robbeets, Martine (eds.), *Copies versus Cognates in Bound Morphology*. Leiden, Boston: Brill, 71–98.

Hajek, John. 2006. Language contact and convergence in East Timor: The case of Tetun Dili. In: Aikhenvald, Alexandra Y. & Dixon, R.M.W. (eds.), *Grammars in contact*. Oxford: Oxford University press, 163–178.

Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.

Stolz, Thomas. 2012. Survival in a niche. On gender-copy in Chamorro (and sundry languages). In: Vanhove, Martine et al. (eds.), *Morphologies in Contact*. Berlin: Akademie, 93–140.

Thomason, Sarah G. 2001. *Language contact. An introduction.* Washington/DC: Georgetown University Press.

Adam James Ross Tallman

The emergence of empty morphs in Chácobo (Pano) and Araona (Takana)

In this paper I consider two domains morphology which have eluded coherent description in two Bolivian Amazonian languages; (i) the apparently polyfunctional "accusative" case suffixes in Chácobo (Pano) which appear on subject pronouns clause-finally as in (compare (1c-d) with (1a-b)) (Córdoba et al 2012; Valenzuela and Iggesen 2007); (ii) the prefix e- in Araona (Takana) which appears as an extended exponent of posture verbs e-…-neti 'standing, present, habitual'; e- … -ani 'sitting, progressive'; e-… ^{m}ba^{n}de 'hanging, initiative'; e-…-ha 'lying, anterior' as in (2a-b) but is also described as marking narrative past tense in examples such as those in (2c-d) (Pitman 1980:46).

(1)a. ~~í~~-a mi tsáya=ki~~ɨ~~
1SG-ACC 2SG see=DEC:PAST
'You saw me.'

b. mí-a no tsáya=ki~~ɨ~~
2SG-ACC 1PL see=DEC:PAST
'We saw you.'

c. ~~í~~-a tsáya=ki mí-**a**
1SG-ACC see=DEC:NONPAST 2SG-EPEN
'You see me.'

d. mí-a tsáya=ki no-**a**
2SG-ACC see=DEC:NONPAST 1PL-EPEN
'We see you.'

(2)a. e-oló-neti
STAND/PRES-fall-STAND/PRES
'It is falling.'

b. e-oló-ani
SIT/PROG-fall-SIT/PROG
'It is falling / is going to fall'

c. (yama) **é-**^{m}ba
(1SG:ERG) EPEN-see
'(I) saw (it).'

d. (yama) **é**-kwe
(1SG:ERG) EPEN-cut_down
'(I) cut (it) down.'

I argue against treating the Chácobo pronominal case suffixes and the Araona e-prefix as morphemes with a bi-unique mapping from form to function across all of their occurences. Rather I show that in certain cases (in bold above) these formatives must be analyzed as epenthetic empty morphs that are inserted to satisfy phonological minimality constraints co-opted from constructions where they have a transparent semantic function (underlined above). Data for this study are based on 24 months of original fieldwork on Chácobo and 8 months of original fieldwork on Araona.

The first part of this paper considers the evidence for minimality constraints in Chácobo and Araona. In Chácobo (Pano) lexical verb roots and noun roots undergo lengthening (ka~ka: 'go') and apocope processes (ínaka~ína 'dog'; ʃĩnó ~*ʃĩ 'monkey') respectively that obey bimoraic minimality. In Araona, the distribution of apocope inside and outside of compounding constructions provides evidence for bisyllabic minimality (etí 'penis'; zotóti 'jaguar penis').

The second part of this paper considers the evidence that the forms in bold are empty morphs and not instances of homophony or polysemy. In Chácobo, the appearance of the epenthetic case formatives varies according to whether there is right adjacent clitic material, not according to semantic function as illustrated in (3) below.

(3)a. tsirí-ʂɨni mi=kato
laugh-ADVZ:HAB 2SG=NMLZ/REL
'You are a compulsive laugher'

b. tsirí-ʂɨni=ka mi-a
laugh-ADVZ:HAB=REL/NMLZ 2SG-EPEN
'You are a compulsive laugher.'

Furthermore the epenthetic case formatives do not all have the same form as the transparently accusative forms (e.g. no-kí '1 plural accusastive' no-a 'first plural epenthetic). Furthermore, other function words in Chácobo alternate between monomoraic and bimoraic forms as a function of their syntactic distribution just as the epenthetic forms do (e.g. =ʂó~=ʂóna 'same A prior'; =ka~=kato 'relativizer').

In Araona the epenthetic e- is only ever systematically associated with the past tense when combined with monosyllabic verb roots: otherwise e- only occurs where a posture verb also occurs (see (2a-b) above). As soon as any morphology is added to such monosyllabic verb roots, the e- prefix can no longer occur as in (4-b). Furthermore, bare

verb roots receive past tense interpretations as in (4c-d), suggesting that the apparent association between e- and the past is spurious.

(4)a. [m]ba-ta
see-3A
'S/he saw (him/her/it).'

b. kwé-ta
cut_down-3A
'S/he cut (it) down.'

c. náwi
bathe
'S/he bathed.'

d. séo
cut
'It was cut'

I argue that in both Chácobo and Araona the best synchronic analysis treats these formatives as epenthetic empty morphs that appear to satisfy phonological minimality rather than as a consequence of the realization of straightforward semantic functions. The third part of this paper provides a brief sketch of the diachronic emergence of such null morphs and the extent to which they illustrative the discriminative function of morphology and its relative autonomy with respect to syntax (see Booij 1997; Maiden 2005; Blevins 2016, *inter alia*).

References

Blevins, J. P. (2016). *Word and Paradigm Morphology.* Oxford: Oxford University Press.

Booij, G. (1997). Allomorphy and the Autonomy of Morphology. *Folia Linguistica, 31*, 25-56.

Córdoba, L., Valenzuela, P. M., & Villar, D. (2012). Pano meridional. In M. Crevels, & P. Muysken (Eds.), *Lenguas de Bolivia, Vol. 2, Amazonía* (pp. 27-69). La Paz: Plurales Editores.

Iggesen, O., & Valenzuela, P. M. (2007). El desarrollo de un marcador suprasegmental. In A. Romero-Figuero, A. Fernández Garay, & A. Corberi Mori (Eds.), *Lenguas indígenas de América del Sur. Estudios descriptivo-tipológicos y sus contribuciones para la linguística teórica* (pp. 187-199). Caracas: Universidad Catolica Andres Bello.

Maiden, M. (2005). Morphological autonomy and diachrony. In G. Booij, & J. v. Marle (Eds.), *Morphology Yearbook 2004* (pp. 137-175). Dordrecht: Springer.

Pitman, D. (1980). *Bosquejo de la Gramatica Araona.* Riberalta: Summer Institute of Linguistics.

Yuta Tatsumi

A compositional analysis of deverbal compounds in Japanese

Issues: Sugioka (2002) argues that Japanese deverbal compounds (henceforth VCs) can be classified into two types: direct argument VCs and adjunct VCs. (1a) is an example of direct argument VCs, and (1b) is an example of adjunct VCs.

(1) a. *Taro-ga tegami-kaki-o sita.*
Taro-NOM letter-write-ACC did
'Taro wrote a letter.'

b. *Taro-ga tegami-o pen-gaki sita.*
Taro-NOM letter-ACC pen-write did
'Taro wrote a letter with a pen.'

In (1a), the non-head element (i.e. *tegami*) is an internal argument of the base verb. In (1b), the non-head element (i.e. *pen*) is interpreted as a non-argument. One piece of evidence for this classification of Japanese VCs comes from *Rendaku* (sequential voicing). When a non-head item is an argument of the base verb, VCs do not show Rendaku, as in (2b) and (3b). In contrast, when a non-head item is a non-argument, VCs show Rendaku as in (2b) and (3b).

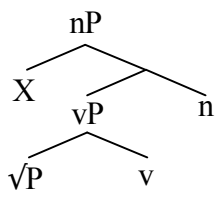
(2) a. *imo-hori* 'potato digging' b. *kikai-**bori*** 'digging by machine'

(3) a. *te-huki* 'hand towel' b. *kara-**buki*** 'a wipe with a dry cloth'

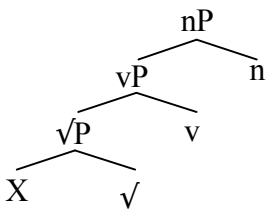
This difference between the two types of VCs has already been observed by Okumura (1955). Although there is a certain number of exceptions to this tendency, as has been pointed out by many researchers (Kindaichi 1976), there is no adjunct VCs without sequential voicing, as pointed out by Sugioka (2002). In other words, adjunct VCs must exhibit sequential voicing.

Proposal: I argue that Sugioka's insight is best analyzed in terms of the phase-based model of Distributed Morphology developed by Arad (2003). As for deverbal compounds with sequential voicing, I offer the structure (4a). In (4a), the non-head item (X) appears in Spec,nP. As for deverbal compounds without sequential voicing, I assume (4b), in which the non-head item (X) is directly combined with the root of the base verb.

(4) a.



b.



The proposed analysis can capture the patterns of sequential voicing in terms of the Phase Impenetrability Condition (Chomsky 2000, 2001). Following Arad (2003), I assume that both the verbalizer and the nominalizer are phase heads. In (4a), the dependent and the base verb are separated by an intervening phase head (n and v). There is no local compound in which sequence voicing rule applies in (4a). Recall that direct argument VCs optionally show sequence voicing,

whereas adjunct VCs obligatorily exhibit sequential voicing. Under the proposed analysis, this means that direct arguments can appear in both local and non-local positions of a base verb in deverbal compounds, while adjuncts are base-generated in the non-local position of a base verb. When a non-head item is an argument of the base verb, it can appear either in (4a) or in (4b). In (4b), the resulting VC exhibits Rendaku.

Support: The proposed analysis can account for a connection between sequential voicing and the verbal use of VCs. Some deverbal compounds can be used as verbs, as shown in (5b) and (6b).

(5) a. *kara-buri* b. *kara-buru*
empty-swing.CONJ empty-swing.PRES
'a swing and a miss' 'to swing and miss' [Adjunct VCs]

(6) a. *ne-biki* b. *ne-biku*
price-pull.CONJ price-pull.PRES
'a discount' 'to discount' [Argument VCs]

I observe that there is a relationship between sequential voicing and the verbal use of Deverbal compounds. The relevant generalization is given in (7).

(7) If a noun-verb Deverbal compound is used as a verb, the Deverbal compound shows sequential voicing unless it violates Motoori-Lyman's Law.

The proposed analysis can capture the generalization in (7). When VCs are used as a verb, it is taken as the complement of T. Since the T head selects a verbal element as its complement, a nominalizer cannot appear between a verbalizer and T. In other words, T cannot take the structure in (4a) as its complement. Under the verbal use, VCs should have the structure (4b), in which the non-head item appears in the local position of the root. VCs thus show sequential voicing in the verbal use.

Selected References: Arad, M., 2003. Locality constraints on the interpretation of roots: the case of Hebrew denominal verbs. *NLLT* 21, 737-778. Sugioka, Y., 2002. Incorporation vs. modification in deverbal compounds. In: Hoji, H. (Ed.) *JK* 10, CSLI, Stanford, pp. 496–509.

(692 words, 5030 characters including blanks and references)

Bogáta Timár and Jeremy Bradley

The expression of epistemic modality in the Volga-Kama Sprachbund

Epistemic modality indicates a speaker's evaluation (judgement/confidence/etc.) of a statement (van der Auwera & Ammann 2013). It differs from situational (or dynamic, deontic, see Nuyts 2012) modality which indicates that an action can or must take place due to internal or external factors such as basic human needs or social conventions, and evidentiality, the grammatical marking of an information source (Aikhenvald 2004) without an implied subjective evaluation (though the categories of evidentiality and epistemic modality are frequently subsumed in language surveys). Epistemic modality does not refer to factors allowing or restricting an action, but the speaker's degree of certainty in a statement (Nuyts 2016).

Epistemic modality can be expressed through different mechanisms (van der Auwera & Ammann 2013): verbal affixes, periphrastic verbal constructions, mental state predicates (e.g. *I think that …*), and lexical modifiers (e.g. *maybe*, Russian *naverno*). Usually languages utilize more than one strategy. Epistemic modality often overlaps with situational modality, as some markers may be used to express both. The paper at hand aims to study the expression of epistemic modality in the Uralic and Turkic languages of the Volga-Kama Sprachbund (cf. Helimski 2003: 15). These languages (Uralic: Mari, Udmurt, Komi, Erzya, Moksha; Turkic: Chuvash, Tatar, Bashkir) are well documented from a formal perspective (Riese et al. 2019, Johanson & Csató 1998, Kel'makov & Hännikäinen 2008, etc.) and markers of situational modality (possibility and necessity) – often auxiliary verbs – are easy to find in existing descriptions. Few reference materials, however, provide explicit information on the expression of epistemic modality, as the myriad mechanisms used to express it do not constitute a coherent category in a form-based language description.

The presentation at hand will focus on three questions:

1) How is epistemic modality expressed morphosyntactically in each language of the Volga-Kama Sprachbund?
2) Do strategies used epistemic modality overlap with strategies used to express situational modality in individual languages?
3) What similarities and differences cann be observed cross-linguistically in the region?

While the survey of the expression of epistemic modality in an individual language is a worthwhile endeavour, an areal cross-linguistic survey promises to be especially fruitful. Markers with the same basic function can be found in different languages, but their exact usage may differ. For example, in Meadow Mari, one of the most commonly used markers to express epistemic neccessity is the future necessitive participle *-šaš*, which is also used to express situational neccessity. A different periphrastic construction consisting of INF + *kül-* 'to be necessary' (the person needing to do something can be marked in the dative) on the other hand marks only situational necessity, but never epistemic necessity. In Udmurt however, the etymologically related structure INF + *kule* 'to be necessary' construction is used to express both situational and epistemic neccessity.

Speakers of any language will require means to express certainty or doubt in a statement, but due to the structurally diverse means used to do this, form-based reference materials are a suboptimal tool in this domain. A survey of this domain (extracting implicit information from

reference materials and utilizing corpora and native speaker consultations) seems worthwhile and can render information both on how individual languages in the area work, and how certain structural features and semantic shifts spread through the area.

References:

Bereczki, Gábor 1983. A Volga – Káma vidék nyelveinek areális kapcsolatai. in: Balázs János (szerk.) *Areális nyelvészeti tanulmányok*. Tankönyvkiadó. Budapest.

van der Auwera, Johan & Andreas Ammann . 2013. Overlap between Situational and Epistemic Modal Marking. In: Dryer, Matthew S. & Haspelmath, Martin (eds) *The World Atlas of Language Structures*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at wals.info/chapter/76, Accessed on 2018-08-01.)

Helimski, Eugene 2003. Areal groupings (Sprachbünde) within and across the borders of the Uralic language family: A survey, *Nyelvtudományi Közlemények* 100: 156–167. (Available online at www.nytud.hu/nyk/100/helimski.pdf, Accessed on 2018-09-13.)

Johanson, Lars & Éva Á. Csató (eds) 1998. *The Turkic Languages*. Routledge. London & New York.

Kel'makov, Valentin & Sara Hännikäinen 2008. *Udmurtin kielioppia ja harjoituksia*. Suomalais-Ugrilainen Seura. Helsinki.

Nuyts, Jan 2012. Modality: Overview and linguistic issues. In: Frawley, William (ed.) 2012. *The Expression of Modality*. The Expression of Cognitive Categories ECC 1. Mouton de Gruyter. New York.

Nuyts, Jan. & Johan van der Auwera (eds) 2016. *The Oxford Handbook of Modality and Mood*. Oxford University Press. Oxford.

Riese, Timothy; Bradley, Jeremy; Schötschel, Monika; Yefremova, Tatiana 2019. *Mari (марий йылме): An Essential Grammar for International Learners* [draft version]. University of Vienna. Vienna. [published online at grammar.mari-language.com]

Jelena Vujić

On N+N Nominalizations with the Invariable Modifying Constituent in Serbian

Unquestionably, the component of morphology which has received the least attention in regard to borrowing is compounding. (Gardani, 2018). In this talk, we address the N+N (M+H) nominalizations with the first invariable constituent in Serbian (1).

(1) *master student, master rad, auspuh servis, sendvič kifla, sendvič šunka, pica majstor, paradajz salata, fedora šešir, gej zajednica, tuna salata, fejsbuk prijatelj, čia semenke.*

(2) *đul-bašta, sa(h)at-kula, sikter-kafa, skandal-majstor, doboš-torta, šah-klub, veš-mašina, leptir-mašna, trač-partija, koncert-majstor.*

Traditional normative grammars of Serbian /Serbo-Croatian (Barić, 1980; Stevanović, 1986; Stanojčić & Popović, 1992; Vukićević, 1995; Klajn, 2002; Piper & Klajn, 2014) classify the examples as those in (2) as semi-compounds, two-word (hyphenated) constructions denoting a single unique concept with each constituent retaining its primary stress (Piper & Klajn, 2014: 249). The pattern (2) is a well-adapted and legitimate Serbian WF pattern which originates from Turkish. While nominalizations as those in (2) are acknowledged by both scholars and lexicographers, the (mostly) novel nominalizations like (1) are typically viewed as non-canonical lexicalized phrasal (syntactic) constructions which result from either a) poorly adapted calques of English N+N attributive compounds (e.g. sandwich bun > *sendvič kifla* instead of *kifla za sendvič* or tuna salad > *tuna salata* instead of *salata od tune*) or b) ungrammatical application of the English (N+N) compounding pattern in Serbian (e.g. *master rad* instead of *masterski rad* or *klima servis* instead of *servis za klima uređaje*). Therefore, despite their high type frequency which has been on the rise in recent years, novel N+N nominalizations with the first invariable constituent are (not quite favorably) regarded by Serbian linguists as a structural novelty imported from English (Prćić, 2005: 218) as part of so called "informal functional style" (Prćić, 2005:22).

The main tenet of this talk is that in spite of their orthographic differences, both novel formations (1) and listed semi-compounds (2) share the same properties and exhibit the same restrictions such as the following: i) morpho-phonological properties (e.g. a tendency to have monomorphemic, mono- or disyllabic lexemes as first constituents; both constituents retain their primary stress), ii) lexical properties (e.g. a tendency to have at least one of the constituents which is a loanword, with the prevailing number of examples where the first constituent is a loanword;) iii) semantic properties (e.g. the largest number of examined examples mean one of the following: N_2 is made from/with/of N_1; N_2 resembles N_1; N_2 is used for N_1) and iv) syntactic properties (argument structure). To support our hypothesis, our study offers the description of the two given groups of N+N nominalizations from morpho-phonological, lexico-semantic and syntactic aspect so as to compare and contrast them. The aim of this analysis is to point to striking similarities that exist between the fully adapted and domesticated WF pattern found in (2) and the recently imported English compounding pattern (1) so as to show that the much criticized borrowed English pattern is not that "foreign and disturbing" to Serbian WF system, but in fact leans against the existing domesticated pattern borrowed from Turkish. In both cases, N+N nominalizing pattern in Serbian affects eligible inputs, in other words, listed lexemes of the lexicon of Serbian. Therefore, we argue that both cases of nominalizations (1) and (2) are the instances of the same construction whose (borrowed) formation pattern can be represented with the single

general schema $[[a]_{Nx}[b]_{Ny}]_N$ _ N_y with relation R to N_x ' and on which speakers rely for (de)composition of N+N nominalizations with the invariable first constituent in Serbian.

Selected references:

Bach, E. (1983). On the relation between word-grammar and phrase-grammar. *Natural Language and Linguistic Theory I,* 65-89.

Babić, S. (1986). *Tvorba riječi u hrvatskome književnom jeziku*: *nacrt za gramatiku*. Razred za filologiju. Zagreb: Jugoslavenska akademija znanosti i umjetnosti.

Barić, E. (1980). *Imeničke složenice neprefiksalne i nesufiksalne tvorbe*. Zagreb: Liber.

Boyadzhieva, E. (2007). Reflections on a new word-formation pattern in Bulgarian newspaper language. In Alexandra Bagasheva (ed.), On Man and language. Papers in honor of Prof. Maya Pencheva on the occasion of her 60th anniversary, 232–244. Sofia: University Publishing House "St. Kliment Ohridski"

Giegerich, H. (1992). 'The limits of phonological derivation: Spelling pronunciations and schwa in English', *Linguistische-Berichte*, 1992, 413–43.

Giegerich, H. (2004). 'Compound or phrase? English noun-plus-nouns constructions and the stress criterion', *English Language and Linguistics*, 8, 1–24.

Klajn, I. (2002). *Tvorba reči u savremenom srpskom jeziku.* Prva knjiga. Beograd: Zavod za udžbenike i nastavna sredstva.

Lieber, R. & P. Štekauer, (2009). *The Oxford Handbook of Compounding.* Oxford: Oxford University Press.

Piper, P. & I. Klajn (2014). *Normativna gramatika srpskog jezika.* Novi Sad: Matica srpska.

Pižurica, M, M. Pešikan, J. Jerković (2017). *Pravopis srpskoga jezika.* Novi Sad: Matica Srpska.

Prćić, T. (2005). *Engleski u srpskom.* Novi Sad: Zmaj.

Rakić, S. (2005). „O strukturi složenica u srpskom jeziku", *Naučni sastanak slavista u Vukove dane, 35/1,* 367-376.

Rečnik srpskoga jezika.(2011). Novi Sad: Matica srpska.

Stanojčić, Ž. & Lj. Popović (1992). *Gramatika srpskog jezika.* Beograd/ Novi Sad: Zavod za udžbenike i nastvana sredstva/ Zavod za izdavanje udžbenika.

Stevanović, M. (1986). *Savremeni srpskohrvatski jezik.* Knjiga I (Uvod, fonetika, morgologija). Beograd: Narodna knjiga.

Trips, C. & J. Kornfilt (2017). *Further Investigations into the Nature of Phrasal Compounding.* Berlin: Language Science Press.

Vukićević, D. (1994). „O značenju imeničkih složenica", *Zbornik Matice srpske za filologiju i lingvistiku,* 37, 145–149.

Vukićević, D . (1995). „Imeničke složenice u savremenom srpskom književnom jeziku", *Zbornik Matice srpske za filologiju i lingvistiku,*, 38/1, 127–174.

Workshops

Implications of psycho-computational modelling for Morphological Theory

Marcello Ferro, Claudia Marzi, Vito Pirrelli
Institute for Computational Linguistics, National Research Council, Pisa, Italy

Goals

The workshop intends to offer an international forum for discussing interdisciplinary prospects of integration between advances in computer modelling of word knowledge and novel theoretical approaches to morphology.

Motivation and background

Following the advent of connectionism (Rumelhart & McClelland 1986), linguistic theory, cognitive models of human language processing and computer language modelling have increasingly been sharing research questions and goals. The rationale for this convergence was presciently epitomized in the early 1980s by Marr's (1980) hierarchy of levels of understanding of a complex processing system. Accordingly, linguistic theory approaches language issues mostly at Marrian level 1 ("what do speakers do when they use language?"), while cognitive psychology and computational linguistics are chiefly concerned with level 2 issues ("how do they use it?"), and neurosciences with level 3 questions ("how is such behavior implemented in the brain?"). Although David Marr originally introduced his hierarchy to emphasize that explanations at different levels can be investigated independently of each other, over the last 25 years there has been growing interest in the potential for between-level interaction, with a view to investigating the methodological conditions for their interdisciplinary unification. Advances in computer sciences and neuro-imaging technology have provided the level of material continuity between linguistic functions (level 1), algorithmic operations (level 2) and neuro-functional correlates (level 3) that is a necessary pre-condition to successful integration of neighboring disciplines along Marr's hierarchy (Alvargonzáles 2011).

This trend represents a challenge and an opportunity for Morphological Theory. Computer simulations can spawn novel explanatory paradigms. The idea that linguistic structure can emerge through self-organization of unstructured input is nowadays key to our understanding of language acquisition (Bybee & Hopper 2001;

Ellis & Larsen-Freeman 2006; MacWhinney 1999; MacWhinney & O'Grady 2015). Nonetheless, it had to await the challenging test of successful computer simulations before being given wide currency in the psycholinguistic (Baayen et al. 2011) and theoretical literature (Blevins 2016).

A recent reconceptualization of morphological generalization as the "Cell Filling Problem" (Ackerman & Malouf 2013) hinges on modelling the implicative structure of morphological paradigms through conditional entropy, an information-theoretic measure of inferential complexity that proves to correlate significantly with speakers' behavior (Ferro et al. 2018; Milin et al. 2009a, 2009b). The task is carried out successfully with either deep learning architectures (Malouf 2017; Cardillo et al. 2018) or linear mappings (Baayen et al. 2018), showing that multiple inferences from a set of paradigmatically-related forms can further reduce the complexity of inflection learning (Bonami & Beniamine 2017).

Time-honored approaches like analogy-based synchronic descriptions of language systems and historical accounts of language change got a new lease of life when analogical relations and their cognitive implications were successfully operationalized in the machine learning literature (Albright 2002, 2009; Albright & Hayes 2003; Daelemans & van den Bosch 2005; Keuleers et al. 2007; Pirrelli & Yvon 1999).

In addition, computer models prove to be instrumental in breaking traditional theoretical deadlocks. To illustrate, the categorical subdivision between regularly and irregularly inflected forms advocated by dual models of word processing (Pinker & Ullman 2002), as well as Hockett's (1954) distinction between Item-and-Arrangement and Item-and-Process approaches to morphology, both rest on the assumption that storage and processing are two independent functions of the human language faculty. This assumption, however, is challenged by integrative, connectionist models of short-term and long-term memories, implemented as two different temporal dynamics of the same underlying process (Marzi & Pirrelli 2015).

Statistical language modelling has recently been used to test competing theoretical frameworks on a quantitative basis. For example, statistical analyses and computer simulations of speakers' reaction times in visual word recognition challenged evidence of amorphous, holistic representations in the speakers' mental lexicon (Lignos & Gorman 2012; Oseki et al. 2019; Virpioja et al. 2018).

Advances in distributional semantics (Baroni & Lenci 2010; Padó & Lapata 2007) have thrown in sharp relief the role of lexical semantics in morphological processing, particularly for compounding and derivation (Marelli et al. 2017; Marelli & Baroni 2015; Günther & Marelli 2018), while helping draw a measurably graded distinction between derivation and inflection (Bonami & Paperno 2018).

The issues

This is the right time to take stock of the implications of current computational models of word processing for morphological theory. We hope that the range of issues raised by the workshop will advance our understanding of issues spanning the entire Marr's hierarchy, from theoretical aspects to neuro-functional ones. In particular, we invite authors to address and discuss the following questions:

What are the optimal representation units of human morphological competence and how are they acquired? What role do they play in the way speakers process and store words? Do speakers combine these units in a linear way, as in chaining Markov models, or rather structure them hierarchically, as suggested by the literature on sentence processing? Do they store them in their long-term lexical repository economically, or rather multiply them redundantly, as a function of their context and use? In addition, are these units represented as independent items, or are they mutually related as nodes in a network of paradigmatic relations? What is the contribution of lexical semantics to this picture, and what type of influence is exercised on lexical units by the communicative context where they are used referentially?

What is the status of the processes combining these units into larger units? Are they implemented by a single mechanism? Or should we rather hypothesize that more than one mechanism is in place? What evidence do we have of the anatomical and functional localization of different combinatorial mechanisms in the brain? And in what ways do their neural implementations differ? Given the mounting evidence that children learn words in chunks and that ready-made stretches of assorted words are committed to the long-term memory by speakers, what does this evidence tell us about the separation between Morphology and Syntax for language learning? Can computer modelling sharpen our current understanding of issues of morphological complexity

and their impact on lexical acquisition? What is its potential for modelling language learning, contact and change in multilingual contexts?

References

Ackerman, F. & Malouf, R. (2013). Morphological organization: the low conditional entropy conjecture. *Language* 89, 429–464.

Alvargonzález, D. (2011). Multidisciplinarity, Interdisciplinarity, Transdisciplinarity, and the Sciences. *International Studies in the Philosophy of Science* 25(4), 387-403.

Albright, A. (2009). Modeling analogy as probabilistic grammar. In Blevins, James P. and Blevins, Juliette (eds.), *Analogy in Grammar*, 185-213. Oxford: Oxford University Press.

Albright, A. (2002). Islands of reliability for regular morphology: Evidence from Italian. *Language* 78, 684-709.

Albright, A. & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90, 119–161.

Baayen, H., Chuang, Y.Y., & Blevins, J.P. (2018) Inflectional morphology with linear mappings. *The Mental Lexicon* 13(2), 230-268.

Baayen H., Milin, P., Đurđević, D.F., Hendrix, P. & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118(3), 438-481.

Baroni, M., & Lenci, A. (2010). Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics* 36(4), 673-721.

Blevins, J.P. (2016). *Word and paradigm morphology*. Oxford: Oxford University Press.

Bonami, O. & Paperno, D. (2018). Inflection vs. derivation in a distributional vector space, *Lingue e Linguaggio* XVII(2), 173-196.

Bybee, J. & Hopper, P.J. (eds.) (2001). *Frequency and the Emergence of Linguistic Structure*. John Benjamin Publishing Company.

Cardillo, A., Ferro, M., Marzi, M. & Vito Pirrelli (2018). Deep Learning of Inflection and the Cell-Filling Problem. *Italian Journal of Computational Linguistics* 4, 57-77.

Daelemans, W. and van de Bosch, A. (2005). *Memory-based Language Processing*. Cambridge, UK: Cambridge University Press.

Ellis, N.C. & Larsen-Freeman D. (2006). Language emergence: Implications for applied linguistics - Introduction to the special issue. *Applied Linguistics* 27, 558-589.

Ferro, M., Marzi, C. & Pirrelli, V. (2018). Discriminative word learning is sensitive to inflectional entropy. *Lingue e Linguaggio* XVII(2), 307-327.

Günther, F. & Marelli, M. (2018). Enter Sandman: Compound Processing and Semantic Transparency in a Compositional Perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition* http://dx.doi.org/10.1037/xlm0000677.

Hockett, C.F. (1954). Two models of grammatical description. *Word* 10, 210-231.

Keuleers, E., Sandra, D., Daelemans, W., Gillis, S., Durieux, G. & Martens, E. (2007). Dutch plural inflection: The exception that proves the analogy. *Cognitive Psychology*, 54, 283–318.

Lignos, C., & Gorman, K. (2012). Revisiting frequency and storage in morphological processing. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society* 48(1), 447-461. Chicago Linguistic Society.

MacWhinney, B. (ed.). (1999). *The emergence of language*. Lawrence Erlbaum Associates.

MacWhinney, B. & O'Grady, W. (eds.) (2015). *The Handbook of Language Emergence*. Wiley Blackwell.

Malouf, Robert (2017). Abstractive morphological learning with a recurrent neural network. *Morphology* 27(4), 431–458.

Marelli, M., Gagné, C. & Spalding, T. (2017). Compounding as Abstract Operation in Semantic Space: Investigating relational effects through a large-scale, data-driven computational model. *Cognition* 166, 207-224.

Marelli, M., & Baroni, M. (2015). Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological review* 122(3), 485.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Henry Holt, New York.

Marzi, C. & Pirrelli, V. (2015). A neuro-computational approach to understanding the mental lexicon. *Journal of Cognitive Science* 16(4): 493-534.

Milin, P., Đurđević, D.F., & del Prado Martín, F.M. (2009a). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language* 60(1), 50-64.

Milin, P., Kuperman, V., Kostić, A., & Baayen, R. H. (2009b). Words and paradigms bit by bit: An information theoretic approach to the processing of paradigmatic structure in inflection and derivation. In J.P. Blevins & Blevins, J. (Eds.) *Analogy in grammar: Form and acquisition*, 214-252. Oxford University Press.

Oseki, Y., Yang, C. & Marantz, A. 2019. Modeling Hierarchical Syntactic Structures in Morphological Processing. In *Proceedings of Cognitive Modeling and Computational Linguistics* (CMCL) 2019, Minneapolis.

Padó, S. & Lapata, M. (2007). Dependency-based construction of semantic space models. *Computational Linguistics* 33 (2): 161–199

Pirrelli, V. & Yvon, F. (1999). The hidden dimension: a paradigmatic view of data-driven NLP. *Journal of Experimental and Theoretical Artificial Intelligence* 11(3), 391-408.

Virpioja, S., Lehtonen, M., Hultén, A., Kivikari, H., Salmelin, R. & Lagus, K. (2018). Using statistical models of morphology in the search for optimal units of representation in the human mental lexicon. *Cognitive science* 42(3), 939-973.

The acquisition of Semitic morphology in Hebrew and Arabic: Developmental cross-modal analyses of corpora

Dorit Ravid, Elinor Salegh-Haddad, Lior Laks, Sharon Armon Lotem, Elitzur Dattner, Ronit Levie, Orit Ashkenazi, Elisheva Salmon and Osnat Kandelshine-Waldman

The acquisitional path of Semitic morphology has challenged researchers ever since the publication of McCarthy's 1981 paper on nonconcatenative morphology. Semitic morphology is synthetic / fusional, and can be characterized as 'rich' in several senses, with implications for acquisition (Ravid, 2003, 2012; Saiegh-Haddad, 2018). First, it encodes many semantic notions, both inflectional and derivational - in word-internal format. For example, a Hebrew verb such as hidlakt 'you, Fm lit' encodes the notions of lighting (root d-l-k), transitivity (verb pattern Hif'il), past tense (pattern vowels and suffix –t), second person, singular, feminine (suffix –t). Children growing up in morphology-oriented languages learn to seek meaning within the word. Second, Semitic morphology is rich in the systemic sense, as it uses at least two major structural systems to encode these notions: (1) the nonlinear (nonconcatenative) root and pattern device - e.g. Arabic kasar 'broke', inkasar 'broke (intransitive)', kassar 'broke, Tr into pieces'; and (2) the linear (concatenative) device (e.g. Arabic busta:n 'garden' / busta:n-ji 'gardener', Hebrew iton 'journal' / iton-ay 'journalist'). Children acquiring Semitic languages learn to think about their morphology in terms of a systematic, complex apparatus and to use morphological structures as pointers to word category and possible meaning. And finally, Semitic morphology involves many morpho-phonological changes within the word and the root. For example, under morphological operations, Hebrew noun stems undergo systematic changes such as vowel deletion or change, and stop / spirant alternation, as in iparon / efron-ot 'pencil / pencil-s'. Children growing up in a language where morphemes keep changing form, yet systematically retain the same meaning, learn to look for patterns of complex meaning / structure relationships.

Despite their common ancestry, Hebrew and Arabic differ along several dimensions, including inflectional and derivational systems. One example that comes to mind is the structure of noun plurals. Hebrew and Arabic also differ in the distance between the spoken and written versions of the language, which in Arabic takes the extreme form of diglossia, namely the existence of a spoken vernacular of Arabic as the language of everyday informal speech alongside Modern Standard Arabic. The linguistic distance between the two varieties of Arabic is evident in all areas of structure and usage, including lexicon, phonology, morphology and syntax, and this distance has been shown to impact language representation and processing in native Arabic speaking children and adolescents (Saiegh-Haddad & Spolsky, 2014).

In the last decade, new spoken and written corpora have been recorded, transcribed, coded and analyzed at Tel-Aviv and Bar-Ilan universities in Israel, yielding new insights on morphological structures and notions typical of Modern Israeli Hebrew, Palestinian Arabic, and Modern Standard Arabic in children, adolescents and adults. The proposed symposium includes 12 papers (two introductory papers and 10 studies) presenting new research focusing on the acquisition and development of Semitic morphology. For Hebrew, we present spoken corpora of approximately 500,000 words, including dyadic mother-child conversations and peer talk from age 2 to 12 years produced by native speakers from high and low socio-economic status. These analyses reveal developmental patterns of usage of verb inflection and derivation (roots, binyan patterns, temporal categories, and subject-verb agreement), as well as the acquisition of prepositions and their pronominal inflections. For Arabic, we present two corpora of approximately 50,000 words, consisting of spoken and written texts produced by native-speaking children and adolescents with typical and atypical development, as well as adults. These analyses reveal developmental patterns of usage of verb and noun inflection and derivation, focusing on roots, patterns and case marking, as

well as the interface of verbs and prepositions. The Arabic analyses underscore the notion of linguistic distance between spoken and written words and morphemes, as well as the role that distance has in morphological awareness.

In both corpora, morphological acquisition is shown to be facilitated by word and morpheme frequency, type frequency and salience of categories and systems, and is mediated by cognitive factors such as language impairment, socio-economic background, and the communicative setting of the discourse.

References

McCarthy, J. J. 1981. A prosodic theory of nonconcatenative morphology. Linguistic Inquiry, 26, 373-418.

Ravid, D. 2003. A developmental perspective on root perception in Hebrew and Palestinian Arabic. In Y. Shimron (Ed.), Language processing and acquisition in languages of Semitic, root-based morphology (pp. 293-319). Amsterdam: Benjamins.

Ravid, D. 2012. Spelling morphology: the psycholinguistics of Hebrew spelling. New York: Springer.

Saiegh-Haddad, E. 2018. MAWRID: A model of Arabic word reading in development. Journal of Learning Disabilities, 51, 454-462.

Saiegh-Haddad, E. & Spolsky, B. 2014. Acquiring literacy in a diglossic context: Problems and prospects. In Saiegh-Haddad, E. & Joshi, M. (Eds.), Handbook of Arabic Literacy: Insights and perspectives (pp. 225-240). Springer: Dordrecht.

Word formation in diachrony

Maria Silvia MICHELI & Giorgio Francesco ARCODIA

Deadline for submission: **30 September 2019**

Workshop description:

Word Formation (henceforth: WF) represents by now a well-established domain in morphological studies. Many inquiries have investigated the mechanisms exploited by languages to create new words (see Štekauer & Lieber 2006; Müller et al. 2015), especially compounding and derivation, according to different approaches, as e.g. cognitive approaches, onomasiological approaches (see Lieber & Štekauer 2009), constructionist approaches (Construction Morphology; see Booij 2015), etc. Most of the above have focused on the semantic and formal properties of derived words and/or compounds from a synchronic perspective (see e.g. Bauer 2017; Schäfer 2018 on compounding).

Conversely, less attention has been given to the diachrony of WF, although a diachronic perspective is crucial for highlighting many aspects of WF, e.g. productivity of affixes or compound constituents, changes affecting the formal and/or semantic dimension of morphologically complex words, competition between two or more WF strategies, etc. These topics have so far been addressed by a relatively limited number of studies, which however have shown that diachrony is a promising perspective to improve our understanding of how WF mechanisms emerge, develop, compete with each other, and also of how they disappear from use (see e.g. Hilpert 2013: 110-154 and Hüning 2019 on diachronic changes in WF).

Against this background, we invite authors to submit original, unpublished research papers on (but not limited to) the following topics of interest:

- The productivity of WF mechanisms from a diachronic perspective;
- Changes in WF (e.g. semantic or formal changes, cases of grammaticalization, etc.);
- Competition between WF mechanisms from a diachronic perspective.

Contributions on any language(s), language familie(s) and area(s), in any theoretical perspective, are equally welcome.

Please send an abstract of no more than 500 words (excl. references) to maria.micheli@unimib.it and giorgio.arcodia@unimib.it.

References

Bauer, L. (2017). *Compounds and compounding* (Vol. 155). Cambridge: Cambridge University Press.

Booij, G. (2015). Word formation in Construction Grammar. In Müller, P. O., Ohnheiser, I., Olsen, S., & Rainer, F. (Eds.). *Word Formation. An International Handbook of the Languages of Europe*. Berlin: De Gruyter Mouton, Volume 1, 188-202.

Hilpert, M. (2013). *Constructional change in English: Developments in allomorphy, word formation, and syntax*. Cambridge: Cambridge University Press.

Hüning, M. (2019). Morphological Theory and Diachronic Change. In Audring, J., & Masini, F. (Eds.). *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press.

Lieber, R., & Štekauer, P. (Eds.). (2009). *The Oxford handbook of compounding*. Oxford: Oxford University Press.

Lieber, R., & Štekauer, P. (Eds.). (2014). *The Oxford handbook of derivational morphology*. Oxford: Oxford University Press.

Müller, P. O., Ohnheiser, I., Olsen, S., & Rainer, F. (Eds.). (2015). *Word-formation: an international handbook of the languages of Europe* (Vol. 40). Berlin: Walter de Gruyter.

Schäfer, M. (2018). *The semantic transparency of English compound nouns*. Berlin: Language Science Press.

Štekauer, P., & Lieber, R. (Eds.). (2006). *Handbook of word-formation*. Cham: Springer.

Posters

Nino Amiridze

Borrowing feminine marking in Middle vs. Modern Georgian

(Oral or poster)

Grammatical gender is understood as only those classification devices that involve agreement [1]. To understand the dynamics of gender in contact situations, [9] deviates crucially from the definition of gender in [1] and adopts a broader perspective allowing to account for systems in which formal distinctions are established on nominals that do not trigger agreement. Such categories are labeled as *proto-genders* [9], suggesting that they might develop into genders which meet requirements of [1].
Additionally, languages, such as Tagalog (1), that acquired gender-based marking via contact and display gender-marked nominals triggering agreement are labeled as languages with *marginal gender* [9].

(1) Tagalog (donor: Spanish), [9, p. 99]
a. *Komik-a* *si* *Linda.*
funny-FEM DET Linda
"Linda is funny."
b. *Komik-o* *si* *Fred.*
funny-MASC DET Fred
"Fred is funny."

While following [9], I will compare borrowed feminine marking in two periods of Georgian (Kartvelian), a genderless language [4], under the influence of Greek and Russian, both possessing grammatical gender. Namely, I will discuss

- the feminine suffix *-a* in literary texts of Middle Georgian (XI–XVIII centuries) under the cultural influence of Old Greek [2], [3], and
- the borrowed feminine suffix *-k.a* in Modern spoken Georgian under the direct contact with Russian (XIX–XX centuries).

I will argue that, in both cases, Georgian represents a language with marginal gender, as it displays gender-marked nominals triggering agreement phenomena. Namely, in the case of Georgian-Greek contact, not only there was a *matter borrowing* [5], [8] of the Greek feminine suffix *-a* (cf. (2a) vs. (2b)) but, unlike Greek, a verb agreement marker for female referents has also been created (cf. (3a) vs. (3b)):

(2) Middle Georgian, [7, p. 360]
a. *moxucebul-a*
elderly.person-(Greek)FEM
"elderly female person"
b. *moxucebul-i*
elderly.person-NOM
"elderly male person"

(3) Middle Georgian, [7, pp. 497]

a. *movida-j*
came-FEM
"She came."

b. *movida*
came
"He came."

In Georgian-Russian language contact, the borrowed suffix *-k.a* (4) represents matter borrowing (cf. Russian feminine nominative ending *-k-a* [6, pp. 204]) and marks adjectives to agree with nouns in gender (5a):

(4) Modern Georgian

nagav-i	*da*	*nagav-k.a.*
trash-NOM	and	trash-(Russ.)FEM.NOM

"A male and a female trash." (a Facebook user about two politicians, 2010)

(5) Modern Georgian

a. From https://forum.ge/?showtopic=34844521&view=findpost&p=47329727

xat.mc.er-i-a	*es*	*debil-k.a*
icon.writer-NOM-COP	this.NOM	moron-(Russ.)FEM.NOM

gogo.
girl.NOM
"This moron girl is an icon painter."

b.

xat.mc.er-i-a	*es*	*debil-i*	*bič'-i.*
icon.writer-NOM-COP	this	moron-NOM	boy.NOM

"This moron boy is an icon painter."

Both cases of borrowing feminine marker characterize Georgian as a language with marginal gender. However, in the Georgian-Greek contact, the initiative to introduce female markers after Greek was mainly shared by writers but left no traces in standard Georgian or its dialects. On the other hand, the Georgian-Russian contact brought various contact phenomena to Georgian, including the borrowed *-k-a*, spread in the spoken language, while being condemned by normative grammarians and excluded from grammars and dictionaries.

The comparison of the two contact cases illustrates how in the absence of direct language contact between speech communities, attempts to make changes in language structure solely on cultural motivations fail. It is the bilingualism of a community that leads to contact-induced changes, rather than cultural and/or political motivations and efforts of language reformers.

References

[1] Corbett, G. (1991). *Gender*. Cambridge textbooks in linguistics. Cambridge University Press.
[2] Danelia, K. (1986). An attempt to express grammatical gender in Georgian. *Macne,* IV. (In Georgian).
[3] Dondua, K. D. (1967). Feminizing vowel in Georgian. In *Stat'i po obschemu i kavkazskomu jazykoznaniju*, pages 60–70. Nauka, Leningrad. (In Russian).
[4] Hewitt, B. G. (1995). *Georgian: A structural Reference Grammar*, volume 2 of London Oriental and African Language Library. John Benjamins, Amsterdam/Philadelphia.
[5] Matras, Y. and Sakel, J. (2007). Introduction. In Matras, Y. and Sakel, J., editors, *Grammatical borrowing in cross-linguistic perspective*, pages 1–13. Mouton de Gruyter, Berlin and New York.
[6] Musatov, V. N. (2010). *Russian Language. Morphemics, Morphonology, Word Formation*. Moscow, Nauka.
[7] Orbeliani, S.-S. (1991). *Georgian Dictionary*, volume I. Merani, Tbilisi. (Edited by E. Metreveli and Ts. Kurtsikidze, in Georgian).
[8] Sakel, J. (2007). Types of loan: Matter and pattern. In Matras, Y. and Sakel, J., editors, *Grammatical borrowing in crosslinguistic perspective*, pages 15–29. Mouton de Gruyter, Berlin and New York.
[9] Stolz, T. (2012). Survival in a niche. On gender-copy in Chamorro (and sundry languages). In Vanhove, M., Stolz, T., Urdze, A. S., and Otsuka, H., editors, *Morphologies in Contact* (Studia Typologica 10), Empirical Approaches to Linguistic Typology (EALT); 35, pages 93–140. Akademie Verlag, Berlin.

Sabine Arndt-Lappe and Melanie J. Bell

Linguistic experience shapes word-formation patterns - evidence from novel formations by native and non-native speakers of English

What mechanisms do speakers use to form a new word when there are several competing possibilities, with no default? Recent work on competition between rival word-formation processes has emphasised that speakers' choices are influenced by patterns among existing words in the lexicon (e.g. Rainer 2018, Bonami & Thuilier 2018, Arndt-Lappe 2014). Most of this work has used data from corpora or dictionaries, i.e. from resources that contain existing words but probably do not reflect what any individual speaker knows. However, if the lexicon is important, then language users will differ in their behaviour depending on the structure of their individual lexicons, which have been shaped by their individual linguistic experiences (cf. Dabrowska 2008). To test this prediction, we examined novel English word formations by two groups of speakers, native speakers of British English and German L2 learners of English, for whom we assume more within-group similarity than between-group similarity in terms of experience of English, and hence in the make-up of their English lexicons.

As the test case for our study, we used non-human patient nouns (NHPNs, Lieber 2015), which are deverbal nouns that can be paraphrased as 'something that is verbed'. According to Lieber (2015), there is no default process in English that derives NHPNs; instead, a variety of processes can be used, as exemplified in (1):

(1) *attachment* < attach, 'something that is attached'

additive < add, 'something that is added'.

We elicited 21 novel NHPNs from 39 native speakers of British English and 59 German learners of English, who differed in L2 exposure and proficiency. The task involved a gap filling exercise, in which verbal bases were presented in short contexts; participants were required to fill each gap with an appropriate derivative, whose meaning was inferable from the context, as shown in (2):

(2) Every time he read the poem out loud, he mangled the rhythm so badly that it hardly made any sense. But the _____ (mangle) themselves became famous, as true pieces of sound art.

The most common strategies used by both native and L2 speakers were *–ing* suffixation and conversion, reflecting general lexical patterns as gleaned from type frequency counts in standard corpora. Furthermore, both groups were sensitive to the length of the base word in terms of the number of syllables; for example, conversion often occurred with monosyllabic bases, but was significantly less frequent with simplex disyllabic bases. On the other hand, L2 learners and native speakers differed significantly in their usage of *–ment* and *–er*. The suffix *–ment* was used more often by learners than by native speakers (e.g. *manglement*, 'something that is mangled'), whereas the reverse was the case with *–er* (e.g. an *imbiber*, 'something that is imbibed'). Interestingly, these two forms also have different distributions across the registers of English. Whereas *–ment* is particularly type-frequent in written English, inanimate *–er* is particularly type-frequent in colloquial spoken English. We therefore interpret the observed differences between learners and native speakers as reflecting asymmetries in their experience with written and colloquial language, assuming that learners have proportionately more experience of the former. Finally, within the group of L2 learners, less proficient learners used a greater variety of strategies than more proficient learners. We take this to reflect differences in the amount of linguistic experience. All the relevant strategies are sufficiently frequent in standard corpora for us to assume that examples will have been encountered even by learners with relatively little exposure to English. However, we further assume that greater exposure is required to develop sensitivity to distributional differences between the strategies.

In sum, our findings provide evidence that the productivity of rival word-formation patterns for any given speaker is shaped by their linguistic experience. The findings challenge the view that word-formation rules are independent of usage once they have been acquired.

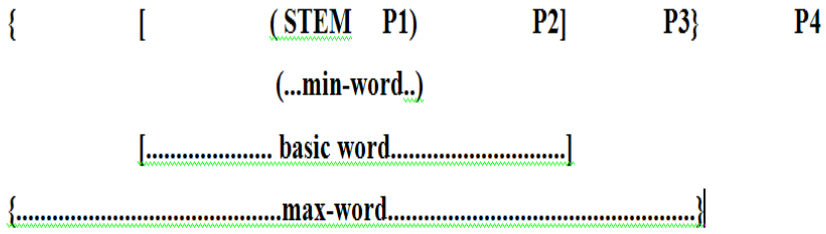
We depart from Basbøll's (e.g. 2005, 2014) general model of integration of suffixes into word structure, testing its potentials to account for derivational morphology in Danish, and contrasting it with a very comprehensive description of Danish word formation, viz. Rajnik 2011 (who, however, only considers nominal derivatives and does not treat phonological aspects).

The primary criterion of Basbøll's model asks whether the suffix is added to new words, and there are three possible answers: (a) as default; (b) only sometimes; or (c) not. Then two additional binary criteria are used: is the suffix added to a word (rather than just to an abstract stem)? and is the suffix signalled phonotactically as an ending? This gives, for the general model, five degrees of integration of suffixes.

Different languages (with suffixes) can grammaticalise differently, just like, e.g., number can be grammaticalised differently (with or without duals, etc.), and the vowel space can be phonologised differently (with a different number of phonemes, different boundaries between them, etc.). Danish grammaticalises so that the two extreme degrees in the general system are distinguished, but the three in the middle are not, amounting to the following three degrees of integration: *minimal integration*: suffixes that are both added as default, *and* added to a word, not just to an abstract stem; *maximal integration*: suffixes that are *not* added to new words, *and not* signalled phonotactically as an ending. This leaves a category with *medium integration* that is heterogeneous morphologically but not with respect to phonology (including prosody).

The three different degrees of integration of suffixes define three positions in word-structure: **P1, P2** and **P3** leading to three phonological domains (fig. 1) that are relevant for phonological rules, viz. vowel shortening before CC, approximant-drop and stød (a laryngeal syllable rhyme prosody, cf. Fischer-Jørgensen 1989, Grønnum & Basbøll 2007). Derivational and inflectional suffixes behave differently phonologically, even though they follow the same basic principles with respect to the model.

Figure 1. Word structure in Danish, based upon the four different positions derived from Basbøll's general model for integration of suffixes. P4 is the position for clitics, outside the max-word. Min-word, basic word and max-word are domains for different phonological (incl. prosodic) rules (from Basbøll 2009: 20)



The focus of the talk is derivational morphology, in particular: (1) similarities and differences between derivational and inflectional suffixes in Danish, and between derivation and compounding (cf. Kjærbæk & Basbøll 2017); (2) the role of lexicalisation for the distinction between derivation and inflection, and specification of the lexical information needed for derivational morphemes in

Danish; (3) the prosodic characteristics of derivational morphemes in Danish, and how we can account for the word prosody of Danish derivatives (stress and stød).

We shall end by discussing some acquisitional aspects of the two models treated here, viz. Basbøll's and Rajnik's, based upon our ongoing investigation of the development of derivatives in Danish-speaking children's spontaneous speech.

References

Basbøll, Hans (2005). *The Phonology of Danish*. Oxford: Oxford University Press.

Basbøll, Hans (2009) "Is nothing something, or isn't it anything? Zeroes in Danish prosodic morphology". In Fraser & Turner (eds.). *Language in life, and a life in language: Jacob Mey – a festschrift*. Oxford: Emerald, 19–24.

Basbøll, Hans (2014) "Danish stød as evidence for grammaticalization of suffixal positions in word structure". *Acta Linguistica Hafniensia* 46.2, 137–158.

Fischer-Jørgensen, Eli. 1989. "Phonetic analysis of the stød in standard Danish". *Phonetica* 46. 1–59.

Grønnum, Nina & Hans Basbøll. 2007. "Danish Stød: Phonological and Cognitive Issues". *Experimental Approaches to Phonology*. In Solé, Beddor & Ohala (eds.), 192–206. Oxford: Oxford University Press.

Kjærbæk, Laila & Hans Basbøll (2017). "Compound nouns in Danish child language". In Dressler, Ketrez & Kilani Schoch (eds.). *Nominal Compound acquisition*. John Benjamins, 39–62.

Rajnik, Eugeniusz (2011). *Wortbildung des Substantivs im Dänischen – explizite und implizite Derivation.* University Adam Mickiewicz.

Oral presentation preferred!

Laura Becker

When verbal markers go nominal: from IRREALIS to NONSPECIFIC and REALIS to SPECIFIC (oral or poster)

Overview While word class flexibility in the lexical domain is a widely discussed (though not uncontroversial) phenomenon (cf. Rijkhoff and Lier 2013), flexibility of grammatical markers is attested but has received less attention in the literature (other than the more evident cases of number or agreement markers that can be applied to both nouns and verbs). Rijkhoff (2008) mentions markers that can occur in both clauses and noun phrases, but does not discuss their development / diachronic relation further. In this talk, I propose two grammaticalization paths from mood to referential markers: IRREALIS > NONSPECIFIC and REALIS > DEFINITE/SPECIFIC. This development is typologically very rare and has a clear areal bias towards North America and Papunesia (especially but not exclusively in Siouan, Mayan, and Oceanic languages, 19 mostly related languages that I am aware of in total).

IRREALIS > NONSPECIFIC I argue that the extension of verbal irrealis markers to nominal nonspecific markers involves the following steps (cf. examples (1) to (3) from Q'anjobal, Mayan): ① the marker *-oq* on a verb marking an event as IRREALIS, e.g. in a conditional context (1); ② *-oq* occurs on a nominal predicate, allowing for an ambiguous reading between an IRREALIS predication and a NONSPECIFIC referential reading of the referent (2); ③ *-oq* occurs on the determiner of noun in argument position, marking the referent as NONSPECIFIC (3).

REALIS > DEFINITE/SPECIFIC The second path from realis to definite is similar but involves certain relative clause structures instead of nominal predicates, as shown with data from Hidatsa (Siouan) in (4) to (6). ① The emphatic declarative marker *-s(d)* on the verb emphasizes the realis status of the predication (4), ② its use in a headless relative clause allows for the ambiguous interpretation between a realis reading of the proposition in the clause or a definite / given reading of the referent of the semantic head of the relative clause (5), ③ the marker *-s(d)* is used with nouns in argument position without any clausal modifiers to indicate the definiteness of the referent.

Factors that favor the development of flexible grammatical markers All languages with this pattern share the following properties that can be argued to be relevant to the development in question: morphologically unmarked nominal predication, internally-headed or headless relative clauses, an elaborate determiner system, broad applicability of relative clauses (as arguments, conditional expressions, etc.) and of existential clauses (e.g. for the expression of possession).

(1) q-q'anjab' ayach ta q-ach q'anjab'-**oq**
POT-talk to.you COND POT-2PL talk-**IRR**
'X will talk to you, if you talk.' Q'anjobal (Mateo Toledo 2017: 538)

(2) man anima-**oq** hach
NEG person-**IRR** 2SG
'You are not a person.' Q'anjobal (Mateo Toledo 2017: 551)

(3) asi' yul [jun-**oq** tuktuk]
go.IMP in INDEF-**NONSPEC** tuktuk
'Let's take a (any) tuktuk.' Q'anjobal (primary data)

(4) hiirahbí-dhaa agá-waa-**sd**.
difficult-NEG suppose-1.CAUS.DIR=**EMPH.DECL**
'I didn't think it would be hard.' Hidatsa (Park 2012: 232)

(5) agu-aghá**à**ga-gsá-aci-**s** m-íì-hee.
REL-be.late-USI-COMPR-**EMPH.DECL** 1-PRO-EMPH
I'm the one who's always late. Hidatsa (Park 2012: 407)

(6) masúga-**s** adá**à**si-hgua núdhi-Ø.
dog-**DEF** outdoors-LOC tie-IMPER.SG
'Tie **the** dog up outside!' Hidatsa (Park 2012: 365)

Becker, Laura (2018). "Articles in the World's Languages". PhD thesis. University of Leipzig.
Mateo Toledo, Eladio (2017). "Q'anjob'al". In: *The Mayan Languages*. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. London: Routledge.
Park, Indrek (2012). "A Grammar of Hidatsa". PhD thesis. Bloomington, IN: Indiana University.
Rijkhoff, Jan (2008). "Synchronic and Diachronic Evidence for Parallels between Noun Phrases and Sentences". In: *Interdependence of Diachronic and Synchronic Analyses*. Ed. by Folke Josephson and Ingmar Söhrman. Studies in Language Companion Series 103. Amsterdam: Benjamins, pp. 13–42.
Rijkhoff, Jan and Eva Helena van Lier (2013). *Flexible Word Classes Typological Studies of Underspecified Parts of Speech*. Oxford: Oxford University Press.

Gilles Boyé

INFORMATION THEORY AND MORPHOLOGY: SOME CAVEATS

Inflectional morphology descriptions usually adopt a top-down perspective using, for example, a partition of the lexicon into more or less fine-grained inflectional classes and describing the different classes (e.g. Network Morphology: Corbett and Fraser 1993, Brown and Hippisley 2012 or Natural Morphology: Kilani-Schoch and Dressler 2005), or a set of stems for lexemes and rules of realizations for feature bundles (e.g. A-Morphous Morphology: Anderson 1992 or Paradigm Function Morphology: Stump 2001). With Blevins (2006) and Ackerman *et al.* (2009), a different type of description with an *abstractive* approach has appeared built around Information Theory (Shannon, 1948), word-based and revolving around the Paradigm Cell Filling Problem (PCFP) in (1).

(1) Given exposure to an inflected wordform of a novel lexeme, what licenses reliable inferences about the other wordforms in its inflectional family?

This has led to a new line of word-based descriptions hingeing on implicative relations between forms, for example the *dynamic principal parts* of Stump and Finkel (2013), or the *joint predictiveness* of Bonami and Beniamine (2015). In the wake of this movement, defining the complexity of inflectional systems has become a central question. The goal is to measure the difficulty of solving the PCFP, what Ackerman and Malouf (2015) call *integrative complexity*. Various answers have appeared some based on set-theoretic descriptions (e.g. Stump and Finkel 2013) or information-theoretic analysis (e.g. Sims and Parker 2016) but in the same way that Bonami (2014, pp97–98) showed that the entropy calculations proposed by Ackerman *et al.* (2009) were biased by not taking into account type frequency, we claim that this type of analyses should not consider all the paradigms of all known lexemes to calculate complexity but rather limit themselves to data readily available to speakers because computing the complexity of an inflectional system should measure the difficulty of filling the missing cells, not the average (or maximal) entropy between all cells.

In practice, frequency lexicons show that full paradigms are the exception rather than the rule. For example, in French, Lexique3 (New, 2006) lists 76348 forms of 6399 verb lexemes appearing in its reference corpus. Considering Lexique3 as a base sample for abstraction, it would contain only 25% of the forms in the full paradigms in general, and moreover only a handful of lexemes with complete exemplary paradigms (ÊTRE, AVOIR, FAIRE) defining inflection classes.

To construct an analysis on the complete dataset rather than on a ecological subset makes the strange prediction that speakers will be able to make the same generalizations whatever the frequency distributions are between lemmas and inflectional cells. To illustrate our point, we constructed a Lexique3 alternate by shuffling the original lexemic indices at random:

- every original lexeme was mapped to a different one
- every original paradigm was shifted to forms of the new lexeme

This manipulation did not change the number of lexeme and forms in the new lexicon, only the mapping. As a result, the size of the known lexicon compared to the full lexicon, the *lexical coverage*, did not change but the the number of cell to cell analogies present in the sample compared to the total analogies in the complete inflection class system, the *analogy coverage*, did change. The consequence is simple. The abstractable system is essentially the same except for the original high frequency, high complexity inflectional classes which become regularized or defective.

For example, the 50 forms of AVOIR ('to have') were mapped to the same 50 forms of FLÂNER ('to stroll') while, in turn, the isolated form of BRIFER ('to gobble'), the 2nd plural imperative /brife/, was mapped to AVOIR, /ɛje/. In this case, verbs of the first French conjugation class (same type as FLÂNER) benefited from an almost complete exemplary paradigm while AVOIR, belonging to an isolated micro-class of its own, in theory, was bundled with the first conjugation class as the unexisting AYER because none of the analogies necessary to build its original paradigm were available in the alternate lexicon.

This shows that integrative complexity cannot be computed by looking at the exhaustive data usually used by linguists to describe inflectional systems. We argue that it should at least take into account analogy coverage of ecological data to answer a somewhat modified version of the PCFP:

(1') Given exposure to ~~*an*~~ inflected wordform**s** of a ~~*novel*~~ lexeme, what licenses reliable inferences about the other wordforms in its inflectional family?

- its known wordform set
- the cell to cell analogies present in his ecological data

References

Ackerman, F. and Malouf, R. (2015). The no blur principle effects as an emergent property of language systems. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*.

Ackerman, F., Blevins, J. P., and Malouf, R. (2009). Parts and wholes: Implicative patterns in inflectional paradigms. In J. P. Blevins and J. Blevins, editors, *Analogy in grammar: Form and acquisition*, pages 54–82. Oxford Scholarship Online.

Anderson, S. R. (1992). *A-Morphous Morphology*. Number 62 in Cambridge Studies in Linguistics. Cambridge University Press.

Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics*, **42**, 531–573.

Bonami, O. (2014). *La structure fine des paradigmes de flexion : Études de morphologie descriptive, théorique et formelle*. Ph.D. thesis, Université Paris-Diderot.

Bonami, O. and Beniamine, S. (2015). Implicative structure and joint predictiveness. In V. Pirrelli, C. Marzi, and M. Ferro, editors, *Word Structure and Word Usage: Proceedings of the NetWordS Final Conference*.

Brown, D. and Hippisley, A. (2012). *Network morphology: A defaults-based theory of word structure*, volume 133 of *Cambridge Studies in Linguistics*. Cambridge University Press.

Corbett, G. G. and Fraser, N. M. (1993). Network morphology: a datr account of russian nominal inflection. *Journal of Linguistics*, **29**(01), 113–142.

Kilani-Schoch, M. and Dressler, W. U. (2005). *Morphologie naturelle et flexion du verbe français*. Gunter Narr Verlag, Tübingen.

New, B. (2006). Lexique 3: Une nouvelle base de données lexicales. In *Actes de la Conférence Traitement Automatique des Langues Naturelles (TALN 2006)*.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423 & 623–656.

Sims, A. D. and Parker, J. (2016). How inflection class systems work: On the informativity of implicative structure. *Word Structure*, **9**(2), 215–239.

Stump, G. and Finkel, R. A. (2013). *Morphological typology: From word to paradigm*, volume 138. Cambridge University Press.

Stump, G. T. (2001). *Inflectional Morphology*. Cambridge University Press.

Integration of English verbs in Italian: Competing Morphological Realizations

Author, University affiliation

The integration of English terms into other languages is a widely studied phenomenon, including in Romance linguistics (cf. Bombi 2009; Jansen 2005). While nouns can simply be borrowed from English without further adaptations, the integration of verbs is more complex and a number of linguistic strategies for the integration of verbs have been explored in the literature (cf. Wohlgemuth 2009). For Romance languages, they can be described as a continuum between two opposing poles (Wohlgemuth 2009, 160; Haugen 1950): On the one extreme, English verbs are integrated and adapted to the inflectional pattern of the target language (cf. 1a). On the other extreme, new verbs, which are semantically close to the original, but morphologically based on the target language, are created, as exemplified in (1b).

(1) a. *to format* > *formattare* (It)
b. *to scan* > *numériser* (F)

The situation is, however, seldom as straightforward as in (1a). In Italian, several of these linguistic strategies can be applied in the integration of a single English verb. The integration of English *to scan* is a prime example of an English verb leading to multiple morphological realizations in Italian, involving both conversion (2c-e) and derivation with different affixes (2a-b):

(2) a. *to scan* > *scanner+izz+are*
b. *to scan* > *?scann+eggi+are*
c. *to scan* > *scann+are*
d. *to scan* > *scanner+are*
e. *to scan* > *scansion+are*

The present paper proposes a case study to shed more light on the integration of English verbs in Italian and the competing morphological realizations. The verbs studied are *to scan*, *to google* and *to buffer*, which all have multiple morphological realizations in Italian. The aim of this paper is twofold: on the one hand, I examine the processes used in the integration of the three verbs and how the morphological variants are used by speakers. On the other hand, I investigate how the selection of integration processes is affected by frequency, semantics and morpho-phonological properties.

For this study, the three English verbs and their Italian counterparts are identified, analyzed and categorized according to the integration strategy. The analysis is informed partly by dictionary entries (Cannella and Lazzarini 2018; Istituto dell'Enciclopedia Italiana 2018) and partly by existing analyses of the verbs in question (Dardano 2009). In a second step, a variety of corpora are consulted to study the usage of the morphological variants in terms of frequency across different genres. The corpora consulted are the Italian TenTen 2016 corpus (Jakubíček et al. 2013), the Timestamped JSI net corpus 2014–2019 (Bušta et al. 2017) and the Italian component of the Europarl7 corpus (Koehn 2005). In a last step, the integration strategies applied will be compared with general frequency, semantic and morphophonological properties of the corresponding word formation processes.

This study aims to shed light on how different morphological realizations can occur in the process of integration and which factors influence the creation of these competing forms. In doing so this paper contributes to studying how English verbs are integrated into Italian, and to studying integration methods in Romance languages.

References

Bombi, Raffaella. 2009. *La linguistica del contatto: tipologie di anglicismi nell'italiano contemporaneo e riflessi metalinguistici*. 2nd ed. Roma: Il calamo.

Bušta, Jan, Ondřej Hermann, Miloš Jakubíček, Simon Krek, and Novak Blaž. 2017. "JSI Newsfeed Corpus." In *9th International Corpus Linguistics Conference CL*. University of Birmingham.

Cannella, Mario, and Beata Lazzarini, eds. 2018. *Lo Zingarelli 2019. Vocabolario della lingua italiana*. 12th ed. Roma: Zanichelli.

Dardano, Maurizio. 2009. *Costruire le parole*. Bologna: Il Mulino.

Haugen, Einar. 1950. "The Analysis of Linguistic Borrowing." *Language* 26 (2): 210–31.

Istituto dell'Enciclopedia Italiana. 2018. "Neologismi | Treccani." Treccani Il portale del sapere. 2018. http://www.treccani.it/magazine/lingua_italiana/neologismi/searchNeologismi.jsp [10/06/2019].

Jakubíček, M., A Kilgarriff, V. Kovář, P Rychlý, and V. Suchomel. 2013. "The TenTen Corpus Family." In *7th International Corpus Linguistics Conference CL*, 125–27.

Jansen, Silke. 2005. *Sprachliches Lehngut im "world wide web": Neologismen in der französischen und spanischen Internetterminologie*. Tübinger Beiträge zur Linguistik ; 484. Tübingen: Narr.

Koehn, Phillip. 2005. "Europarl: A Parallel Corpus for Statistical Machine Translation." *MIT Summit*.

Wohlgemuth, Jan. 2009. *A Typology of Verbal Borrowings*. Berlin, Boston: De Gruyter Mouton.

ADOPTION OF RUSSIAN SUFFIXES BY THE CHUVASH LANGUAGE

At present, the Russian language has a strong influence on the Chuvash language in all levels of the language system. Thus, in phonetic terms, in Chuvash words, in an intervocal position, there is a vocalization of voiceless consonants to the level of the Russian voiced ones; the Chuvash speech is full of macaronisms; in syntax the Chuvash language more often uses free word order contrary to the natural property of Chuvash words to have a strict position in the sentence, synesis becomes important. The influence of the Russian language has also spread to the morphemic level.

One of the living morphological processes in the field of Chuvash-Russian contacts is the adoption of the Russian verb conjugation system by the Chuvash language in the present tense.

This process, in particular, is observed in children's speech and supported by their desire to learn Russian:

(1) *larahayu* 'I am sitting' < Chuv. *laratăp* 'I am sitting' + Russ. *-u;*

(2) *vulahayeš'* 'you are reading' < Chuv. *vulatăn* 'you are reading' + Russ. *-eš'*.

Thus, the entire paradigm of the Russian conjugation in all time forms is adopted, which implies a combination of Chuvash bases and Russian suffixes with the help of a complex combination of infixes.

Thus, the entire paradigm of the Russian conjugation in all time forms is mastered, which implies a combination of Chuvash bases and Russian suffixes with the help of a complex combination of infixes.

(3) *śiyu* < Chuv. *śiyetĕp / śiyep* 'I am eating' + Russ. *-u.*

The morphemic composition of the considered forms will look as follows:

Chuvash	Russian
(4) lar-ah-a<y>u sit<?>PRS<INF>1SG 'I am sitting'	(5) sižu sit-PRS.1SG 'I am sitting'
(6) śi<y>u eat<INF>1SG 'I am eating'	(7) kuša<y>u eat<INF>1SG 'I am eating'

In example (4) the root *lar-* 'sit' is saved from the Chuvash form, the combination *-ah* (apparently, asemantic and only formally coinciding with the form of amplification, cf: *kay-ah* 'leave already') is added to it, the indicator of the present tense is *-a-*. Interfix *y* and the ending *-u* are already Russian in origin. The boundary between the Chuvash and Russian elements in the word form takes place before the infix. In this example, the agglutinative essence of word existence inherent in the Chuvash language is preserved.

Example (6), built exclusively on the Russian model and excluding the agglutinative properties of the Chuvash language, is rather interesting. According to the system implemented in examples (1) and (2), the word *śi-* should be presented in the form **śi-eh-a<y>u.* The departure from syngarmonism, an important principle of word design in the Chuvash language, is unexpected in the reconstructed example.

The processes mentioned above are marginal.

Borrowing of Russian word-formation suffixes is in the same vein, cf.:

(8) *-izm: šuχăšlavizm* 'aphorism' (*šuχăšlav* 'thinking' + *-izm*);

(9) *-izaci: maskălizaci* 'mockery' (*măškăl* 'mockery + *-izaci*);

naχalizaci 'mockery' (*naχal* 'impudent'+ *-izaci*);

(10) *-skiy:* → *шупашкарскиӳ* 'owned by Cheboksary' (*Šupaškar* 'Cheboksary' + *-skiy*).

As a rule, such formations are used for comic effect.

Meanwhile, the Chuvash language has the case of systemic change of the negative form of verbs, which was formed under the influence of the contacting language: at a certain period of development it borrowed the Finno-Ugric negative particle *an*, which replaced the Turkic suffix -

м(as) in the initial (imperative) form, while the suffixical organization of the verbal forms remained traditionally Turkic, cf:

Basic form	Conjugation
(11) an yurat NEG love-IMP 'do not love'	(12) yurat-mas-t-ăp love-NEG-PRS-1SG 'I do not love'

Thus, the condition for borrowing Russian word-modifying suffixes by the Chuvash language is the developed Chuvash-Russian bilingualism. Adoption of Russian affixational morphemes is a living process that allows variability. It is subject to the basic principles of Chuvash grammar - agglutination remains, but it refuses to comply with the principle of syngarmonism, an important factor in the phonetic design of Chuvash words.

Francesca Franzon, Alessia Zampieri and Davide Crepaldi

The role of semantics in learning morphological systems. An artificial lexicon experiment.

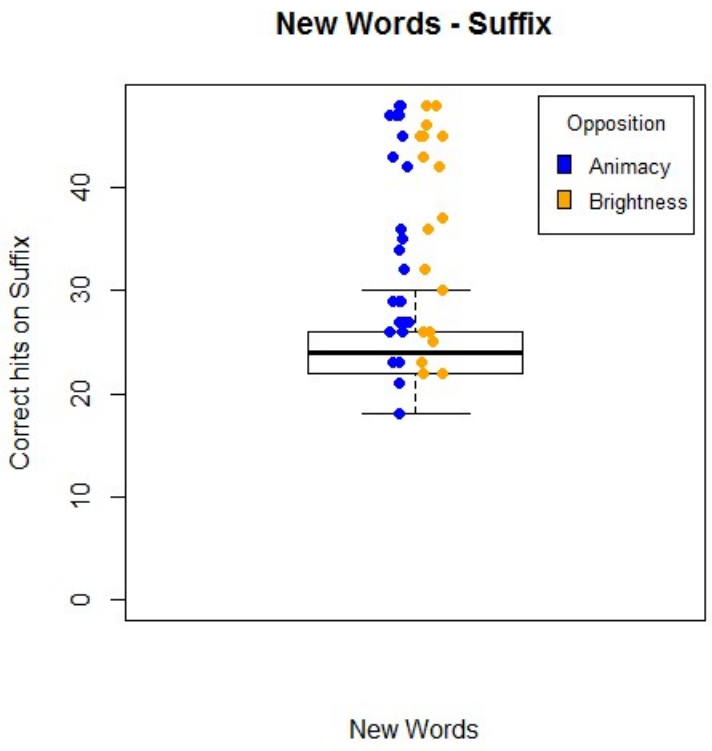
Introduction. Natural languages encode into morphology only a restricted and typologically consistent set of semantic features, such as numerosity and animacy (Corbett 1991, 2000). Only cognitively salient information seems to be grammaticalized, suggesting some sort of universal "cognitive primacy" principle behind morphological systems (Franzon, Zanini & Rugani, 2018; Strickland, 2017). We assess this cognitive primacy hypothesis with respect to inflectional morphology, in an artificial lexicon learning experiment. We will compare the learning of a semantic feature that is very often coded in morphological systems across languages (animacy) to the learning of one that is never grammaticalised instead (brightness), in order to assess possible biases for the former. We will test adult speakers of Italian, where neither feature is marked morphologically.

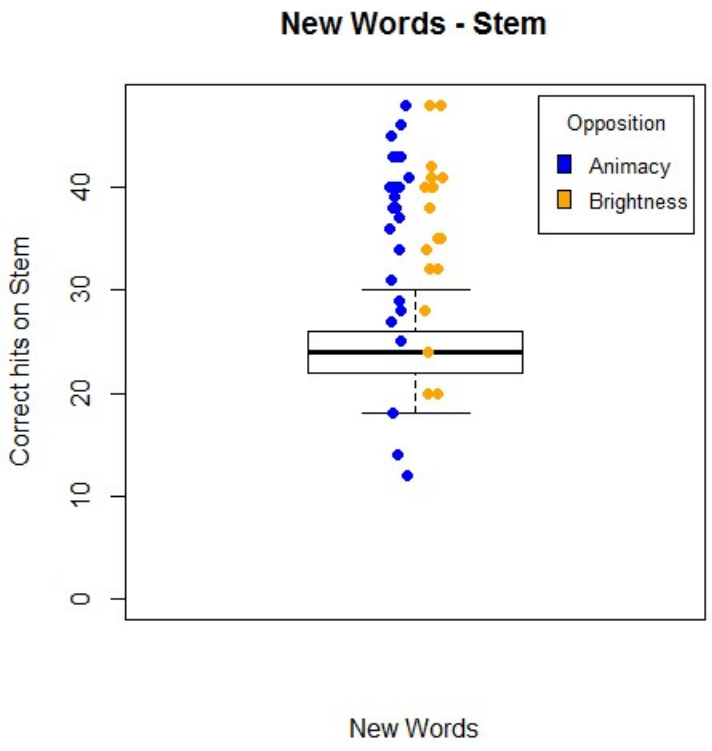
Materials and Methods. Entirely novel figures, words, stems and suffixes were used in these experiments, in order to avoid competition with any existing lexical or semantic knowledge. Participants had to learn associations between 32 made-up words (e.g., BARGIZ) and 32 made-up figures. Sixteen of the figures represented animate creatures, 8 bright-colored and 8 dark-colored; 16 figures represented inanimate objects, 8 bright-colored and 8 dark-colored. All the words were made up of a stem and a suffix (BARG-IZ). Each figure was associated with a unique stem, which was attached to only one of two suffixes (IZ vs. EB). In experiment 1, suffixes coded for the animacy opposition — nouns for animate creatures ended in IZ and nouns for inanimate objects ended in EB. In experiment 2, suffixes coded for the color opposition—each bright-colored entity ended in IZ and each dark-colored object ended in EB counterbalanced across participants. Therefore, the two experiments are perfectly identical, except for the semantic feature that is grammaticalised in the suffixes. During the Learning Phase, participants had to learn the names of 8 of the figures. In the Testing Phase (24 trials),

the 8 trained pictures were presented with four words. The participants had to choose the correct one (e.g., BARGIZ). The three competing words were composed by correct stem + wrong suffix (BARGEB); or by congruent suffix + wrong stem (NIDRIZ); or wrong stem + wrong suffix (NIDREB). In the Generalization Phase (72 trials), 24 previously unseen pictures were presented, each with four words. The participants were asked to choose the one that "fitted better" the image. The target word had a novel stem, which the participants were never exposed to; and a morpheme congruent with the semantic feature (e.g., COPLEB). One distractor had the same novel stem, + wrong morpheme (COPLIZ); another distractor had the congruent morpheme + stem that must be wrong, because it was associated with one of the learned words (BARGEB); the third distractor had the wrong suffix + the same must–be–wrong stem as above (BARGIZ). Twenty-four Italian native speakers took part to Experiment 1; 19 took part to Experiment 2.

Results and discussion. In E1, all participants were able to recognize the correct word in the Testing Phase significantly above chance ($p<.05$); 95% were able to do so in E2. The majority of the participants also learned the morphological opposition behind the suffixes: 62.5% selected the appropriate one significantly above chance in E1, and 68,4% in E2. Crucially, this pattern also held when objects and words were completely novel to the participants, showing that they truly learned a morphological system, rather than individual items. Also crucially, no significant difference emerged between learning the animacy and the brightness opposition, as assessed by comparing the performance in E1 and E2 trough a mixed-effects model (χ^2[1]=0.3429 (1) p=0.55).

Conclusions. Speakers can infer a systematic relation between a morphological opposition and a semantic opposition after being exposed to a few novel lexical items, referring to entities that are not present in the lexicon of their native language. This implicit learning process can take place even when the semantic feature mapped in the opposition is not marked in the speakers' native language, or in any language at all. The morphological opposition can be used as a cue to assign newly learned words to a class. Note that these experiments presented a fully consistent system to the participants (there were no exception to the morphological associations), which were trained in the absence of any competing cue — whether these results generalise to a more realistic scenario where morphological ties aren't entirely consistent is the subject of current research in the lab.


New Words - Suffix
Correct hits on Suffix
0
10
20
30
40
Opposition
Animacy
Brightness
New Words


New Words - Stem
Correct hits on Stem
0
10
20
30
40
Opposition
Animacy
Brightness
New Words

Francesca Franzon, Chiara Zanini and Valentina Pescuma

Form, function, meaning. A study on the distribution of inflectional morphemes in Italian.

INTRODUCTION

Inflectional morphology can encode semantic features, such as numerosity or sex, and, at the same time, it plays a functional role. The agreement of morphological features disambiguates the relations between constituents in sentence parsing, and reduces processing effort by favouring word predictions (Dye et al. 2017; Wicha et al. 2004). Ideally, these processes should be facilitated by consistency between form and features. Instead, the inflectional categories, such as Gender and Number, often show allomorphy and syncretism, and only rarely surface in consistent morphophonological values (such as -s for plural nouns in English; from now on referred to as morphemes).

We will discuss how the distributions of inflectional features and morphemes can comply with general coding and processing principles. We will also suggest that the encoding of referential information can still partially affect the distributions of inflectional features, even in a fundamentally functional category like inflection. Finally, we will point to some questions that arise from the observed data.

METHODS

We will report data on the distribution of the inflectional features and morphemes of Italian nouns, a language that marks Gender (masculine and feminine) and Number (singular and plural). A sample of nouns was obtained by merging two freely available resources. Noun types were obtained from Morph-It!, a morphologically tagged lexicon containing approximately 500,000 word forms (Zanchetta & Baroni 2005). Token frequencies of nouns were obtained from ItWaC, a 1.9 billion token corpus based on web-collected data (Baroni et al. 2009). We excluded homograph forms that could belong to more than one declensional class (e.g. cameriere, masculine singular 'waiter' or feminine plural 'waitresses'), obtaining a total of 210,325,942 tokens of 22,638 noun types.

First, we counted the type and token frequency of nouns for Gender and Number. Second, using the last character of each word form, we counted how many different types of morphemes could be linked to each of the inflectional features; we counted the type and token frequency of the nouns ending in each of them.

RESULTS AND DISCUSSION

Noun types are quite evenly distributed across the inflectional features. The sample contains 4,615 (20%) feminine singular nouns, 4,550 (20%) feminine plural nouns, 6,724 (30%) masculine singular nouns, and 6,749 (30%) masculine plural nouns. The almost maximal entropy of the distribution of types (H=1.974) supports the hypothesis that the inflectional features are optimised to reduce uncertainty in sentences by linking constituents.

However, when tokens are counted, the information decreases to 1.881, due to the increased proportion of singulars (feminine singular 34%, feminine plural 13%, masculine singular 35.6%, masculine plural 17.4%). This likely corresponds to the fact that, within Number features, the singular can be used as a default value to express no semantic interpretation about numerosity while still providing a feature for functional operations. Conversely, the greater information in plurals corresponds to their less ambiguous encoding of a semantic interpretation about a numerosity (Arcara et al. 2019).

Each of the features shows a prevalence in the association with one morpheme; the other morphemes with which each feature is associated decrease in type frequency following a zipfian-like distribution (Figure 1). Within each feature, the persistence of morphemes with a lower type frequency could be accounted for by their higher token frequency, as predicted in discriminative learning accounts (Blevins, Milin & Ramscar 2017).

Likewise, one morpheme can be linked to more than one feature: for example, the most frequent marking for feminine singular (-a) can be frequently associated with the other features as well; -e in the singular is ambiguous between masculine and feminine. Whereas the observed diversity of morpheme types within feature is likely to emerge as a result of communicative and learning pressures (Ramscar et al., 2013), the fact that a same morpheme is associated to more than one feature seems undesirable. Notably, plurals, which are more informative (see paragraph above), are represented by less types of morphemes than singulars are, therefore showing a more stable association between form and meaning. This suggests that allomorphy is a property of more functionalized features. Thus, semantic interpretability seems to affect morpheme distribution as well.

We are quantifying the uncertainty of the relation between phonological forms of morphemes, inflectional features, and their semantic interpretability, with the aim of assessing how their ambiguity affects on-line processing of sentences.

References

Arcara, G., Franzon, F., Gastaldon, S., Brotto, S., Semenza, C., Peressotti, F., & Zanini, C. (2019). One can be some but some cannot be one: ERP correlates of numerosity incongruence are different for singular and plural. Cortex, 116, 104-121.

Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. Language resources and evaluation, 43(3), 209-226.

Blevins, J. P., Milin, P., & Ramscar, M. (2017). The Zipfian paradigm cell filling problem. In Perspectives on Morphological Organization (pp. 139-158). Brill.

Dye, M., Milin, P., Futrell, R., & Ramscar, M. (2017). A functional theory of gender paradigms. In Perspectives on Morphological Organization (pp. 212-239). Brill.

Ramscar, M., Dye, M., & McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of" mouses" in adult speech. Language, 760-793.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. Journal of cognitive neuroscience, 16(7), 1272-1288.

Zanchetta, E., & Baroni, M. (2005). Morph-it! A free corpus-based morphological resource for the Italian language. In IN: PROCEEDINGS OF CORPUS LINGUISTICS, HTTP://DEV. SSLMIT. UNIBO. IT/LINGUISTICS/MORPH-IT. PHP.

Figure 1.

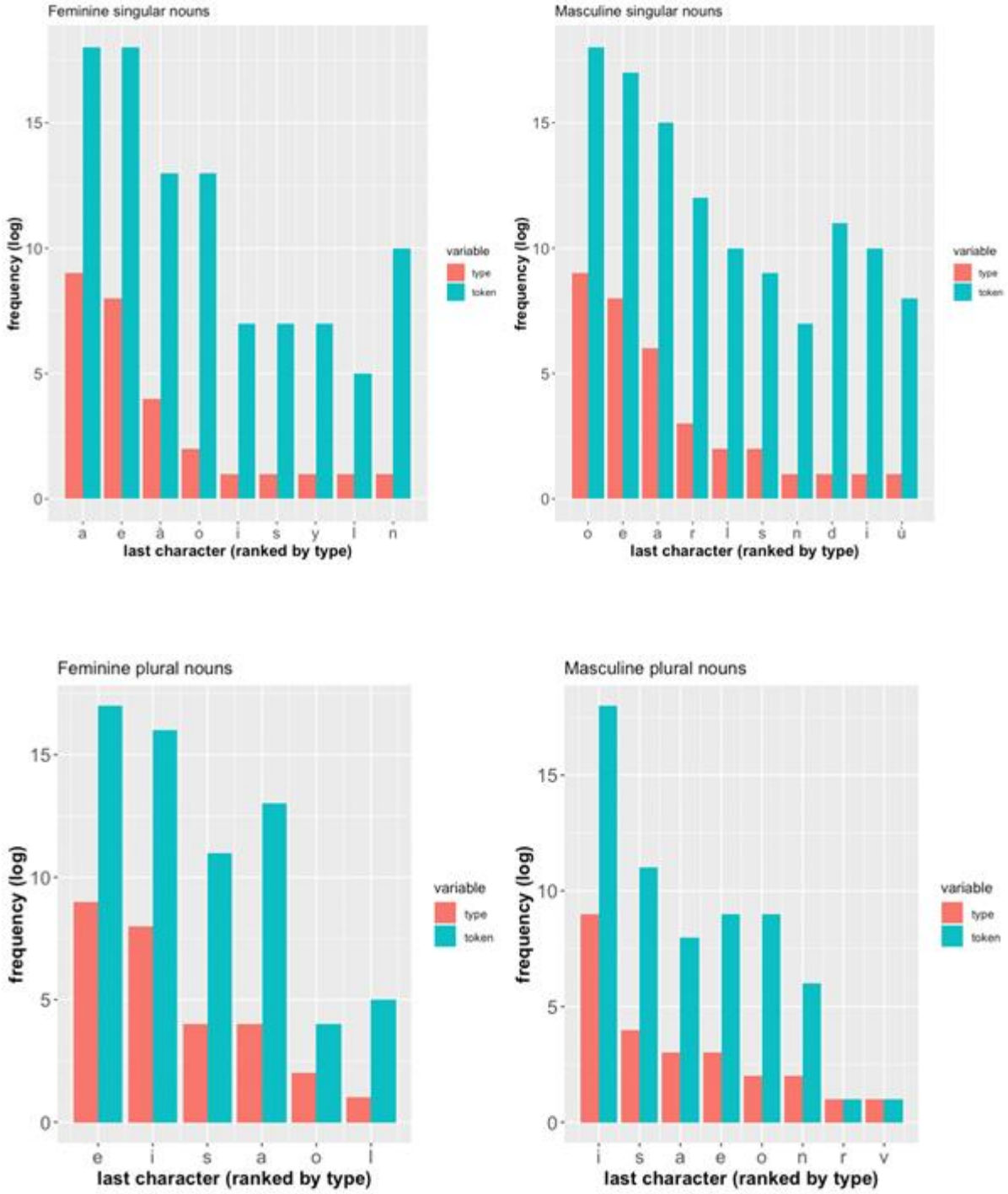
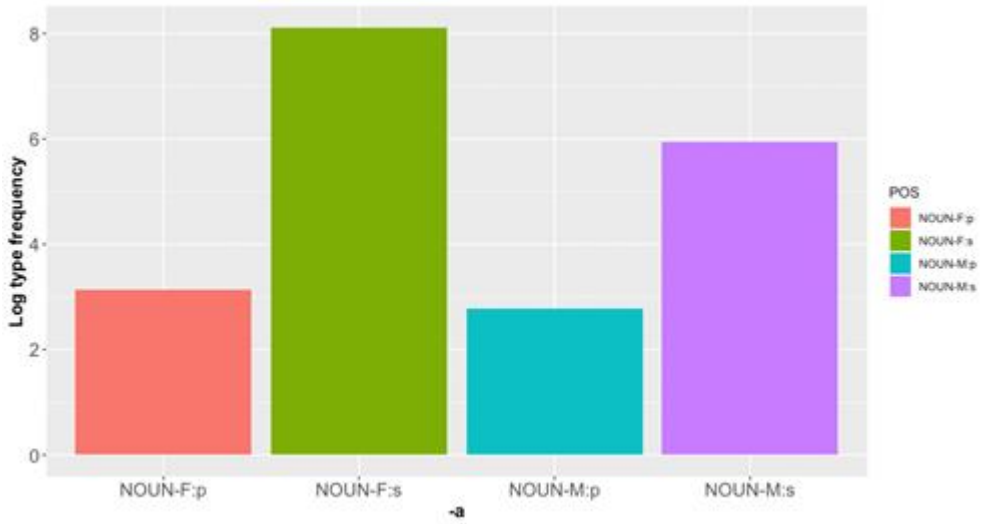

Feminine singular nouns
frequency (log)
15
10
5
0
a e à o i s y l n
last character (ranked by type)
variable
type
token
Masculine singular nouns
frequency (log)
15
10
5
0
o e a r l s n d i ù
last character (ranked by type)
variable
type
token
Feminine plural nouns
frequency (log)
15
10
5
0
e i s a o l
last character (ranked by type)
variable
type
token
Masculine plural nouns
frequency (log)
15
10
5
0
i s a e o n r v
last character (ranked by type)
variable
type
token

Figure 2.


Log type frequency
8
6
4
2
0
NOUN-F:p
NOUN-F:s
NOUN-M:p
NOUN-M:s
-a
POS
NOUN-F:p
NOUN-F:s
NOUN-M:p
NOUN-M:s

Adriana Rosalina Galván Torres, Katharina Korecky-Kröll and Michael Schmid

Language contact with English influences learners' production of German comparatives and superlatives: evidence from adult native speakers of Mexican Spanish
(Oral presentation)

This contribution investigates how L3 learners of German deal with German adjective gradation, after having learned a genetically related language, namely English, as their L2. Our theoretical framework is Natural Morphology (Dressler et al. 1987), particularly the subtheory of universal markedness.

German adjective gradation shows high variation and competition between forms (e.g., Nowak 2017). Most comparatives and superlatives are formed synthetically from their positive forms via suffixation. Some also undergo a stem vowel change (umlaut), either obligatorily or optionally (e.g. comparative *schmal-er/schmäl-er* 'more narrow'), others deletion of the *e* schwa in comparative (e.g. *böse – bös-er* 'more angry'), and others may undergo an *e* epenthesis in superlative (e.g. *am schlau-sten/schlau-est-en* 'most clever').

Additionally, token frequency of comparatives and especially superlatives in corpora of spoken and written speech (e.g., Ford et al. 2003) is very low: Forms that are rarely heard are more susceptible to uncertainties in learners' minds.

Natural Morphology assumes certain universal preferences in humans for more natural (i.e., less marked) forms according to the following preference parameters:

(a) Constructional iconicity: Most comparatives and superlatives may be classified as diagrams: The use of suffixes results in longer forms with addition of the meaning "more" to the positive form. In addition, superlatives are usually longer than comparatives, which is also an iconic relationship. Comparatives and superlatives with suffixes and umlaut may be classified as a combination of diagrams and metaphors and therefore as more iconic than pure suffix forms – however, this is disputed in the literature (e.g., Seifert 1988).

(b) Morphotactic transparency is variable: Comparatives and superlatives that add only suffixes are more transparent than forms with umlaut.

(c) Morphosemantic transparency describes the relation between form and meaning: Derivational suffixes are usually more transparent than pseudosuffixes which do not have a clear meaning (e.g., *stachel-ig* ‚spiny', derived from *Stachel* ‚spine' vs. *wichtig* ‚important').

(d) Perceptual salience describes the ease of perceptibility of different linguistic units. The positional preference for final and initial elements leads to a preference for suffixes over prefixes. Likewise, stressed suffixes and suffixes containing full vowels are preferred. Whereas comparatives are less salient due to their *a* schwa marker (i.e., one phoneme corresponding to the two graphemes *er*), superlatives are more salient because the marker *–st-e(n)* used in predicative and adverbial position (as in the present study) does not only contain a vowel plus several consonants, but it even corresponds to a syllable.

(e) Binarity and optimal form of units: The preference for binary relations leads to a preference of suffixes or prefixes over circumfixes as well as of one single affix over several affixes. A grammatical morpheme has the optimal length of one syllable, and a word with an

affix should ideally not be longer than a trisyllabic foot. This holds for most comparatives and superlatives found in our data, with a few exceptions (e.g., *stachel-ig-st-en* 'spiniest').

From a comparative perspective, German and English adjectives show considerable similarities, not only in the lexicon (e.g., *voll = full*), but also in morphology: Like English synthetic comparatives, all German comparatives take an *-er* suffix (e.g., *voll-er = full-er*). Although superlatives differ slightly in their forms (e.g., *am voll-st-en = the full-est*), their resemblance is nevertheless evident. However, there is no umlaut in English adjective gradation (apart from a few fossilized exceptions, such as *the elder lady*).

To test the effects of the L2 (English) on the production of the L3 (German) in 42 adult native speakers of Mexican Spanish with low intermediate to good levels of English and A2.1 level of German, we conducted a sentence completion task, in which the positive forms as well as the sentence frame were provided, to which participants were asked to produce appropriate comparative and superlative forms.

Results show strong influences from English, especially in superlative formation, where a main strategy of many participants was just to add the English suffix (e.g. **voll-est*). Some participants also produced entire English comparatives and superlatives.

Overall, more natural forms (e.g., forms without umlaut or comparatives) were clearly preferred over less natural forms according to the predictions of universal markedness. This stresses the importance of universal preference parameters in learner varieties.

Finally, results also confirm that transfer is an important phenomenon in less advanced L3 learners, especially in domains where the L2 and the L3 core grammars show many similarities.

References

Dressler, Wolfgang U., Willi Mayerthaler, Oswald Panagl & Wolfgang U. Wurzel. 1987. *Leitmotifs in Natural Morphology*. Amsterdam: Benjamins.

Ford, Michael A., William D. Marslen-Wilson & Matthew H. Davis. 2003. Morphology and frequency: Contrasting methodologies. In R. Harald Baayen & Robert Schreuder. eds. *Morphological structure in language processing*. Berlin: De Gruyter, 89–124.

Nowak, Jessica. 2017. Klar – klärer – am klärsten? Umlaut comparison as a doubtful case in contemporary German. *Yearbook of the Poznań Linguistic Meeting* 3, 77-99.

Seifert, Katharina 1988. *Ikonizität von Pluralformen: eine Untersuchung zur psychologischen Realität der linguistischen Theorie der „natürlichen Morphologie"*. Wien: VWGÖ (= *Dissertationen der Universität Wien* 188).

Anthony Grant

Bound verbal morphology and borrowing: limits and possibilities.

The borrowing of verbal inflectional morphology is regarded as a rarity across the world's languages. Exceptions do occur of languages which have productively borrowed verbal infletional items or paradigms (partial or total: Cech 2006 for Dolenjska Romani in Slovenia), or which have taken over and productively employed affixes from other languages (Merlan 1982 for interaction between Mangarrayi [then spelt *Mangarayi*] and Jawoyn in Arnhem Land). A number of scholars, including Jeffrey Heath (notably Heath 1978) and Yaron Matras (for example Matras 2007) have set up hierarchies of probability of borrowing, including that of verbal morphology. An important recent survey is Seifart (2017).

Examining a series of case studies using data from languages across the globe I evaluate these claims. It appears that patterns of borrowing, and dispreferred patterns in borrowing, are both strong and applicable as rules of thumb. They are. , however, not without exception, and I aim to show why, with especial emphasis on the role of complementary or convergent morphological patterns on shaping these outcomes.

REFERENCES

Cech, Petra, 2006. *Dolenska Romani: the language of the Dolenjski Roma in Nove Mesto and Bela Krajina, Slovenia.* Munich: Lincom Europa.

Heath, Jeffrey. 1978. *Linguistic Diffusion in Arnhem Land*. Canberra: Australian Institute of Aboriginal Studies.

Matras, Yaron. 2007. The borrowability of structural categories. *Grammatical borrowing in cross-linguistic perspective*, edited by Yaron Matras and Jeanette Sakel, 31-74. Berlin: de Gruyter Mouton.

Merlan, Francesca C. 1982. *Mangarayi*. Amsterdam: North-Holland.

Seifart, Frank. 2017. Patterns of affix borrowing in a sample of 100 languages. *Journal of Historical Linguistics* 7(3). 389–431

Camiel Hamans

Blends: Category at the Crossroads of Morphology and Phonology

Traditionally blends are seen as idiosyncratic phenomena with no importance for the study of morphology. Marchand (1968: 451) for instance claimed that blending only has stylistic status and relegated the study of blends to lexicology. With the rise of optimality theory and especially of prosodic morphology, the picture has changed. Blends are now studied seriously (see for instance Piñeros 2000, 2002, Bat-El 2006, Bet-El and Cohen 2012, Trommer and Zimmermann 2012). However, Bauer, Liber and Plag (2013) who give an overview of the state of affairs, are not able to draw a coherent picture of English blends. Mattiello (2013) does not go further than a taxonomic inventory. This paper wants to present an in depth analysis of blends, which will lead to the conclusion that blends have one leg in phonology and the other in morphology. Data from English, German and Dutch are analyzed; that is why no universal claim is made.

Bauer, Liber and Plag (2013, 458) distinguish two types of blends: AC blends and AD blends.
AC blends combine the first part of both source words, as in *sitcom* from *situation comedy* or *misper* from *missing person.*
AD blends are concatenations of the first part of the first source word and the final part of the second source word as in *smog* from *smoke* and *fog* or *boatel* from *boat* and *hotel* or *stagflation* from *stagnation* and *inflation*.

First it will be shown that AC concatenations differ from AD blends essentially. AC concatenations all have stress on the first part, whereas this is not necessarily the case with AD blends. In addition the right part of AC concatenations appears to be the semantic and formal head which leads to the conclusion that AC concatenations are compounds, be it of clipped words.

The remaining of the presentation will be devoted to an analysis of AD blends.
AD blends also have a formal head as the data in (1) show. The head is the right hand part.

(1a) English:	barkitecture (**N**) 'design of doghouses'	< bark (V)	+ architecture (N)
(1b) German:	Naktivist (**N**) 'naked activist'	< nackt (Adj)	+ Aktivist (N)
(1c) Dutch:	het potel (**neuter**) 'hotel for Polish workers'	< de Polen (common g.)	+ het hotel (neuter)

In this respect blends behave as compounds. However, when it comes to stress, it appears that blends tend to copy the stress pattern of the source word underlying the head, the right hand part (Piñeros 2000, 2002, Bat-El 2006, Bet-El and Cohen 2012, Trommer and Zimmermann 2012). Stressed vowels are italicized.

(2) boat	+ hot*é*l	→	boat*é*l
frapp*é*	+ cappucc*í*no	→	frappucc*í*no
fl*ú*stered	+ frustr*á*ted	→	flust*á*ted

Even when the segmental material of the second source word is not preserved, the suprasegmental prosodic feature stress of this second source word survives:
(3) pr*é*stinant < prest*í*gious + d*ó*minant

Unlike compounds, which consist of two prosodic words, blends appear to consist of only one prosodic word, although they are concatenations of parts of two words.

Stress is not the only aspect blends copy from their second source word, also the syllabic structure of blends is a copy of it (cf. Arndt-Lappe and Plag 2012). The segments of the source words which ultimately form the blend are italicized.

(4)	*br*eakfast	+ l*unch*	→ brunch	onset
	*Spa*nish	+ E*nglish*	→ Spanglish	onset + nucleus
	*stag*nation	+ in*flation*	→ stagflation	σ
	*adver*tisement	+ edi*torial*	→ advertorial	σσ

The data presented in (4) demonstrate that it is the syllabic structure of the second source word which determines the syllabic structure of the blend. Exactly as much syllabic segmental material that is truncated from the second source word may be extracted from the first source word and inserted into the open spot(s) in the syllabic structure of the head.

The analysis summarized here will show that blends are fully systematic but cross the border between phonology and morphology or maybe better, show how fluid this border is.

References

Arndt- Lappe, Sabine and Ingo Plag (2012): *Phonological Variability in English Blends*, paper presented at the Conference "Data-Rich Approaches to English Morphology: From corpora and experiments to theory and back". Wellington New Zealand, 4-6 July 2012.

Bat-El, Outi (2006): "Blend". Keith Brown (ed.) *Encyclopedia of language and linguistics*. 2nd edition. Oxford: Elsevier: vol.2: 66-70.

Bat-El, Outi & Evan-Gary Cohen (2012): "Stress in English blends: A constraint based analysis". In : Vincent Renner, François Maniez & Pierre Arnaud (eds.), C*rossdisciplinary Perspectives on Lexical Blending*. Berlin: De Gruyter Mouton: 193-211.

Bauer, Laurie, Rochelle Liber & Ingo Plag (2013): *English Morphology.* Oxford: Oxford University Press.

Marchand, Hans (1969): *The Categories and Types of Present-Day English Word-Formation.* München: Beck. Second, completely revised and enlarged edition.

Mattiello, Elisa (2013): *Extra-grammatical morphology in English. Abbreviations, blends, reduplicatives, and related phenomena*. Berlin & Boston: De Gruyter Mouton.

Piñeros, Carlos E. (2000): *Word-blending as a case of non-concatenative morphology in Spanish.* ROA (Rutgers Optimality Archive), 343-0999.

Piñeros, Carlos E., (2002). *The creation of portmanteaus in the extragrammatical morphology of Spanish*. ROA (Rutgers Optimality Archive), 562-0602, also (2004). *Probus* 16,2, 203-240.

Trommer, Jochen and Eva Zimmermann (2012): "Portmanteaus as generalized templates". Vincent Renner, François Maniez & Pierre Arnaud (eds.), *Crossdisciplinary Perspectives on Lexical Blending*. Berlin: De Gruyter Mouton : 233-258.

Gary Libben, Sarah Macdonald, Learose Pinkham, Mira Reisinger, Michaela Ringseis, Kaitlin Goertz, Katharina Korecky-Kröll and Wolfgang Dressler

A multi-method investigation of morphological processing in German

Recent research (e.g., Schmidke, Van Dyke & Kuperman (2018) has suggested that the processing of multimorphemic words is characterized by considerable variability. Variation in morphological processing may have many sources (Gagné, 2017) and may change in a single individual over time (Ramscar, 2014 et al.). Libben (2017) has claimed that this variation is made possible by the fundamental nature of morphology as a cognitive phenomenon. It is claimed that words do not have fixed morphological structures. Rather, morphological knowledge takes the form of morphological superstates. This suggests that the understanding of morphological processing requires that we account for the variability of morphological configurations enabled by a morphological superstate and the extent to which these can be affected by the characteristics and demands of the language processing environment.

We report on a series of lexical processing experiments carried out with 41 native speakers of German that was designed to uncover the nature and influencers of morphological variability. Stimuli consisted of 75 German words: 25 triconstituent compounds (e.g., *Fussballspiel*) and 50 suffixed words ending in either the suffixes *bar+keit*, *ig+keit*, *lich+keit*, or *sam+keit*

A key feature of the study was the use of repeated measures. It was designed so that a single participant, tested over a four-week period, would supply data from four experimental paradigms. We used progressive demasking (Grainger & Segui, 1990) as a measure of word recognition We used typing (Will, Nottbusch & Weingarten, 2006) as a measure of word production and we created a morpheme boundary selection task as a measure of conscious morphological segmentation. Crucially, we created presentation conditions that biased the viewer toward a particular reading of the three-morpheme string. This was accomplished by using coloured fonts so that each word was seen in one of three conditions: (1) full morphological decomposition, (2) split between the first and second morphemes, (3) split between the second and third morphemes, and (4) split at a non-morphological boundary.

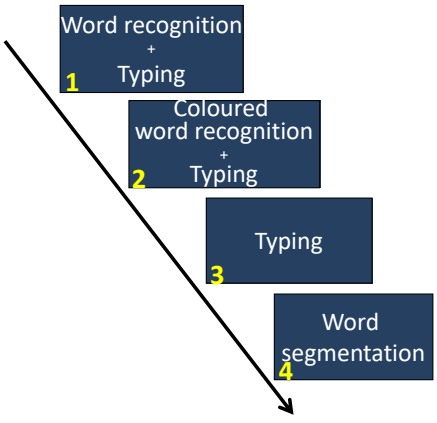


Figure 1. The overall multi-task design.

The typing paradigm constituted our key task. Per-letter typing times served as the primary dependent variable. Analysis of those typing times indicated that morphological boundary effects are much stronger for compounds than for suffixed words. We found that these

morphological effects in typing production were uninfluenced by the colored font manipulation. Recognition latency was, however, strongly influenced by the coloring manipulation. This underlines the extent to which the typing paradigm taps internalized representations. It is not simply a 'repetition' of the stimulus.

A particularly interesting finding was the convergence of online measures and the lack of convergence between inline and off-line measures. We found that word recognition latencies predicted typing times. However, we did not see alignment of conscious segmentation choices and online recognition patterns.

Our finding that morphological patterns persist in the typing record suggests that morphological effects are not limited to word recognition paradigms that depend on participants being surprised by stimuli. Rather, they seem to reflect the nature of internal representations. We discuss the manner in which these internal representations can change over time, both as a result of stimulus manipulation and as a result of repeated exposure to individual lexical items and morphological patterns. We claim that the multi-method repeated measures approach that we have employed in this study provides a privileged perspective on the variability and malleability of morphological processing.

References

Gagné, Christina. (2017), Psycholinguistic Approaches to Morphology, in M. Aronoff (ed), *Oxford Research Encyclopedia of Linguistics,* Oxford: Oxford University Press.

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics*, 47, 191–198.

Libben, Gary (2017), The quantum metaphor and the organization of words in the mind, *Cultural Cognitive Science*, 1, 49-55.

Ramscar, M., Hendrix P., Shaoul C., Milin P., and Baayen, R. H. (2014). The Myth of Cognitive Decline: Non-Linear Dynamics of Lifelong Learning. *Topics in Cognitive Science*, 6, 5-42.

Schmidtke, D., Van Dyke, J. A., & Kuperman, V. (2018). Individual variability in the semantic processing of English compound words. *Experimental Psychology: Learning, Memory, and Cognition*, 44(3), 421–439.

Will, U., G. Nottbusch & R. Weingarten. 2006. Linguistic units in word typing: Effects of word presentation modes and typing delay. *Written Language & Literacy 9*. 153–176.

Fujio Nakamura

An exploration into why the establishment of the negative contraction *doesn't* was delayed in American English (Oral presentation)

In the history of British English (BE), even educated people seem not to have shown an aversion to the usage of the third person singular present [3SG] *don't*. X (2012) elucidated the period and chronological order in which negative contractions became established and diffused in BE. X (2013) explored the transition from *he don't know* to *he doesn't know* in the mid-nineteenth century, and also (2018) sought the reason why *doesn't* and the past-tense group (such as *didn't* and *couldn't*) were established 100-150 years later than the present-tense group (such as *don't* and *can't*). The usage of the 3SG *don't* continued up to the early twentieth century, when it was finally established in its "vulgar" (*OED*^{2} on CD-ROM, s.v. do. *v*., 2c) or "non-standard" (Denison 1998: 195) speech-level form, or "current" "conversational grammar" (Biber, et al. 1999: 1123).

In American English (AE), however, the negative contraction *doesn't* was established as late as the second half of the twentieth century (Bloomfield and Newmark 1963: 26), except for Southern dialects and the west part of the State of New York (Mencken 1919 [1977]: 542; Trask 2004: 199). No one seems to have clearly explained the reason why the establishment of *doesn't* was delayed in AE. However, the reason for this prolongation of *doesn't* seems to be simple: as stated above, *doesn't* was rarely used even in BE until the mid-nineteenth century, and accordingly, this negative contraction was little known and unfamiliar to the multitudes of immigrants coming into the United States before that time. They had no choice but to continue to use the non-contracted *does not* or else the 3SG *don't* as normal usage until the mid-twentieth century. Thus, the current presentation attempts to respond to the question of why the establishment of the negative contraction *doesn't* was delayed until around the mid-twentieth century in AE, even though its earliest attested example in BE appeared in a drama written in 1674.

All of this is demonstrated in the current presentation, based upon the examples of *doesn't* and 3SG *don't* collected from 129 volumes of BE diaries and correspondence written primarily between 1600 and 1950, a heterogeneous mixture of 324 volumes of electronically logged BE texts in the years 1351-1950 (101 MB) and 410 volumes of electronically logged AE texts in the years 1751-1950 (121 MB), such as biographies, essays, journals, letters and novels, ARCHER corpus, version 3.2 (BE 1600-1999; AE 1750-1999), all of the citations in the *OED*^{2} on CD-ROM, version 3.1, six present-day corpora such as LOB (1961), FLOB (1991-1992), Brown (1961), Frown (1991-1992), Kolhapur (1978) and ACE (1986), and four historical corpora such as ICAMET (1386-1688), CEECS (1418-1680), Lampeter (1640-1740), and Newdigate (1673-1692). The tabulation of the data for variations between *doesn't* and 3SG *don't* is presented according to syntactic and stylistic properties such as functions of sentences (declaratives / tag-questions / interrogatives), kinds of subjects (personal pronouns / substantives / demonstratives / etc.) and types of clauses (superordinate clauses / subordinate clauses).

It will be shown that one of the morpho-syntactic mysteries about present-day AE can be solved more easily than expected by examining historical BE.

References

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education.

Bloomfield, Morton W. and Leonard Newmark. 1963. *A Linguistic Introduction to the History of English*. New York: Alfred A. Knopf.

Denison, David. 1998. "Syntax", in S. Romaine, ed., *The Cambridge History of the English Language*, vol. IV, *1776-1997*. Cambridge University Press, Cambridge, 92-329.

Mencken, H. L. 1919 [1977]. *The American Language: An Inquiry into the Development of English in the United States*. One-Volume Abridged Edition. New York: Alfred A. Knopf.

Trask, Robert L. 2004 [1995]. *Language: The Basics*. 2nd ed. London & New York: Routledge.
X, Y. 2012. "The period of establishment of tag-questions", . . .
X, Y. 2013. "A history of the third person singular present *don't*: Transition from *he don't know* to *he doesn't know*", . . .
X, Y. 2018. "A history of negative contractions: Seeking the reason why *doesn't* and the past-tense group (such as *didn't* and *couldn't*) were established 100-150 years later than the present-tense group (such as *don't* and *can't*)", . . .

word-count 719
characters incl. blanks and references 4,569

Maria Rosenberg

Adjectives as a case of word-class changing inflection in the early course of Swedish language development

3) Oral or poster

This study investigates the emergence and use of adjectives in Swedish child language in relation to the issue of word-class changing inflection. In most languages, adjectives are less frequent (types, tokens) than nouns and verbs (Tribushinina et al. 2015). Swedish nouns have two genders, arbitrary assigned, common or neuter (70–80% are common, Bohnacker 2003), Attributive and predicative adjective agreement occurs. In singular NPs, adjectives are unmarked with common nouns (*en stor bil* 'a big car') but agree with neuter nouns (*ett stor-t hus* 'a big.AGR house'). Moreover, the neuter, singular form (with a final *-t*) of most Swedish adjectives can function as an adverbial (Hultman 2003). In this case, the *t*-ending is rather inflectional than derivational (Söderbergh 1968). Haspelmath questions "the claim that word-class-changing affixes always have derivational status" (1996:44) and posits instead that inflectional affixes can alter the word-class; inflection and derivation being two ends of a continuum. Few studies have focused on adjectives in Swedish children, but studies on the problems with NP processing (Leonard et al. 2001; Hallin & Reuterskiöld 2018) proposed the arbitrary assignment of noun gender (favouring the common gender) and the complexity of adjective agreement to be challenges. However, it is important to consider that Swedish children meet many adjectives in neuter forms as adverbials, something that could compensate for the common gender advantage.

The study aims to elucidate whether the adjective and adverb distinction can be claimed to have psycholinguistic reality for the child, by addressing in what frames Swedish children use adjectives and adverbs respectively? It is based on longitudinal and cross-sectional spontaneous production data from typically developing, monolingual Swedish children: (i) diary notes of one girl, ages 1;9–3;6; (ii) longitudinal recordings of four children, ages 1;8–3;6 (Strömqvist et al. 1993, at CHILDES MacWhinney 2000); (iii) single recordings, 20–30 min, of nine children, ages 2;0–2;6, and six children, ages 3;0–3;6. Utterances containing adjectives and adverbs are extracted and analysed as to form and function (also CDS when available).

The data show that the first adjectives tend to be produced in isolation and rarely before age 2, and also that agreement is mastered gradually and not before age 3. In the cross-sectional data, agreement errors, as well as omissions of obligatory articles and copulas in constructions with adjectives, still occur between the ages 3;0–3;6. One salient frame is the copula construction [*den/det är* ADJ/ADV] 'it.C/it.N is ADJ/ADV' (cf. Diessel & Tomasello 2000). Another frame, used by all children in the longitudinal data, is the intensifying [*jätte-*ADJ/ADV] (e.g *jättebra* 'very-good'/jättefort 'very-fast'). Hence, it could be argued that for Swedish children, at least, the reality presents little evidence for a firm form-function distinction between adjectives and adverbs (semantic aspects left aside).

In conclusion, acquiring adjectives is a challenging task that progresses gradually. For Swedish, however, the fact that many adjectives appear as adverbials is a factor that favours both the type and token occurrence in the input, seeing that, as Haspelmath (1996) emphasizes, inflectional forms differ from derivational forms in being more frequent in discourse.

References

Bohnacker, U. (2003). Nominal Phrases. In G. Josefsson, C. Platzack & G. Håkansson (Eds.), *The Acquisition of Swedish Grammar* (195–260). Amsterdam/Philadelphia: Benjamins.

Diessel, H. & Tomasello, M. (2000). A new look at the acquisition of relative clauses. *Language* 81: 882–906.

Hallin, A. E. & Reuterskiöld, C. (2018). Effects of frequency and morphosyntactic structure on error detection, correction, and repetition in Swedish-speaking children. *Applied Psycholinguistics*, 39: 1189–1220.

Haspelmath, M. (1996). Word-class-changing inflection and morphological theory. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology* 1995 (43–66), Dordrecht: Kluwer Academic Publishers.
Hultman, T. G. (2003). *Svenska Akademiens språklära*. Stockholm: Svenska Akademien.
Josefsson, G. (1998). *Minimal words in a minimal syntax: Word formation in Swedish*. Amsterdam/Philadelphia: Benjamins.
Leonard, L. B., Salameh, E.-K. & Hansson, K. (2001). Noun phrase morphology in Swedish-speaking children with specific language impairment. *Applied Psycholinguistics*, 22: 619–639.
MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum.
Strömqvist, S., Richthoff, U., & Andersson, A.-B. (1993). *Strömqvist's and Richthoff's corpora: A guide to longitudinal data from four Swedish childre*n. Gothenburg Papers in Theoretical Linguistics, 66.
Söderbergh, R. (1968). *Svensk ordbildning* [Swedish word-formation]. Stockholm: Norstedts.
Tribushinina, E., Voeikova, M. D. & Noccetti, S. (2015). *Semantics and morphology of early adjectives in first language acquisition.* Newcastle upon Tyne: Cambridge Scholars.

Thomas Stewart

Morphological construction schemas distribute initial consonant mutation in Scottish Gaelic

The morphologized patterns of initial consonant alternation in Celtic, the so-called INITIAL MUTATIONS, have been analyzed from the perspectives of multiple grammatical components (e.g., Green 2007, Hannahs 2013). In extensive research with members of a Scottish community in which Gaelic was in clear retreat relative to English, Dorian (1977) documented uneven attrition rates in the use of an ostensibly unitary mutation pattern in the language, namely LENITION. The differences in performance correlated fairly directly with speaker's fluency in Gaelic, but Lenition also varied within the less fluent ("semi-") speakers' use across different grammatical contexts, in all of which contexts the Lenition mutation is conventionally obligatory (1-4, in descending level of semi-speakers' accuracy of use):

	Root-based (Radical) forms				Lenition forms			Contexts	Accuracy
(1)	***b**ris*	[**p**riʃ]	'break!'	≈	***bh**ris*	[**v**riʃ]	'broke'	(past V)	90+%
(2)	***c**earc*	[**k**^{hj}ar^{h}k]	'hen'	≈	*a' **ch**earc*	[**ç**^{j}ar^{h}k]	'the hen'	(def., fem. N)	<50%
(3)	***f**uar*	[**f**uar]	'cold'	≈	*glè **fh**uar*	[uar]	'very cold'	(quantified A)	<50%
(4)	***S**eumas*	[**ʃ**emas]	'James'	≈	*a **Sh**eumais*	[**h**emiʃ]	'(o) James!'	(vocative N)	<20%

Dorian raised and refuted several functional reasons for the fragmenting of Lenition, ultimately without settling on a clear convincing cause. Subsequently, Stewart (2013) showed that Lenition in Scottish Gaelic is not in fact the single coherent pattern that its unitary, phonologically-derived label implies (cf. Hannahs 2013) and that served to make Dorian's observations especially striking.

The present analysis first represents the hierarchical nature of the array of partial formal similarities found in Gaelic Lenition (cf. Janda 1982), and then relates each of the five distinct sub-patterns to the collections of morphological contexts that respectively call for them. Concretely, initial consonant mutations are distributed synchronically as integral prosodic elements of definable morphological constructions (Zwicky 1990; Booij 2010, Masini & Audring 2019). A further implication of the present analysis is that the mutation sub-patterns are likely MORPHOMIC (Aronoff 1994), in that they collect and state formal relationships that are used in a regular manner in the realization of contexts that are not (necessarily) featurally or lexically coherent.

In the current context, an empirical advantage of this less monolithic account of initial mutation-types is its consistency with differential retention and only partial collapse, such as that described by Dorian, exacerbated in light of dialect endangerment, compromised transmission, and attrition.

References

Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.

Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Dorian, Nancy C. 1977. A hierarchy of morphophonemic decay in Scottish Gaelic language death: The differential failure of lenition. *Word* 28. 96-109.

Green, Antony D. 2007. *Phonology limited*. Potsdam: Univeritätsverlag Potsdam.

Hannahs, S. J. 2013. Celtic initial mutation: Pattern extraction and subcategorisation. *Word Structure* 6(1). 1-20.

Janda, Richard D. 1982. Of formal identity and rule-(un)collapsibility: On lost and found generalizations in morphology. West Coast Conference of Formal Linguistics *WCCFL* 1. 179-197.

Masini, Francesca, & Jenny Audring. 2019. Construction Morphology. In Jenny Audring & Francesca Masini (eds.), *The Oxford handbook of morphological theory*, 365-389. Oxford: Oxford University Press.

Stewart, Thomas W. 2013. The sub-types of initial lenition in Scottish Gaelic. In Janet Cruickshank & Robert McColl Millar (eds.), *After the storm: Papers from the Forum for Research on the Languages of Scotland and Ulster triennial meeting, Aberdeen 2012*, 100-16. Aberdeen: FRLSU.

Zwicky, Arnold M. 1990. Syntactic representations and phonological shapes. In Sharon Inkelas & Draga Zec (eds.), *The phonology-syntax connection*, 379-397. Chicago: University of Chicago Press.

Alina Villalva, Sydelle de Souza and Carina Pinto

Word Associations: Analyzing activation pathways

Word association tasks (WAT) have been instrumental in research pertaining to the mental lexicon (Clark, 1970; Precosky, 2011). The present study is based on a carefully designed WAT, that aims to determine (a) if WATs are constrained by grammatical factors, (b) the role of the morphological structure of stimulus words, and (c) lexical activation pathways favored by L1 Portuguese speakers in WATs that focus on complex words.

The stimulus set consisted of 152 Portuguese verbs (e.g. *aceitar 'to accept'*) and their corresponding 152 deverbal action nouns, formed by the suffix *-ção* (e.g. *aceitação* 'the acceptance'), organized by the number of syllables (2-6) and frequency values (according to the *Corpus de Referência do Português Contemporâneo*). The subject sample consisted of 22 Portuguese speakers, university students, M_{age}= 20 ±1.95, with no language pathologies. The test was executed through Google Forms and lasted approximately two hours with all subjects present in the same room concurrently, with instructions to enter the first word that occurred to them on reading the stimuli. The responses were manually coded according to their relationship with the stimulus as follows:

A. Belongs to the word family of the stimulus (i.e. a base word, a derivative of the stimulus, a derivative of the base word, etc.).
B. Semantic relation.
C. Syntactic relation.
D. Alternate relations:
    D1. phonetic resemblance
    D2. no systematic relation.
E. The subject does not know the word or has not replied.
F. The response is not a word.

A preliminary analysis of the results shows that all subjects have frequently activated a morphologically (A), semantically (B) or syntactically (C) driven pathway. Grammatically random (D) and invalid responses (E and F) correspond to barely 6.3% of the total amount of responses:

	Same Word Family (A)		Semantic Response (B)		Syntactic Response (C)		Other (D+E+F)		Total	
N	1211	36.2%	1464	43.8%	456	13.6%	213	6.4%	3344	100%
V	1053	31.5%	1633	48.8%	447	13.4%	211	6.3%	3344	100%
T	2264	33.9%	3097	46.3%	903	13.5%	424	6.3%	6688	100%

Table 1: Global results

The slight asymmetry between responses to nominal and verbal stimuli in A and B is related to the fact that the verbs include some morphologically simplex structures, whereas all the nouns are complex. The following table shows that the values for complex verbs are closer to the values for complex nouns than for simplex verbs:

	Same Word Family (A)		Semantic response (B)		Syntactic response (C)		Other (D+E+F)		Total	
Simplex V	260	20.0%	717	55.2%	252	19.4%	69	5.3%	1254	100%
Complex V	793	38.8%	940	44.8%	205	9.5%	144	6.9%	2090	100%
Total V	105	31.5%	1633	48.8%	447	13.4%	211	6.3%	3344	100%

Table 2: Simplex and complex verbs

These results suggest that the morphological nature of the stimuli tends to constrain the response. Simplex words predominantly trigger a higher semantic activation, whereas complex words increase the probability of triggering a morphological activation.

It was also observed that words with a higher value of D were split into two classes: (a) phonetically similar to the stimulus and (b) not systematically associated with the stimulus. The following table also shows values for invalid responses:

	Grammatically Unrelated Responses				Invalid Responses (E+F)		Total	
	D1		D2					
N	37	35,2%	39	37,2%	29	27,6%	105	100%
V	24	25,3%	50	52,6%	21	22,1%	95	100%
T	61	30,5%	89	44,5%	50	25%	200	100%

Table 3: Grammatically unrelated and invalid responses

Further examination suggests that these responses are distributed among a small group of words that may have been unfamiliar to most subjects (e.g. *atemorizar* 'to frighten', *beatificar* 'to beatify', *beatificação* 'beatification', etc.).

In sum, this off-line WAT consolidates the hypothesis that speakers establish lexical relationships between words by prioritizing a semantic activation. These semantic relationships are morphologically dominated for morphologically complex stimulus words. Additionally, the concentration of a high score of random and invalid responses observed in some stimuli may help to devise a strategy, in terms of lexical knowledge, to contrast and isolate words more frequently known by an average speaker from words familiar only to more proficient speakers.

References:

Clark, H. H. (1970). Word Associations and Linguistic Theory. *New Horizons in Linguistics*, 271–286. Retrieved from https://web.stanford.edu/~clark/1970s/Clark, H.H. _Word associations and linguistic theory_ 1970.pdf (Last access: 28/08/2019)

Corpus de Referência do Português Contemporâneo. URL http://alfclul.clul.ul.pt/CQPweb/portugal/ (Last access: 29/08/2019)

Precosky, K. (2011). Exploring the mental lexicon using word association tests : How do native and non-native speakers of English arrange words in the mind? University of Birmingham. Retrieved from https://www.birmingham.ac.uk/Documents/college-artslaw/cels/essays/lexis/PrecoskyKristaLexisWAT2011.pdf (Last access: 28/08/2019)

Martina Werner

Domain adverbs in German: rivalry or split? A corpus study of the semi-suffixes *-technisch* and *-mäßig*

Domain adverbs (DAs) are a special type of adverb: they are semantically valid for a certain domain (1a), where they act as the topic, and, hence, semantically obligatory (1b) (Frey/Pittner 1998). Due to their domain semantics, *if*-sentences serve as equivalents (Pittner 1999:118) or metalinguistic (Ruge 2004: 39) phrases like "in terms of" (German *gesehen/betrachtet* lit. 'seen') can be added without a change in meaning (1c).

(1a) *Einkaufsmäßig steht Ljubljana Wien in nichts nach.* 'In terms of shopping, Ljubljana is in no way inferior to Vienna.'
(1b) = **(Einkaufsmäßig) steht Ljubljana Wien in nichts nach.*
(1c) *(Wenn man es) einkaufsmäßig betrachtet, steht Ljubljana Wien in nichts nach.* '(If it is) seen in terms of shopping, Ljubljana is in no way inferior to Vienna.'

As opposed to other adverbs (cf. Schäfer 2013: 65ff for more details), DAs are located at the morphology-syntax interface (Ramaglia 2011: 26ff, Marchis 2010, Marchis Moreno 2015) and they have an adjectival counterpart represented by the relational adjectives (RAs, also classifying or pseudo-adjectives), which are a frequent phenomenon for many languages (such as Romance, cf. e.g. Ramaglia 2011) and also for German (recently, see e.g. ten Hacken 2019). Despite their long-standing description in grammatical theory (seminal Bally 1944: 96f), RAs are characterized as having the „morphological shape of an adjective but behave in many respects like nouns" (Fábregas 2007: 3) in which they clearly differ from quality adjectives (QAs, also: characterizing adjectives) because of their non-qualitative (2a-b), non-scalable (2c) or non-polar character (2d; cf. Mravlag 2013). In addition, RAs cannot be nominalized (2e; cf. e.g. Holzer 1996, Frevel/Knobloch 2005, Fábregas 2007: 4, Zifonun 2011, Ganslmayer 2012: 138, among others).

(2a) *(Die) ärztliche Praxis* '(the) medical practice'
(2b) *#Die Praxis ist ärztlich.* lit. 'The practice is medical-RA.'
(2c) *#(Die) ärztlichere Praxis* '(the) more medical practice'
(2d) *#(Die) Praxis läuft (sehr) (un-)ärztlich.* 'The practice runs (non-)medically-RA.'
(2e) *#Die Ärztlichkeit der Praxis* 'the medicality of the practice'

When RAs are used as QAs, RAs (as in 3a) shift to denominal QA-readings with the form 'like a N / typical for a N' as in (3b-c) (see Dornseiff 1921, Motsch 2004: 196f; Trost 2006: 15f; pace Frevel/Knobloch 2005):

(3a) *studentisches Lernen* 'student's learning, learning of students'
(3b) *studentischeres Lernen* 'learning the more student-like way'
(3c) *Peter ist/wohnt studentisch.* 'Peter is/lives like a student.'

Morphologically, some suffixes (especially *-isch, -lich, -ig*) have been more predestinated for the formation of DAs and RAs than others in the history of German (e.g. *-sam, -haft*; Zifonun 2011: 103). For present-day German (PDG), new, serial formations of present-day German of RAs (4a-b) or DAs respectively (4c) are based on semi-suffixes (cf. 1, 4a, 4b) or pseudo-compounding (4c) in PDG (Hotzenköcherle 1968, Kann 1974, Inghult 1975, Ruge 2004).

(4a) *lehrertechnisches Unterrichten* 'teachers' teaching', > Lehrer 'teacher', *-technisch*
(4b) *Arbeitsmäßig geht es ihm gut, aber beziehungstechnisch hat er nur Pech gehabt*. 'Work-wise everything is ok, but family-wise he has problems.'
(4c) *reinigungskraftgestütztes Putzen* 'cleaning staff supported cleaning'

The talk presents a corpus study of DAs derived by the semi-suffixes *-mäßig* and *-technisch* in PDG *Falter corpus* of Austrian Standard German, a subcorpus of Cosmas II (see https://www.ids-mannheim.de/cosmas2/) in order to investigate potential differences in the distribution of the two morphemes for the first time. The results show slightly different morphological properties between both morphemes as regards the particular word class of the base involved (verbs vs. nouns) and as regards the underlying word-formation processes within the nominal domain (especially compounding, derivation). As opposed to previous observations in the literature (cf. e.g. Duden 2016: 768), the results show that *-technisch* is more widely distributed than *-mäßig* while the latter is formally more restricted and tends to be attested with more QA-readings in PDG. As it will be argued, both developments are part of a bigger picture, namely of an ongoing process of grammaticalization in which *-technisch* is the more promising candidate in this rivalry which can be seen by the dynamics of its selectional properties.

Cited references:

Bally, C. (1944): Linguistique générale et linguistique française. Berne: A. Francke.

Duden (2016): Die Grammatik. 9. Aufl. Mannheim u.a.

Fábregas, A. (2007): The internal structure of relational adjectives. In: Probus 19/1, 1-36.

Frey, W. & Pittner, K. (1998): Zur Positionierung der Adverbiale im deutschen Mittelfeld. LB 176, 489-534.

Frevel, C. / Knobloch, C. (2005): Das Relationsadjektiv. In: Knobloch, C. / Schaeder, B. (eds.): Wortarten und Grammatikalisierung. Perspektiven in System und Erwerb. (Linguistik – Impulse & Tendenzen 12). Berlin / Boston: de Gruyter, 151–175.

Ganslmayer, C. (2012): Adjektivderivation in der Urkundensprache des 13. Jahrhunderts: Eine historisch-synchrone Untersuchung anhand der ältesten deutschsprachigen Originalurkunden. (Studia Linguistica Germanica 97). Berlin / Boston: de Gruyter.

Gunkel, L. / Zifonun, G. (2008): Constraints on relational-adjective noun constructions. A comparative view on English, German and French. In: Zeitschrift für Anglistik und Amerikanistik 56.3, 283–302.

Holzer, P. (1996): Das Relationsadjektiv in der spanischen und deutschen Gegenwartssprache. Wilhelmsfeld: Egert.

Hotzenköcherle, R. (1968): Gegenwartsprobleme im deutschen Adjektivsystem. In: Neuphilologische Mitteilungen 69, 1–28.

Inghult, G. (1975): Die semantische Struktur desubstantivischer Bildungen auf "-mässig": eine synchronisch-diachronische Studie. Stockholm: Almgvist & Wiksell.

Kann, H.-J. (1974): Belege zum Wortbildungsmuster Substantiv + technisch. In: Muttersprache 309-313.

Marchis, M. (2010): On the morpho-syntactic properties of relational adjectives in Romanian and Spanish. In: Bucharest Working Papers in Linguistics XII.1, 77–92.

Marchis Moreno, M. (2015): Relational adjectives as interfaces. In: Studia Linguistica 69.3, 304–332.

Motsch, W. (2004): Deutsche Wortbildung in Grundzügen. 2nd ed. Berlin: de Gruyter.

Mravlag, H. (2013): Relationsadjektive im Deutschen, Französischen und Russischen. Innsbruck: University Press.

Pittner, K. (1999): Adverbiale im Deutschen: Untersuchungen zu ihrer Stellung und Interpretation. Tübingen: Stauffenburg.

Ramaglia, F. (2011): Adjectives at the Syntax-Semantics Interface. München: Lincom.

Ruge, N. (2004): Das Suffixoid „-technisch“ in der Wortbildung der deutschen Gegenwartssprache. In: Muttersprache 1/2004, 29-41.

Schäfer, M. (2013): Positions and Interpretations. German Adverbial Adjectives at the Syntax-Semantics Interface. Berlin, Boston: Mouton de Gruyter.

Ten Hacken, P. (2019). Relational adjectives between syntax and morphology. In: SKASE Journal of Theoretical Linguistics 16(1), 77–92.

Trost, I. (2006): Das deutsche Adjektiv. Untersuchungen zur Semantik, Komparation, Wortbildung und Syntax. Hamburg: Buske.

Wilmanns, W. (1896): Deutsche Grammatik. Gotisch, Alt-, Mittel- und Neuhochdeutsch. Vol. 2: Wortbildung. Straßburg: Trübner.

Zifonun, G. (2011): Relationale Adjektive – ein „klassisches“ Muster im europäischen Vergleich. In: Deutsche Sprache 9.2, 98–112.

Jelena Zivojinovic

The so-called “gerund” in Ladin

The traditional view in morphology defines gerund as a specific form made of a root + morphological suffix derived from Latin. In this talk, I will claim that this etymological / historical definition no longer reflects the the present-day state in Ladin.

Firstly, traditional grammars, such as Chiocchetti & Iori (2002) signalize the presence of gerund (root + *-an/-en,* e.g. *chantan* ‘by singing’). However, their illustrated examples do not seem to match the traditional view of a gerund derived from Latin and surviving in its adverbial function.

1) L ’è vegnù co la mans scorl-an.
He AUX came with DET hand GER
‘He came empty handed’ (Fassano)

Gerunds in the example 1, despite its glossing, seems to be more likely to have an adjectival function, which is traditionally reserved for the present participle.

Secondly, Casalicchio (2011, 2013) claims that gerund in Ladin is limited to perceptive constructions, restricted to Badioto and Gardenese, the two northern varieties of Ladin, and exclusive to the verbs udëi (‘to see'), audì (‘to hear’), sentì (‘to smell’ or ‘to touch’).

2) Canche l dessënia uciei i aud i sculeies sciblan
when he.CL draws birds them.CL hear DET students whistle.GER
‘When he draws the birds, students can hear them whistling’
(Bels. 1) (Gardenese)

Thirdly, fieldwork data, collected in the areas of Val Badia and Val di Fassa, show that Ladin speakers have a strong preference towards using the traditional gerund only in the context of simultaneity.

3) Vigni dé vai al laur ciantan dedalt t l’ auto
every day go to work sing.GER out loud in DET car
‘Every day I go to work by singing out loud in the car’

To sum up, we find the very same formula (root + *-an*) used in the context of perceptive constructions, holding an adjectival function and holding an adverbial function in contexts of simultaneity. This syncretism forces us to reconsider the borders between non-finite forms. My claim is that gerund is not defined by a morphological marker, but it is rather a set of different functions. In my talk, I will present fieldwork data, along with previous works, in order to support this theory of rethinking non-finite verbal categories.

Selected references

Benincà, P. (1994). L'interferenza sintattica: di un aspetto della sintassi ladina considerato di origine tedesca. In: Benincà, P. *La variazione sintattica*. Bologna: Il Mulino, 89–103.

Casalicchio, J. (2011). *L'uso del gerundio con i verbi di percezione gardenesi*. Istitut Ladin Micurà de Rü.

Casalicchio, J. (2013). *Pseudorelative, gerundi e infiniti nelle varietà romanze: somiglianze (solo) superficiali e corrispondenze strutturali*. Munich: LINCOM.

Casalicchio, J. (2016). The use of gerunds and infinitives in perceptive constructions. *Theoretical Approaches to Linguistic Variation, 234*, 53-87.

Casalicchio, J. (2016). Ricostruire la diacronia della sintassi ladino-dolomitica con l'aiuto di Joppi. Il caso dei costrutti percettivi. In *Ad limina Alpium. VI Colloquium Retoromanistich, Cormons, dai 2 ai 4 di Otubar dal 2014* (pp. 97-126). Società filologica friulana.

Da Milano, F., & Ramat, P. (2011). Differenti usi di gerundi e forme affini nelle lingue romanze. *Vox Romanica*.

Salvi, G. (2000). Il ladino. Schizzo linguistico.