16th International Morphology Meeting (IMM 16)

Budapest, May 29 – June 1, 2014,

Research Institute for Linguistics (Hungarian Academy of Sciences)

*

ABSTRACTS

*

Main session	***May 29 – June 1***
Poster session	***May 31, 13:30-15:00***
Workshop 1: Compounds	***May 30, 14:00-15:30,*** ***June 1, 9:00-16:30***
Wokshop 2: Info-theoretic approaches	***May 30, 16:00-18:00,*** ***May 31, 10:10-12:30, 15:00-17:00***
Workshop 3: Inflection	***May 31, 10:10-12:30, 15:00-17:00,*** ***June 1, 10:10-12:30***

16th International Morphology Meeting (IMM 16)

Budapest, May 29 – June 1, 2014

Organizing committee:

Ferenc KIEFER (Hungarian Academy of Sciences)
Mária LADÁNYI (Eötvös Loránd University)
Huba BARTOS (Research Institute for Linguistics, H.A.S.)
Péter REBRUS (Research Institute for Linguistics, H.A.S.)

Program Committe:

Jóhanna BARÐDAL (Ghent University)
Martin MAIDEN (University of Oxford)
Ingo PLAG (Heinrich Heine University, Düsseldorf)
Gregory STUMP (University of Kentucky)
Ferenc KIEFER (Hungarian Academy of Sciences)
Wolfgang U. DRESSLER (Austrian Academy of Sciences)
Mária LADÁNYI (Eötvös Loránd University)
Franz RAINER (Austrian Academy of Sciences)
Hans Christian LUSCHÜTZKY (University of Vienna)
Huba BARTOS (Research Institute for Linguistics, H.A.S.)
Péter REBRUS (Research Institute for Linguistics, H.A.S.)

Assistant organizers:

Veronika HEGEDŰS (Research Institute for Linguistics, H.A.S.)
Kinga GÁRDAI (Research Institute for Linguistics, H.A.S.)
Ildikó Emese SZABÓ (Eötvös Loránd University)

INVITED SPEAKERS

When is a Nominative an Object?

Jóhanna Barðdal (Ghent University)*

In traditional grammar the nominative has been equated with subject, irrespective of argument structure and position in the clause. This includes nominatives of "inverse" predicates such as 'like' in languages with non-canonically case-marked argument structures. However, beginning in the 1960s with the general theorizing of grammatical structure, behavioral properties of subjects were identified. This led to the recognition that behavioral subjects could be non-canonically case marked, for instance in the accusative, dative and the genitive, and as a corollary that behavioral objects could be in the nominative case.

A long-standing division between Icelandic and German in the literature takes for granted that there are non-nominative subjects and nominative objects in Icelandic, like with the Dat-Nom case frame. At the same time, apparent corresponding structures in German have been analyzed as diametrical opposites, i.e. as exhibiting a Nom-Dat case frame (cf. Zaenen, Maling & Thráinsson 1985, Sigurðsson 1989). This alleged contrariety between Icelandic and German is in particular based on two factors:

a) different behavior of the dative with regard to a subset of subject tests across the two languages
b) the apparent subject behavior of the nominative argument in German

This presentation focuses on the second difference between Icelandic and German, i.e. the apparent subject behavior of the nominative argument in German Dat-Nom constructions and the contradiction in its behavior, alternating between behaving as a subject and an object. This involves an analysis of additional data relevant to grammatical relations and case marking, data that have not figured in the earlier literature and are vital for a deeper understanding of the overarching problem. These data involve alternating predicates, which behave in such a way that either argument, the dative or the nominative, may take on subject properties while the other argument of the argument structure behaves syntactically as an object, irrespective of its case marking. The verbal predicate, however, remains constant.

A comparison between Icelandic and German shows that Icelandic Dat-Nom predicates are of two types, a non-alternating and an alternating type, whereas German seems to exhibit only the alternating type. On this assumption, the apparent subject behavior of the nominative in German is easily explained, as such occurrences in fact involve the Nom-Dat construction and not the Dat-Nom construction. Hence, the subject behavior of the nominative does not invalidate a subject analysis of the dative in Dat-Nom constructions in German, only in Nom-Dat constructions. This, in turn, begs the question of which factors decide on which of the two constructions, Dat-Nom or Nom-Dat, are used with a given verb in a given context.

On a more general note, the concept of alternating predicates, i.e. predicates which systematically alternate between two diametrically-opposed argument structure constructions, in this case Dat-Nom and Nom-Dat, is new to both linguistic theorizing and empirical method. It will be shown how predicates like these can be accounted for in Sign-Based Construction Grammar (Sag 2012), including their case marking, the alternating argument structure constructions, and the object behavior of the nominative.

Sag, Ivan. 2012. Sign-Based Construction Grammar: An Informal Synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar*, 69–202. Stanford: CLSI Publications.
Sigurðsson, Halldór Ármann. 1989. *Verbal Syntax and Case in Icelandic*. Doctoral dissertation. Lund: Department of Scandinavian Languages.
Zaenen, Annie, Joan Maling & Höskuldur Thráinsson 1985, Case and Grammatical Functions: The Icelandic Passive. *Natural Language and Linguistic Theory* 3: 441–483.

*Based on joint work with Thórhallur Eythórsson and Tonya Kim Dewey.

Lexical Synonymy and the Diachrony of Paradigm Structure

MARTIN MAIDEN
Research Centre for Romance Linguistics, University of Oxford
martin.maiden@mod-langs.ox.ac.uk

From detailed observation of patterns of morphological change, mainly but not exclusively in Romance languages, I argue that such superficially disparate phenomena as analogical levelling-out of stem allomorphy, the diachronic persistence of morphomic alternation patterns and the genesis of 'incursive' suppletion invite unified explanation if we recognize the crucial role played in morphological change by lexical synonymy in interaction with the long-acknowledged tendency (e.g., Girard 1718; Clark 1993) for synonymy to be eschewed.

Lexical synonymy obviously exists across the word-forms of inflexional paradigms: indeed it serves as a classic (if obviously inadequate; cf. Spencer 2013) criterion for distinguishing 'inflexional' and 'derivational' morphology. Analogical levelling-out of root-allomorphy in inflexional paradigms can plausibly be assigned to synonymy-avoidance, such that multiplicity of forms sharing identical lexical semantic content is subject to synonymy-avoidance strategies, rooted in principles of iconicity, and favouring the survival of just one form ('one meaning - one form'). One of the striking diachronic characteristics of *morphomic* alternation patterns is that (if they are not themselves levelled out) they overwhelmingly display diachronic *coherence*: any morphological innovation affecting any one of the paradigm-cells over which a morphome is defined always, and identically, affects all the others. I argue that coherence reflects a strategy rendering maximally predictable, and thereby minimizing, the mismatch between multiplicity of form and unitary lexical meaning (i.e., synonymy). Crucial supporting evidence comes from the observation that otherwise wholly robust morphomic structures disintegrate where even slight lexical semantic differences emerge, i.e., where synonymy disappears.

Despite occasional denials that truly synonymous *lexemes* exist (e.g., Bloomfield 1933:145), if we admit Cruse's view of synonyms (2011:142) simply as 'words with construals whose semantic similarities are more salient than their differences' then there are various circumstances under which speakers may find themselves unable to distinguish the meaning of different lexemes. As argued in Maiden (2004) and elsewhere, the emergence of synonymy between historically distinct lexemes (a scenario particularly common under conditions of language contact — cf. Carstairs-McCarthy 2010:ch.4 — but also with a variety of other causes) is a common source of suppletion, the synonymy being resolved ('avoided') by merging or 'accommodating' the lexemes, effectively by assigning them (often, morphomic) complementary domains.

However, my view that lexical suppletion is determined by synonymy-avoidance (and is therefore of a kind with levelling and morphomic coherence) has been interestingly but, I submit, incorrectly, challenged by Börjars and Vincent (2011), who adduce Scandinavian adjectives meaning 'small' as an example of a suppletion (for number) which, allegedly, does not involve synonymy. This type is actually found beyond Scandinavian (occurring, for example, in Middle Breton and Middle Cornish), having a particularly interesting manifestation in Megleno-Romanian (where 'big' also shows suppletion). Through analysis of the Megleno-Romanian data I show that such behaviour in basic 'size' adjectives is in fact strong evidence *for* the role of avoidance of lexical synonymy in determining suppletion, even though that synonymy turns out to be 'local', in that it occurs only in plural cells. We shall see that this is not the only case where 'local' synonymy between lexemes is sufficient to result in suppletion.

In conclusion, I argue that avoidance of lexical synonymy drives and unites apparently disparate forms of morphological change. My conclusions thus very closely support the arguments about the role of synonymy-avoidance in the evolution of morphology developed by Carstairs-McCarthy (e.g., 2010). I shall particularly underscore, however, the role of *distributional predictability* in morphological change, suggesting — perhaps surprisingly — that appeal to the 'autonomy' of morphology in such cases may be open to question.

References

BLOOMFIELD, L. 1933: *Language*. New York: Holt • BÖRJARS, K. and VINCENT, N. 2011. 'The pre-conditions for suppletion', in Galani, A., Hicks, G. and Tsoulas, G. (eds) *Morphology and its Interfaces*. Amsterdam: Benjamins, pp. 239-65 • CARSTAIRS-MCCARTHY, A. 2010 *The Evolution of Morphology*. Oxford: OUP • CLARK, E. 1993: *The Lexicon in Acquisition*. Cambridge: CUP • CRUSE, A. 2011. *Meaning in Language. An Introduction to Semantics and Pragmatics*. Oxford: OUP • GIRARD, G. 1718. *La Justesse de la langue françoise, ou les différentes significations des mots qui passent pour synonymes*, http://visualiseur.bnf.fr/Visualiseur?Destination=Gallica&O=NUMM-50626 • MAIDEN, M. 2004. 'When lexemes become allomorphs: on the genesis of suppletion', *Folia Linguistica* 38:227-56; • SPENCER, A. 2013. *Lexical Relatedness*. Oxford: OUP

Homophony in morphology:
The acoustic properties of morphemic and non-morphemic word-final *s* in English

INGO PLAG
Heinrich-Heine-Universität Düsseldorf
ingo.plag@uni-duesseldorf.de

Recent research on lexemes has shown that homophonous lexemes show striking phonetic differences (e.g. Gahl 2008, Drager 2011), with important consequences for models of speech production such as Levelt et al. (1999). Such models would need to be adjusted to accommodate frequency effects at the lemma level and/or information of fine phonetic detail into lexical representations (e.g. Johnson 1997).

These findings also pose the question of whether similar problems hold for allegedly homophonous affixes (instead of free lexemes), or with affixes and phonologically identical non-morphemic segments. If the same kind of effects were to be found also with affixes it would seriously challenge traditional analyses. A good test case for this is English, which, in the traditional view, has a number of homophonous bound {s} morphemes. Thus, plural, genitive, genitive plural, the auxiliary forms of *has* and *is*, and the pronoun *us* (as in *let's*) can all be realized by /s/ (after voiceless segments) or /z/ after voiced segments (cf. plural *boys*, genitive singular *boy's*, genitive plural *boys'*, or cliticized *boy's*, see e.g. Bauer, Lieber & Plag 2013). The underlying assumption for such analyses is that the form of morphemes is lexically specified only phonologically, i.e. without phonetic specification.

The present investigation challenges this assumption. First, we test Walsh & Parker's (1983) hypothesis that plural /s/ and /z/ differ acoustically from non-morphemic final /s/ and /z/. Our data show a significant contrast between the two groups, with non-morphemic /s/ and /z/ being longer than the plural allomorphs. We then present an investigation of almost 600 words from the Buckeye corpus (Pitt et al. 2007) that additionally includes words with the other {s} morphemes mentioned above. We demonstrate that there are significant acoustic differences between at least some of these morphemes (especially in duration). Mixed effects regression models with a number of pertinent covariates (e.g. speech rate, position in the phonological phrase, frequency etc.) show that, for example, plural {s} is significantly longer than all other {s} morphemes.

At the theoretical level these findings challenge standard assumptions in morphological theory, Lexical Phonology and models of speech production.

References

BAUER, LAURIE, ROCHELLE LIEBER & INGO PLAG. 2013. *The Oxford reference guide to English morphology.* Oxford: Oxford University Press.

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

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

JOHNSON, KEITH. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson (ed.), *Talker variability in speech processing*, 145-165. San Diego, CA: Academic Press.

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

PITT, MARK, LAURA C. DILLEY, KEITH JOHNSON, SCOTT KIESLING, WILLIAM D. RAYMOND, ELIZABETH HUME & ERIC FOSLER-LUSSIER 2007. *Buckeye corpus of conversational speech*, 2nd release. Columbus, OH: Department of Psychology, Ohio State University.

WALSH, THOMAS & FRANK PARKER. 1983. The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics* 11. 201-206.

Polyfunctionality, positional exponence and affix conflation

GREGORY STUMP
University of Kentucky
gstump@uky.edu

The only grammatically significant relation that exists between inflectional markings and morphosyntactic properties is that of exponence (Stump 2001:11); but variable affix ordering and morphotactic conditioning reveal that affixes participate in three sorts of exponence relations: ***inherent***, ***positional*** and ***conflated***. The latter two relations contribute to the polyfunctionality of a language's inflectional morphology, as evidence from Swahili reveals.

In Swahili, the noun-class concord *vi-* is an inherent exponent of the property set {σ} = {{GEND:ki-vi, NUM:plural}}. But when it appears in affix position –3 (see (1) below), it is also a positional exponent of subject agreement{{ARG:sbj}}, as in (2a); and when it appears in affix position –1, it is a positional exponent of object agreement {{ARG:obj}}, as in (2b). Most of the Swahili noun-class concords exhibit this sort of positional exponence, one of the sources of polyfunctionality in Swahili verb inflection.

A related source is the phenomenon of ***affix conflation***, in which two affixes combine to form a complex affix whose content combines that of its component affixes and whose distribution may override the default morphotactics of one or both of these affixes. This phenomenon is exemplified by Swahili relative verb forms, which inflect for the gender and number of a relativized argument. A relative verb form's relative affix is (in most cases) a conflated affix consisting of two parts: (i) the noun-class concord appropriate for the gender and number of the relativized argument; and (ii) the affix *o*, an inherent exponent of relative agreement {{ARG:rel}}. Conflated relative affixes combine the content of these two parts; for example, the noun-class concord *vi-* (an inherent exponent of {σ} = {{GEND:ki-vi, NUM:plural}}) conflates with *o* to produce the relative affix *vyo*, a conflated exponent of {{ARG:rel} ∪ σ}. In the absence of overt marking for tense or negation, a relative affix is by default suffixal, appearing in affix position ±2 (cf. (3a)). But relative affixes themselves conflate with TN2 affixes (affixes of tense or negation associated with position ±2); these conflated affixes appear prefixally in position ±2 (cf. (3b)), overriding the relative affixes' default suffixal morphotactics. Thus, the Swahili noun-class concords are polyfunctional both because they serve as positional exponents of subject and object agreement and because they conflate with *o* to form relative affixes (which may in turn conflate with TN2 affixes).

How are inherent, positional and conflated exponence to be modeled theoretically? A reasonable approach is to distinguish declarations of affixal exponence from rules of affix sequencing (Stump 1993: 174f), the former being the domain of inherent exponence relations, and the latter, the domain of positional exponence relations; together, an exponence declaration and a sequencing rule have the effect of a rule of exponence of the sort familiar in inferential-realizational theories of morphology (Zwicky 1986, Stump 2001). Instances of affix conflation can, correspondingly, be seen as involving the conflation of distinct exponence declarations. Under such an approach, an exponence declaration is the pairing ⟦x, σ⟧ of an affix x with a morphosyntactic property set σ; thus, the Swahili affix *vi-* has the exponence declaration in (4a), which specifies its inherent exponence. This affix is subject to the sequencing rules in (4b), which specify its positional exponence; these rules are interpreted realizationally, as in (4c, d).

The exponence declaration in (4a) is also subject to conflation, where the ***conflation*** of two exponence declarations ⟦x, σ⟧ and ⟦y, τ⟧ is the exponence declaration ⟦xy, ***s***(σ, τ)⟧, whose operator ***s*** is as defined in (5a, b). Thus, according to the conflation rule in (6a), the conflation of ⟦*vi*, {{GEND:ki-vi, NUM:plural}}⟧ and ⟦*o*, {{ARG:rel}}⟧ (= ⟦*vyo*, {{ARG:rel, GEND:ki-vi, NUM:plural}}⟧) is a relative affix, whose sequencing is determined by rule (6b);

and according to rule (7a), the conflation of the TN2 affix ⟦*li*, {TNS:pst}⟧ and the relative affix ⟦*vyo*, {{ARG:rel, GEND:ki-vi, NUM:plural}}⟧ is a TN2 affix ⟦*livyo*, {TNS:pst, {ARG:rel, GEND:ki-vi, NUM:plural}}⟧, whose sequencing is determined by rule (7b).

This approach to modeling Swahili verb inflection correctly entails that by default, a subject-agreement affix has the same form as the corresponding object-agreement affix. Moreover, it correctly entails five facts about the relative affixes: (a) the relative prefixes have the same form as the relative suffixes; (b) the relative suffixes are mutually exclusive with the TN2 prefixes; (c) the relative suffixes are mutually exclusive with the relative prefixes; (d) the relative prefixes require a TN2 prefix; and (e) the relative prefixes are adjacent to the TN2 prefixes. This analysis makes it possible to say that each Swahili noun-class concord has a constant bit of content (of which it is an inherent exponent) but may express additional content according to its position or to its conflation with other affixes.

(1) Affix positions in Swahili verb morphology

negative prefix *ha-*	subject agreement	tense or negative prefix *si-*	object agreement	verb root	final vowel	relative suffix
–4	–3	±2	–1		1	±2

(2) a. *Vi-tabu* *vi-me-anguka.*
σ-book SBJ:σ-COMPL-fall.down
'The books have fallen down.'

b. *U-me-vi-ona* *vi-tabu?*
2SG.SBJ-COMPL-OBJ:σ-see σ-book
'Have you seen the books?'

(3) a. *vi-tabu* *a-vi-soma-vyo*
σ-book 3SG.SBJ-OBJ:σ-read-REL:σ
'the books which he reads'

b. *vi-tabu* *a-li-vyo-vi-soma*
σ-book 3SG.SBJ-PST-REL:σ-OBJ:σ-read
'the books which he read'

(4) a. ⟦*vi*, {{GEND:ki-vi, NUM:plural}}⟧ is a ***noun-class affix***. (read: the noun-class affix *vi-* is an inherent exponent of the property set {{GEND:ki-vi, NUM: plural}}.)

b. For any ***noun-class affix*** ⟦x, {σ}⟧,
Pref(⟦x, {{ARG:sbj} ∪ σ}⟧) is a sequencing rule for Slot –3;
Pref(⟦x, {{ARG:obj} ∪ σ}⟧) is a sequencing rule for Slot –1.

c. Interpretation: Pref(⟦z, τ⟧) applies to the pairing of a stem Y with a property set ρ only if ρ is an extension of τ; the result of applying Pref(⟦z, τ⟧) to ⟨Y, ρ⟩ is ⟨zY, ρ⟩. (Similarly, the application of Suff(⟦z, τ⟧) to ⟨Y, ρ⟩ is ⟨Yz, ρ⟩.)

d. Given two property sets σ, τ: σ is an ***extension*** of τ iff for each x ∈ τ, either x ∈ σ or x is a set such that for some y ∈ σ, y is an extension of x.

(5) a. Where σ and τ are morphosyntactic property sets, ***s***(σ, τ) is smallest well-formed extension of both σ and τ (and is undefined if no such property set exists).

b. Given two property sets σ, τ, τ is ***smaller than*** σ iff either (i) |σ| > |τ| or (ii) |σ| = |τ| and σ is an extension of τ but τ is not an extension of σ.

(6) a. The conflation of a ***noun-class affix*** ⟦x, σ⟧ and the ***relative affix*** ⟦*o*, {{ARG:rel}}⟧ is a ***relative affix***.

b. For every ***relative affix*** ⟦ x, σ ⟧, Suff(⟦ x, σ ⟧) is a sequencing rule for Slot ±2.

(7) a. The conflation of a ***TN2 affix*** ⟦x, σ⟧ and a ***relative affix*** ⟦y, τ⟧ is a ***TN2 affix***.

b. For every ***TN2 affi***x ⟦x, σ⟧, Pref(⟦x, σ⟧) is a sequencing rule for Slot ±2.

References

ASHTON, E. O. 1944. *Swahili grammar*. Essex: Longman. ● LIPPS, JONATHAN. 2011. *A Lexical-Functional analysis of Swahili relative clauses*. MPhil thesis, Oxford University. ● STUMP, GREGORY T. 1993. Position classes and morphological theory. In G. Booij & J. van Marle (eds.), *Yearbook of Morphology 1992*, 129-180. Dordrecht: Kluwer. ● STUMP, GREGORY T. 2001. *Inflectional morphology*. Cambridge: Cambridge University Press. ● ZWICKY, ARNOLD M. 1985. How to describe inflection. *BLS* 11, 372-386. Berkeley Linguistics Society.

MAIN SESSION

Hyponymy in simplex and complex formations and the distinction between skeleton and body

MARIOS ANDREOU
University of Patras
andreoum@upatras.gr

Hyponymy is a relation of inclusion and, in particular, it is the KIND/TYPE/SORT OF semantic relation among words (e.g. *whale* is a hyponym of *mammal*). Contrary to semantic relations such as synonymy, hyponymy is asymmetrical and hierarchical. That is, the relation which holds between *whale* and *mammal* (*whale<mammal*) is not the same as the relation between *mammal* and *whale* (*mammal>whale*).

The study of hyponymy with respect to the Morphology-Lexical Semantics interface is of paramount importance since the hyponymy test is used by scholars for the identification of the semantic head of the word. In this respect, *door knob* is headed by *knob* since the whole compound is a 'kind of' *knob* (*knob>door knob*). The way hyponymy works in morphology, however, (with the exception of Bauer, 1990) has not been studied in any detail. This is mainly due to the fact that a framework which would allow one to conduct such a research has been lacking.

In this paper, I enquire into the use of hyponymy with respect to simplex and complex configurations. In particular, I adopt the framework of Lieber (2004) and raise two questions. The first question is whether hyponymy relates the complete lexical-semantic representations of two items or only parts of them. On the latter assumption, hyponymy may hold between the grammatical skeletons in which categorial information is encoded or between the pragmatic bodies of two lexical items which consist of encyclopedic knowledge. The second question is whether hyponymy should be considered as a relation between words or as a relation among the meanings of words (for a discussion see Murphy, 2003). The former renders hyponymy a paradigmatic lexical relation and, as a result, it should be considered important for lexical organization. On the contrary, the latter shows that hyponymy is a relation among the things lexical items describe and it should, therefore, not be represented in the lexicon. That is, relations of hyponymy and hyperonymy are available as part of our knowledge of the world and should not be part of the organization of the lexicon.

In order to answer these questions, I first examine the way hyponymy relates simplex lexical items since the relations which hold between simplex items are also evident in the relations among complex items (Lieber, 2004) and show that categorial information is not relevant to hyponymy. Consider the lexical semantic representations of *flower* and *rose* in (1):

(1a) *flower*	[+material ([ ])]	(1b) *rose*	[+material ([ ])]
	<-animate>		<-animate>
	<function>		<function>
	{a bloom on a plant}		{a fragrant flower}

Given that hyponymy is an asymmetrical relation, it is not accurate to claim that it relates the skeletal parts of *rose* and *flower*, since the relation (if any) which holds between the skeletons of *rose* and *flower* is the same as the relation between the skeletons of *flower* and *rose*, i.e. both are [+material]. Rather, it seems that the relation of hyponymy holds between the bodies of the two items. That is, based on the information {a fragrant flower}, *rose* is considered a hyponym of *flower*. This is in accordance with the hierarchical and asymmetrical nature of hyponymy since the relation which holds between the bodies of *flower* and *rose* (*flower>rose*) is not the same as the relation between the bodies of *rose* and *flower* (*rose<flower*). That is, *flower* is a hyperonym of *rose* and *rose* is a hyponym of *flower*.

Hyponymy can also hold between members of distinct lexical categories (Lyons, 1977). By way of example, adjectives such as *sweet/bitter* and *happy/sad* can serve as hyponyms of nouns, namely *taste* and *emotion* respectively. In a similar vein, a number of countable nouns may be in a relation of hyponymy with other uncountable nouns. The noun *chair*, for example, is a hyponym of *furniture* (also *knife<cutlery*, *shirt<clothing*). Of importance is that the relation of hyponymy holds between the pragmatic bodies of two lexical items and not between the skeletons in which categorial information is encoded. That is, no hierarchical and asymmetrical relation can be established between *sweet* and *bitter* if we take into consideration that *sweet* is an adjective and *taste* a noun. In a similar vein, no hyponymic relation can be established between *chair* and *furniture* if we consider that *chair* is a mass noun and *furniture* a singular count noun.

Another argument in favor of the proposal that hyponymy relates the bodies of two items comes from the distinction between taxonomic and functional hyponymy (Miller, 1998). A *dog*, for example, can be a taxonomic hyponym of *animal* and a functional hyponym of *pet*. Of relevance to our discussion is that the relations of taxonomic and functional hyponymy are available to us as part of the pragmatic bodies of lexical items.

The proposal that hyponymy and hyperonymy are only established between the bodies of two items has implications for the way we classify complex configurations into headed and headless. In particular, Bauer (2010:167) argues that headless compounds can fail the hyponymy test in three ways: (a) they can fail to show a head element, (b) they may belong to a lexical category which is not the category of their head element, and (c) their head may be of the correct category, but with the wrong meaning. On the assumption that the lexical category of an element (skeleton) is not relevant to hyponymy, these claims should be reassessed.

In more detail, that the hyponymy test applies to the bodies of two items can account for the fact that compounds which exhibit a figurative meaning (e.g. metaphorical and metonymical compounds) fail the hyponymy test. A compound such as *hard hat* (taken from Lieber, 2009:69) can be both headed and headless depending on whether it denotes a 'kind of' *hat* (helmet) or a *person* (worker). Notice that whether *hard hat* is headed or headless depends on whether a relation of hyponymy can be established between the bodies of *hat* and *hard hat*; skeletal information is not relevant since there is no change in the word-class of the compound in any case.

Configurations in which an anti-intersection adjective, that is, an adjective which requires negation of the noun with which it combines, modifies a noun (e.g. Greek *psevd-o-profitis* 'false-LE-prophet') can also greatly inform our discussion. A similar problem arises in the case of denominal derivatives with the prefix *anti-*. Formations such as *anti-fascist* and *anti-hero* denote the 'opposite of X' and fail the hyponymy test (e.g. an *anti-fascist* is not a *fascist*). Of importance is that these formations fail the hyponymy test as a result of their non-heads, *psevdo-* 'pseudo, false, phony' in *psevd-o-profitis* and *anti-* in *anti-fascist*.

Given that hyponymy only relates the pragmatic bodies and not the grammatical skeletons of two items, the proposed analysis argues in favor of the idea that hyponymy should not be represented in the lexicon. Relations of hyponymy and hyperonymy are available as part of our knowledge of the world and should not be part of the organization of the lexicon.

References

BAUER, L. 1990. Be-heading the word. *Journal of Linguistics* 26, 1–31. • BAUER, L. 2010. The typology of exocentric compounding. In S. Scalise and I. Vogel (Eds.), *Cross-disciplinary issues in compounding*, pp. 167–176. Amsterdam: John Benjamins. • LIEBER, R. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press. • LIEBER, R. 2009. A lexical semantic approach to compounding. In R. Lieber and P. Štekauer (Eds.), *The Oxford handbook of compounding*, pp. 78–104. Oxford University Press. • LYONS, J. 1977. *Semantics*. Cambridge: Cambridge University Press. • MILLER, G. A. 1998. Nouns in WordNet. In C. Fellbaum (Ed.), *WordNet: an electronic lexical database*, pp. 23–46. Cambridge, MA: MIT Press. • MURPHY, M. L. 2003. *Semantic relations and the lexicon*. Cambridge: Cambridge University Press.

Contrastive co-reference cross-linguistically: the issue of polyfunctionality in word-formation

MARCO ANGSTER
University of Turin
marco.angster@gmail.com

In the languages of Europe we find complex formations – compounds or **prefixed** words, such as Italian *autodistruzione* or German *Selbstzerstörung* 'self-destruction' –, in which the semantic contribution of the **first** element (e.g. *self-*, *auto-*, *selbst-*) allows to manifest a relation of co-reference between arguments or non-argumental participants – in the proposed example: between the destroyer and the destroyed referent. Besides co-reference, elements like *auto-* suggest a contrast between the specific configuration of action and a more usual configuration in which, for example, the destroyer and what is destroyed are different entities. Because of these two main contributions (co-reference and contrast) I define the function expressed by these marks as **contrastive co-reference** (Angster 2012).

Apart from sparse works on single marks in specific languages (e.g. Mutz 2004 on Italian and French), little attention has been until now devoted to this kind of formations in a cross-linguistic perspective. König (2011), a pilot study based on a cross-linguistic sample, proposes two sets of so-called "reflexive compounds" comparing the functions of intensifiers with functions performed by the elements found as first constituent in such compounds. The first set corresponds to "adnominal reflexive compounds" in which co-reference is connected to the remarkability of a patient (*Selbstironie* 'self-irony' = usually irony is directed toward others); the second set is that of "adverbial reflexive compounds", in which the agent is remarkable (*Selbsthilfe* 'self-help' = usually the help is given by others).

First of all, data issued from 30 European languages (mainly Indo-European, but also from Finno-ugric, Uralo-altaic and Afro-asiatic languages and a language isolate, namely Basque) will support König's conclusion that the main source for contrastive co-reference marks are elements which have in syntax at least one of the functions typical of the formally heterogeneous class of intensifiers (König/Gast 2006).

Secondly I will show that, despite the existence of different strategies connected with the expression of contrastive co-reference, in no language it is possible to find a formal differentiation connected to the functions proposed in König's work.

Thirdly, besides the fact that formal differentiations in this realm are far from being clear-cut, I will claim that, if a formal differentiation arises between two productive word-formation strategies, this differentiation is connected to the semantic class which the second element of a complex formation belongs to. For example in German two competing marks exist: *selbst-* and *eigen-*. *Hand* in *eigenhändig* is an object in Croft's (2001) sense and a formation like *selbsthändig* does not exist and would not be well-formed. This conclusion is supported by other languages in which a competition between different marks exists. For example, beside German and several other Germanic languages, also most Slavic languages, Finnish and Estonian and Modern Greek show a similar contrast between two marks. However most part of the languages in the sample show only a mark related to the realm of contrastive co-reference. In these languages either the CC-marks are not compatible with base-words belonging to the class of objects (for example *auto-* in Romance languages and *self-* in English), or the same mark used with deverbal base-words of the class of actions is applied (see Lithuanian *savi-*, Hungarian *ön-*). An unavoidable difficulty in the study of the contrast between CC-marks and marks of what we could define "contrastive possession" is that the second function is more rare, the relevant processes are less productive and therefore the final lexemes often have a rather idiomatic meaning.

Fourthly, functions that are evident in syntax may be less evident in the realm of word-

formation and what is sufficient to lead to a formal differentiation in syntax may be just a different interpretation of a single formal strategy in word-formation. Defining a set of possible functions confronting related syntactic and morphological strategies might give a road-map for defining functions of affixes and word-formation processes. However, when one works on word-formation on a cross-linguistic basis, only an evident formal differentiation with stable form-function pairings in specific languages gives us a cross-linguistically reliable generalization about a function. In absence thereof, we should just speak of different interpretations of a wider function, even if these interpretations have a formal differentiation in the relevant paraphrases. In this context, contrastive co-reference appears to lack reliable formal differentiations in word-formation connected to different functions detectable instead through a syntactic paraphrasis.

Selected references

ANGSTER, M. (2012) *Marche di coreferenza contrastiva nella formazione delle parole. Una panoramica delle lingue d'Europa*. (unpublished PhD thesis)

CROFT, W. (2001). *Radical Construction Grammar*. Oxford : Oxford University Press.

GAST, V., HOLE, D., KÖNIG, E., SIEMUND, P., & TÖPPER, S. (2007). *Typological Database of Intensifiers and Reflexives. Version 2.0*. http://www.tdir.org.

KÖNIG, E. (2011) "Reflexive nominal compounds" Studies in Language 35:1. 112-127.

KÖNIG, E. & GAST, V. (2006). "Focused assertions of identity: A tipology of intensifiers". Linguistic Typology, 10, 223–276.

MUTZ, K. (2004). "Zur Argumentstruktur der deverbalen Ableitungen von auto-". In: M. HUMMEL and R. KAILUWEIT (eds), *Semantische Rollen in der Romania*. Tübinger Beiträge zur Linguistik 472 , 355–374. Tübingen: Narr.

Holistic properties of morphological constructions: Evidence from Akan

CLEMENT KWAMINA INSAIDOO APPAH
Department of Linguistics, University of Ghana
cappah@ug.edu.gh

In accounting for the properties of morphological constructions, one may adopt a source-oriented view where every property of the whole is expected to be accounted for in the parts. Alternatively, one may adopt a product-oriented view where the whole may have properties that do not necessarily emanate from the properties of the constituents (cf. Haspelmath, 1989; Zager, 1981).

Studies on Akan complex nouns have been largely source-oriented. Scholars have assumed that properties of the whole either emanate directly from the constituents, e.g., the syntactic category of complex nouns(Anderson, 2013; Anyidoho, 1990; Christaller, 1875; Obeng, 2009), or arrived at through processes that modify corresponding properties in the constituentse.g., tonal melody of complex nominals(Abakah, 2004, 2006; Anderson, 2013; Anyidoho, 1990; Obeng, 2009).

Constructionist approaches to morphology have stressed the fact that morphological constructions often have properties that do not emanate from those of their constituents. Such properties are said to be holistic or output properties of the constructions themselves. In this presentation, I will provide evidence of such holistic properties of complex nouns in Akan (Kwa, Niger-Congo) in the form of the syntactic category and tonal melody of Akan compounds as well as the semantics of a particular construction types which Appah (2013) calls personal attribute nominal constructions

First, I will show that the syntactic category of Akan compounds is a constructional property not depending on the syntactic category of the constituents. Thus, notwithstanding the syntactic category of the constituents, the Akan compound is invariably nominal. For example, all the compounds in (1) are nouns although there are two nouns (1a), a noun and an adjective (1b), a verb and a noun (1c), noun and a verb (1d) and two verbs (1e).

(1)	a.	*odwan*	'sheep'	*onini*	'male'	*odwanini*	'ram'	$[N\text{-}N]_N$
	b.	*etiri*	'head'	*bɔne*	'bad'	*etiribɔne*	'bad luck'	$[N\text{-}A]_N$
	c.	*di*	'to take'	*bea*	'place'	*dibea*	'rank'	$[V\text{-}N]_N$
	d.	*asɛm*	'matter'	*bisa*	'to ask'	*asɛmbisa*	'question'	$[N\text{-}V]_N$
	e.	*gye*	'receive'	*di*	'eat'	*gyedi*	'faith/belief'	$[V\text{-}V]_N$

Secondly, I show that the tonal melody of Akan compounds may not necessary be accounted for by tweaking those of their constituents. Dolphyne (1988) identified two types of compounds based on their surface tonal melodies. In the first, all the tone bearing units preceding the last constituent of the compound bear low tone, as in (2). In the second, the constituents of the compound retain their underlying tonal melodies.

(2)	a.	*àbùsùà-bɔ́*	*b.*	*bàkà-nú*	*c.*	*àdè-sùá*
		family-join		lagoon-stirring		thing-learn
		'being a family member'		'fishing in a lagoon'		'education, learning'

The first tonal melody is regarded as a defining feature of Akan compounds and is normally said to result from the spread of low from the left edge of the word. I argue that this tonal melody does not define compoundhood in Akan because this tonal melody characterises other non-compound lexical items like the personal attribute nominal constructions. As such, the tonal melody should be regarded as a lexical tonal melody and a holistic property of lexical(ized) items because it seems to be independent of the tonal melodies of the constituents.

Finally, I show that a large part of the semantics of personal attribute nominal constructions is a holistic property attributable to the constructional schema and not to the individual instantiating nominals. This is because, the constructions as a whole refers to a property predicated of the possessor of the body-part occurring as the first constituent in the construction. Thus, the construction as a whole only bears a non-direct semantic link to one of the constituents.

(3)	a.	*ani-ɛ-den*	b.	*a-ho-ɔ-fɛ*
		eye-be-hard		Pref-self-be-nice
		'haughtiness'		'beauty'

I will present a construction morphology modelling of the lexical tonal melody and the semantics of the personal attribute nominal constructions.

References

ABAKAH, E. N. (2004). *The segmental and tone melodies of Akan.* Ph.D. Dissertation, Norwegian University of Science and Technology, Trondheim.•ABAKAH, E. N. (2006). The Tonology of Compounds in Akan. *Languages and Linguistics, 17*, 1-33.•ANDERSON, J. C. (2013). Verb-internal compound formation in Akan. *Journal of West African languages, XL*(1), 89-104. •ANYIDOHO, A. (1990). *On Tone in Akan Compound Nouns*. Paper presented at the the 19th West African Languages Congress, University of Ghana, Legon. •APPAH, C. K. I. (2013). *Construction Morphology: Issues in Akan Complex Nominal Morphology*. PhD dissertation, Lancaster University, Lancaster. •CHRISTALLER, J. G. (1875). *A Grammar of the Asante and Fante Language called Tshi (Chwee, Twi) based on the Akuapem dialect with reference to the other (Akan and Fante) dialects*. Basel: Basel Evangelical Missionary Society.•DOLPHYNE, F. A. (1988). *The Akan (Twi-Fante) language: Its sound systems and tonal structure*. Accra: Ghana Universities Press.•HASPELMATH, M. (1989). Schemas in Hausa plural formation: product-orientation and motivation vs. source-orientation and generation. *Buffalo Working Papers in Linguistics, 89*, 32-74. •OBENG, S. G. (2009). Akan deverbal nouns. In S. G. Obeng (Ed.), *Topics in Descriptive and African Linguistics: Essays in Honor of Distinguished Professor Paul Newman* (pp. 96-110). Muenchen: LINCOM Europa.•ZAGER, D. (1981). On orientation or, constraining false analogy. *Linguistics, 19*(11-12), 1107-1131.

Interplay of agglutination, cumulation and overabundance: non-canonical case-number paradigm in Adyghe

PETER ARKADIEV
Institute of Slavic Studies of the Russian Academy of Sciences / Russian State University for the Humanities / Sholokhov Moscow State University for the Humanities, Moscow
peterarkadiev@yandex.ru

Circassian languages (Adyghe and Kabardian) belong to the North-West Caucasian family well-known for the complexity of its polysynthetic verbal morphology. By contrast, nominal morphology in Circassian languages is relatively simple. Putting aside possessive prefixation, adverbial and coordinative suffixes, the Circassian nominal paradigm is constituted by two numbers (singular and plural) and there cases (absolutive, oblique and instrumental). Cf. the paradigm from Kabardian (e.g. Colarusso 1992: 51–52; Kumakhov & Vamling 2009: 22–25) in table 1, showing that case suffixes are uniform across numbers and that number exponence is uniform across cases, implying a fully transparent agglutinative organization of inflection; the only complication arises with the instrumental, which is formally based on the oblique.

Table 1. The Kabardian nominal paradigm

	Singular	Plural
Absolutive	*-r*	*-xe-r*
Oblique	*-m*	*-xe-m*
Instrumental	*-m-ɕ'e*	*-xe-m-ɕ'e*

However, this almost ideal transparency and uniformity of the Circassian nominal paradigm is disrupted in the closely related Adyghe (e.g. Paris 1989: 166–167). In this language, both the structure of the paradigm and the inflectional material are identical to those shown in Table 1, but for one exception: the combination of values {plural, oblique} can be expressed, in addition to the agglutinative combination of the plural suffix *-xe* and the oblique suffix *-m*, by the cumulative suffix *-me*, or by the combination of the latter with the plural suffix, cf. table 2.

Table 2. The Adyghe nominal paradigm.

	Singular	Plural
Absolutive	*-r*	*-xe-r*
Oblique	*-m*	*-xe-m,* ***-me, -xe-me***
Instrumental	*-m-ɕ'e*	*-xe-m-ɕ'e*

From the point of view of morphological typology and morphological theory, the Adyghe paradigm in table 2 poses several problems, primarily because different kinds of exponents for the oblique plural cell occur in free variation thus constituting a non-trivial example of the understudied phenomenon of overabundance (Thornton 2012). Though cumulative exponents in otherwise "separatist" paradigms are well attested cross-linguistically (see e.g. Plank 1986, 1999), instances where both the "separatist" and the cumulative expression of the same bundle of morphosyntactic values coexist seem to be very rare or at least underdocumented. The situation in Adyghe is further complicated by the pleonastic combination *-xe-me*, where the value "plural" is expressed twice, first by the "separatist" plural suffix *-xe* and then by the cumulative oblique plural *-me*. Thus from the point of view of Canonical Typology (e.g. Corbett 2011), the Adyghe paradigm is not just non-canonical in that it deviates from the requirement of uniform composition of paradigm cells, but highly non-canonical in that it violates this requirement in several ways at once.

The Adyghe paradigm is problematic for realizational theories of morphology such as Paradigm Function Morphology (Stump 2001), which crucially incorporate the so-called Panini's Principle, whereby among morphological rules competing for the exponence of the same bundle of morphosyntactic features the one whose domain of application is more narrowly specified wins. Indeed, the description of the Adyghe declension should incorporate at least the following realization rules (the rules format is simplified for expository reasons):

(1) {number:plural} → *-xe*
(2) {case:oblique} → *-m*
(3) {number:plural; case:oblique} → *-me*
(4) {case:instrumental} → {case:oblique}*-ç'e*

According to Panini's Principle, rule (3) should always override rules (1) and (2), implying that the oblique plural forms in *-xe-m* would never be derived, unless the application of (3) is specified as optional. Even in this case, it remains unclear how the pleonastic exponence *-xe-me* can be derived, since the rule introducing *-me* should arguably belong to the same block of rules as *-xe*. Further problems are posed by the instrumental taking oblique as its stem: the instrumental plural does not inherit the overabundance of the oblique plural, since potential forms in **-me-ç'e* and **-xe-me-ç'e* are ungrammatical. This implies that the formulation of rule (4) should contain a stipulation that only the oblique suffix *-m* can serve as its input.

The Adyghe nominal paradigm clearly shows that just a single "deviant" cell in a paradigm can render the whole paradigm highly non-canonical and that the phenomenon of overabundance is especially problematic for morphological theory, first because it requires special treatment in formal frameworks, and second because, as in the case of Adyghe, it may bear the large part of responsibility for the non-canonicity of particular morphological paradigms.

References

COLARUSSO, J. (1992). *A Grammar of the Kabardian Language*. Calgary: University of Calgary Press.

CORBETT G.G. (2011). Higher order exceptionality in inflectional morphology. In: H.J. Simon, H. Wiese (eds.), *Expecting the Unexpected: Exceptions in Grammar*. Berlin: Mouton de Gruyter, 107–126.

KUMAKHOV M. & K. VAMLING (2009). *Circassian Clause Structure*. Malmö: Malmö University.

PARIS C. (1989). Esquisse grammatical du dialecte abzakh (tcherkesse occidental). In: B.G. Hewitt (ed.), *The Indigenous Languages of the Caucasus. Vol. 2. The North West Caucasian Languages*. Delmar, N.Y.: Caravan, 154–260.

PLANK FR. (1986). Paradigm size, morphological typology, and universal economy. *Folia Linguistica* 20, 29–48.

PLANK FR. (1999). Split Morphology: How agglutination and flexion mix. *Linguistic Typology* 3.3, 279–340.

STUMP G.T. (2001). *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

THORNTON A.M. (2012). Overabundance in Italian verb morphology and its interactions with other non-canonical phenomena. In: Th. Stolz et al. (eds.), *Irregularity in Morphology (and beyond)*. Berlin: Akademie-Verlag, 251–269.

The work has been supported by the Russian Foundation for the Humanities, grant No. 14-04-00580.

Multifunctionality in Icelandic morphology

KRISTJÁN ÁRNASON
The University of Iceland
kristarn@hi.is

Morphophonemics: stems, suffixes and ablauts

The morphophonemic structure of Icelandic words distinguishes between stems and endings. Thus nominal forms like *hest-ur* 'horse-NOM', *hest-a* 'horse-ACC.PL', *hest-um* 'horse-DAT.PL' show the stem *hest-* with the suffixes *–ur*, *–a* and *–um*. Importantly, the same structure of stems and endings is used in verbal conjugation: *drep-ur* 'kills' – *drep-a* 'they kill' - *drep-um* 'we kill' etc. The permitted inflectional affixes form a phonologically delimited set, and the forms are subject to systematic morphophonemic restrictions. Possible vowels are: /a/, /ɪ/, /ʏ/; possible consonants: /r/, /n/, /m/, /s/, /t/, /ð/. The endings are maximally disyllabic but sometimes segmentable by "fission" (cf. below). The morphophonemic restrictions are part of the language specific word phonological system (the "inventory", cf. McCarthy 2002:70) of Modern Icelandic. Ablaut and umlaut forms are learned as (related) word forms, and there is room for phonotactic principles, e.g. stating that certain endings call for u-umlaut: *land - lönd-um* 'land-DAT.PL' (**land-um*); *fara* 'go' *för-um* 'we go' (**far-um*).

Sample monosyllabic affixes

/-a/: <A,N,GEN.PL>: *dag-a* 'days-GEN', *systr-a* 'sisters-GEN', *góð-a* 'good-ACC.FEM';
<N,FEM,NOM,SG>: *stúlk-a* 'a girl'
<V,PRES.3PL.> *far-a* 'go', *kalla* '(they) call'
<V,INF.> *far-a* 'to go' *tala* 'to speak'

/-na/: <N,FEM,GEN,PL> *kirk-na* 'churches-GEN', *sagn-a* 'stories-GEN',

/-ra/: <A,FEM,GEN,PL>*góð-ra* 'good-GEN.PL'

/-ar/ <N.Nom,Pl> *hest-ar* 'horses'
<N,Gen,Sg> *vin-ar* 'friends'
<V,2./3.Sg, Pres> *tal-ar* 'speak(s)'

/-ɪ/: <N, NOM,MASC> *han-i* 'cock'
<A, NOM,SG,DEF> *góð-i* 'good-DEF'
<V,1.SG,PRES> *dæm-i* 'I judge'
<V, 1./3.SG, PRES,SUBJ> *tak-i* 'take-SUBJ'

/-ɪn/: <PARTICIPLE>: *kom-in* '(have) come'
<N, MASC,ACC,DEF>: *hest-inn* 'the horse-ACC'

/-ɪr/: <N,A, NOM,PL> *gest-ir* 'guests', *góð-ir* 'good-NOM.PL'
<V,2.SG,PRES> *dæm-ir* 'judges'
<V,2.SG,PRES, SUBJ> *tak-ir* 'you take-SUBJ'
<V,2.SG,PRES> *kallað-ir* 'called'

/-ʏ/: <N,NEUT, PL,-GEN> *hjört-u* 'hears-PL'
<N,FEM,SG,-NOM> *stúlk-u* 'girl-OBLIQUE'
<A, NEUT,SG,-NOM> *góð-u* 'good-NEUT.OBLIQE'
<V, 3.PL,PAST> *fóru* 'they went'

/-ʏm/: <N,A,DAT,PL> *hest-um* 'horses-DAT', *ferð-um* 'travels-DAT', *kert-um* 'candles-DAT'
<A,Dat.Sg/Pl> *góð-um* 'goodDAT'
<V,1.PL> *ber-um* 'we carry', *fór-um* 'we went', *kölluð-um* 'we called'

/-s/: <N,GEN,SG,MASC/NEUT> *hest-s* 'horse'

/-t/ <A,NEUT,SG> *gul-t* 'yellow-NEUT', *víst* 'certain-NEUT'

<V,PARICIPLE> *fær-t* ‘moved’

/-ið/ <N,NOM,NEUT,DEF> *barn-ið* ´the child’

<V,PARTICIPLE > *far-ið* ‘gone-NEUT’

/-að/ <V,PARICIPLE> *kall-að* ‘called-NEUT’

Sample polysyllabic (and segmentable) affixes

/ʏr-ɪn/ <N,NOM,SG,MASC> *hest-ur-inn* ‘the horse’

/ɪn-a/ <N,FEM,ACC,DEF> *bók-in-a* ‘the book’

/að-ir/ <V,PAST,2.SG> *kall-að-ir* ‘you called’

Sample ablaut patterns

u-mlaut:

/a/ – /ö/: *land* ‘land-SG’– *lönd* ‘land-PL’, *svartur* ‘black-MASC’ – *svört* ‘black-FEM’

Double exponence: *fara* ‘go’ – *för-um* ‘we go’

i-umlaut:

/a/ – /ɛ/: *taka* ‘to take’– *tekur* ‘takes’

/œ/ - /ɛ/: *köttur* ‘cat’ – *kettir* ‘cats’

Indo-European ablaut:

/i/- /ei/: *líta* ‘to look’– *leit* ´looked’, *fara* ‘go’ – *fór* ‘went’

Morphological markers and disambiguation

The facts summarised above are either not handled or clumsily by generative theories, where endings are inserted into syntactically generated trees (e.g. Halle & Marantz 1993). Thus Rögnvaldsson (1990) assumes 55 noun declensions, whereas Müller (2005) reduces the number by allowing syncretism within certain “domains”, but not between unrelated paradigms, leaving out verbal conjugation. A generalisation is obviously missed which is that a limited set of forms and ablaut relations sampled above can be used across word classes as exponents of abstract mophemes or morphomes (Aronoff 1994). I want to suggest that the suffixes and vowel relations are diacritics added to stems, which means that their function depends on the context, and they do not as such have meaning. The meaning is a morphosyntactc/semantic property derived from the text. The disambiguation of the meanings follows the same principles as any sort of resolution of polysemy, i.e. by reading the context. Thus /-ir/ in the context of a noun stem: *gest-ir* ‘guests’ is interpreted as NOM.PL, and in the context of a verb stem: *dæm-ir* ‘judges’, as 2.SG, much in the same way as the form *á* in the context *á hest* is interpreted as a preposition governing the accusative case: ‘onto a horse-ACC’ (movement), or as part of a verb phrase ‘owns a horse’, but in *á hesti* ‘on horse-DAT’, the meaning is stative. In certain cases the interpretation and disambiguation is segmented (cf. “fission” in Distributed Morphology, see Müller 2005:247 and references), e.g. between PAST and 2.SG in *kall-að-ir* ‘you called’ (*Þú kallaðir* ‘You called’). Here the derived past tense stem *kall-að* (root+affix) takes the ending /-ir/, but the same form occurs in the inflected participle *kall-að-ir* ‘called-PL.MASC’ (*Þeir eru kallaðir* ‘They are called’).

This approach to the analysis of morphological multifunctionality seems to fit into the general framework of usage based grammar, according to which “constructions are the basic units of morphosyntax”, and that grammar is “emergent from experience, ever coming into being rather than static, categorical and fixed” (Bybee 2006: 714).

References:

ARONOFF, M. 1994. *Morphology by Itself*. MIT Press, Cambridge, Ma. • BYBEE, J. 2006. From usage to grammar: The mind‘s response to repetition. *Language* 82:711-733. • HALLE, M. & A. MARANTZ 1993. Distributed Morphology and the pieces of inflection. In: K. HALE and S. J. KEYSER (EDS.) *The View from Building 20*, Cambridge Ma: MIT Press. 111-176. • MÜLLER, G. 2005. Syncretism and Iconicity in Icelandic noun declensions: a Distributed Morphology approach. *Yearbook of Morphology* 2004: 229-271. • RÖGNVALDSSON, E. 1986. Íslensk orðhlutafræði. Reykjavík:Háskóli Íslands.

Motivation in the constructionist lexicon

JENNY AUDRING & RAY JACKENDOFF
University of Amsterdam,Tufts University
j.audring@uva.nl, ray.jackendoff@tufts.edu

Many models of morphology and the mental lexicon manage the amount of redundancy among lexical items by means of inheritance. Typically, inheritance is associated with the "impoverished entry" hypothesis, according to which lexical entries are specified for their idiosyncrasies only -- predictable properties are inferred from nodes higher up in the lexical network. With the help of inheritance, lexical storage is modelled as maximally parsimonious and redundancy-free.

While inheritance is proving useful in computational theories and implementations, psycholinguistic evidence pleads against the assumption that parsimony is a priority of the mind. For this and other reasons, a number of theories have embraced a "full entry" model of the lexicon (first proposed by Jackendoff 1975), stipulating that any lexical entry can be represented with its complete set of properties. A recent full entry theory of morphology is Construction Morphology (Booij 2010).

In a full entry lexicon, inheritance does not serve the purpose of reducing redundancy. Instead, inheritance relations are assumed to ***motivate*** complex words by capturing the non-arbitrariness of their properties and structure (Booij 2012). Thus, a word such as *sleepless* is motivated by the $[X\text{-less}]_A$ schema as well as by its base noun *sleep* (cf. Booij 2010: 26).

In this talk, we suggest that in a full entry lexicon, the notion of inheritance can be enriched to model a wider variety of lexical relations in a uniform way. Taking seriously the assumption that inheritance links capture shared information among items, we wish to draw attention to two theoretical limitations in the usual construal of inheritance:

- inheritance is asymmetric
- inheritance proceeds 'vertically' from a more general to a more specific node.

Instead, we propose that:

- inheritance is bi-directional
- inheritance can also work 'horizontally' between lexical items of equal complexity.

The clearest advantage of this approach lies in the modeling of sister-sister relationships in the lexicon. In a full entry lexicon, several types of relations are recognized between lexical items. The first is the general-to-specific relation that holds between schemas and words. The second is the part-whole relation between a base word and a related complex word. The standard view of inheritance models these two types of relation. However, it does not capture the third, which subsumes the paradigmatic relations between members of inflectional paradigms and among word families, in particular in paradigmatic word formation (Koefoed and Van Marle 1980). In more recent work, paradigmatic relations have also been posited as obtaining between schemas, under the rubric of “second-order schemas” (Nesset 2006, Kapatsinski 2013).

Standard inheritance falls short in this domain of paradigmatic relations between items of equal complexity, since such relations are typically symmetrical. There is often no principled way to tell which of a pair of sisters inherits from which. This point is particularly important for root derivations such as English *ambitious* or *amputate*, which lack a lexical base but nevertheless form (small) families with other words such as *ambition* and *amputee*. Such clusters are

notoriously challenging, in particular for word-based morphological theories. Since *ambit-* and *amput-* are neither base words nor seem to be candidates for a schema, the relation among the words in their families is difficult to state.

A similar problem recurs in second-order schemas such as <[X-ism]$_N$ ≈ [X ist]$_N$ > (Booij and Masini in preparation), instantiated by words such as *socialism* and *socialist*. While the schema expresses the fact that many *-ism* nouns have *-ist* counterparts and vice versa, the relation is nondirectional and non-hierarchical, and therefore not covered by traditional inheritance.

Here we suggest that this variety of lexical relations can be subsumed under an enriched notion of inheritance based on a domain-general cognitive relation proposed in Culicover & Jackendoff 2012. This relation is called SAME-EXCEPT and expresses the notion 'item X is the same as item Y except in respect Z.' It is manifested in a wide range of perceptual and conceptual phenomena, from vision to music to language. SAME-EXCEPT has two variants. In the first, ELABORATION, item X has an extra component that item Y lacks. In the second, termed CONTRAST, item X and item Y differ in a single property.

'Top-down' relations in morphology, such as between the schema [X-less]$_A$ and its instantiation *sleepless* can be characterized as SAME-EXCEPT ELABORATION: the schema is enriched by replacing the variable with a constant. Moreover, we can state that *sleepless* is an elaboration of its base word *sleep*. Thereby, we can avoid having to say that *sleep* is more general or situated 'higher up' in the lexicon than *sleepless*, a slight embarrassment that arises in a traditional inheritance-type explanation.

For cases of 'horizontal inheritance', the notion SAME-EXCEPT CONTRAST offers a precise characterization. Paradigm sisters, root family members, and the parts of second-order schemas are best treated as non-directional relations consisting of shared and contrasting properties. Here, SAME-EXCEPT CONTRAST captures the intuition that sister items in the lexicon motivate each other semantically and structurally, without the existence of a higher-level mother schema.

Thus SAME-EXCEPT enables us to capture the enriched notion of inheritance, unifying various types of lexical relation. Moreover, since it is a domain-general relation, it incurs no extra cost to morphological theory per se.

References

BOOIJ, GEERT (2010). *Construction Morphology*. Oxford: Oxford University Press.

BOOIJ, GEERT (2012). *Inheritance in Construction Morphology*. Paper presented at the workshop on Default Inheritance, U Kentucky, Lexington Ky, May 21-22.

BOOIJ GEERT, and MASINI, FRANCESCA (in preparation). *The role of second order schemas in the construction of complex words*. ms.

CULICOVER, PETER W., and JACKENDOFF, RAY (2012). Same-Except: A domain-general cognitive relation and how language expresses it. *Language* 88, 305-340.

JACKENDOFF, RAY (1975). Semantic and morphological regularities in the lexicon. *Language* 51, 639-671.

KAPATSINSKI, VSEVOLOD (2013). Conspiring to Mean: Experimental and Computational Evidence for a Usage-Based Harmonic Approach to Morphophonology. *Language* 89,110-48.

KOEFOED, GEERT and VAN MARLE, JAAP (1980). Over Humboldtiaanse taalveranderingen, morfologie en de creativiteit van taal. *Spektator* 10: 111–147.

NESSET, TORE (2006). Second-Order Schemas and Active Subschemas: Vowel Reduction in Cognitive Grammar. *Poljarnyj Vestnik* 9.

Floating Morphological Paradigms in Seri

MATTHEW BAERMAN
Surrey Morphology Group, University of Surrey
m.baerman@surrey.ac.uk

The inflectional class system of verbs in Seri (a language isolate of Sonora State, Mexico; Marlett 2009-11) displays a degree of complexity that rivals or surpasses any that has ever been described. This complexity is concentrated in the four-cell suffix paradigm that distinguishes subject number (singular ~ plural) and aspect (perfective ~ imperfective, or effectively single event ~ multiple event). Each cell displays considerable allomorphy (1), with the result that in a corpus of 1001 lexemes taken from Moser & Marlett's (2010) dictionary, there c. 244 distinct paradigm types. There is very little internal predictability between the allomorphs (2). Semantic, syntactic, phonological and morphological classes play little or no observable role in determining the form of the suffixes. Compounding this cross-lexemic unpredictability, few of the suffixes have a fixed function; e.g. in (3a) the distribution of the suffixes -(*t*)*ox* and -(*t*)*am* is the reverse of (3b) (The presence or absence of suffix-initial /t/ is phonologically predictable; Marlett 2009-11). On the face of it, it looks as if the four basic forms must be lexically listed.

The nature of the system is revealed by comparing the verbal paradigm to the nominal paradigm. The verbal suffixes are nearly all drawn from the repertoire of nominal plurals (4). The reason for this is that both subject plurality and event plurality (imperfective aspect) are treated morphologically as simply as plural, their markers drawn from a common set of plural markers, so that there is no overarching morphological distinction between three of the four cells of the paradigm. Nevertheless, a certain systematicity can be observed. Given the morphosyntactic hierarchy:

SG PFV ~ SG IPFV ~ PL PFV ~ PL IPFV

there are more or less rigid suffixal hierarchies that map onto this, but differ in their entry point. E.g. the suffixal hierarchy *-t ~ -(t)ox ~ -(t)ołka* maps onto both (5a) and (5b), but the entry point for (5b) is 'later', and thus the paradigm only gets as far as *-(t)ox*. These differing entry points account for a portion of the complexity of the system. Since the hierarchy has an endpoint, some local implications emerge. E.g. the suffix *-(t)ołka* is the end of the line, thus whatever point it enters the paradigm, it must be retained to the end, so that it appears in the last (6a), last two (6b) or last three cells (6c) of the paradigm.

This shows that in order to find the underlying organization of the Seri paradigm, we need to completely separate the morphological paradigm from the morphosyntactic paradigm (Sadler & Spencer 2001, Stump 2006). The morphological paradigm can then be described as a series of ordered forms without reference to the morphosyntactic values. The system is still complex, but within a typologically familiar range. The real complexity emerges when the morphological paradigm is mapped onto the morphosyntactic paradigm, an operation which itself falls into various subtypes. The resulting mismatch, though allowed by a realizational framework, is quite remarkable typologically.

(1) Examples of suffixal allomorphy

	SG PFV	SG IPFV	PL PFV	PL IPFV
'make small'	-ataasit-ot	-ataasit-im	-ataasit-x	-ataasit-ołka
'make into a fence'	-oots	-ootos-im	-oots-o	-ootos-am
'be pregnant	-iiket	-iikt-o	-iikt-ox	-iikt-ołka
'be connected'	-iti	-iti-tim	-iti-koł	-iti-xam
'cover o.s.'	-ak^{w}t-im	-akot-im	-ak^{w}t-am	-ak^{w}xt-ox

(2) Seri compared to the languages discussed by Ackerman & Malouf (2013)

Language	average conditional entropy
Burmeso	0.000
Kwerba	0.428
Ngiti	0.484
Fur	0.517
Russian	0.538
Arapesh	0.630
Greek	0.644
Mazatec	0.709
Nuer	0.750
Amele	1.105
Seri	2.37
	1.96 (with frequency factored in)

(3)

	SG PFV	SG IPFV	PL PFV	PL IPFV
a. ‘shoot’	-axkapt	-axkapot-im	-axkapot-**ox**	-axkapot-**am**
b. ‘lap up’	-oaala	-oaala-tim	-oaala-**tam**	-oaala-**tox**

(4)

		SG PFV	SG IPFV	PL PFV	PL IPFV
verb	‘go away’	-ntita	-ntita-**t**	-ntita-**tox**	-ntita-**tołka**
		SINGULAR	PLURAL		
	‘plant’	ʔeʔe	ʔeʔe-**t**		
nouns	‘estuary’	χtaasi	χtaasi-**tox**		
	‘camp’	ikaʔeme	ikaʔem-**tołka**		

(5)

	SG PFV	SG IPFV	PL PFV	PL IPFV
a. ‘go away’	-ntita	-ntita-**t**	-ntita-**tox**	-ntita-**tołka**
b. ‘be used up’	-eeme	-eeme	-eeme-**t**	-eeme-**tox**

(6)

	SG PFV	SG IPFV	PL PFV	PL IPFV
a. ‘go away’	-ntita	-ntita-t	-ntita-tox	-ntita-**tołka**
b. ‘die’	-ooxi	-ooxya-t	-ooxya-**tołka**	-ooxya-**tołka**
c. ‘make soft’	-aʔixpχ-ax	-aʔixpχ-**ołka**	-aʔixpχ-**ołka**	-aʔixpχ-**ołka**

Nominals with the suffix *–ač* in Croatian

MATEA BIRTIĆ
Institute of Croatian Language and Linguistics
mbirtic@ihjj.hr

Although the interplay between nominal and verbal properties in deverbal and similar nouns has been widely studied and recognized as an important morphological issue in the literature and within different theoretical research, I will try to approach this interesting topic by investigating the derivation of agentive/instrumental nouns with the suffix *-ač* in the Croatian language.

The suffix *-ač* is mostly a deverbal suffix from which nouns with agentive and/or instrumental interpretation are derived. Three main properties of the suffix *-ač* are commonly known (Babić 1991). Firstly, *-ač* mainly selects deverbal bases; secondly, it is mostly attached to imperfective bases and thirdly, its interpretation varies between agentive and instrumental.

Birtić (2008) describes a fourth property of the suffix *-ač*: similar to the English suffix *-er*, it can attach to unergative intransitive bases without restrictions, and mostly not to unaccusative intransitive bases (although it can attach to all transitive bases):

1. pliv-ač ‘swimmer’
2. *ton-ač ‘one who sinks’

It has been claimed that English *-er* nominals correspond to the external argument of the base verb (Fabb 1992, Rappaport Hovav & Levin 1992, Marantz 2001), but other research shows that this is not completely accurate (Lieber 2004).

Firstly, I will try to determine, on the basis of a large corpus (*Croatian Language Repository*), the correctness of the claim that the suffix *-ač* selects only verbs with the external arguments. Apparently, there are counterexamples, such as the noun *padač* 'one who falls down', and I will attempt to identify other meanings of nouns with this suffix.

Secondly, my aim is to investigate whether the aspect of the verbal base (perfective or imperfective) has any impact on the interpretation of nouns with *-ač* (agentive, instrumental or other). Although the suffix *-ač* mostly selects imperfective verbal bases, there are some exceptions to the rule:

3. osigurač ‘fuse’ < osigurati (perf.) ‘secure, insure’

The fact that the lexical aspect of the base verb is recognizable in the derived noun (but does not have the meaning 'perfective' or 'imperfective' as in the case of *-nje* nouns) and that the categories of perfectivity/imperfectivity play a role in the selectional restrictions of this suffix will serve as guidelines for the coming discussion of the syntactic vs. the lexicalist approach to morphology.

References

BABIĆ, S. 1991. *Tvorba riječi u hrvatskom književnom jeziku.* Zagreb: Hrvatska akademija znanosti i umjetnosti – Globus • BIRTIĆ, M. 2008. *Unutarnja struktura odglagolskih imenica u hrvatskome jeziku.* Zagreb: Institut za hrvatski jezik i jezikoslovlje. • FABB, N. 1988. English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory*, 6, 527–539. • LIEBER, R. 2004. *Morphology and Lexical Semantics*. Cambridge: Cambridge University Press. · MARANTZ, A. 2001. *Words*, abstract, WCCFL 20, USCL, 2/2001. •RAPPAPORT HOVAV, M. & B. LEVIN 1992. *-ER* Nominals: Implications for the theory of Argument structure. *Syntax and Semantics 26,* 127–153.

The role of second order schemas in word formation

GEERT BOOIJ
Leiden University
g.e.booij@hum.leidenuniv.nl

Word formation can be characterized in terms of schemas that specify the systematic relationships between form and meaning of complex words. These schemas are based on paradigmatic relationships between sets of words. For a proper characterization of word formation we also need second order schemas, that is, schemas that are paradigmatically related. For example, the relation between English nouns of the form [x-*ism*] and those of the form [x-*ist*] is such that the meaning of the words in *–ist* is often a compositional function of the words in *–ism*, even thought the morpheme *–ism* is lacking in the words in *–ist*. Hence we need a second order schema of the following type (Booij 2010):

(1) $<[\text{x-ism}]_{Ni} \leftrightarrow \text{SEM}_i > \approx < [\text{x-ist}]_j \leftrightarrow [\text{person involved in SEM}_i]_j>$

The necessity of second order schemas has also been argued for by Kapatsinski (2013) and Booij (2010, to appear) for inflection, and by Nesset (2008) for phonology.

In my presentation, I will present various types of evidence for the necessity of second order schemas for a proper account of word formation processes., with a focus on Dutch word formation processes such as the formation of elative compounds and the nominalization of particle verbs.

References

BOOIJ, GEERT. 2010. *Construction morphology*. Oxford: Oxford University Press.

BOOIJ, GEERT. to appear. Construction Morphology. In: A. Spencer and A. Zwicky (eds.), *The Handbook of Morphology*, second edition. London: Blackwell [downloadable from http://geertbooij.word.press.com].

KAPATSINSKI, VSEVOLOD. 2013. Conspiring to mean: experimental and computational evidence for a usage-based harmonic approach to morphophonology. *Language* 89.110-48.

NESSET, TORE. 2008. *Abstract phonology in a concrete model: Cognitive linguistics and the morphology-phonology interface*. Berlin: Mouton de Gruyter.

Incidental acquisition of (ir)regularity from syntactically simple and complex texts in L1 and L2 German

DENISA BORDAG, AMIT KIRSCHENBAUM, ANDREAS OPITZ, MARIA ROGAHN, ERWIN TSCHIRNER
Leipzig University
denisav@uni-leipzig.de

Though L2 grammar acquisition has been in the focus of researchers for several decades, surprisingly little is known about its incidental acquisition during reading. It has been claimed that incidental acquisition substantially contributes to the vocabulary growth both in L1 and L2 (Grabe, 2009). However, despite the general consensus that "knowing a word involves more than being able to recall the meaning (or L1 translation) of a presented word form" (Nation, 1994, p. 121), little research has gone beyond the acquisition of the meaning of new words. In the present study, we use a novel experimental paradigm to explore the incidental acquisition of (ir)regularity status of novel verbs in syntactically simple and complex contexts. The results are compared to those on acquisition of subcategorization (SC) - transitive vs. intransitive verbs. Differences in the result patterns of L1 and L2 participants as well as the influence of syntactic complexity are discussed.

Method: 72 advanced learners of German and 80 native Germans read 20 short texts (plus fillers), each with one occurrence of a novel verb in infinitive and two in preterit (e.g. *belfen - balf*). The novel verb (constructed according to German phonotactic rules) was replacing a low frequency existing German verb and its meaning could be derived from the context. Half of the novel verbs were conjugated regularly, the other half irregularly. The syntactic complexity of the texts was manipulated yielding two different versions of otherwise identical texts: one syntactically simple and the other syntactically complex (the latter containing longer sentences with more embedded structures). After each text, several additional sentences were presented. One of them contained the novel verb in perfect tense, conjugated either consistently with the forms occurred in the text (e.g. irregular *hat gebolfen*, plausible condition), or differently (e.g. regular *hat gebelft*, implausible condition). Participants read these sentences in a self-paced manner with a moving window. The L1 and L2 experiments on SC were constructed analogically: The texts contained intransitively used novel verbs and in the implausible self-paced reading condition these verbs were then presented as transitive with the direct object “es” (it).

Results and Discussion: In the ***SC experiment***, both L1 and L2 participants acquired the SC frame from the texts. Reading times were slower in the implausible condition reflecting participants’ sensitivity to the SC violation. The implausibility effect in the L2 experiment appeared later (at the spill-over region) and only if the novel verb was presented in a syntactically complex context. In the ***Irregularity experiment***, the results of the *L2 participants* are similar to those on SC: The implausibility effect was observed only in the complex condition when the novel word was presented as irregular. We assume that when syntactic complexity leads to difficulties in reading comprehension (as in the case of the L2 learners), it triggers “conscious registration of attended specific instances of language”, which corresponds to Schmidt’s (2012:32) definition of noticing. We further assume that noticing can be triggered also by irregular or otherwise phonologically or morphologically conspicuous word forms. For advanced learners who are familiar with the German verb system, morphological irregularity can trigger noticing, since advanced learners know that acquisition of irregular verbs requires additional effort including memorization. Being thus sensitized to irregular verbs as a “learning problem”, learners attend to them more closely when encountering them during reading. Along these lines, both the claim that irregular forms are salient and are not salient can

be valid: They are salient in the sense that they are morphologically conspicuous to learners on advanced levels and thus attract their attention (cf. Godfroid & Uggen, 2013 for beginning learners), but they are possibly not salient with respect to the particular phonological changes that occur on their stems. The pattern of the results of the *L1 participants* gave no indication that they acquired the regularity status of the novel verbs in the texts since reading times were not slower in the implausible condition in general. Instead, they responded with slower reading times and higher outlier rates whenever the presented novel verb form in the self-paced sentence was an irregular one – irrespective of the regularity status of the verb in the previous text. This sharply contrasts with the results on SC, when the L1 participants reacted with slower reading times in the spillover regions when the SC frame of the novel verb was violated in the implausible condition compared to how the novel verb was presented in the text. We assume that this difference is due to the differences in the linguistic characteristics of the two grammatical features: Different theoretical frameworks converge on the assumption that regular and irregular verbs are stored and processed differently and assume that the regular conjugation is the default type. So called "dual-mechanism models" implement this idea in form of two principally different systems: The default system splits regular forms into their stem morphemes and affixes (rule based route) and the memory-based system stores and retrieves all exceptions to the default as whole forms (lexical route) (e.g. Dual-Route Model: Clahsen, 1999). On the other hand, most of the studies on SC do not propose that one type of verbs (transitive vs. intransitive) would be expected as the default representation. If they make such assumptions, then transitive verbs are considered the default class: Syntax always provides a complement position for verbs, but the verb's selectional properties determine whether it is filled by an explicit argument (Cummins & Roberge 2005). In our study we saw that both L1 and L2 participants readily acquired the intransitive feature of the novel verbs and perceived the transitive usage as a violation, supporting the claims that intransitive and monotranisitive verbs do not essentially differ in their prominence. We further observed that while the L2 learners can acquire the (ir)regularity status of novel irregular verbs and perceive their regularization as violation, native German speakers perceive as violation any novel irregular form and seem to classify all new verbs as regular irrespective of the evidence that they receive through input. We interpret this finding in terms of a "learning by unlearning" effect: The L1 participants learnt that regular conjugation is a productive, default type and that the set of irregular verbs is a rather small, closed group of verbs whose all members they already know. Having learnt this, they stop acquiring regularity from texts and instead assume per default that all new verbs are regular. If an unknown irregular form appears, they consider it implausible. Seemingly paradoxically, the reduced amount of knowledge and experience that the L2 learners have, keeps them more open to turning "input into intake" and to learning new forms of all types. On the other hand, the more profound previously acquired knowledge of L1 native speakers makes them much more reluctant to acquire idiosyncratic morphosyntactic features of novel words.

References

CLAHSEN, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and brain sciences*, 22(6), 991–1013. • CUMMINS, S., & ROBERGE, Y. (2005). A modular account of null objects in French. *Syntax*, 8(1), 44–64. • GODFROID, A., & UGGEN, M. S. (2013). Attention to irregular verbs by beginning learners of German. *Studies in Second Language Acquisition*, 35(02), 291–322. • GRABE, W. (2009). *Reading in a Second Language.* Cambridge, New York:Cambridge University Press. • NATION, P. (1994). New Ways in Teaching Vocabulary. *New Ways in TESOL Series: Innovative Classroom Techniques*. ERIC. • SCHMIDT, R. (2012). Attention, awareness, and individual differences in language learning. *Perspectives on Individual Characteristics and Foreign Language Education*, 6, 27-50

BETWEEN NOUNS AND VERBS: BORROWED ADJECTIVES AS INTERMEDIATE CATEGORIES

EVAN-GARY COHEN LIOR LAKS

Tel-Aviv University Bar-Ilan University

evan@post.tau.ac.il Lior.Laks@biu.ac.il

This talk examines the strategies of the morphological adaptation of English adjectives that are borrowed into Modern Hebrew (MH). Borrowed words can undergo both morphological and phonological adaptation. In morphological adaptation, the borrowed word receives morphological features of the target language, i.e. one of the native word formation strategies is applied (Schwarzwald 1998, 2013). We argue that borrowed adjectives are an intermediate category, applying morphological adaptation strategies of both verbs and nouns.

Borrowed adjectives demonstrate three main patterns of adaptation (Ravid 1992).

Linear formation: Borrowed adjectives tend to end with the suffix *–i*. This is done by three main strategies.

1. Suffixation: Borrowed adjectives receive the native MH suffix *–i*, although some of them already end with an English adjectival suffix (*efectiv**i*** 'effect**ive**', *noʁmal**i*** 'normal').
2. Truncation: Adjectives that end with the English suffix *–ic* or *–ate* undergo deletion of this suffix or part of it, resulting in an adjective that ends with *–i* (*stat**i*** 'stat**ic**').
3. Zero conversion: Adjectives that end with the English suffix *–i* do not undergo any morphological process (*elementaʁ**i*** 'elementar**y**', *kitʃ**i*** 'kitch**y**').

These three patterns conspire towards the same goal: having an adjective that ends with *–i*, making it appear as a MH adjective. This is because the suffix *–i* is one of the most productive word processes in MH adjective formation (e.g. *jaldut* 'childhood'→*jalduti* 'childish' (Ravid & Shlezinger 1987)).

Non-concatenative formation: The participle *meCuCaC* form is used both for the formation of adjectives and the present forms of *CuCaC* passive verbs. Some foreign adjectives are formed directly in this template, e.g. *medupʁas* 'depressed' and *meʃnutsal* 'having the shape of a schnitzel', *mefuksal* 'pixellated'. Such formations are independent of verb formation; for example, the adjective *meʃnutsal* has no verbal counterpart (**ʃnitsel* 'give X the form of a schnitzel'). This is also true for MH adjectives that have no verbal counterpart, e.g. *menumas* 'polite' but **numas*, **nimes*.

Zero conversion: There are adjectives that are borrowed as is without any morphological adaptation. These are mostly adjectives that end with *–ing* or *–ed* (e.g. *ameziŋg* 'amazing', *dedikeited* 'dedicated') but also other forms like *laʁdʒ* 'large' and *kul* 'cool'. Note that unlike the forms in (1c), these do not end with *–i* like MH adjectives. In addition, unlike foreign adjectives that undergo morphological adaptation, these adjectives are also not inflected for gender and number, but remain morphologically frozen. The adjective *dedikeited*, for example, is used for feminine and does not take the suffix *–it* (**dedikeitedit*), the feminine suffix typical of the inflection of borrowed words.

Conclusions: Borrowed adjectives function as an intermediate category between nouns and verbs, as they are the only group where such a wide array of formation strategies are found.

Verbs are formed exclusively by non-concatenative morphology, as every verb in the language must conform to one of the existing verbal templates (Berman 1978, Bolozky 1978), e.g. *fikses* 'fax' in the *pi'el* template and *hiklik* 'click' in the *hif'il* template. Nouns are mostly borrowed as is, undergoing only phonological adaptation, e.g. *templet* 'template' and *neʁativ* 'narrative' (Cohen 2010). Adjectives, as we have shown, demonstrate different patterns: taking a MH suffix, similarly to MH adjectives, templatic formation, similarly to verbs and borrowing with no morphological adaptation, similarly to borrowed nouns.

This correlates with the behaviour of adjectives as an intermediate category in general (Smith 2001, 2011, Anttila 2002, Bat-El 2008). Previous studies have shown that adjectives tend to share phonological and morphological properties and processes that are typical to both verbs and nouns. Morphological adaptation appears to be scalar with respect to adjectives, from a highly typical MH formation to no morphological adaptation whatsoever.

Selected References

ANTTILA, A. 2002. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20:1-42.

BAT-EL, O. 2008. Morphologically conditioned V–Ø alternation in Hebrew: Distinction among nouns, adjectives and participles, and verbs. In S. Armon-Lotem, G. Danon, and S. Rothstein (eds.) *Current Issues in Generative Hebrew Linguistics*. Amsterdam: John Benjamins. pp. 27-60.

BERMAN, R. A. 1978. *Modern Hebrew Structure*. Tel-Aviv: University Publishing Projects.

BOLOZKY, S. 1978. Word formation strategies in MH verb system: denominative verbs. *Afroasiatic Linguistics* 5:1-26.

COHEN, E. G. 2010. *The Role of Similarity in Phonology: Evidence from Loanword Adaptation in Hebrew*. Doctoral dissertation, Tel-Aviv University.

RAVID, D. 1992. The absorption of foreign words in the nominal and verbal systems. In M. Muchnik (ed) *The foreign Influences on Contemporary Hebrew*. Tel Aviv: The Open University of Israel. pp. 11-20.. [in Hebrew].

RAVID, D. AND Y. SHLESINGER. 1987. On the classification and derivation of adjectives with the suffix -i in Modern Hebrew. *Hebrew Computational Linguistics* 25:59-70.

SCHWARZWALD, O. R. 1998. Word foreignness in Modern Hebrew. *Hebrew Studies* 39:115-142.

SCHWARZWALD, O.R. 2013. The Typology of Nonintegrated Words in Hebrew. *SKASE Journal of Theoretical Linguistics* 10(1):41-53.

SMITH, J. L. 2001. Lexical category and phonological contrast. In R. Kirchner, J. Pater and W. Wikely (eds.), *Papers in Experimental and Theoretical Linguistics 6: Workshop on the Lexicon in Phonetics and Phonology*. Edmonton: University of Alberta. pp. 61-72.

SMITH, J. L. 2011. Category-specific effects. In M. van Oostendorp, C. Ewen, B. Hume and K. Rice (eds.), *The Blackwell Companion to Phonology*. Malden: Wiley-Blackwell. pp. 2439-2463.

Categorial multifunctionality and stress-(in)sensitivity of affixes

AVA CREEMERS, JAN DON, PAULA FENGER
Universiteit van Amsterdam
ava.creemers@student.uva.nl, j.don@uva.nl, paula.fenger@uva.nl

In the literature derivational affixes have been viewed in different ways. In lexical models of morphology affixes are analyzed as elements on a par with stems: they are phonologically and semantically specified and have a syntactic category (a.o. Lieber 1980, Strauss 1982, Kiparsky 1982). In syntactic approaches of morphology (e.g. Distributed Morphology) the syntactic information of affixes is separated from their phonological realization. Affixes are thus seen as the mere spell-out of morphosyntactic information. Within this theoretical framework it has been proposed quite recently that derivational affixes have no syntactic information whatsoever, reducing them to categoriless 'roots' (Lowenstamm 2010, De Belder 2011). We claim that these latter approaches are not precise enough and consequently, empirically insufficient; instead, we propose that only a subset of the derivational affixes is indeed not specified for categorial information (i.e. roots) - implying that these affixes are syntactically multifunctional - whereas the second type of affixes is the realization of a syntactic head. This distinction correlates with stress-behavior of these affixes and their distributional properties.

Lexical models explain the distinction between stress-sensitive affixes and stress-neutral affixes (SPE), as well as the so-called Affix Ordering Generalization (Siegel 1974) in terms of level-ordering (Kiparsky 1982). However, such models have received critique in the late eighties (e.g. Fabb 1988, Gussmann 1988), pointing at both empirical and theoretical problems. Moreover, it turns out that these models are problematic for the syntactically multifunctional affixes that *seem* to realize different syntactic heads. For if affixes are specified for a syntactic category, it is not expected that they can be multifunctional in the category they denote (cf.(1a) as observed by Lowenstamm 2010 for English and (1b) by De Belder 2011 for Dutch).

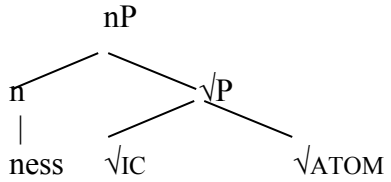
(1)	a.	*Noun*	*Adjective*	b.	*Noun*	*Adjective*
		librar-ian	reptil-ian		person-eel	univers-eel
					'staff'	'universal'
		function-ary	legend-ary		kandid-aat	desper-aat
					'candidate'	'desperate'

Syntactic models of morphology can more easily account for the observation in (1), because in such approaches affixes are roots, i.e. underspecified elements not realizing syntactic categorial information. Therefore, affixes occur wherever we may expect roots to occur. However, in such syntactic approaches it is not clear how a difference can be made between stress-sensitive and stress-neutral affixes. Lowenstamm (2010) has tried to account for this difference in a distributed approach; a key-element in his analysis is that all affixes are roots not denoting any categorial information. The different stress-behavior follows from a difference in feature-specification on these affixal roots.

This paper aims to show that an interesting improvement of Lowenstamm's analysis is possible. We provide a syntactic analysis (couched in terms of Distributed Morphology) for both the (non-)flexibility and the stress-behavior of affixes and propose that there are in fact two different types of derivational affixes: affixes that are unspecified for a category, thus roots ('root-affixes') and affixes that spell out a categorial head ('head-affixes'). The present hypothesis predicts a correlation between the flexibility of affixes and the stress-sensitivity of affixes: root-affixes can be multifunctional and are always stress-sensitive, whereas head-affixes are always non-multifunctional and stress-neutral. Root-

affixes are in the same, first phase as other root elements and therefore are predicted to be stress-sensitive. Head-affixes are in a different phase, because they are the realization of a categorial head, thus in a phase determining node (Marantz 1997), and cannot influence the stress-pattern of the word they attach to, see (2).

(2)



We show that this hypothesis makes the correct predictions with regard to both the English and Dutch derivational affixes. This paper further shows that stress-sensitive affixes (being roots) can only attach to (complex) roots, and stress-neutral affixes (being the spell-out of categorial heads) cannot occur inside a root-phrase. Finally, it turns out that there is a third type of affix that does contain categorial information ánd is stress-sensitive. At first sight this type of affixes seems to contradict our hypothesis since it breaks down the correlation between stress-sensitivity and flexibility. However, a closer look at the interaction between phasal heads and the interpretative components in fact predicts the existence of such elements: these affixes ('first heads') are always the first categorial head above the root phrase and precisely this structural position has a somewhat special status in a particular theory of phasal spell-out (Embick 2010) that we adopt.

Our analysis shows that there are three types of derivational affixes, whereas previous approaches (lexical approaches following SPE) only recognized two types of affixes. Moreover, these approaches failed to see a correlation between stress-sensitivity and multi-categoriality. The division in root-affixes, head-affixes and 'first heads' gives a better insight in the different behavior of derivational affixes. This three-way distinction is problematic both for level-ordering approaches and for the recent proposals within Distributed Morphology. In our proposal the existence of three types of affixes comes out as the only possible option, given the interaction between syntax and spell-out.

References

BELDER, MARIJKE DE (2011) *Roots and Affixes, eliminating Lexical Categories from Syntax*. LOT: Utrecht.

CHOMSKY, NOAM AND MORRIS HALLE (1968) *The Sound Pattern of English*. New York: Harper and Row.

EMBICK, DAVID (2010) *Localism versus globalism in morphology and phonology*, Cambridge: MIT Press.

FABB, NIGEL (1988) 'English suffixation is constrained only by selectional restrictions', *Natural Language and Linguistic Theory* 6, p. 527 – 539.

GUSSMANN, EDMUND (1988) Review of Mohanan (1986). *Journal of Linguistics* 24: 232-239.

KIPARSKY, PAUL (1982) 'From Cyclic Phonology to Lexical Phonology', in: Hulst, H. van der and N. Smith (eds.) *The Structure of Phonological Representations* (I), p. 131-175.

LIEBER, ROCHELLE (1980) *On the Organization of the Lexicon*, PhD diss. Univ. of New Hampshire, rep. by IULC.

LOWENSTAMM, JEAN (2010) Derivational affixes as roots (phasal spellout meets English stress shift). Ms.,Univeristé Paris-Diderot & CNRS.

MARANTZ, ALEC (1997) 'No Escape from Syntax: Don't try morphological analysis in the privacy of your own lexicon', in: Dimitriadis, A. et al. (eds.) Proceedings of the 21st Annual Penn Linguistics Colloquium: *Penn Working Papers in Linguistics* 4, p. 201-225.

STRAUSS, STEVEN (1982) *Lexicalist Phonology of English and German*, Foris: Dordrecht.

Deconstructing exuberant exponence in Batsbi

Berthold Crysmann
CNRS, Laboratoire de linguistique formelle, U Paris-Diderot
crysmann@linguist.jussieu.fr

Extended exponence provides one of the core arguments in favour of word-and-paradigm approaches (Matthews, 1974). Harris (2009) has placed the case of Batsbi exuberant exponence on the research agenda, a phenomenon which constitutes a particularly challenging case of multiple extended exponence where one and the same marker expressing the same morphosyntactic property may wind up being realised up to 4 times in a morphological word. According to Harris, class markers in Batsbi define a paradigm with 16 cells, expressing combinations of 8 genders in 2 numbers (singular/plural). Choice of class markers is controlled by an ergative pattern, i.e. they mark gender/number of intransitive subjects and transitive direct objects. Furthermore, the properties this marker expresses may overlap person/number agreement in 1st and 2nd person.

Morphotactically, these class markers can surface in several positions within a complex word: immediately preceding (i) the root (or roots in the case of compounding), (ii) the transitive/intransitive suffixes *-i, -al, -is*, and (iii) evidential suffixes. Presence vs. absence of the marker on the root is morphologically conditioned: while all roots featuring the (consonantal) markers are essentially vowel-initial, there are still roots that match the phonological condition yet fail to take the marker. Similarly, presence of post-root class markers is licensed exclusively by the suffixes listed above, to which they are left adjacent. Thus, presence of class markers is morphotactically conditioned by the class of their right-adjacent morph, a suitable root or suffix. As a consequence of which suffixes and stems appear in a complex word, one may find one of four logically possible situations (see Harris, 2009, tab. 2, p. 276): both pre-stem and pre-suffix markers (1), pre-stem markers and no suffix markers (2), no pre-stem markers but pre-suffix markers (3), as well as no class markers at all (4).

(1) don-e-v taylz-i **d**-ek'-**d**-iy-e^{n}
horse(b/d)-OBL-ERG saddlebags(/d)-PL.ABS CM-fall-CM-TR-AOR
'The horse threw off the saddlebags.' (Harris, 2009, p. 274)

(2) xen-go-ħ potl-i **d**-ek'-i^{n}
tree-ALL-LOC leaf(d/d)-PL.ABS CM-fall-AOR
'The leaves of the tree were falling.' (Harris, 2009, p. 274)

(3) kuyrc'l-e-x qečqečnayre^{n} daq'r-i lal-**d**-iy-e^{n} makaħŏ
wedding-OBL-CON various food(d/d)-PL.ABS go-CM-TR-AOR above
'At the wedding [they] passed around various foods.' (Harris, 2009, p. 274)

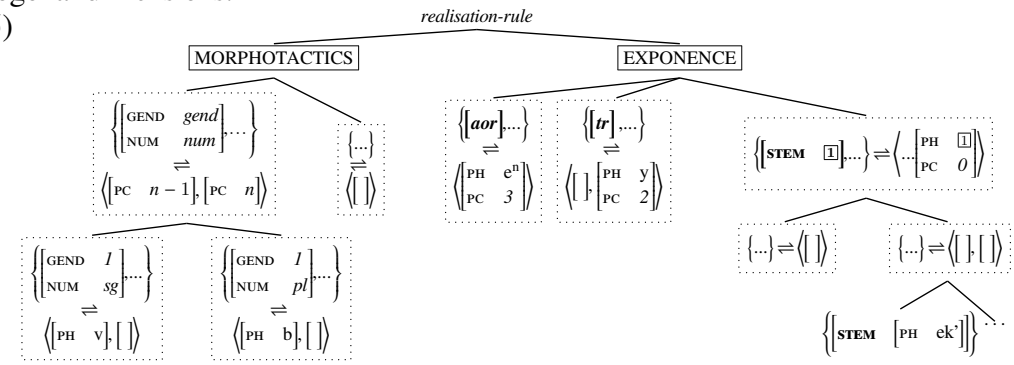
(4) k'alam xerc-n-as
pen(d/d).ABS change-AOR-1SG.ERG
'I change my pen.' (Harris, 2009, p. 275)

Based on the systematic application of a series of diagnostic tests (word-final reduction, choice of agreement trigger, intervention of particles, coordination and gapping), Harris (2009) shows convincingly that exuberant marking does indeed occur within a single morphological word. She further emphasises the problems exuberant extended exponence presents for lexical and incremental theories of inflection. Harris (2009) therefore proposes a word-and paradigm approach to Batsbi exuberant exponence that crucially relies on the notion of exemplars and schemata. Under her approach, combinations of class markers with stems and affixes are stored as paradigms, and abstracted into schemata such as [**d**-MORPH] for morphs with a dependent class marker and [MORPH] for those without. In order to capture the relatedness between cells in these paradigms, Harris builds on the concept of analogy to derive new forms from existing listed paradigms of inflected words. However, under this exemplar-based solution the shared morphotactic properties,

as well as the identity of shapes are pictured as coincidental, placing these generalisations outside the scope of the grammar.

In this talk, I shall investigate how a constructional approach to the introduction of exhuberant marking can be formalised in such a way, as to explicitly integrate into the grammar the description of shared properties of exponence and morphotactics. Moreover, under the approach I shall adopt no paradigms of fully inflected words must be listed, but they can rather be inferred on the fly by means of systematic combination of underspecified descriptions, thereby constructing complex expressions from more primitive ones. To this end, I shall build on recent proposals by Bonami and Crysmann (2013) who propose an inferential-realisational model of morphology that crucially makes use of online type construction (Koenig, 1999) by means of cross-classification between orthogonal dimensions.

(5)



The main intuition behind the present analysis is that class markers are syntagmatically dependent: as illustrated by the type hierarchy of realisation rules, rules of exponence are distinguished as to whether they introduce a single morph (as with the aorist), or rather two morphs (as with the transitive), where only the shape of the second morph is constrained. Similarly, stem introduction rules are equally segregated into an open (default) class that exclusively introduces stem phonology (PH) and position class (PC), and a closed class that requires the presence of a pre-stem marker, again, without constraining that morph's identity. In order to form a well-formed realisation rule (cf. Koenig, 1999; Bonami and Crysmann, 2013), types from the EXPONENCE dimension must be combined, under unification, with a compatible type from the MORPHOTACTICS dimension, which provides a default type for single morph introduction, as well as types for combinations of multiple morphs. As shown in (5), these latter types constrain the shape of the initial morph, together with an adjacency requirement, couched in terms of position class (PC) indices. Breaking down morph combinations into syntagmatically dependent realisation rules directly expresses generalisations about shape and position of individual morphs, whereas their combinatorial aspect is handled by the logic of orthogonal dimensions. Finally, given the inferential-realisational character of the underlying morphological theory, extended, or even exhuberant exponence can be accounted for straightforwardly.

To conclude, I have shown how the organisation of realisation rules into multi-dimensional type hierarchies lends itself quite naturally to capture the dependent nature of exuberant exponence in Batsbi: by abstracting out the introduction of class markers into a dimension of its own, we can account for both morphotactics, i.e. the dependent nature of the markers, and the pairing of morphosyntactic features with exponents. Thus, online construction of realisational rules from underspecified rule types makes it possible to deconstruct the class marking construction into its constituting parts, providing a formal grammar-internal interpretation of analogy.

References

Bonami, O. and Crysmann, B. (2013). 'Morphotactics in an information-based model of realisational morphology'. In S. Müller (ed.), *Proc. of the 20th Intl' Conference on HPSG, Berlin*. 27–47. • Harris, A. C. (2009). 'Exuberant exponence in Batsbi'. *Natural Language and Linguistic Theory*, 27:267–303. • Koenig, J.-P. (1999). *Lexical Relations*. Stanford: CSLI publications. • Matthews, P. H. (1974). *Morphology: An Introduction to the Theory of Word-Structure*. Cambridge: Cambridge University Press.

Opacity, lexicalisation, recomposition and phonological adjustment in prefixation

ANDRÁS CSER
Pázmány Péter Catholic University & Research Institute for Linguistics, H.A.S.
cser.andras@btk.ppke.hu

This paper discusses certain morphophonological phenomena displayed by prefixed forms in Latin, in particular the role of internal factors (e.g. sonority) vs. external factors (e.g. frequency) in assimilations at morphological boundaries, the scope of regularities such as the Place Condition with respect to the same boundaries, and specific dissimilatory tendencies (e.g. that of [r]) that determine the possibility of certain morphologically complex forms.

In terms of productivity, transparency and phonological interference, prefixed forms can be arranged on a cline with strongly lexicalised and opaque forms at one end and transparent formations at the other. The difference can be exemplified with *dēgere* 'to live' (usu. with an object like 'life' or 'time'), which is composed of *dē* 'from' + *agere* 'to do, drive', vs. *perpolitus* 'highly polished', in which the adjective *politus* 'polished' is combined with *per*, which has two meanings, 'through' and 'highly, very'. The transparent nature of *perpolitus* is not only seen in its semantics and in the lack of phonological interference between the two component parts, but also in some instances of its use, where it is used next to other *per*-prefixed adjectives with similar meanings. This cline is related, though certainly not in an isomorphic fashion, to the diachronic emergence of these forms (for data see Prinz 1949–50 and 1953, Leumann 1977:181–219, Buck 1899, García González 1996).

Prefixation led in many cases to lexicalisation, which in turn resulted in drastic phonological modifications at the prefix–stem boundary as well as within the stem. The pace and the extent of lexicalisation, however, was highly variable. Furthermore, prefixation also involved recomposition in all periods of the documented history of Latin. An early case of recomposition is seen in *perjūrare* 'to forswear', which is the recomposed variant of the older form *peierare* ([pejjeraːre], same meaning). Later recompositions can be reconstructed on the evidence of Romance languages; it is well known that reflexes of forms like *rétinet* 'he keeps' (< *re* + *tenet*) often derive not from the inherited Classical Latin forms but from recomposed variants such as **reténet* (> Fr *retient* etc.).

The varying pace of lexicalisation and the varying degree of transparency coupled with the phonological processes that took place at prefix–stem boundaries resulted in an often indirect relation between written forms and probable phonological variants. This is further complicated by the fact that beginning with the 1st century AD etymologically oriented habits of spelling began to gain ground, but did not affect all words of a similar composition to the same extent. Since, however, many of the characteristic traits of the variation that appear in the texts are clearly phonologically based, it can be assumed that the picture is not badly distorted — that is, with all the necessary provisos.

The analytical goal of the paper is to flesh out the observations introduced above (the role of frequency, place of articulation and sonority in morphophonological processes at boundaries) and the scope of regularities involving assimilation and dissimilation with respect to morphological boundaries. The conceptual goal of the paper is to explore the relation between lexicalisation, recomposition and phonological adjustment in the history of prefixed forms in Latin.

References

BUCK, CARL DARLING (1899) Notes on Latin orthography. *The Classical Review* 13:116–9, 156–167.

GARCÍA GONZÁLEZ, JUAN JOSÉ (1996) Asimilación de prefijos en inscripciones latinas. In: A. Bammesberger, F. Heberlein (eds) *Akten des VIII. Internationalen Kolloquiums zur Lateinischen Linguistik*. Heidelberg: Winter.

LEUMANN, MANU (1977) *Lateinische Laut- und Formenlehre*. München: Beck.

PRINZ, OTTO (1949–50) Zur Präfixassimilation im antiken und im frühmittelalterlichen latein. *Archivum Latinitatis Mediae Aetatis* 21.87–115.

PRINZ, OTTO (1953) Zur Präfixassimilation im antiken und im frühmittelalterlichen latein. *Archivum Latinitatis Mediae Aetatis* 23.35–60.

On present subjunctive inflection, between morphosyntax and pure morphology: some clues from Ibero-Romance dialects

MARTINA DA TOS
Università degli Studi di Padova
martina.datos@unipd.it

The affixal inflection of the present subjunctive displays some unexpected developments in some dialectal varieties of Ibero-Romance. In providing a possible interpretation of data, this contribution casts light on the kind of information, both morphosyntactic and purely morphological in nature, which should be assumed to be associated with present subjunctive affixes in the verbal systems of the observed varieties.

In the analysis of Spanish verbs, present subjunctive forms are traditionally regarded as involving a Tense/(Aspect)/Mood marker, whose actual shape may be influenced by inflectional class so as to give rise to a binary opposition. Thus first conjugation verbs, characterised by the affix *–e*, are distinguished from second and third conjugation verbs, sharing the affix *–a* (cf., inter alia, Alcoba 1999, Aguirre and Dressler 2001): Table 1.

In Carstairs-McCarthy's 1994 terms, the marker *–e* can therefore be analysed as a 'class-identifier', while *–a* would be the 'class-default'.

Dialectal varieties may be found to depart from this binary opposition according to either of two almost opposite types.

The so-called 'Type-A' varieties show a tendency towards neutralization of the binary opposition, such that some present subjunctive cells end up displaying one affix for the inflection of all verbs, regardless of conjugation (a 'superstable marker', Wurzel 1989). In the variety of San Ciprián de Sanabria (Krüger 1923), for instance, the affixes *–jes* and *–jen* are invariably used for the inflection of 2nd singular and 3rd pl(ural) present subjunctive forms, respectively: e.g. first conj(ugation) *cánt-**jes***, *cánt-**jen*** 'that you (sg.)/they sing', second conj. *vénd-**jes***, *vénd-**jen*** 'that you (sg.)/they sell', third conj. *dúrm-**jes***, *dúrm-**jen*** 'that you (sg.)/they sleep', as opposed to, for instance, 1st singular *cánt-**i*** versus *vénd-**a***, *dúrm-**a***.

In 'Type-B' varieties, on the other hand, not only is the binary opposition described above basically retained throughout the present subjunctive, but, within this basic distinction, the inflectional material used for inflecting second and third conjugation verbs is found to multiply. Thus in upper Aller (Rodriguez-Castellano 1952) slightly different endings compete for the inflection of both 1st pl. and 2nd pl. forms, e.g. *dem-**jamos***, *dem-**jais*** 'that we/you (pl.) harvest', *tus-**jamos***, *tus-**jais*** 'that we/you cough', versus *diš-**ámos***, *diš-**áes*** 'that we/you say'. Notice that this may also occur in the inflection of one and the same verb: *ker-**jámos***, *ker-**jáis*** beside *kiš-**ámos***, *kiš-**áis*** 'that we/you (pl.) want'. In one of the most striking cases (the Asturian dialect of Cabranes, as described by Canellada 1944), some second and third conjugation verbs display two competing sets of endings for the whole present subjunctive: Table 2.

The existence of the so-called 'Type-A' and 'Type-B' inflectional patterns can be accounted for by acknowledging, as suggested above, that present subjunctive affixes convey some morphosyntactic information fused with some purely morphological information (actually, 'Present Subjunctive + Class Identifier / Class Default'). The idea, at this point, is that in Type-A varieties the morphosyntactic information takes over, though at the cost of sacrificing the morphological information related to inflectional class. In Type-B varieties, by contrast, the purely morphological information strengthens, as it manifests itself not only on the paradigmatic dimension (the relevant distinction being between 'class-identifier' and 'class-default' markers, as already observed) but also on the syntagmatic dimension, with 'class-default' markers being used as indices of different co-occurring stems (Carstairs-McCarthy and Cameron-Faulkner 2000). What differentiates the stems at issue, in fact, is the pattern of

paradigmatic distribution (the 'morphome', in Aronoff's 1994 terms) with which they are associated. Returning to the data of Table 2, for instance, it is worth noting that the stem of the forms in I is a 'special' stem, i.e. a stem with a diphthong, normally marking the so-called 'N-pattern' (Maiden 2011), while the stem in II is the 'default stem', i.e. the stem which is not associated with any morphomic pattern, and normally appears in the infinitive (cf. Malkiel 1969).

The idea that inflection-class membership should be counted as part of the information possibly conveyed by inflectional affixes (Carstairs-McCarthy 1994) proves useful in providing an explanation of the desinential allomorphy characterising the present subjunctive in several Romance varieties. Starting from this assumption, the analysis of the Ibero-Romance data carried out above suggests two things: first, the information about inflectional class is something that fusional affixes may, but need not, convey, as proven by the cases in which the inflection-class distinction is neutralized. Second, when there is, actually, a formal opposition between 'class-identifier' and 'class-default' affixes, it is possible for the latter to be differentiated on the syntagmatic dimension, as indices of (distributionally) different co-occurring stems.

Table 1 – Present Subjunctive (Standard Spanish)

	First Conjugation CANTAR 'sing'	Second Conjugation TEMER 'fear'	Third Conjugation PARTIR 'leave'
1sg.	*cánt-**e***	*tém-**a***	*párt-**a***
2sg.	*cánt-**e**s*	*tém-**a**s*	*párt-**a**s*
	etc.	*etc.*	*etc.*

Table 2 – Present subjunctive, competing endings (Cabranes)

	ARRESPONDER 'answer'	
	I	II
1sg.	*arrespuend-**a***	*arrespond-**ia***
2sg.	*arrespuend-**a**s*	*arrespond-**ie**s*
3sg.	*arrespuend-**a***	*arrespond-**ia***
1pl.	*arrespuénd-**a**mos*	*arrespond-**iá**mos*
2pl.	*arrespuend-**a**es*	*arrespond-**iá**es / arrespond-**ié**s*
3pl.	*arrespuend-**a**n*	*arrespónd-**ie**n*

References:

AGUIRRE, C. & W.U. DRESSLER. 2001. 'On Spanish Verb Inflection'. *Folia Linguistica* XL/1-2, pp.75-91.

ALCOBA, S. 1999. 'La flexión verbal'. In: Bosque I. and V. Damonte (eds.), *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe, pp. 4915-91.

ARONOFF, M. 1994. *Morphology by Itself. Stems and Inflection Classes*. Cambridge (MA): The MIT Press.

CANELLADA, M.J. 1944. *El bable de Cabranes*. Madrid: RFE, Anejo XXXI

CARSTAIRS-MCCARTHY, A. 1994. 'Inflection Classes, Gender and the Principle of Contrast'. *Language 70*:737-88.

CARSTAIRS-MCCARTHY, A. & T. CAMERON-FAULKNER. 2000. 'Stem alternants as morphological signata: evidence from blur avoidance in Polish nouns'. *Natural Language and Linguistic Theory 18*: 813-35.

KRÜGER, F. 1923. *El dialecto de San Ciprián de Sanabria*. Madrid: RFE, Anejo IV.

MAIDEN, M. 2011. 'Morphophonological innovation'. In: Ledgeway, A. et al. (eds.), *The Cambridge History of the Romance Languages*. Cambridge: Cambridge University Press, pp. 216-67.

MALKIEL, Y. 1969. 'The five sources of epenthetic /j/ in Western Hispano-Romance'. *Hispanic Review 37*: 239-75.

RODRIGUEZ-CASTELLANO, L. 1952. *La variedad dialectal del alto Aller*. Oviedo: La Cruz.

WURZEL, W.U. 1989. *Inflectional morphology and naturalness*. Dordrecht – Boston – London: Kluwer Academic Publishers.

Economy and Multifunctional Affixes

ANNA MARIA DI SCIULLO
Université du Québec à Montréal
di_sciullo.anne-marie@uqam.ca

I take multifunctional affixes to be part of functional structures legible at the interfaces with the external systems, the semantic and the sensorimotor systems. I discuss the properties of Italian and English Numerals, Adjectives, Demonstratives and Determiners including multifunctional affixes. I argue that these functional elements are derived by the operations of the language faculty and subject to principles reducing complexity. I bring theoretical, empirical and diachronic evidence in support of my hypothesis.

1. Focus. A multifunctional affix has a basic phonetic form and more than one formal and semantic properties. I discuss the properties of multifunctional affixes in the nominal extended projection, (1), adapted from Cinque (2005). I focus on *-esimo* in Italian and *-th* in English. I argue that these functional items are derived by morphological merger, as defined in Asymmetry Theory (Di Sciullo 2005, 2014), and subject to economy principles, including *Minimal Search*, (2a), and *Pronounce the Minimum*, (2b), Chomsky (2013).

(1) [Quniv . . . [Dem . . . [Numord . . . [RC . . . [Numcard . . . [Cl . . . [A . . .N]]]]]]]
(2) a. Minimal Search: Limit the search space.
b. Pronounce the Minimum: Limit the externalization.

2. Determiners, Ordinal numbers and Adjectives. 2.1 Ordinal numbers and Adjectives. I argue that in Italian, the merger of the multifunctional affix *-esimo*, with a complex cardinal yields an ordinal number, (3). The merger of *-esimo* with the identity affix *med* yields an emphatic identity adjective in (4), (5), an emphatic reflexive pronoun in (6) and an anaphoric pronoun in (7). The derivation of complex functional elements with the multifunctional affix *-esimo* is restricted to the extended projection of nominal constituents.

(3) undicesimo, ….ventesimo , … millesimo, … millionesimo, …. (It)
(4) medesimo, medesima, medesimi (It) (5) Mi ha detto le medesime cose. (It)
'same sg.' 'same fem sing', same pl' 'He told me the same things.'
(6) Il presidente medesimo era presente. (It) (7) Il presidente era il medesimo. (It)
'The president himself was present.' 'The president was the same one.'

Thus, according to this analysis, cardinal numbers and emphatic reflexives are derived on the basis of the merger of the multifunctional affix *-esimo* with complex cardinal numbers on the one hand and the adjectival identity affix *med* on the other. Both the complex numeral and the pronoun have a binary branching structure.

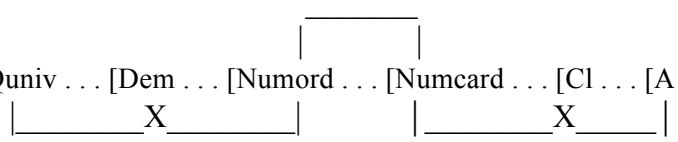
2.2 Ordinal numbers and Determiners. In English, an ordinal number is formed of a cardinal number and the multifunctional affix *-th*, (8), apart for *first, second* and *third*. I extend the granular approach to the multifunctional *th-* affix to the definite determiner and demonstrative pronouns, (9). I argue further that the definite determiners and demonstratives are the result of the merger of the *th-* multifunctional affix and a restricting vocalic constituent, including a spatial proximate or distal component (Di Sciullo 2005).

(8) eleventh, … twentieth, ... one hundredth, … one thousandth, …
(9) a. th-e book on the table b. th-is book on the table c. th-at book on the table

3. Principles reducing derivational complexity. The derivation of functional elements is the effect of the interaction of Merge and economy principles. While morphological merger derives sets, economy principles ensure efficient derivations by limiting the search space. It is predicted that functional affixes may not merge with categories that are not in their immediate local domain. This can be seen by the difference in acceptability of (10a,b) versus (10c,d). The locality restriction is represented in (11).

(10) a. #omni-th (En)
b. #omni-esimo (It)
c. one hundredth (En)
d. cento-unesimo (It)

(11) [Quniv . . . [Dem . . . [Numord . . . [Numcard . . . [Cl . . . [A . . .N]]]]]]



4. Multifunctional affixes at the interfaces. I argue further that principles reducing complexity also ensure that the externalization of the functional structure is reduced to the minimum at the sensorimotor interface, while preserving the fine-grained conceptual structure, including the configurations in (12) at the semantic interface.

(12) a. ($\text{Relation}_{\text{RANK}}$ ($\text{Space}_{\text{Numords}}$))
b. ($\text{Relation}_{\text{IDENTIY}}$($\text{Space}_{\text{Objects}}$))
c. ($\text{Relation}_{\text{PROXIMITY}}$ ($\text{Space}_{\text{Places}}$))

The externalization of linguistic expressions and the variation between languages is a function of the properties of the language faculty and laws reducing complexity (Hauser, Chomsky and Fitch 2002, Chomsky 2005, Di Sciullo 2011, 2012).

I provide diachronic evidence showing the effect of the principles reducing complexity in the development of multifunctional affixes. For example, *-esimo* subsumes multiple diachronic developments, including the following. Latin *-esimo* (identical form in masc. and fem.) derives complex nouns, e.g. *protestante, protestantesimo*. Latin *-esimus* (masc.) gives rise to *-esimo* (It) *-ième* (Fr) in ordinals, e.g. *ventesimo* (It), *vingtième* (Fr). In Vulgar Latin *metipsimus* is formed of *met* (ego), *ipse* (self) and *-issimus* (superlative). The Romance-Latin *medisme* gives rise to *medesimo* (It), *meïsme* and *mesme* in (OFr), to *mismo* in (Sp) and to *mesmo* in (Po).

5. Consequences for the model of the language faculty. The minimalist Asymmetry Theory approach to multifunctional affixes offers both a principled and a fine-grained account of the very existence of these affixes, their derivation and their interface legibility.

References

CHOMSKY, N. 2013. Problems of Projection. *Lingua* 130: 33-49.

CHOMSKY, N. 2005. Three factors in language design. *Linguistic Inquiry* 36:1-22.

CINQUE, G. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36:315-332.

DI SCIULLO, A.M. 2014. Minimalism and I-Morphology. In P. Kosta, S. Franks and L. Schürcks (eds), *Minimalism and Beyond: Radicalizing the interfaces*, 267-286. Amsterdam: John Benjamins.

DI SCIULLO, A.M. 2012 Perspectives on Morphological Complexity. In F, Kiefer, M. Ladanyi & P, Siptar (eds), *Morphology. (Ir)regularity, Frequency, Typology*, 105-135. Amsterdam: John Benjamins.

DI SCIULLO, A.M. 2011. A Biolinguistic Approach to Variation. In A.M. Di Sciullo & C. Boeckx (eds), *The biolinguistic Entreprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, 305-328. Oxford University Press.

DI SCIULLO, A. M. 2005. *Asymmetry in Morphology*. Cambridge, Mass: The MIT Press.

HAUSER, M., N. CHOMSKY, AND T. FITCH. 2002. The faculty of language: What is it, who has it and how did it evolve? *Science* 298:1569-1579.

Exploring the meaning and productivity of a polysemous prefix: The case of the Modern Greek prepositional prefix *para-*

ANGELIKI EFTHYMIOU
Democritus University of Thrace, Greece
aefthym@eled.duth.gr

GEORGIA FRAGAKI
University of Athens, Greece
efraga@phil.uoa.gr

ANGELOS MARKOS
Democritus University of Thrace, Greece
amarkos@eled.duth.gr

This paper aims at exploring the meaning and productivity of the Modern Greek prepositional prefix *para-*, which creates adjectives (e.g. *para-kratikós* 'extra-governmental'), nouns (e.g. *par-onímio* 'byname'), adjectival (e.g. *para-zalizménos* 'bemused') and adverbial participles (*para-pléondas* 'sailing by'), verbs (e.g. *para-káno* 'to overdo') and adverbs (e.g. *para-ékso* 'further out'). *Para-* derives mainly from the ancient Greek preposition *pará* 'close to (but falling short of)' (Bortone 2010: 291, cf. Triandafyllidis Dictionary 1998). *Para-* is also used in neological loan translations (e.g. *para-stratiotikos* 'paramilitary'). Moreover, Triandafyllidis Dictionary (1998) has two different homonymous lemmas, one for *para*-1, and another for *para*-2; the latter denotes the meaning of excess and derives from the adverb *pára*, which in turn originates from the ancient Greek preposition *pará*.

Given that *para-* derives from Ancient Greek *pará*, its morphological status is often described as ambiguous and the formations in which it participates can be considered as either compounds or derivatives (for discussion, see e.g. Ralli 2004, 2005). On the other hand, according to Ralli (2004), prepositional prefixes like *para-* should not be considered as autonomous words, but rather as bound morphemes which are similar to derivational prefixes. As regards verbal formations with *para-*, according to Ralli (2004) *para-* displays a dual character, since it behaves like an internal prefix (in the sense of Di Sciullo 1997) when it affects the meaning of the base and like an external prefix when it brings only external specifications to it. More specifically, she claims that *para-* is internal when it expresses the basic idea of proximity or parallelism to the meaning of the base (e.g. *méno* 'to stay, to live' - *para-méno* 'to stay on, to remain') and external in its excessive meaning (*tróo* 'to eat' - *para-tróo* 'to over-eat'). Moreover, Ralli (2004) claims that the 'excessive' *para-*, which was not present in Ancient Greek, is highly productive in Modern Greek.

Our aim in this paper is to study the meaning and productivity of the prefix *para-* in a large corpus of Modern Greek and in particular: (a) to identify the different meanings of the prefix on the basis of authentic data, (b) to examine which of these meanings are more or less productive, (c) to investigate if the excessive meaning is productive in Modern Greek, (d) to compare the productivity of the prefix across spoken and written registers, and especially across the 14 text types of the corpus, (e) to measure the productivity of *para-* across the grammatical categories of the items derived by it, and (f) to find correlations between the grammatical categories of the derived items and the meanings of *para-*.

The data used for the measurement of the productivity of *para-* comes from the Corpus of Greek Texts (CGT), a synchronic corpus of Modern Greek, including approximately 28 million words from a variety of spoken and written text types (see Goutsos 2010). We follow our previous work (Efthymiou, Fragaki & Markos 2012), which has investigated the frequency and morphological productivity of Greek verb-forming suffixes with a corpus-based methodology, similarly to earlier research on other languages (e.g. Baayen & Lieber 1991, Baayen 1992, Plag, Dalton-Puffer & Baayen 1999, Gaeta & Ricca 2003). Token frequency, type frequency and hapax legomena were used to calculate the potential productivity of the prefix, both in the corpus as a whole and within each subcorpus. In order to account for the differences in size of subcorpora, we resorted to the family of Large-Number-of-Rare-Events (LNRE) models, which are appropriate for modelling word frequency distributions (see Baayen 2001). An LNRE model attempts to estimate the expected number of types (the vocabulary size) both at smaller sample sizes (interpolation)

and at larger sample sizes (extrapolation), based on the counts of low frequency types in the corpus (the frequency spectrum). Currently, three major models are available: Generalized Inverse Gauss-Poisson (GIGP; Baayen 2001), finite Zipf-Mandelbrot and Zipf-Mandelbrot (fZM and ZM; Baroni & Evert 2006). These models are implemented in the package zipfR (Baroni & Evert 2006), a tool for lexical statistics in the R language, which is used in this study.

In our data *para-* expresses several meanings such as: (a) proximity (*parathalásios* 'seaside'), (b) parallel, subsidiary or accessory role (*parádromos* 'side road', *paramána* 'nanny'), (c) divergence, error or violation (e.g. *paravlépo* 'to overlook', *parerminévo* 'to misread', *parátipos* 'irregular'), (d) excess (e.g. *paracimáme* 'to oversleep', *paraxaidévo* 'to overindulge', *parefthís* 'straight away'), (e) resemblance (e.g. *parafrázo* 'to paraphrase'), (f) temporal continuity or duration (e.g. *paraméno* 'to stay on, to remain', *paratherízo* 'to spend the summer'), (g) transmittal (e.g. *paradído* 'to deliver'). The most productive meanings are divergence, error or violation and proximity, followed by excess and parallel, subsidiary or accessory role. The least productive meanings are those of transmittal, resemblance and temporal continuity or duration.

The prefix *para-* is more productive in written than spoken registers, following the pattern found with other Greek prefixes (Efthymiou, Fragaki & Markos 2013). In terms of particular text types, *para-* is more productive in opinion articles and popularized non-fiction texts, followed by academic texts, news and literature. The prefix is least productive in public speech and law and administration. It is also most productive in nouns, followed by verbs, adjectives and adjectival participles. Our results show that both locational (e.g. proximity) and non-locational meanings (e.g. divergence, error or violation) are productive in Modern Greek prefixation with *para-*. It is interesting that the new meaning of excess is also very productive, especially in verbs. There also seems to be a correlation between productivity in meaning and productivity in grammatical categories, a finding which will be further explored.

References

EFTHYMIOU, A., FRAGAKI G. & A. MARKOS. 2012. Productivity of verb forming suffixes in Modern Greek: A corpus-based study. *Morphology* 22 (4): 515-545. ● EFTHYMIOU, A., FRAGAKI G. & A. MARKOS. 2013. Aspects of productivity of the Modern Greek prefix *iper-*: A corpus-based study. Paper presented at the *11th International Conference of Greek Linguistics*, 26-29 September 2013, University of the Aegean, Rhodes. ● BAAYEN, H. 1992. "A quantitative approach to morphological productivity." In Geert Booij, and Jaap van Marle (eds), *Yearbook of Morphology 1991*, 109-149. Dordrecht: Kluwer. ● BAAYEN, H. & R. LIEBER. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29: 801-843. ● BAAYEN, R. H. 2001. *Word Frequency Distributions*. Dordrecht: Kluwer. ● Baroni, M. & S. Evert. 2006. The zipfR package for lexical statistics: A tutorial introduction. http://zipfr.r-forge.r-project.org/materials/zipfr-tutorial.pdf. ● BORTONE, P. 2010. *Greek Prepositions. From Antiquity to the Present*. Oxford: Oxford University Press. ● DI SCIULLO, A. M. 1997. Prefixed-verbs and adjunct identification. In A. M. Di Sciullo (ed.), *Projections and Interface Conditions.* New York: Oxford University Press, 52-73. ● GAETA, LIVIO & D. RICCA. 2003. Frequency and productivity in Italian derivation: A comparison between corpus-based and lexicographical data. *Rivista di Linguistica* 15 (1): 63-98. ● Goutsos, D. 2010. The Corpus of Greek Texts: A reference corpus for Modern Greek. *Corpora* 5 (1): 29-44. ● PLAG, I., DALTON-PUFFER, CH. & H. BAAYEN. 1999. Morphological productivity across speech and writing. *English Language and Linguistics* 3 (2): 209-228. ● RALLI, A. 2004. Stem-based versus word-based morphological configurations: The case of Modern Greek preverbs. *Lingue e Linguaggio* 2004/2: 241-275 ● RALLI, A. 2005. *Morphology*. Athens: Patakis. [In Greek]. ● TRIANDAFYLLIDIS DICTIONARY. 1998. *Standard Modern Greek dictionary*. Thessaloniki: Institute of Hellenic Studies, Aristotle University of Thessaloniki. [In Greek].

On How Fusional and Agglutinative Patterns Emerge from a Learning Constraint

ADRIANA FASANELLA SELIGRAT
Universitat Autònoma de Barcelona
afasanella@gmail.com

The aim of this study is to argue that a learning constraint independently motivated for acquisition on how the learner builds morphological paradigms is responsible for the appearance of fusional and agglutinative patterns. Given the representation and acquisition of morphological knowledge in paradigms as in Pinker (1984), the following constraint, active during the process of language acquisition, operating as a simplifying complexity device (Chomsky 2005, Roberts 2007), is proposed:

Minimise Paradigms Constraint (MPC). *The learner hypothesises just one general paradigm when affixes in cells show a systematic syntagmatic relationship.*

A more informal way of reformulating the MPC is stating that the learner prefers to hypothesise the fewest possible paradigms when affixes show a concrete pattern detected by the Language Acquisition Device, that of being in a *systematic syntagmatic relationship*. Two morphs show that configuration when they always appear adjacently, one cannot appear without the other and no other material can appear in between. In Pinkerôs original model, if the learner encounters a specific paradigm (with the stem, X, included) like the one in (1), he builds, after extracting the phonetic material in common, the following two agglutinative general paradigms (2), (3). However, once we introduced the modification in Pinkerôs model in order to capture the effects of the MPC, when the learner encounters two morphs that show a systematic syntagmatic relationship through a paradigm, he builds just one general paradigm, that is, he takes all affixal information in cells and let it the same as a general paradigm (4).

(1)

Xab	Xac	Xad	Xae	Xaf

(2)

b	c	d	e	f

(3)

a

(4)

ab	ac	ad	ae	af

Due to the effects of the MPC, two relevant predictions about the degree of fusion (Sapir 1921) in verbal morphology can be made. *Prediction 1* states that when some morphs show a systematic syntagmatic relationship in a paradigm and consequently the learner builds just one general paradigm for all the affixes, these affixes will (eventually) show a fusional pattern in subsequent intances of the language. The logic behind prediction 1 is that, once affixes are put together in the same paradigm, they will begin to show inconsistencies among forms, suffer morphophonological erosion and finally fusionalisation, due to the systematic contiguity of. For that reason, where on a first stage there are two different morphs instantiating two morphemes, in subsequent stages of the language there will be just one fusionalised morph. *Prediction 2* states that when two morphs do not show such a systematic

relationship and consequently the learner builds as many general paradigms as needed, the affixes will show an agglutinative pattern in subsequent intances of the language. Prediction 2 captures the observation that potentially discontiguity between morphs block morphophonological erosion and favours agglutination. The logic behind this proposal is that the learner's analysis during language acquisition can influence the I-language that at the end of the process he will end up acquiring. If learners' analyses of a given generation are consistent, their (modified with respect to the previous one) language will serve as input to the following generation of learners, who will acquire an already modified languagel (Roberts 2007).

In order to illustrate the emergence of *fusional patterns* we will focus on Latin verbal paradigms and their Romance counterparts in Catalan, Spanish, Italian and French. We will argue that, because of some independent changes in the Latin passive voice system, two morphs, the Tense-Aspectual-Mood (TAM) marker and the personal desinence, became always adjacent in Latin and, as an effect of the MPC, verbal paradigms underwent an important fusionalisation, observable in different degrees in Romance. We will pay attention to morphological paradigms of different tenses in all conjugations in order to show the increasing difficulty in distinguishing TAM markers from personal desinences as two different units in Romance languages.

Regarding *agglutinative patterns*, we will pay attention to the structure of verbal complexes in Bantu languages. We will argue that the agglutinative nature of Bantu morphs is due to the non-adjacency of mandatory elements, as it can be observed in the traditional schema of the full structure of the verbal pieces in Bantu languages, as in Meeussen (1967): *(preinitial) initial (postinitial) (preradical) radical (prefinal) final (postfinal).* Given that distribution, the MPC cannot be used by the learner in these contexts and the agglutinative patterns are derived. We will concentrate on data in Chichewa from Mchombo (2001) and other Bantu languages, and on some Turkish data (Korn 1997).

The present approach can be considered a step forward in predicting how morphological change happens inasmuch as it establishes specific morphological contexts where the learnerôs analises are going to change input representations. As far as the author is aware, this is the first attempt in the literature to relate the use of distributional properties, which are extensevely managed by language learners (Redington et al. 1998), with the discovery of concrete semantic aspects of morphs, namely, if they encode only one semantic distinction or more.

Selected References

Chomsky (2005) Three Factors in Language Design. *Linguistic Inquiry*; Korn (1997) *Turkish*. Routledge; Mchombo (2001) Chichewa (Bantu). *The Handbook of Morphology*. Oxford University Press; Meeussen (1967) Bantu grammatical reconstructions. *Africana linguistica*; Pinker (1984) *Language Learnability and Language Development*. Harvard University Press; Redington et al. (1998) Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*; Roberts (2007) *Diachronic Syntax*. Oxford University Press; Sapir (1921) *Language*. Harcourt, Brace and World.

Combining gender and classifiers: a typological perspective

SEBASTIAN FEDDEN & GREVILLE. G. CORBETT
University of Surrey
s.fedden@surrey.ac.uk g.corbett@surrey.ac.uk

Many languages systematically categorize their nominal vocabulary. This may involve a gender system, as in Italian, where nouns are assigned to either the masculine or the feminine gender. Another possibility is a system of classifiers, as in the Austronesian language Kilivila, which distinguishes at least 177 categories (Senft 1996), based on semantic properties, for example long and flexible objects, pots, or wooden objects. In the main a language has only one system or the other. However, there are languages which combine two systems. The interaction and functioning of the different systems shed light on the typology of noun categorization. We here present new data on a language with two categorization systems, and use it to refine the typology of such systems.

The data come from Mian, a Papuan language of Papua New Guinea (Fedden 2011). Mian has a gender system with four values: masculine, feminine, neuter 1 and neuter 2. In addition there are six classifiers: the M-classifier (for male referents and some inanimates), the F-classifier (for female referents and many inanimates), the bundle classifier, the classifier for objects which serve as covers, and a residue classifier. Agreement of the clitic article (=*e*) in gender and number with the controller *fút* 'tobacco (N1)' and the use of the classifier *tob-* 'long object (SG)' are both illustrated in example (1):

(1) *nē* *fút=e* *tob-ò-n-i=a*
1SG tobacco=ART.SG.N1 3SG.LONG.OBJ-take-SS-1SG.SBJ=and
'I take the long tobacco leaf and then I …'

The classifiers are prefixes on the verb. They make a singular-plural distinction and they are restricted to occurring on verbs of object handling or movement, such as 'give', 'take', 'throw' and 'fall'.

Languages like Mian require us to further develop the typology of categorization systems. Earlier work often treated the properties of gender systems and classifier systems as opposed to each other. Dixon (1982; 1986) gives helpful criteria for contrasting the two, and he was followed by Aikhenvald (2000) and Grinevald (2000). Since then, further instances of intermediary systems have been found, for instance, Miraña (Seifart 2005). Data from these languages suggest that factors which are grouped together in the more familiar languages can vary independently of each other, and that we need to look at these variables individually. Given this situation, it is appropriate to take a Canonical Typological approach (Brown, Chumakina & Corbett 2013). This approach involves examining the variability within and across languages to establish the parameters of variation which are relevant to setting up a theoretical space of possibilities. Only then do we examine how real examples are distributed in the space. We find that there is a clear point of convergence in this space, and this points us to defining a canonical gender system. We can do this largely on the basis of previous

work on canonical morphosyntactic features (Corbett 2013). Like other canonical morphosyntactic features, a canonical gender system (and its gender values) is clearly distinguished by formal means, its use is determined by simple syntactic rules and it is realized by canonical inflectional morphology. Where gender differs from the other morphosyntactic features is in the availability of values: in the canonical system, controllers have one gender value, while targets have all values. From previous work on canonical agreement (Corbett 2006), we take over the criterion that a canonical agreement system involves lexical features, and gender is the key type of lexical feature.

Given this idealized picture provided by the canon, we can demonstrate its usefulness by calibrating the facts of Mian against it. We find that the Mian gender system is close to canonical in most respects. On the other hand, the classifier-like elements have some properties close to a canonical gender system (they are realized as verbal prefixes reminiscent of the agreement affixes found in gender systems) and other properties which are distant from the canonical ideal (they are highly restricted in the lexical items involved).

We conclude that:

- languages like Mian are crucial to the understanding of the typology of noun categorization
- categorization systems can be understood better by calibrating them from a canonical ideal
- adopting Canonical Typology allows us to integrate into our typology the systems close to canonical (Italian) as well as the intermediate cases we find both in single systems and in combined systems (as in Mian)

References

AIKHENVALD, A.Y. 2000. *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.● BROWN, D., M. CHUMAKINA & G.G.CORBETT. 2013. *Canonical morphology and syntax*. Oxford: Oxford University Press.● CORBETT, G.G. 2006. *Agreement*. Cambridge: Cambridge University Press. ● CORBETT, G.G. 2013. *Canonical morphosyntactic features*. In: Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 48-65. Oxford: Oxford University Press. ● DIXON, R.M.W. 1982. *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin: Mouton. ● DIXON, R.M.W 1986. Noun classes and noun classification in typological perspective. In Colette Craig (ed.), *Noun classes and categorization*, 105-112. Amsterdam: Benjamins. ● FEDDEN, S. 2011. *A grammar of Mian*. Berlin: De Gruyter Mouton. ● GRINEVALD, C. 2000. A morphosyntactic typology of classifiers. In Gunter Senft (ed.), *Systems of nominal classification*, 50-92. Cambridge: Cambridge University Press. ● SEIFART, F. 2005. The structure and use of shape-based noun classes in Miraña (North West Amazon). Radboud University PhD dissertation. ● SENFT, G. 1996. *Classificatory particles in Kilivila*. Oxford: Oxford University Press.

Deverbal nominalizations and the 'means' interpretation

BERNARD FRADIN
Laboratoire de Linguistique Formelle, CNRS & U Paris Diderot
bernard.fradin@linguist.univ-paris-diderot.fr

1. Introduction Deverbal nominalizations (NZNs) that express a 'means' have repeatedly been overlooked in the literature (e.g. Grimshaw 1990) or confounded with NZNs having an Instrument interpretation (Vendler 1968). Recently, Melloni distinguished them from other nominalizations and proposed to group them into a subclass dubbed 'entities in state', which includes derived nouns denoting concrete entities associated with a specific state viz. result e.g. *translation* and psych nominals e.g. *preference* (Melloni 2011: 115). This talk aims to make explicit the conditions that determine when a given NZN may denote a means. It will discuss Melloni's proposal and show that means NZNs have to be set apart from Instrument NZNs and are orthogonal to Result NZNs. The data will mainly be taken from English, French and Italian.

2. Properties A distinctive feature of means nominalizations is their capacity to head an NP occurring as the subject of their base-V, as *heating* does in (1). From this feature comes the formula and paraphrase (2) regularly associated with them:

(1) Our new central heating heats the whole basement.

(2) λx. **Bse-V**'(e, x, y) 'X such that X base-V Y' e.g. *heating* = 'X such that X heats Y'.

What makes Means NZNs special among nominalizations is the fact that they share formula (2) with deverbal nouns that denote an Agent e.g. *seller* = 'X who sells Y', although their exponent is primarily used to express the eventive meaning e.g. *discuss-ion*, *beat-ing*, *spill-age*. It is so because they denote extensional entities (they are "R nominals"), which are either concrete e.g. *wrapping* or abstract e.g. fra *empêchement* 'hindrance'. As expected, NZNs denoting an event never satisfy (2), hence the ungrammaticality of (3)

(3) *Our thorough examination examines the data.

This situation follows from the fact that Means NZNs are derived from Vs heading a stative construction, which involves either a spatial e.g. (4) or a causative relation e.g. fra *équipement* in (5), whence (6) (Fradin 2012)(the representations given here are rough approximations).

(4) **surroundings'**(x) ≡ **surround'**(e,x,y) ∧ LOC(s, x, CIRCUM(y)) ∧ FIG(x) ∧ GRND(y)

(5) *L'équipement des monocoques surpasse celui des catamarans.*
'the equipment of mono-hulls overpasses that of catamarans'

(6) **équipement'**(x) ≡ **equip'**(e_1,x,y) ∧ CAUSE(e_1, e_2) ∧ **function_normally**'(e_2, y), where e = event, s = state

The stativity of the constructions in question is shown by the fact that their base-V is odd with the progressive and gets a non-habitual reading in the present tense (cf. (1))(Dowty 1979).

(7) ??Our new central heating is heating the whole basement.

3. Distinctions The following criteria allow us to distinguish Means (MNS) from Instruments (INS): (i) INSs always involve a dynamic situation, (ii) INSs are not licenced as subjects of a sentence denoting an actual eventuality that is not agent driven e.g. (8)(Schlesinger 1989), (iii) INSs denotes countable objects which must exist as a separate entity, before and after the

event in which they are used has been completed, which is not necessarily the case of Means cf. (9), (iv) INSs need not comply with (2).

(8) ?*This plane planed the floor.

(9) These junction obstructions can even obstruct tiny blood vessels.

4. Independency The occurrence of the Means interpretation hinges on the possibility to derive a NZN from a V exhibiting properties (2) and (4)/(5). But this NZN may itself result from nominalizing a creation verb, the outcome of which is a noun denoting a means e.g. ita *prolungamento* 'continuation' = 'part of an X (road) which has been extended (but did not exist before)' as in (10).

(10) *Un prolungamento di 5 chilometri prolunga la stada*.
'a 5 km extension continues the road'

Hence the three way interpretation for *prolungamento*: eventive, result and means. On the other hand, there are NZNs which only have a means interpretation e.g. fra *renseignement* 'piece of information'← *renseigner* 'give some information'. This NZN cannot denote an event e.g. (11), let alone a result, since *renseigner* heads a verbal construction denoting an activity or more precisely a so-called 'degree achievement' (Beavers 2009) and only telic (creation / duplication / modification) verbs can be the base of Result nominalizations (Bisetto & Melloni 2007).

(11) **Le renseignement du touriste a eu lieu a midi.*
'The tourist has been informed at noon'

Many NZNs with a means interpretation pattern like *renseignement* e.g. fra *consolation* 'consolation', *garantie* 'guarantee'.

5. References

BEAVERS John. 2009. Aspectual Classes and Scales of Change. Austin: The University of Texas at Austin.

BISETTO Antonietta & Chiara MELLONI. 2007. "Result Nominals: a Lexical-Semantic Investigation". In *On-line proceedings of the Fith Mediterranean Morphology Meeting (MMM5) Fréjus 15-18 September 2005*, Booij G., L. Ducceschi, B. Fradin, E. Guevara, S. Scalise & A. Ralli (eds). http://mmm.lingue.unibo.it/. Bologna: University of Bologna.

DOWTY David R. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

FRADIN Bernard. 2012. "Les nominalisations et la lecture 'moyen'". *Lexique* **20**:129-156.

GRIMSHAW Jane. 1990. *Argument Structure*. Cambridge (Mass.) / London: The MIT Press.

MELLONI Chiara. 2011. *Event and Result Nominals*. Bern: Peter Lang.

SCHLESINGER I. M. 1989. "Instruments as Agents: On the Nature of Semantic Relations". *Journal of Linguistics* **25** 1:189-210.

VENDLER Zeno. 1968. *Nominalizations and Adjectives*. The Hague / Paris: Mouton.

Mass and count: a *gradatum* resulting from two combining features

FRANCESCA FRANZON & CHIARA ZANINI
Università degli Studi di Padova
francescafranzon7@gmail.com; chiara.zanini.2@unipd.it

Introduction. The mass property and the count property have been traditionally described as inherent properties of the lexeme that trigger some grammatical constraints (*i.a.*: Renzi *et al.*, 2001). With respect to morphology, a mass noun is said not to have a plural form. As a consequence, in the literature the phenomenon of mass nouns admitting plural is treated as some kind of semantic shift (Renzi *et al.*, 2001). This description does not really explain why the plural is acceptable for a mass noun and is not predictive for what concerns the contexts in which such a possibility takes place. Moreover, the traditional mass-count dichotomy is not exhaustive about the range of possibilities that happen to surface, for example, in Italian. With respect to nouns behavior, it **seems** that we face more than two categories: A. *Sangue*(blood)-like nouns: they refer to substances. They can occur only at the singular form: the plural is not grammatical; B. *Anello*(ring)-like nouns: they refer to individuated entities: at the singular they refer to one, at the plural to more than one; C. *Vino*(wine)-like nouns: they are usually singulars. When they surface as plurals, they can get only a "kind" interpretation; D. *Mela*(apple)-like nouns: they refer to individuated entities: at the singular they refer to one, at the plural to more than one. At the singular -under certain conditions- they can refer to the substance of which the entities they refer to are made (e.g.: *c'era molta mela nel frullato*, lit: 'there was a lot of apple in the smoothie'). At the plural -under certain conditions- they can get a "kind" interpretation; E. *Pizza*-like nouns: they may realize either the possibilities of the *mela*(apple) or of the *vino*(wine)-like noun type. Jackendoff (1991) and Chierchia (2010) have explained this range of possibilities by assuming a complex system of features intrinsic on the lexemes and operations performed on them.

Main proposal. Differently from the previous interpretations, we propose that there is no a priori label attached to the lexeme to define a noun as mass or count. Thus, with respect to Jackendoff (1991) and Chierchia (2010), there is no need to perform operations starting from on a fixed status to change the position of a noun in a classification. We propose that a nominal expression can be said to refer to a mass or count entity only when it has undergone a computation through the levels that interface between the perceptual reality and the lexicon, the morphology, the syntax. To give account of this, we will show: (i) the necessary and sufficient features whose combinatory possibilities are needed to individuate the mass interpretation or the count interpretation - as well as the other possibilities listed above - of a nominal expression; (ii) that a mass or count interpretation is not a matter of a lexical specification but the offspring of a computation at the levels of perceptual, morphologic and syntactic interfaces as well; (iii) that a nominal expression can surface alternatively with more than one feature configuration. In this presentation we will focus on the interface between the referential world and lexical morphology.

[±u; ±k]. At a semantic level, an entity may have –or not – a boundary. If an entity is defined within its boundaries, it is recognizable as an individual. This semantic property is linked with the perceptual level: if an entity has a physical and natural boundary, the noun referring to it will likely be associated with a semantic [UNIT] interpretation. However, it doesn't follow from this that the symmetric condition is as likely: the semantic property must not be confused with a physical property of the entity to which the noun refers to. An entity that has no physical boundaries per se can be associated with a semantic [UNIT] interpretation. For example, the noun *birra* 'beer' may designate the substance: in this case, there is no possible [UNIT] interpretation. Rather, it may designate a portion of the substance as well: this portion comes in precise boundaries (e.g., a glass or a bottle) and thus the interpretation [UNIT] is

available. There is another possibility to get an individuality interpretation: it could not refer to the entity defined by the boundary, instead by its kind (e.g., a particular kind of beer). 'A kind of beer' shares all the properties of the substance 'beer', but at the same time it is recognizable as an individual as it shows some properties of its own that allow to distinguish it among the other kinds of beer. Here, the individuality property takes from the [KIND] interpretation. The semantic level must interface on the one side to the perceptual, referential level and on the other side with the morphological level. At the morphology-semantic interface the individuality in terms of unit or kind (or, as a consequence, the non-individuality) is codified at the level of this interface into binary features [± unit], from now on [± u], and [± kind], from now on [± k]. Every entity linked (or not linked) to an individuality property, [UNIT] or [KIND], at the semantic level puts into availability the feature system, [±u; ±k], that make semantically interpretable the number morphology at the semantic-morphologic interface.

features		Semantic interpretation	Morphological number	Example	Behaves like:
-u	-k	[SUBSTANCE]	Singular only	*sangue*/**sangui* (blood/*bloods)	A, C, E, D
-u	+k	[KIND]	Sing/plur	*vino*/*vini* (wine/wines)	C, E, D
+u	-k	[UNIT]	Sing/plur	*anello*/*anelli* (ring/rings)	B, D, E
+u	+k	Not possible	Ø	Ø	Ø

From what we have stated and illustrated above, we can draw the following generalizations: (i) in the computation, at least one of the two features must assume a negative value; (ii) only one of the two features may assume a positive value. Roughly speaking, the two generalizations state that if an individuality property is available, it is possible to activate a counting/quantification processing focusing on the individual or on the kind; but it is never possible to count on the unit and on the kind at the same time.

Conclusions. The different behavior of the nouns does not actually reflect a categorization among five different groups A, B, C, D, E. Instead, it is the direct consequence of the fact that a noun is linked to a particular combination of features. This links basically depends on the speaker's "encyclopedic" knowledge of the referent and on the communicative context in which this latter is referred to. The consequence of this is the fact that a noun can be associated to different combinations of features; each of them is equally well-formed if placed in the appropriate communicative context. The fact that a noun has been traditionally associated with a particular class depends on the frequency with which it is associated to a specific configuration of features. In other words, a noun as *'burro'* (butter) has been traditionally classified as a mass noun since the contexts in which it surfaces with the configuration [-u; -k] are more frequent than those in which it surfaces with the configuration [-u; +k] or [+u; -k]. But, at least in theory, this does not mean that such a noun cannot be linked to the configuration [-u; +k] or [+u; -k], *i.e.*, cannot be a count noun at all.

Potentially, a lexeme can have more than one configuration of features at the morphology-lexical interface and then undergo different syntactic computations. In this view, the unmarked condition for a nominal expression is the one which allows the larger number of combinations in the larger number of contexts. In the present study we considered a language (Italian) that bears an overtly marked distinction singular vs. plural in number morphology. Nevertheless, our analysis is based on a feature system that, at least in theory, should be applicable cross-linguistically. We leave these surveys open to further examinations.

Selected references

CHIERCHIA, G. (2010). Mass nouns, vagueness and semantic variation. S*ynthese* 174, 99-149. • CORBETT, G. (2000). *Number*. Cambridge: Cambridge University Press. • STUMP, G. (2001). *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press. • JACKENDOFF, R. (1991). Parts and Boundaries. *Cognition* 41, 9-45. • RENZI, L., SALVI, G. & CARDINALETTI, A. (2001). *Grande grammatica italiana di consultazione*. Bologna. Il Mulino.

Word formation in a balanced diachronic corpus of Italian

Livio Gaeta — Università di Torino — livio.gaeta@unito.it
Claudio Iacobini — Università di Salerno — ciacobini@unisa.it

A characteristic of recent morphological research is the ever growing importance of the use of corpora. The broadening of the empirical basis for morphological analysis has allowed scholars to provide descriptions and explanations of phenomena which would not have been possible some years ago. Within a large research project funded by the Italian government and aimed at investigating Italian word formation, a new research tool is being developed, namely MIDIA (*Morfologia dell'Italiano in DIAcronia*). This diachronic corpus of Italian fills an important lacuna: although a considerable amount of texts are available electronically with special regard to the earlier stages of Italian, a balanced diachronic corpus was until now a desideratum. MIDIA was explicitly designed to cover all periods of the linguistic history of Italian, divided in five homogeneous time spans and collecting texts for seven different text types, going from prose and poems to chronicles, scientific essays, etc. For each time span and text genre 25 text chunks of about 8000 tokens have been taken and acquired into a database. This amounts to a final corpus of about 7 million tokens, which is hopefully large enough to perform a number of investigations on productivity of single word formation patterns (cf. Cowie & Dalton-Puffer 2002, Gaeta & Ricca 2006, in press). Clearly, the database is independently searchable for any of the parameters indicated, namely time span and text genre.

The software utilized for building the database comes from the TreeTagger family as already employed for the *Repubblica* corpus (cf. Baroni et al. 2004), enriched by more than 170,000 wordforms collecting all sort of allomorphs attested during the long history of Italian (about 26,000 nominal, 11,500 adjectival, and 98,000 verbal forms plus about 15,600 forms hosting clitics) and including most grammatical words (about 1,800) and proper names (about 21,000). All of this considerably enlarges the basic lexicon already present in the TreeTagger version of the *Repubblica* corpus, and substantially improves its efficiency as a syntactic parser. On the basis of a golden standard test, it was possible to verify a significant improvement of the MIDIA-parser with regard to the *Repubblica*-parser increasing from about 89% of correctness of PoS assignment to about 96% (cf. Gaeta et al. 2013). Moreover, the user's friendly searching mask allows any possible research on single as well as on sequences of tokens or PoS-tagged lemmas (or parts of them), employing all common logical symbols and extracting the data into any sort of tables and graphs.

On this basis, we want to exploit the empirical potential of MIDIA by carrying out two different investigations sharing the same perspective, namely to explore two onomasiological domains taking seriously into consideration the range of grammatical polysemy or polyfunctionality, including homonymy, characterizing multifunctional affixes.

First, we will focus on the domain of spatial verbal prefixes. In the transition from Latin to the Romance languages, and in the subsequent phase of development of the latter up until the present day, there has been a reduction in the number of preverbal prefixes and the types of meanings that can be expressed by these prefixes (cf. Iacobini 2004, Iacobini & Masini 2006). Contemporary Italian is characterized by a low productivity of spatial verbal prefixes and the reduction of their number compared to Latin. Verbal prefixation is productively used almost exclusively to express privative, reversative, and iterative values. We intend to examine the reasons that determined this situation, moving in two converging directions: a) the identification of the verbal prefixes used productively in the different temporal phases of the Italian language and the description of the types of meaning they

express; b) the identification over time of the possible types of prefixed verbs (starting from the verbal bases available in each period).

Second, we will have a look at the domain of agentive suffixes (cf. Gaeta 2010, 2012). In particular, we will try to understand whether data extracted from dictionary-based investigations are well supported on the textual basis provided by MIDIA, in particular with regard to the competing productivity of masculine and feminine suffixes (*-tore* and its allomorphs vs. *-trice*) for designating agents and instruments (cf. Rainer 2011). For this goal MIDIA provides data from different text types, going from poems to legal texts, which allows us to keep under control the textual variation relating to genre. Moreover, as with verbal prefixes, a comparison will be made with different and competing coding strategies for expressing agent / instrument meanings coming from syntax. This will give us significant cues for understanding the way how the process of lexical entrenchment favoring complex derived words over syntactic units concretely takes place in the course of time.

References

Baroni, Marco et al. (2004), Introducing the la Repubblica corpus: A large, annotated, TEI(XML)-compliant corpus of newspaper Italian. *Proceedings of LREC 2004*, 1771-1774. Lisbon: ELDA.

Cowie, Claire S. & Christiane Dalton-Puffer (2002), Diachronic word-formation and studying changes in productivity over time: Theoretical and methodological considerations. In: Javier E. Díaz Vera (ed.), *A Changing World of Words. Studies in English Historical Lexicography, Lexicology and Semantics*, 410-437. Amsterdam: Rodopi.

Gaeta, Livio (2010), On the viability of cognitive morphology for explaining language change. In: Alexander Onysko, Sascha Michel (eds.), *Word Formation from Cognitive Perspectives*, 75-95. Berlin / New York, Mouton de Gruyter.

Gaeta, Livio (2012), Strumenti come agenti: a metà tra natura e storia. In: Valentina Bambini, Irene Ricci, Pier Marco Bertinetto (eds.), *Linguaggio e cervello - Semantica / Language and the brain – Semantics. Atti del XLII Congresso Internazionale di Studi della Società di Linguistica Italiana (Pisa, SNS, 2008)*, Vol. 2, Paper II.A.1, 1-14. Roma: Bulzoni.

Gaeta, Livio & Davide Ricca (2006), Productivity in Italian word-formation: A variable-corpus approach. *Linguistics* 44.1: 57-89.

Gaeta, Livio & Davide Ricca (in press), Productivity. In: Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen, Franz Rainer (eds.), *Handbook of Word-Formation. An International Handbook of the Languages of Europe*. Berlin/New York: Mouton de Gruyter.

Gaeta, Livio, Claudio Iacobini, Davide Ricca, Marco Angster, Aurelio De Rosa & Giovanna Schirato (2013), MIDIA: a balanced diachronic corpus of Italian. Paper presented at the *21st International Conference on Historical Linguistics*, Oslo 6.-9.8.2013.

Iacobini, Claudio (2004), Prefissazione. In: Maria Grossmann, Franz Rainer (eds.) *La formazione delle parole in italiano*, 97-163.Tübingen: Niemeyer.

Iacobini, Claudio & Francesca Masini (2006), The emergence of verb-particle constructions in Italian: locative and actional meanings. *Morphology* 16.2: 155-188.

Rainer, Franz (2011), The agent-instrument-place "polysemy" of the suffix -TOR in Romance. *Sprachtypologie und Universalienforschung* 64.1: 8-32.

On the person agreement suffixes of gerunds in Udmurt

EKATERINA GEORGIEVA
University of Szeged
katina.geo@gmail.com

In this talk I discuss the agreement suffixes of two gerunds in Udmurt, a Uralic language spoken in the Russian Federation by about 350 000 speakers. I will focus on the gerunds, formed with the suffixes *-ku* and *-toź*, both expressing temporal relations. The data comes mainly from my own fieldwork with 27 native speakers of Udmurt.

The two gerunds can be inflected with agreement suffixes indicating the person/number features of the subject (1), or can have an overtly expressed subject (2). Edygarova (2010) claims that overt subjects and agreement suffixes are in complementary distribution (3).

(1) *[Korka-me* *pukty-ku-m]* *mynym*
house-1SG.ACC build-NMLZ-1SG 1SG.DAT
nokin *no* *ẹz* *jurtty.*
nobody[NOM] PART NEG.PST.3 help[SG]
'When I was building my house, nobody helped me.'

(2) *[Koľa otyn uly-toź],* *Kaťa* *bazar-yś* *vyľ* *derem* *baśt-i-z.*
PN there be-NMLZ PN market-EL new dress[ACC] buy-PST-3SG
'While Kolya was there [in the library], Katya bought a new dress at the market.'

(3) (Edygarova 2010: 87)
**Ton* *[mon* *škola-je* *myny-ku-m]* *eš-te* *pumita-d.*
2SG 1SG school-ILL go-NMLZ-1SG friend-ACC.2SG meet-PST.2SG
'While I was going to school, you met your friend.'

The type of agreement affixes that cannot co-occur with an overt subject have often been considered to be pronominal in nature. Pronominal affixes are widely discussed topic in the literature on agreement from both typological and theoretical perspectives (Siewierska 1999, 2004, Corbett 2006, Jelinek 1984, Bresnan&Mchombo 1987, Baker 1996, Fuß 2005 among others). Several criteria have been proposed for distinguishing pronominal affixes from grammatical agreement. The most important one is the co-occurrence of the agreement marking with an overt DP. In contradiction to Edygarova's claim, the results of my fieldwork with 27 native speakers of Udmurt show the person agreement of the two gerunds in question is subject to variation – speakers of the Northern (and some of the Southern) dialects do accept examples like (3). Thus, for those speakers, the agreement marking is not strictly pronominal since it can be doubled by an overt DP. In my talk I will also discuss other tests for distinguishing pronominal affixes from grammatical agreement, e.g. the co-occurrence of the agreement marking with indefinite, quantified and questioned DP-s and the degree of referentiality of the agreement marking. I will demonstrate that the person agreement suffixes on the gerunds have mixed properties: despite their similarity to the free pronouns in terms of referentiality they can co-occur with different types of overt DP-s (for some speakers), which indicates that they should not be considered pronominal suffixes. I will argue that there is an ongoing morphosyntactic change in Udmurt, namely the former pronominal suffixes are being reanalyzed as an instance of pro-drop (or ambiguous marking in Siewierska's terms). The change might have been induced by the fact that Udmurt is generally pro-drop language (see Fuß 2005 for a recent theoretical proposal on how the pro-drop properties of a language might enforce the reanalysis of the former pronominal affixes/clitics into grammatical agreement).

To sum up, in this talk I deal with two Udmurt gerunds and more specifically, with their agreement. It has been stated in the literature that this agreement is in complementary distribution with overt subjects. My fieldwork with native speakers of Udmurt demonstrates

that this is not always true. I suggest that there is an ongoing morphosyntactic change in some varieties of Udmurt, according to which the former pronominal suffixes are turning into grammatical agreement. In this way the talk sheds light on the some unexplored issues of the Udmurt morphosyntax and contributes to our general knowledge on the Udmurt language.

References

BAKER, MARC C. 1996. *The Polysynthesis Parameter*. Oxford Studies in Comparative Syntax. New York: Oxford University Press.

BRESNAN, JOAN & SAM A. MCHOMBO. 1987. Topic, pronoun and agreement in Chichewa. *Language 63*. pp741−782.

CORBETT, GREVILLE G. 2006. *Agreement*. Cambridge: CUP.

EDYGAROVA, SVETLANA. 2010. *Kategoria posessivnosti v udmurtskom yazyke*. Dissertationes Philologiae Uralicae Universitatis Tartuensis 7. Tartu: University of Tartu.

FUß, ERIC. 2005. The Rise of Agreement. A formal approach to the syntax and grammaticalization of verbal inflection. Amsterdam/Philadelphia: John Benjamins.

JELINEK, ELOISE. 1984. Empty Categories, Case, and Configurationality. *Natural Language and Linguistic Theory 2*:39-76.

SIEWIERSKA, ANNA. 1999. From anaphoric pronoun to grammatical agreement marker: why objects don't make it. *Folia Linguistica* 33/2:225–51.

SIEWIERSKA, ANNA. 2004. *Person*. Cambridge: CUP.

German Nominalization Patterns in a Diachronic Perspective: Constructional Change and Functional Reorganization

STEFAN HARTMANN
University of Mainz
stefan.hartmann@uni-mainz.de

This paper investigates the diachronic development of two highly productive German word-formation patterns, namely nominalization in the suffix *-ung* (e.g. *Landung* 'landing', *Behandlung* 'treatment') and Infinitival Nominalization (e.g. *das Singen* 'singing', *das Warten* 'waiting'). Both word-formation patterns have previously been studied in a variety of different frameworks ranging from Generative Grammar to Cognitive Linguistics and Construction Grammar. However, to date, no systematic quantitative corpus analysis has been conducted to address the key questions concerning their diachronic development. The present study aims at filling this gap with an extensive corpus-based investigation of both patterns. The results lend support to a usage-based and constructionist view of word-formation and word-formation change.

Previous research suggests that *ung*-nominalization is subject to an increase in semantic word-formation constraints, which is mirrored in a decrease in morphological productivity (cf. Demske 2000). Infinitival Nominalization comes in as a "replacement process" (Barz 1998, cf. also Werner 2010). At the same time, *ung*-nominalization is ubiquitous in Present Day German. In addition, a variety of cultural factors leads us to assume that the overall number of nominalizations has increased, for example a) the need to name concepts previously conveyed only in Latin (cf. e.g. Eichler 1996), b) the drive for "brevity" and "language economy" that can be observed as a tendency in 17th and 18th century German word formation (cf. Schmidt 2007: 153).

Adopting Baayen's (2009) measures of productivity, we can thus predict (1) an increase in realized productivity (i.e. type frequency) for both *ung*-nominalization and Infinitival Nominalization over the past centuries, while we expect (2) a decrease in potential productivity, i.e. the relation of hapax legomena to the sum total of tokens belonging to the word-formation patterns in question, for *ung*-nominalization. For Infinitival Nominalization, by contrast, we predict (3) an increase in potential productivity, since Nominalized Infinitives (NIs) have to "replace" terms previously coined by means of *ung*-nominalization.

Apart from these overall developments, Demske (2000) has detected noticeable changes concerning the constructions in which *ung*-nominals tend to occur. For example, in Present Day German, they occur more frequently with a determiner or in the pluralized form than in Early New High German. By contrast, the construction in which *ung*-nominals serve as complements of a preposition (e.g. *in Lesung des Gedichts* 'in reading the poem') comes out of use. This construction can be seen as profiling the progression of an event rather than its boundaries, hence evoking a more 'verby' construal, while pluralization and the determiner construction highlight the 'nouny' properties of nominalizations by evoking a 'count noun' construal. Since Infinitival Nominalization, by and large, seems to have retained its 'verbal' properties (cf. e.g. Ehrich 2002), we expect a corpus analysis investigating the three aforementioned constructions to show significant changes only for *ung*-nominalization, but not for Infinitival Nominalization.

A corpus analysis using the Mainz Early New High German Corpus – an as yet unpublished corpus covering the time span from 1500 to 1710 – and the GerManC Corpus (Durrell et al. 2007), which covers the years from 1650 to 1800, confirms the hypotheses outlined above. Further examination of the data drawing on different variants of collostructional analysis (cf. Stefanowitsch 2013) leads to a more fine-grained characterization of the word-formation pat-

terns in question, which in turn has important implications for properly describing the constraints affecting these patterns as well as their diachronic change.

I will argue that the diachronic development of *ung*-nominalization and Infinitival Nominalization can be accounted for in a usage-based and constructionist framework. Word-formation patterns are seen as constructions, i.e. form-meaning pairings emerging via generalizations over actual language use (cf. e.g. Traugott & Trousdale 2013). Since *ung*-nominals are subject to lexicalization at the semantic level, their increasingly frequent use in lexicalized reading variants brings about changes in the generalizations language users make over this input. Hence, the constructional schema itself (i.e. the word-formation pattern) is modified. Thus, more and more *ung*-nominals are replaced by NIs. Infinitival Nominalization becomes the default word-formation pattern for deriving action nouns, while *ung*-nominalization is now subject to considerable constraints. Hence, the diachronic development can be seen as a case of functional reorganization in the German word-formation system caused by a series of interdependent constructional changes in the domains of semantics (lexicalization), morphology (changes in word-formation constraints), and syntax (change in the prevalent syntactic patterns in which the respective word-formation products are used).

References

BAAYEN, R. Harald (2009): Corpus Linguistics in Morphology. Morphological Productivity. In: Lüdeling, Anke; Kytö, Merja (eds.): Corpus Linguistics. An International Handbook. Vol. 2. Berlin, New York: De Gruyter (HSK, 29.2), 899–919.

BARZ, Irmhild (1998): Zur Lexikalisierungspotenz nominalisierter Infinitive. In: Barz, Irmhild; Öhlschläger, Günther (eds.): Zwischen Grammatik und Lexikon. Tübingen: Niemeyer, 57–68.

DEMSKE, Ulrike (2000): Zur Geschichte der ung-Nominalisierung im Deutschen. Ein Wandel morphologischer Produktivität. In: Beiträge zur Geschichte der deutschen Sprache und Literatur 122, 365–411.

DURRELL, Martin; Ensslin, Astrid; Bennett, Paul (2007): The GerManC Project. In: Sprache und Datenverarbeitung 31, 71–80.

EHRICH, Veronika (2002): On the verbal nature of certain nominal entities. In: Kaufmann, Ingrid; Stiebels, Barbara (eds.): More than Words. A Festschrift for Dieter Wunderlich. Berlin: Akademie-Verlag (studia grammatica, 53), 69–89.

EICHLER, Birgit (1996): Fachlich konnotierte Sprachstrukturen in frühneuzeitlicher Wissensliteratur. In: Kalverkämper, Hartwig; Baumann, Klaus-Dieter (eds.): Fachliche Textsorten. Komponenten - Relationen - Strategien. Tübingen: Narr (Forum für Fachsprachen-Forschung, 25), 271–291.

SCHMIDT, Wilhelm (2007): Geschichte der deutschen Sprache. Ein Lehrbuch für das germanistische Studium. 10th ed. Stuttgart: Hirzel.

STEFANOWITSCH, Anatol (2013): Collostructional Analysis. In: Hoffmann, Thomas; Trousdale, Graeme (eds.): The Oxford Handbook of Construction Grammar. Oxford: Oxford University Press, 290-302.

TRAUGOTT, Elizabeth Closs; Trousdale, Graeme (2013): Constructionalization and Constructional Changes. Oxford: Oxford University Press (Oxford Studies in Diachronic and Historical Linguistics, 6).

WERNER, Martina (2010): Substantivierter Infinitiv statt Derivation. Ein ‚echter' Genuswechsel und ein Wechsel der Kodierungstechnik innerhalb der deutschen Verbalabstraktbildung. In: Bittner, Dagmar; Gaeta, Livio (eds.): Kodierungstechniken im Wandel. Das Zusammenspiel von Analytik und Synthese im Gegenwartsdeutschen. Berlin, New York: De Gruyter (Linguistik - Impulse und Tendenzen, 34), 159–178.

Overabundance in Zurich German

ANJA HASSE
University of Zurich
anja.hasse@ds.uzh.ch

THORNTON (2011) deals with overabundance, a kind of non-canonical inflectional morphology. She distinguishes three types of overabundance: "forms built according to different means of exponence in the same cell, […] forms built on two different stems in the same cell, […] forms with different inflectional endings in the same cell." Each type is a kind of deviation from canonical inflection and by itself a class of canonical overabundance (THORNTON 2011:361). Zurich German shows an example for the two latter deviations. The indefinite article shows cell-mates in the dative in all three genders, cf. Table 1.

	masc.	**neutr.**	**fem.**
nom. / acc.	*en*	*es*	*e*
dat.	*emene, eme*	*emene, eme*	*enere, ere*

Table 1. The Zurich German indefinite article, cf. WEBER (1923:168)

Concerning the feminine paradigm, the emergence of the two forms can diachronically be traced back to earlier stages of German. Synchronically, they do not show the same stem or the same suffix either. In the masculine paradigm the overabundance is due to a suffix {ne} which is not obligatory. This suffix is not yet attested in the inflection of the Early New High German (ENHG) indefinite article. It is homophonous with the suffix in the dative plural of the pronominal inflection and it builds a new form based on the already inflected form *eme*. Generally, it seems possible that an inflected form is reanalyzed as a stem and another suffix is added to the stem, as JOSEPH (2013) shows for Ancient Greek pronouns. I suggest a similar process in the emergence of *emene*: the form *eme* has been expanded to *emene*. There is in fact some dialectological evidence for this derivation. Still, the origin of this suffix {ne} in *emene* remains a matter of dispute. The traditional view, however, that *emene* is a phonological variant of *eneme* showing metathesis, has recently been regarded as not convincing, cf. SEILER (2003).
Overabundance in the dative cells of the indefinite article is already evident in earlier stages of German, cf. Table 2**.** In ENHG, there are two different types of dative forms of the indefinite article in the feminine and masculine / neuter.

masc. /neutr. dat.	**fem. dat.**
einem, eineme, eime	*einer, einere, eire*

Table 2. The Early New High German indefinite article (dative), cf. REICHMANN / WEGERA (1993:220)

According to this, overabundance in this cell is diachronically well attested. It is, however, very difficult to state how overabundance is organized in older periods of a language, cf. THORNTON (2011:363). The mere fact that it appears in a grammar of ENHG does not mean that the overabundant forms are not subject to diatopic, diaphasic, diachronic or other factors. In modern languages further conditions can be found. Thus, overabundance can be determined much more precise, and even intrapersonal variation is attested. In fact, the Zurich German indefinite article does not only represent an instance of overabundance within a paradigm, but is also subject to intrapersonal variation. The conditions for the variants to occur are still unknown, though. Four kinds of factors have to be taken into account: syntactical (e.g. syntactical position and function), phonological (e.g. rate of speaking, metre), semantic-functional (e.g. specificity) and socio-linguistic conditions (e.g. age). In my previous research I could not identify a determining factor for the usage of the forms. This makes

Zurich German overabundance a rather canonical type with regards to THORNTON's (2011) classification.
As far as we know, only the dative case is affected by overabundance. This, too, makes it a highly canonical instance of overabundance since overabundance occurs in one cell only. The question why it is restricted to the dative remains unanswered.
In my presentation, I will discuss the Zurich German data with respect to the factors presumably affecting the occurrence of the variants. I will show that the occurrence of the variants is unpredictable and therefore I can present a canonical case of overabundance. Finally, I will discuss some hypotheses on how the dative overabundance could have developed.

Selected references

JOSEPH, Brian (2013): Deixis and person in the development of Greek personal pronominal paradigms. In: Kragh, Kirsten Jeppesen and Jan Lindschouw (eds.): Deixis and pronouns in Romance Languages. Amsterdam: John Benjamins (= Studies in Language Companion Series 136), 19-32. • REICHMANN, Oskar / WEGERA, Klaus-Peter (eds.) (1993): Frühneuhochdeutsche Grammatik. Tübingen: Max Niemeyer. • SEILER, Guido (2003): Präpositionale Dativmarkierung im Oberdeutschen. Stuttgart: Steiner (= Zeitschrift für Dialektologie und Linguistik, Beihefte 124). • THORNTON, Anna M.: (2011): Overabundance (Multiple Forms Realizing the Same Cell): A Non-canonical Phenomenon in Italian Verb Morphology. In: Maiden, Martin et. al. (eds.): Morphological Autonomy. Perspectives from Romance Inflectional Morphology. New York: Oxford University Press, 358-381. • WEBER, Albert (1923): Die Mundart des Zürcher Oberlandes. Frauenfeld: Huber & Co.

Ambivalent affixes: the *-idz(o)* verbalizer in Griko

NIKOS KOUTSOUKOS & ANGELA RALLI
University of Patras
nkoutsoukos@upatras.gr, ralli@upatras.gr

Morphemes are usually divided into groups according to their meaning/function. For example, affixes which derive verbs out of nominal bases are usually called ***verbalizers*** and are grouped together. However, there are affixes which have more than one function and thus can be defined as multifunctional. ***Multifunctional affixes*** are formatives which occur in a multitude of contexts conveying a number of different features (Himmelmann, 2006). These affixes have been considered as an interesting topic in morphological theory since they pose a number of problems such as the asymmetric relation between form and meaning/function (Mel'cuk, 1973).

In this paper, we focus on the formative *-idz(o)* (a form variant of *–iz(o)* in Standard Modern Greek) in Griko, a Greek based dialect of South Italy. This affix is primarily used as a verbalizer, but it has developed some new functions in a number of contexts. The aim is to explore the different functions of this affix, the motivation behind its multifunctionality and the grammatical process underlying the correlations between the different functions. The formative *-idz(o)* attaches to nominal bases in order to form verbs of a specific inflectional class (see Ralli, 2013 for a description of the ICs in Modern Greek):

(1) [[[θer]-idz]-o] 'to reap' < [[θer]-os] 'reaping' (Griko)
BASE-VERBALISER-1SG BASE-NOMINATIVE.SG

However, in a significant number of cases this formative shows up in bases which are already specified as to the verbal category and the inflectional class (IC) and forms verbal doublets:

Table (1)

A. Form with stem extension	**B. Form without stem extension**
[[akkoumb]-o] 'to lean' BASE(V)-1SG	[[[akkoumb]-idz]-o] 'to lean' BASE-IDZ-1SG
[[krat]-o] 'hold' BASE(V)- 1SG	[[[krat]-idz]-o] 'hold' BASE-IDZ-1SG
[[lyp]-o] 'to mourn, feel sad' BASE(V)- 1SG	[[[lyp]-idz]-o] 'to mourn, to feel sad' BASE-IDZ-1SG

In these data we notice that *-idz(o)* does not keep its principal derivational character since it attaches to verbal stems which are already specified as to both the verbal category and the IC. In this paper we examine these cases and try to answer the following questions:

(a) *How does this affix arise in the first place?*

It has been claimed that similar elements may appear in order to show a semantic opposition or an aspectual difference between the two members of the verbal pair. A classic example is the inflectional element *-isc-* which appears in the 4th verbal conjugation in Italian and is a continuation of the Latin derivational suffix -sc-. As Maiden (2003) points out, the broad consensus about Latin -sc- is that it carried an 'ingressive' value, i.e. expressing the meaning of 'becoming/entering a state', but nowadays it does not have any function or particular semantic nuance, either syntagmatically or paradigmatically (see the discussion in Maiden 2003). Data in table 1 show that there is no semantic opposition or aspectual difference

between the forms with and without the formative *-idz(o)* (cf. Karanastasis, 1997). Thus, in order to find the motivation behind this phenomenon we need to go beyond the classic explanations and to examine the inflectional system of the dialect.

(b) *What is the function of this affix in these particular formations?*

In Griko, the inflectional classes of the verbs differ with respect to their productivity (Karanastasis, 1997). This difference triggers cross-paradigmatic levelling which results in the reshape of IC2 in favour of IC1 (most productive paradigm in Griko). Verbs of IC2 show a phonological similarity in the Aorist with verbs which have the formative *-idz(o)* but belong to IC1. We assume that this phonological similarity leads to paradigmatic interference and through the process of proportional analogy *-idz(o)* has developed new properties, that is, it can be added to formations as an explicit formal marker of the inflectional class. In this view, we claim that it has acquired a rather stem-forming (morphomic, in the sense of Aronoff, 1994) status.

(c) *One or two different affixes?*

If looked from a broader perspective, the formative *-idz(o)* serves a dual function in the system as both a verbalizer and a stem-forming suffix and thus is ambivalent between an inflectional and a derivational status. Similar cases can be found in IE data where stem-formatives may have the function of either aspect or mode of action which are separate grammatical and lexical categories, respectively (Kastovsky, 2005: 39). The critical question with respect to these data is whether one should assume that *-idz(o)* is one single affix or two different affixes. We would like to propose that along with the old pattern, *-idz(o)* has developed a new pattern which is used in different contexts.

This case study is important for two different reasons: first, it sheds light on the different functions a multifunctional affix may have, and more importantly, it shows that an affix which is purely derivational may develop into a formally explicit marker of the inflectional class. Second, it shows that multifunctional affixes can emerge from reasons other than semantic opposition or aspectual difference. We propose that paradigmatic levelling may be seen as the driving force behind the rise of a multifunctional affix. Last but not least, it opens the way to discuss the relation between inflection and derivation in the diachronic development of affixes.

References

ARONOFF, M. (1994). *Morphology by itself: Stems and Inflectional Classes* (Linguistic Inquiry Monographs 22). Cambridge, Massachusetts: MIT Press.

HIMMELMANN, N. (2006). How to miss a paradigm or two: multifunctional *ma-* in Tagalog. In: F. Ameka, A. Dench & N. Evans (Eds.), *Catching Language* (487-526). Berlin: Mouton de Gruyter.

KARANASTASIS, A. (1997). *Grammar of the Greek dialects of Southern Italy* (in Greek). Athens: Academy of Athens.

KASTOVKSY, D. (2005). Conversion and/or zero: word-formation theory, historical linguistics, and typology. In: Bauer, L. and S. Valera (Eds.), *Approaches to conversion/zero derivation* (31-50). Münster: Waxmann.

MAIDEN, M. (2003). Verb augments and meaninglessness in Early Romance morphology. *Studi di Grammatica Italiana* 22: 1-61.

MEL'CUK, I. A. (1973). The structure of linguistic signs and possible formal-semantic relations between them. In: R-D. Josette (ed.), *Recherches sur les systèmes signifiants. Symposium de Varsovie* 1968 (103-135). The Hague, Paris: Mouton.

RALLI, A. (2013). *Compounding in Modern Greek*. Dordrecht: Springer.

On a realistic LFG treatment of the periphrastic *irrealis* mood in Hungarian

TIBOR LACZKÓ
Department of English Linguistics, University of Debrecen
tibor_laczko@yahoo.com

1. Introduction. Consider the following table showing the 2SG indefinite and definite parts of the conditional and irrealis mood paradigms in Hungarian.

CONDITIONAL, INDEF. 'would see' + Oid	**CONDITIONAL, DEF.** 'would see' + Od	**IRREALIS, INDEF.** 'would have seen' + Oid	**IRREALIS, DEF.** 'would have seen' + Od
lát-ná-l see-COND-2SG	*lát-ná-d* see-COND-2SG	*lát-t-ál vol-na* see-PAST-2SG be-COND	*lát-t-ad vol-na* see-PAST-2SG be-COND

As these examples demonstrate, conditional verb forms are synthetic and irrealis verb forms are systematically analytic (= periphrastic). The latter use the following two-word pattern: the first word is the conjugated past tense form of the lexical verb and the second verb is the combination of one of the stems of the copula *van* 'be' (*vol-*) and the conditional marker (*-na*) invariably. In other words, Hungarian encodes irrealis mood periphrastically via the combination of two morphosyntactic features: PAST and CONDITIONAL.

Bartos (2000) shows that *volna* is an independent syntactic atom, see his examples in (1-4).

(1) *%vár-t is volna*
wait-PAST.3SG too VOLNA
'he would also have waited'

(2) *%vár-t-ál csak volna*
wait-PAST-2SG only VOLNA
'you would only have waited'

(3) *%vár-t-ál-e volna?*
wait-PAST-2SG-QM VOLNA
'would you have waited?'

(4) *én megsüt-ött-em, te pedig mege-tt-ed volna*
I fry-PAST-1SG.DEF you by.contrast eat-PAST-2SG.DEF VOLNA
'I would have fried and you, in turn, would have eaten (it)'

For a considerable number of speakers (but not for all speakers, hence the % symbol), the two verbal elements can be separated by unquestionably independent words (*is* 'too' and *csak* 'only', as in (1) and (2), respectively), and by the yes/no question marker (*-e*), as in (3), which, under normal circumstances, attaches to finite, fully conjugated verb forms (e.g. *vár-t-ál-e tegnap?* wait-PAST-2SG.INDEF-QM yesterday 'did you wait yesterday?). For the other speakers, these three elements have to follow *volna* immediately. Moreover, these forms can produce right-node-raising effects, as in (4). This construction is acceptable for all speakers.

The challenge then is the development of an appropriate and explicit treatment of this fully and invariably periphrastic irrealis mood paradigm. My framework is Lexical-Functional Grammar (LFG). In the paper, first I will outline an account using the technical apparatus of classical LFG and then I will show how the inferential-realizational approach to paradigms (also containing periphrastic forms) advocated by Ackerman & Stump (2004) and Ackerman et al. (2011) can be formally accommodated in this model.

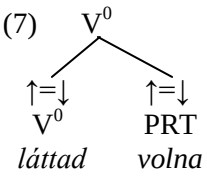
2. Two possible LFG analyses. It is a crucial and shared syntactic aspect of both accounts to be presented below, that, inspired by Laczkó & Rákosi's (2011) treatment of Hungarian non-compositional particle verb constructions, PVCs, I assume a non-projecting syntactic category: PRT (particle) and claim that, in addition to preverbs (particles) in Hungarian PVCs, *volna* is another PRT in its use in the paradigm under investigation.

2.1. A classical LFG solution. This account is morpheme-based. It employs two distinct lexical entries for *volna* (PRT) and the finite, past-tense-marked verb form conjugated for subject agreement and definiteness. Consider the following lexical and syntactic representations in (5)-(7). (5) is the familiar lexical entry for this particular past tense verb form except for the MOOD annotational disjunction: in the regular past tense use, this form contributes the indicative value for the mood feature, or, alternatively, it constrains the mood

to be irrealis. The non-projecting word, *volna*, contributes the specification for irrealis mood, at the same time constraining the tense form of the inflected lexical verb to encode past tense, see (6). PRT, being non-projecting, is head-adjoined to the verbal head (and the two elements are functional co-heads, each making its own contribution to the f-structure of the sentence), see (7).

(5) *láttad,* V 'see <(↑SUBJ) (↑OBJ)>'
(↑TENSE)= PAST
(↑SUBJ PERS)= 2
(↑SUBJ NUM)= SG { (↑MOOD)= INDICATIVE
(↑OBJ DEF)= + | (↑MOOD) =$_C$ IRREALIS }

(6) *volna*, PRT
(↑TENSE)=$_C$ PAST
(↑MOOD)= IRREALIS

(7)


2.2. An inferential-realizational paradigmatic LFG treatment. The basic idea here is that a finite (lexical) verb form like (5) has two, more radically different lexical entries (contra the previous approach), because it is involved in two distinct paradigms: in the regular past tense paradigm and in the irrealis mood paradigm.

(8) *láttad,* V 'see <(↑SUBJ) (↑OBJ)>'
a. (↑TENSE)= PAST
(↑MOOD)= INDICATIVE
(↑SUBJ PERS)= 2
(↑SUBJ NUM)= SG
(↑OBJ DEF)= +

b. (↑MOOD) = IRREALIS
(↑SUBJ PERS)= 2
(↑SUBJ NUM)= SG
(↑OBJ DEF)= +
(↑PRT FORM)=$_C$ VOLNA

(9) *volna*, PRT
(↑PRT FORM)= VOLNA

(8a) is a single, synthetic form encoding a particular (finite, past tense) paradigmatic slot. (8b) is the crucial lexical form from our current perspective. It is one of the two elements of a periphrastic (analytic) mode of expressing a particular irrealis mood paradigmatic slot. Notice that ALL the relevant features characterizing this slot are encoded in this lexical entry, see the first four equations AND it also constrains that this form has to co-occur with the *volna* PRT. At the same time, the lexical entry for *volna* in (9), the second analytic element of this periphrastic expression, has been "impoverished": it no longer contributes any morphosyntactic features or constraints; instead, it only has a FORM feature with the VOLNA value (just like idiom chunks in the classical LFG treatment). The functional co-head annotations of V^0 and PRT remain the same as in (7).

This approach has the following **advantages**. **(1)** It spells out the general and programmatic paradigmatic (inferential-realizational) approach to periphrasis advocated by Ackerman & Stump (2004) and Ackerman et al. (2011) in an LFG framework. **(2)** It leaves the classical view of lexical encoding in LFG intact: by using an appropriate checking and cross-referencing mechanism in the relevant lexical forms, it can avoid recourse to multiple word lexical entries, which would pose rather severe problems for LFG's general morphological assumptions as well as for implementation). **(3)** The devices it employs can be argued to be motivated and justified independently, again, see Laczkó & Rákosi (2011) for the treatment of derivational processes in the case of non-compositional PVCs.

References

ACKERMAN, FARRELL; STUMP, GREGORY T. & WEBELHUTH, GERT. 2011. Lexicalism, periphrasis, and implicative morphology. In Borsley, Robert D. & Börjars, Kersti. eds. *Non-Transformational Syntax: Formal and Explicit Models of Grammar*. Oxford: Blackwell. 325-358. ● ACKERMAN, FARRELL & GREGORY STUMP. 2004. Paradigms and periphrastic expression. In Sadler, Louisa & Andrew Spencer. *Projecting Morphology. Stanford Studies in Morphology and the Lexicon.* Stanford: CSLI Publications. 111-157. ● BARTOS, HUBA. 2000. Az inflexiós jelenségek szintaktikai háttere. In Kiefer, Ferenc. ed. *Strukturális magyar nyelvtan 3. Morfológia*. Budapest: Akadémiai Kiadó. 653-762. ● LACZKÓ, TIBOR & RÁKOSI, GYÖRGY. 2011. On particularly predicative particles in Hungarian. In Butt, Miriam & King, Tracy H. eds. *Proceedings of the LFG '11 Conference*. On-line publication. Stanford: CSLI Publications, ISSN 1098-6782. 299-319.

THE COST OF CHANGE: PLURAL FORMATION OF LOAN WORDS IN ARABIC

LIOR LAKS

Bar-Ilan University
Lior.Laks@biu.ac.il

This study examines the formation of plural forms of borrowed nouns in Palestinian. Arabic. Arabic has two types of plurals, suffix-based sound plural (e.g. *mat ʕa:r – mat ʕa:ra:t* 'airport sg.–pl.') and template-based broken plural (e.g. *maktu:b – maka:ti:b* 'letter sg.–pl.'). These two types of plurals are found also in borrowed nouns, as exemplified in (1).

(1) Plural forms

	Singular	**Sound plural**	**Broken plural**	
a.	faks	faks**a:t**		'fax'
b.	blog	blog**a:t**		'blog'
c.	film		afla:m	'film'
d.	folder		fala:dir	'folder (in a computer)'
e.	malyo:n	malyo:n**a:t**	mala:yi:n	'million'

The nouns in (1a-b) take the sound plural and those in (1c-d) take the broken plural. However, the noun in (1e) takes either plural. The question addressed in this study is: How does the morphological mechanism select between the two formation strategies? I show that the selection of plural forms can be partially predicted, based on morpho-phonological constraints, as well as a semantic criterion.

The study is based on a collection of 153 examples of foreign plural nouns provided by native speakers of Palestinian and Jordanian Arabic. While in most cases the *-a:t* suffix is selected as the default plural marker, there is a noticeable number of cases where plural templates are selected. From the morpho-phonological point of view, I argue that properties of the loan word like stress, vowels and the number of syllables, as well as resemblance to existing Arabic nouns, determine which word formation strategy is selected and within the templatic plural forms, which ones are favored over others. On the semantic dimension, I show that when the foreign noun denotes human being, broken plural are almost exclusively selected.

Plural formation in Arabic. There are two main plural formation strategies in Arabic: Suffixed-based 'sound plural' (SP) and template-based 'broken plural' (BP). There are two SP suffixes in Arabic. The suffix *-i:n* is attached to masculine nouns that denote humans only. It is also restricted with regard to the form of the singular base. This suffix is attached to participle forms like *maCCu:C* as well as other templates like *CaCCa:C*. Feminine formation in Arabic is performed by attaching the suffix *-e* to the masculine form, e.g. *mʕallem-mʕalleme* 'teacher'. The plural formation of such nouns is performed by attaching the SP suffix *-a:t* (*mʕallem*a:t 'teachers fem.'). The broken plural involves internal modification of the singular base, e.g. *maktab-maka:tib* 'office'. BP formation manifest diversity, but at the same time there is a great extent of regularity that allows partial prediction about the selection of one template and not another (Levy 1971, McCarthy 1983, Hammoud 1988). McCarthy & Prince (1990) show that there are structural properties that are drawn from the singular base, which cannot be attributed to the root or the template. This is attested when properties like vowel length are transferred from the singular stem to the plural and when derivational morphemes survive derivational processes. This study also supports their claim, showing that most cases where BP templates are selected for foreign words can be explained based on properties on the loan singular base.

Plural formation of loan words in Arabic. Out of 153 foreign plural nouns such forms, 108 (71%) take the *-a:t* suffix, 34 (22%) are formed is a template and 11 (7%) have two possible plural forms. While Arabic loan words prefer the sound plural *-a:t* as their plural marker (71%), the cases where BP is used does not seem accidental and can be accounted for by faithfulness constraints and a semantic constraint.

Faithfulness constraints. Faithfulness constraints require identity between a base and its derived form. I show that they play an important role in the selection of a plural form.
Selecting the SP suffix *-a:t* is the default strategy, as it allows keeping the structure of the base as is without any internal modifications. Syllable structure, number of syllables and the vowels and consonants of the base all remain intact (*kombyuter- kombyutera:t* 'computer').
The selection of BP templates occurs almost exclusively in cases where the foreign base is mono-syllabic (*film-afla:m* 'film') or bi-syllabic (*banšer-bana:šer* 'puncture'). In such cases the loan word has a syllabic structure of Arabic words that take themselves BP template. The word bank, for example (*CVCC*) has the same structure as the Arabic noun *bayt* 'house'. The latter has the BP form *byu:t* (*CCV:C*) and the plural of *bank* takes the same template (*bnu:k*). The existence of a *CVCC- CCV:C* paradigm in plural formation allows the implementing the same strategy on loan words. When the loan nouns exceed the minimal word size, none of the BP are selected. This is because the morphological mechanism cannot adjust such words into any of the Arabic templates without deletion of syllable and some of the consonants. A quarto-syllabic word like *karburater* 'carburetor' cannot fit any of the existing BP templates.
Stress preservation also plays a role in the selection of BP forms. The suffix *-a:t* always takes stress and therefore changes the location of stress. Selecting some of the BP forms allows keeping the same type of stress. The selection of one of the BP forms *CaCa:CiC* and *CaCa:Çi:C* is not random. *CaCa:Ci:C* is selected when the base has ultimate stress (*karto:n-kara:ti:*n 'cartoon'). The type of stress in the plural form remains in tact, as well as original location of stress in the loan. When stress is penultimate, *CaCa:CiC* is selected (*tánker-tanáːkir* 'tanker'). The location of stress does not remain intact, as it is on the first syllable in the loan. However, the type of stress remains the same. Adding the SP suffix would alter the stress pattern completely as it fall on the suffix (e.g. *fólder -*folder**á:t***, instead of *fala:dir*).

Such cases demonstrate the central role of stress in word formation. Generally, the findings provide further support to the claim that properties of the base are taken into account in word formation. The morphological mechanism tries to keep the derived form as faithful as possible, and in the case of plural formation of loan words, the competition is between two faithfulness constraints: keeping the base structure in tact or keeping the same stress pattern, while taking into considerations templatic restrictions on word formation in Arabic. This is supported by previous studies, where faithfulness constraints have been shown to play a role in adapting foreign words (Bolozky 1978, 1999, Bat-El 1994, 2002, Ussishkin 1999, 2005).

The [+/- human] distinction. The SP *-i:n* suffix is rather unproductive within Arabic existing words. Its use is restricted to nouns and adjective of certain templates like *mCaCCeC*. Since the *-a:t* suffix is never used for plural of [+human] masculine nouns, it cannot be used for the foreign nouns in (2). The only option that is left for plural formation is the use of the BP forms that are also used in Arabic words and that can also denote human being. This provides further evidence that the +/- human (or animate) distinction is relevant for the grammar, as have been shown in various studies.

(2) Plural forms of [+human] nouns

Singular	Plural	
kabtin	kaba:tin	'captain'
daktu:r	daka:tra	'doctor'
imbrato:r	aba:tira	'emperor'

Variation in plural formation. Plural formation of Arabic loan nouns is subject to variation, where the same noun takes both a SP and a broken plural form. The noun *ballo:n* 'balloon', for example, takes both the *-a:t* SP suffix (*ballo:na:t*) and the BP *CaCa:CiC* template (*bala:li:in*). This variation is found between dialects of Arabic and in some cases within the same dialect. This happens mainly in nouns with ultimate stress. The existence of few plural forms result from the two competing faithfulness constraints discussed above.
Examining the morphological behavior of loan nouns provides direct access to the process of word formation and shows how different types of constraints are taken into consideration. Specifically, this study provides further support for a word-based approach, as it demonstrates another case where elements from the base, in addition to the consonants, are transferred to the derived plural form.

What are you afraid of?
A semantic analysis of the French neoclassical element -(o)phobie

MARINE LASSERRE
CLLE-ERSS & University of Toulouse
marine.lasserre@univ-tlse2.fr

Neoclassical compounds contain elements which were originally lexemes in Ancient Greek or Latin but are bound forms in modern languages (for a description, see among others Amiot and Dal (2007)). It is tempting to classify them as an homogenous class, but this hypothesis turns out to be difficult at a purely synchronic analysis. Rather, we consider that neoclassical elements which can form long set of lexemes (for example in French *-(o)logue*, *-(i)cide*, *-(o)crate*, ...) are closer to affixes than elements like *anthrop-* or *pyr-* which can be analyzed as the suppletive stems of lexemes HOMME, 'man' and FEU, 'fire' (see Amiot and Dal (2007) for an analysis of these latter elements). These productive final elements are analyzed here as exponents of constructions (Booij, 2010): they are the formal manifestation of semantic and syntactic operations applied to lexemes, just as affixes are. These operations can be more or less close to the original meaning of the neoclassical element. The final neoclassical element *-(o)phobie* is different from the others in that it can be used as an autonomous lexeme and therefore might be interpreted as the second element of a canonical compound. Nevertheless, we argue here that its meaning is more restricted than when it is used as an autonomous lexeme. In fact, the latter only displays the original meaning of the Greek noun *φοβός*, 'panic fear', already used in Ancient Greek as two bound forms, *-φοβός* to form adjectives with the meaning of 'afraid of' and, more rarely, the suffixed bound form *-φοβια*, with the meaning 'fear of'. The semantic treatment of the French bound form *-(o)phobie* is quite different from both Greek bound forms and French autonomous form. We indeed inventoried at least two distinct meanings that the X-phobie construction can take.

(1) X-phobie constructions: a. 'fear/dread of X'
b. 'aversion/hostility to X'

Meaning (1-a) refers to a psychopathological condition and meaning (1-b) to an inclination to combat. Lexemes which were coined first and/or are the most frequent ones (e.g. *xénophobie*, *agoraphobie*, *claustrophobie*, *homophobie*) play the role of leader words and behave as attracting poles for the construction of the meaning of new lexemes (concerning the notion of leader word, see, among others, Roché (2011)). Hence, the meaning of each new lexeme is built in connection with the meaning of one or more actual lexemes. The question here is about the way in which this connection is realized. In order to answer to this question, we built a corpus based on both dictionaries (*Le Grand Robert* (online) and *Le Trésor de la Langue Française informatisé* (TLFi)) and Google. It is constituted of 529 lexemes, of which 73 (13.8%) come from dictionaries.

A first way to determine which construction we are dealing with is to observe the bases of the constructed words. We can indeed predict that when the base is [+ human], the whole lexeme will contain the feature 'hostility' instead of 'fear', by analogy with leader words *xénophobie* or *homophobie*. However, an analysis only built on the observation of bases is not sufficient: a same lexeme can display different interpretations depending on its context.

The distributional hypothesis of Harris or Wittgenstein (see Sahlgren (2008) or Lenci (2008) for a description and an application to semantics and computational linguistics) states that there is a correlation between the meaning of a word and its distribution. If we implement this for derivational morphology, we can consider that it is necessary to take into account the context in order to determine the meaning or meanings of constructed lexemes. We applied

this hypothesis on our corpus in two ways. First, we investigated the distribution of the ten most frequent[i] lexemes of our corpus in different resources: a Web sample represented by the two first pages of Google results, all the distributions given in Frantext (a corpus of French based on literary texts) and the results given in Les Voisins de Le Monde (http://redac.univ-tlse2.fr/voisinsdelemonde/) and Les Voisins de Wikipédia (http://redac.univ-tlse2.fr/voisinsdewikipedia/), two resources based respectively on ten years of the French newspaper Le Monde and the French Wikipedia. The two latter are tagged for syntactic contexts and are presented via an interface which allows searching the syntactic relations a word enters into. After this first step, the most relevant contexts were applied to the whole corpus by Google searches; for example, a context like "souffrir de", 'suffer from', which is used in several resources with claustrophobie 'claustrophobia', was searched with all the lexemes of our corpus, by typing each entry of our corpus preceding by the sequences "souffrir de", "souffre de" and "souffres de" in Google, in order to have a representative sample of different verbal forms.

	'fear, dread'		'aversion, hostility'	
Google searches	"souffrir de X" 'suffer from X'	"soigner * X" 'treat * X'	"manifester contre * X" 'protest against * X'	"X ambiante" 'prevailing X'
X= christianophobie < chrétien 'Christian'	1	3	66000	2760
X = arachnophobie < araignée 'spider'	491	200000	0	2

Table 1 Example of results for a distributional analysis of -phobie

Obvious results, as those presented in Table 1, were highlighted. However, not every lexeme could be interpreted in this way and a second analysis was necessary. To do so, we constituted a sample of contexts for each lexeme in the corpus by listing its distribution in the two first pages of results provided by Google. When it was possible, these contexts were indexed according to semantic features (e.g. "traitement de X", 'treatment of X' contains obviously the feature 'medicine, pathology', whereas "lutter contre", 'fight against' can be applied either to a disease or an inclination) and all these distributions were then cross-referenced in order to interpret each lexeme relative to each other. This method is representative of how we view the organization of the lexicon and the creation of new words: a new word is coined in connection with those already present in the lexicon and its meaning is constructed thanks to several factors, including its distribution. The distributional hypothesis applied here allowed us to distinguish different meanings for the outputs of an apparently unique construction. This implies that distribution is as relevant as the formal construction itself for the formation of new lexemes. Analogy takes into account not only the form but also the possible contexts of occurrence. Distribution can also distinguish the metaphorical uses of an unique output.

References

AMIOT, D. and DAL, G. (2007). Intergrating neoclassical combining forms into e lexeme-based morphology. *On-line Proceedings of Fifht Mediterranean Morphology Meeting*, ed. Geert Booij and Luca Ducceschi and Bernard Fradin and Emiliano Guevara and Angela Ralli and Sergio Scalise:323-336.

BOOIJ, G. (2010). *Construction Morphology*. Oxford: Oxford University Press.

LENCI, A. (2008). Distributional semantics in linguistic and cognitive research. *Italian Journal of Linguistics*, 20(1):1-31.

ROCHÉ, M. (2011). Des unités morphologiques au lexique. In Roché, M., Boyé, G., Hathout, N., Lignon, S., and Plénat, M. editors, *Des unités morphologiques au lexique*, pages 69-143. Hermès Science.

SAHLGREN, M. (2008). The distributional hypothesis. *Italian Journal of Linguistics*, 20(1):33-54.

[i] Frequency was calculated on the basis of the pages containing each lexeme indexed by Google

The rise and fall of a multifunctional affix: the case of *-hAtÓ* (-able) in Hungarian

ANIKÓ LIPTÁK & ISTVÁN KENESEI
Leiden University / Research Institute for Linguistics, H.A.S.
A.Liptak@hum.leidenuniv.nl / kenesei.istvan@nytud.mta.hu

The issue The received opinion on the Hungarian *–hAtÓ* (-able) affix holds that it is an unifunctional derivational affix that always attaches to passivized transitive verbal roots and derives adjectives from them (Kiefer & Ladányi 2000, Kiefer 2008, see also Wasow 1977 on English *-able*): $[_V\ V_{pass}] + –hAtÓ = [_{Adj}V_{pass}–hAtÓ\ ]$. Contrary to received wisdom, we will show that this view is not tenable for the reason that the category of *ható*-affixed words is in many cases demonstrably not adjectival: applying a battery of diagnostic tests for the adjectival category only gives positive results for some but not all cases of *ható* expressions. The availability of the comparative morpheme (*-bb*) for example singles out *érthető* 'understandable' but not *leültethető* 'seatable' as an adjective, while both forms are perfectly acceptable as prenominal modifiers, as the following data illustrate.

(1) a. az ért-hető(-bb) szövegek *adjectival* **ható**
the understand-ABLE-(COMP) texts
b. a leültethető-(*bb) gyerekek *non-adjectival* **ható**
the seat-ABLE-(COMP) texts

We will argue that the difference observed is that of syntactic complexity. While the adjectival forms are indeed *lexical adjectives*, the non-adjectival forms are syntactic objects corresponding to *non-finite clauses*. Although this might suggest that *ható* is a multifunctional affix – in one of its uses a derivational affix (attaching to verbal roots in the lexicon), in another an inflectional one (deriving non-finite clauses in the syntax) – appearances are deceptive here. *Ható* is in fact an unifunctional affix and inflectional in all its occurrences: it always derives non-finite clauses, some of whose verbal heads are lexicalized as adjectives.

Arguments for distinguishing adjectival vs. clausal *ható* forms The two groups of *ható* forms referred above can be distinguished on the basis of syntactic, semantic and morphophonological evidence of various sorts. To cherry-pick a handful of these, next to their compatibility with the comparative morpheme (cf. 1), the two forms also differ in (i) their compatibility with intensifiers (cf. 2); (ii) their serving as input for further derivational processes, like nominalizations (cf. 3); (iii) the retention of the event and argument structure of the underlying verbal predicate (cf. 4), and (iv) the morphophonological realization of plural morpheme they appear with (cf. 5).

(2) a. egy {kifejezetten / nagyon} hosszú / ért-hető szöveg
a particularly very long understandable text
b. a problémára { * kifejezetten / jól} alkalmaz-ható megoldás
the problem.ONTO particularly easily apply-ABLE solution

(3) ért-hető-ség / * leültet-hető-ség
understand-ABLE-NOM seat-ABLE-NOM
'understandability' 'seatability'

(4) a. a *(gyerekeknek) ad-ható könyvek
the children.DAT give-ABLE books
'the books that can be given to children'

b. a (*Mari által) beszámít-ható emberek
the Mari by count-ABLE people
'people who are sane to/for/by Mari'

(5) a. A piros színű gombák nem e-hető{–k/–ek}.
the red coloured mushroom-PL not eat-ABLE-PL
'Red mushrooms are not edible.'

b. A legfiatalabb asszisztensek nem bíz-ható{–k/?*–ak} meg nehezebb feladatokkal.
the youngest assistants not trustable-PL PV difficult.COMP tasks.WITH
'The youngest assistants cannot be trusted with more difficult tasks.'

Further evidence for the syntactic clausehood of non-adjectival *ható* forms and the lexical wordhood of *ható* adjectives will be provided from the realm of preverb placement, where *ható* clauses show the same possibilities as finite (and some non-finite) clauses in Hungarian. After focus and negation (and in some elliptical contexts not illustrated here), the preverb (PV) belonging to the base verb is necessarily split off the verb and may appear at a distance from it (see verb and preverb underlined), cf. (6). With adjectival *ható* expressions, such splitting is never possible, which is predicted if these forms are lexical words and syntax does not have access to their subparts (cf. Lexical Integrity; Selkirk 1982, Booij 1985).

(6) János nem <u>bíz-ható</u> ma <u>meg</u> ezzel a feladattal. *clausal* ható
J. not trust-ABLE today PV this.WITH the task.WITH
'János cannot be trusted today with the task.'

(7) * János nem <u>számít-ható</u> ma <u>be</u>. *adjectival* ható
J. not count-ABLE today PV
(intended meaning) 'János is not sane' (lit. 'countable on')

***Ható* as a unifunctional inflectional affix** To explain the above noticed differences, we will argue that *ható* clauses are non-finite clauses with passive and modal semantics, and that *ható* is in fact bimorphemic, containing *-hAt* and *-Ó*. *-Ó* spells out the [–finite] T head, while *-hAt* lexicalizes Modality, which in turn selects a VoiceP >*v*P projection. The derivation of *ható* clauses is thus akin to that of non-finite clauses and explains why *ható* clauses necessarily retain their base verb's event and argument structure (cf. 4), and exhibit syntactic independence between the verb and its preverb (cf. 6).

It will be shown that adjectival *ható* forms originate from syntactically derived *ható* clauses, too, but they acquire their lexical status/adjectival category via the process of lexicalization. Supportive evidence for lexicalization comes from the fact that *ható* adjectives (unlike *ható* clauses) cannot be productively formed and that such a lexicalization process is also attested with other non-finite verb forms that give rise to adjactives.

This novel account allows us to define *ható* as an inflectional affix in all its occurrences, and this in turn leads to the conclusion that even though it gives the impression of an affix that can attach to both word-level and phrase-level entities, it is in fact a unifunctional morpheme.

References

BOOIJ, G. E. 1985. Coordination reduction in complex words: A case for prosodic phonology. In: H. van der Hulst, N. Smith (eds), *Advances in nonlinear phonology*, 143–60. Foris, Dordrecht. • KIEFER, F. 2008. A melléknevek szótári ábrázolásáról. *Strukturális magyar nyelvtan 4.* 505-538. • KIEFER F. & M. LADÁNYI. 2000. A szóképzés. In: F. Kiefer (ed.) *Strukturális magyar nyelvtan 3, Morfológia,* Akadémiai Kiadó, Budapest, 137-164. • SELKIRK, E. 1982. *The Syntax of Words*. Cambridge MA, MIT Press. • WASOW, T. 1977. Ttransformations and the lexicon. In P. Culicover, A. Akmajian and T. Wasow (eds), *Formal syntax*. NY: Academic Press.

Polyfunctional pronominal markers in South-Western Mande

DARIA MISHCHENKO
Langage, Langues et Cultures d'Afrique Noire, CNRS
zenitchiki@yandex.ru

The paper deals with the syntactic functions and morphosyntactic features of polyfunctional (PF) pronominal markers in South-Western Mande languages (SWM). This group includes 6 languages spoken in Guinea, Liberia and Sierra Leone. The central purpose of the paper is to provide a comparative analysis of PF pronominal markers in SWM. This analysis is based on the synchronous grammar descriptions and manuals on SWM languages as well as on the data collected during the author's fieldwork with Woi-Bhalaga, a Guinean dialect of Looma. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-LABX-0083)".

It is typical for Mande languages to have several sets of pronominal markers opposed according to their syntactic functions and/or morphological status. One of them is often PF: for example, in Goo (Southern Mande) several sets of subject pronouns are opposed to a set of PF pronouns occurring in non-subject positions (Vydrin 2013). Traditionally, PF markers in SWM are regarded as personal pronouns, too. However, here they are not autonomous word forms but rather inseparable and intransposable bound morphemes. Moreover, the paradigms of PF markers include two suprasegmental morphemes, these expressing the 1st and the 3rd person singular (except for Mende, where the 1st person singular is expressed by means of a segmental morpheme *ɲá-*); they have no segment exponent and manifest themselves on the following morpheme through the change of its initial consonant and tonal alternations. On this evidence, V. Vydrin in (2011) proposes to name such elements in SWM personal prefixes. Here, this term will be avoided in order not to refer to their morphological status.

The specificity of these morphemes consists in their transcategorial nature: they can be attached to a verb (1), a noun (2) or a postposition (3):

LOOMA

(1) *Gè* *ká-nì.*
1sg.base 3sg.pi\see-aor
'I saw him'.

(2) *kɛ̀ ɛ̀*
1sg.pi\father
'my father'

(3) *Fòlòmò* *ɣà* *níkɛ́-y* *wódì-zù* *má.*
Folomo 3sg.npst ref\cow-def sell-ipfv 3sg.pi\on
'Folomo is selling the cow to him'.

Particular functions which can be fulfilled by PF pronominal markers in SWM vary. The standard set includes a direct object (4), an argument of a qualitative verb (5), a possessor in the situation of inalienable possession (6) and an indirect object with a postposition (7):

MENDE

(4) *ŋgá* *(ˋ)-píé* *lɔ̀* *fóló* *gbí* *mà.*
1sg.ipfv 3sf.pf-do cop ref\day every ipfv
'I do it everyday'. (Innes 1971: 150)

BANDI

(5) *Í-lùkpú-ŋgɔ* *lè́.*
2sg.pf-short-qual cop
'You are not tall'. (Bandi, ms.: II, 17)

ZIALO

(6) *(́)-tókó-y*
3sg.pf-hand-def
'his hand' (Babaev 2011: 60)

KPELLE (LIB.)

(7) *Gáàloŋ* *èì* *kɛ̀* *ɓɛ́* *èì* *tɔŋ̂* *pù* *(̀)-máà.*
ref\chief.def 3sg.cond2 be here 3sg.cond2 ref\law put 1sg.pf-on
'If only the chief was here, he'd punish me'. (Leidenfrost & McKay 2005: 77)

At the same time, there are a lot of smaller differences in the use of PF pronominal markers in SWM. Thus, for example, SWM languages vary according to whether PF markers are necessary even with a full NP present or not. For example, in Looma in an intransitive resultative construction, verbal agreement with a plural subject is facultative (8), while in Zialo it is obligatory (9):

LOOMA

(8) *Zúnú-y-tì* *tí-ɓɛ́tɛ́-vɛ́* */zúnú-y-tì* *kpɛ́tɛ́-vɛ́.*
ref\man-def-pl 3pl.pi-be.wealthy-res /ref\man-def-pl be.wealthy-res
'The men are wealthy'.

ZIALO

(9) *Pélé* *fèlè-gɔ̀-y-tì* *tí-w̃ɛ́-yàà* *lè.*
ref\road two-def-def-pl 3pl.pf-meet-res exi
'The two roads have been linked together'. (Babaev 2011: 129)

Especially interesting cases represents the use of PF markers for marking inalienable possession on the head noun. In this case their use may depend on a whole range of factors, such as the type of a noun representing possessum, the meaning of the category of number expressed by a NP in the possessor's position, its referential status, etc. In the report, a detailed analysis of such peculiarities of PF pronominal markers in SWM will be proposed.

References

(A Bandi Grammar). Unpublished Manuscript. Monrovia: The Institute for Liberian languages vol. I-II.

BABAEV, K. 2011. Zialo Language: Grammar Sketch and Dictionary. Moscow: HSSU. (in Russian)

INNES, G. 1971. A Practical Introduction to Mende. London: School of Oriental and Classical Studies.

LEIDENFROST, T., MCKAY, J. 2005. Kpelle–English Dictionary with a grammar sketch and English-Kpelle finderlist. Moscow (USA): Palaverhut Press.

VYDRIN, V. 2011. Ergative/Absolutive and Active/Stative alignment in West Africa: The case of Southwestern Mande. In: Studies in Language 35:2, 409-443.

VYDRIN, V. 2013. Verbal construction with a progressive meaning in Goo: Transformation of non-subject pronouns to subject ones. Handout, The Third conference "Typology of Morphosyntactic Parameters", Moscow, 16-18 October 2013. (in Russian)

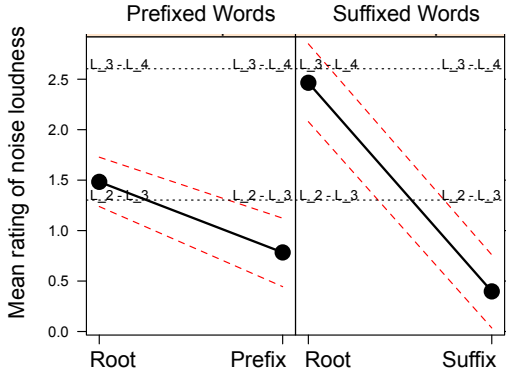
Subjective perception of affixation: A test case from Spanish

ANNE PYCHA
University of Wisconsin, Milwaukee
pycha@uwm.edu

Affixation changes our perception of a word, but in ways that we do not fully understand. Typological studies point to the prevalence of suffixation over prefixation as evidence that suffixed words provide perceptual advantages, because they place the root in initial position (Hawkins & Cutler 1988). But this idea is at odds with the experimental literature, which shows that listeners "strip" prefixes from multimorphemic words (Taft & Forster 1975), and provides evidence for a prefixation advantage: the root *pay* and the prefixed word *prepay* facilitate recognition of *payment*, but the suffixed word *payable* inhibits it (Marslen-Wilson et al. 1994; Zwitserlood, & Roelofs 1991, Feldman & Larabee 2001). Sorting out the perceptual consequences of affixation is important for models of word recognition, and could help explain problems in phonology, such as the lack of prefix-triggered alternations on roots (Hyman 2008).

The current study investigates this issue using a noise-rating task in Spanish. Participants heard Spanish phrases that had been partially overlaid with white noise, and assigned a rating indicating how loud the noise sounded, on a scale from 1 (softest) to 5 (loudest). The phrases consisted of verbs plus pronouns; the noise, indicated by strikethrough, coincided with either the prefix (~~*me*~~ *patea* 'she kicks me', *se* ~~*me*~~ *pisa* 'I am stepped on'), the root, or the suffix (*patéa*~~*me*~~*,* 'kick me', *pisa*~~*me*~~*lo* 'step on it for me'). Spanish offers a good test case for our question because the same personal pronoun clitics, which behave phonologically like affixes (Hualde 2005, 2012), can precede or follow the verb root, and need not occur at word edges – properties that allow us to focus on affixhood per se. The noise-rating task permits a measurement of listeners' subjective experience of the word, with the premise that listeners will experience the overlaid noise as softer when the spoken word itself seems perceptually clearer (Jacoby et al. 1988).

A speaker of Colombian Spanish recorded the words. The intensity of each root and affix was calculated separately and white noise was added accordingly at one of three signal-to-noise ratios: +24 dB, +17 dB, or +10 dB. Forty-four native Spanish-speaking participants each heard 36 target phrases, plus 80 fillers, and were asked to attend to phrase meanings during the rating task. Results, analyzed with proportional odds logistic regression, showed an interaction between Noise location and Affix type ($\beta = -1.37$, $t = -7.10$, $p < 0.01$).



Participants rated noise on suffixed roots as very loud, suggesting diminished perceptual clarity, but rated noise on prefixed roots as relatively soft, suggesting perceptual enhancement. These findings confirm the previously reported prefixation advantage, and extend it by suggesting that prefixes facilitate priming of related words precisely because they

provide listeners with enhanced perception of the root. Interpreted in light of listener-based theories of diachronic change (Ohala 1993), this finding could help to explain the lack of prefix-root phonological alternations cross-linguistically: just as certain sequences of vowels and consonants encourage misperception more than others, so do certain sequences of morphemes.

References

FELDMAN, L. B., & LARABEE, J. (2001). Morphological facilitation following prefixed but not suffixed primes: Lexical architecture or modality-specific processes?. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 680.

HAWKINS, J. A., & CUTLER, A. (1988). Psycholinguistic factors in morphological asymmetry. *Explaining language universals*, 280-317.

HUALDE, J. I. (2005). *The sounds of Spanish*. Cambridge University Press.

HUALDE, J. I. (2012). Stress and Rhythm. *The Handbook of Hispanic Linguistics*. In Hualde, J. I., Olarrea, A., & O'Rourke, E. (Eds.). *The Handbook of Hispanic Linguistics* (Vol. 69). (pp. 153-172). Wiley.

HYMAN, L. M. (2008). Directional asymmetries in the morphology and phonology of words, with special reference to Bantu. *Linguistics*, *46*(2), 309-350.

JACOBY, L. L., ALLAN, L. G., COLLINS, J. C., & LARWILL, L. K. (1988). Memory influences subjective experience: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 240.

MARSLEN-WILSON, W., TYLER, L. K., WAKSLER, R., & OLDER, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological review*, *101*(1), 3.

OHALA, J. J. (1993). The phonetics of sound change. In Charles Jones (ed.), *Historical Linguistics: Problems and Perspectives* (pp. 237-278). London: Longman.

SCHRIEFERS, H., ZWITSERLOOD, P., & ROELOFS, A. (1991). The identification of morphologically complex spoken words: Continuous processing or decomposition?. *Journal of Memory and Language*, *30*(1), 26-47.

TAFT, M., & FORSTER, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of verbal learning and verbal behavior*, *14*(6), 638-647.

Gender neuterisation in the light of language contact: the case of Pontic & Heptanesian

ANGELA RALLI, MARIANNA GKIOULEKA & VASILIKI MAKRI
University of Patras
ralli@upatras.gr, mgkiouleka@upatras.gr & vasmakri@upartras.gr

From a theoretical perspective, Corbett (1991) argues that gender is an inherent property of nouns and its assignment depends on semantic, pragmatic, phonological and morphological factors. Ralli (2002) maintains that gender constitutes a lexically-specified feature, actively involved in the word-formation process of Standard Modern Greek (hereafter SMG), where a tripartite gender distinction into masculine, feminine and neuter characterizes nouns, adjectives and determiners.

In this paper, we examine grammatical gender assignment in loanwords of two different contact-induced dialectal systems of Modern Greek (hereafter Greek), Pontic and Heptanesian. Pontic has been affected by the agglutinative Turkish, while Heptanesian has been influenced by the semi-analytic Romance (Italian and Venetian). Data drawn from these dialects reveal similarities as well as deviations in gender assignment compared to that in SMG, and confirm the significance of gender as an organizational feature of grammar (Corbett 1991, Ralli 2002). More particularly, we show that in both Pontic and Heptanesian, gender assignment in loanwords is primarily conditioned by language-internal properties of the recipient system, that is, an old attested tendency for shifting to the neuter value of certain types of nouns (Hatzidakis 1907) and the prevalence of the semantic feature of humanness, while a phonological similarity between the inflectional endings of the donor and those of the recipient language also play a role. We argue that paradigms of loanwords are formulated accordingly, except for some rare but interesting cases, and try to interpret a number of peculiarities that arise in gender agreement mechanisms involving a determiner and a noun, by appealing to a language contact effect.

First in Pontic, although the tripartite grammatical gender distinction is maintained, there is a dominant tendency to neutralize gender distinctions in the context of plural in favor of the neuter value, affecting both native (1c-d) and loan words (1a-b), while the donor language, i.e. Turkish, is gender-neutral, at least on the morphological level. Interestingly, when gender leveling occurs, it may affect both the determiner and the noun (1a) or only the determiner (1d) (Papadopoulos 1955). As a consequence, there are repercussions in the agreement patterns of the particular dialect (occasional absence of agreement between the determiner and the noun (1d)), as well as a reformulation of the inflectional paradigms.

(1)***a.*** *i giurulti ~ ta giurultia* [the.FEM.NOM.SG noise.FEM.NOM.SG ~ the.NEU.NOM.PL noise.NEU.NOM.PL] < Turkish *gürültü* : 'the noise(s)' • ***b.*** *i orospi ~ i orospiδes* [the.FEM.NOM.SG prostitute.FEM.NOM.SG ~ the.FEM.NOM.PL prostitute.FEM.NOM.PL] < Turkish *rospı* : 'the prostitute(s)' • ***c***. *i mana ~ i manaδes / ta manaδes* [the.FEM.NOM.SG mother.FEM.NOM.SG ~ the.FEM.NOM.PL mother.FEM.NOM.PL / the.NEU.NOM.PL mother.FEM.NOM.PL] (Ofis Pontic): 'the mother(s)' • ***d***. *i kossara ~ ta* kossaras [the.FEM.NOM.SG chicken.FEM.NOM.SG ~ the.NEU.NOM.PL chicken.FEM.NOM.PL]: 'the chicken(s)' • ***e***. *o minas ~ ta minas* [the MASC.NOM.SG month MASC.NOM.SG ~ the NEU.NOM.PL month MASC.NOM.PL] 'the month(s)

Second, the Heptanesian data also depicts a certain prominence of the neuter gender, contrary to Romance, where there are only masculine and feminine gender values expanding even on native examples (2c). This prominence is shown in both the determiner and the noun and in both number values, singular and plural. Contrary to Pontic, however, agreement is preserved between the determiner and the noun, as depicted in (2a, d).

(2)***a***. *i belandza/to belandzi ~ i belandzes/ta belandzia* [the.FEM.NOM.SG weighing.scale.FEM.NOM.SG/ the.NEU.NOM.SG weighing.scale.NEU.NOM.SG ~ the.FEM.NOM.PL wheighing.scale.FEM.NOM.PL / the.NEU.NOM.PL wheighing.scale .NEU.NOM.PL] < Italian *la*

bilancia : 'the scale(s)' • ***b***. *to tsekini* [the.NEU.NOM.SG venetian.gold.coin.NEU.NOM.SG] < Venetian *lo zéchin*: 'venetian gold coin'• ***c***. *to mayerjo ~ ta mayerja* [the.NEU.NOM.SG kitchen.NEU.NOM.SG ~ the.neu.NOM.PL kitchen.NEU.NOM.PL] : 'the kitchen(s)' • ***d***. *to sokors-o* [the.NEU.NOM.SG assistance.NEU.NOM.SG] < Venetian *il socórs-o* : 'assistance' • ***e***. *i indovino ~ i indovines / ta indovina* [the.FEM.NOM.SG clairvoyant.NEU.NOM.SG ~ the.FEM.NOM.PL clairvoyant.FEM.NOM.PL/ the.NEU.NOM.PL clairvoyant.NEU.NOM.PL] < Italian *l'indovino ~ gli indovini* : 'the clairvoyant(s)'

In order to incorporate loanwords of Turkish origin, Pontic activates semantic, phonological and morphological mechanisms, which do not differ from those employed by SMG, Heptanesian and the majority of Greek dialects. For instance, there is an innate tendency to assign masculine and feminine values to male and female nouns respectively, which may be enhanced by the interference of the semantically-implied genderness (1b-c) (covert gender according to Aronoff 1998) but overtly genderless Turkish. Generally, grammatical gender assignment in Pontic confirms Ralli's (2002) suggestion that the lexically-specified gender is not part of the features of the inflectional suffixes but belongs to those of stems, since gender in Pontic words is morphologically realized on the stem and/or the determiner. Nevertheless, a question arises why the Pontic data conflict with the corresponding SMG one as far as agreement and the predominance of the neuter value are concerned, as well as why shift to neuter gender occurs only in the plural number.

A semantic effect also applies in Heptanesian, which bears witness to the effect of humanness on gender, since -human loanwords are predominantly neuter (2b). It is also corroborated by a tendency to shift to the neuter value loanwords which belong to a different gender, e.g. feminine, as in (2a), for phonological reasons, that is, because of a form similarity between the endings of the donor (Italian/Venetian) and the recipient language (Greek/Heptanesian). This tendency seems to be stronger than in Pontic, since it may alter even +human nouns (2e), affects both the determiner and the noun, and is spread throughout the entire paradigm of singular and plural. In addition, like Pontic, Heptanesian brings to the forefront the linguistic property of Greek to have the stem and/or the determiner lexically specified for gender. Again, for loanwords, this property seems to be enhanced by a similar tendency in the donor language, i.e. Italian/Romance, as stated by Thornton (2003).

Interestingly in Pontic, the shift of feminine and masculine nouns to the neuter gender seems to be related to the semantic +/-human feature, and has an impact on the agreement between the determiner and the noun. In certain areas of Pontus (e.g. Ofis), an innovative absence of agreement is systematically attested in the plural of +/-human feminine nouns (1c), while in other areas, this absence is sporadic, related to -human masculine and feminine ones (1d,e). We believe that contact with the genderless Turkish must have played a role into it. In contrast, Heptanesian, which has been in contact with a language with gender distinctions, shows a systematic agreement in both singular and plural, and whenever gender shifts to neuter, the noun is combined with a determiner marked for neuter.

To sum up, this paper demonstrates that the integration of loans in a recipient linguistic system which bears an overtly marked gender may offer insights into grammatical gender assignment. It also provides evidence that in contact morphology, this feature results from the interplay of language-internal factors of both the recipient and the donor languages.

References

ARONOFF, M. 1998. Gender Agreement as Morphology. *Proceedings of the 1stMediterranean Morphology Meeting*, ed. by G. Booij, A. Ralli & S. Scalise. Mytilene 1997, 7-18. • HATZIDAKIS, G. 1907. *Meseonika ke nea Elinika* [Medieval and Modern Greek], v. 2. Athens: Sakellariou. • CORBETT, G. 1991. *Gender*. Cambridge: Cambridge University Press.• PAPADOPOULOS, A. 1955. *Istoriki Grammatiki tis Pontikis Dialectou* [Historical Grammar of the Pontic Dialect]. Athens: Epitropi Pontiakon Meleton.• RALLI, A. 2002. The Role of Morphology in Gender Determination: Evidence from Modern Greek. *Linguistics* 40: 519-551.• THORNTON, ANNA M. 2003. L'assegnazione del genere ai prestiti inglesi in italiano [Gender assignment in English borrowings in Italian]. *Italiano e inglese a confronto*, ed. by Anna-Vera Sullam Calimani. Firenze: Cesati, 57-86.

Approaching case polysemy from the lexicon:
The case of the Kabardian instrumental

DARIA RYZHOVA
NRU Higher School of Economics
daria.ryzhova@mail.ru
MARIA KYUSEVA
NRU Higher School of Economics
mkyuseva@gmail.com
PETER ARKADIEV
Institute of Slavic Studies (Russian Academy of Sciences), Russian State University for the Humanities, Sholokhov Moscow State University for the Humanities
peterarkadiev@yandex.ru

The instrumental case in Circassian languages Adyghe and Kabardian (North-West Caucasian language family) is highly polysemous, its uses ranging from typical instrument and means to the encoding of spatial and temporal relations and even such functions as iudicantis and stimulus (Kuznetsova & Serdobolskaya 2009). Though it is fairly obvious that not all functions of the instrumental arise when the case marker combines with any lexical noun, semantic interaction of the instrumental case with nominal lexemes has not been studied in any detail. Likewise, this issue has not been raised in the detailed cross-linguistic studies of the polysemy in the domain of the instrumental such as Narrog & Ito (2007) and Lehmann & Shin (2005), and general works on case polysemy, e.g. Malchukov & Narrog don't treat it either.

We follow Aristar (1997), Rakhilina (2000) and Luraghi (2003) in assuming a crucial role of nominal lexical semantics in the functioning and interpretation of case, and empirically test the hypothesis that the functions of the highly polysemous Circassian instrumental case are non-randomly distributed according to the lexical-semantic classes of nominals. The data for our research come from the Besleney dialect of Kabardian spoken in the village Ulyap in Republic of Adygheya (Russia), collected during the field-trip July 2013.

Our investigation is based on a questionnaire containing fifty nominal lexemes; we asked native speakers to compose naturally sounding sentences containing the target noun in the instrumental case. The results of the experiment show that Besleney nouns fall into several classes according to their co-occurrence with the meanings of the instrumental case.

1) Nouns which resist co-occurrence with the instrumental altogether, i.e. those for which our consultants experienced difficulties of providing naturally sounding contexts; these are e.g. nouns denoting animals ('fly', 'fox').

2) Nouns co-occurring with a wide range of the meanings of the instrumental, e.g. those denoting humans. Thus, the instrumental form of 'father' can have the functions of iudicantis ('it's good for father'), spatial direction ('go towards father'), stimulus ('be content with father'), specification ('paternal uncle', lit. 'uncle father-INS').

3) Nouns naturally co-occurring with just one meaning of the instrumental, like nouns denoting prototypical instruments ('hammer') or food ('to feed with porridge').

4) Nouns regularly combining with several meanings of the instrumental. These fall into two subgroups.

4.1) Nouns combining with semantically similar meanings of the case. E.g. with nouns denoting spatial objects the instrumental can, depending on the type of verb, express motion towards, from or via the landmark, while the combination of the instrumental with temporal nouns express temporal extent ('run home in half an hour'), temporal distance ('return in a year') and temporal location ('at night').

4.2) Nouns which due to their lexical semantics are compatible with heterogeneous semantic roles. E.g. nouns denoting cattle when combined with the instrumental can express exchange

equivalent (‘exchanged a cow for a horse’) and means of transport (‘ride a horse’), while body-part nouns can be used as instruments (‘take with the hand’), point of application (‘hit one’s head’) or beneficiary (‘good for the eyes’).

On the basis of this study we draw the following conclusions:

1) Individual functions of the instrumental case are distributed across lexical classes of nominals in a principled manner, which suggests that each function selects those lexical classes which are most semantically compatible with it. This closely resembles the interaction between verbs and argument structure constructions (Goldberg 1995) or prefixes (Janda et al. 2013), which form radial networks where particular types of verbs combine with particular constructional or prefixal meanings.

2) According to their compatibility with the instrumental, nominals can be divided into “core” and “peripheral” (cf. Aristar 1997). The “core” are those which are robustly associated with a single function or with a family of related functions of the case. To the “periphery”, in addition to those nominals which do not readily combine with the instrumental, belong those whose instrumental form can yield a wide range of meanings, usually more generalized and less “selective” than the core meanings of the instrumental, cf. such functions as beneficiary or stimulus. Both the wide range of possible interpretations of such nouns in the instrumental and inter-speaker variation in their interpretation are probably due to the low frequency of instrumental forms of the nouns of these lexical classes.

We conclude that explicitly taking into account and controlling for lexical semantics is necessary for an adequate characterization of case polysemy and that looking at the meanings of case from the perspective of lexical semantic classes of nominals it combines with is a promising direction of empirical and cross-linguistic research of morphological case and related functional domains.

References

ARISTAR A.R. (1997). Marking and hierarchy types and the grammaticalization of case-markers. *Studies in Language* 21, 313–368.

GOLDBERG A. (1995). *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago, London: The University of Chicago Press.

JANDA L., A. ENDERSON, J. KUZNETSOVA, O. LYASHEVSKAYA, O. MAKAROVA, T. NESSET, S. SOKOLOVA (2013). *Why Russian Aspectual Prefixes Aren’t Empty. Prefixes as Verb Classifiers*. Bloomington: Slavica.

KUZNETSOVA J. & N. SERDOBOLSKAYA (2009). Dvojnoe padežnoe markirovanie: unikal’nyj slučaj adygejskogo jazyka [Double case marking: A unique case in Adyghe]. In: Ya.G. Testelets et al. (eds.), *Aspekty polisintetizma: Očerki po grammatike adygejskogo jazyka*. [Aspects of Polysynthesis: Studies in the Grammar of Adyghe] Moscow: RGGU, 166–200.

LEHMANN CHR., SHIN Y.-M. (2005). *The Functional Domain of Concomitance. A Typological Study of Instrumental and Comitative Relations*. Berlin: Akademie Verlag.

LURAGHI S. (2003). *On the Meaning of Prepositions and Cases. The Expression of Semantic Roles in Ancient Greek*. Amsterdam, Philadelphia: John Benjamins.

MALCHUKOV A., H. NARROG (2009). Case polysemy. In: A. Malchukov & A. Spencer (eds.) *The Oxford Handbook of Case*. Oxford: Oxford University Press.: 518–534.

NARROG H., ITO SH. (2007). Re-constructing semantic maps — The Comitative-Instrumental area. *Sprachtypologie und Universalienforschung* 60, 273–292.

RAKHILINA E.V. (2000). Semantika russkogo tvoritel’nogo i taksonomija [The semantics of the Russian instrumental and nominal taxonomy]. In: E.V. Rakhilina. *Kognitivnyj analiz predmetnyx imën: semantika i sočetaemost’*. [Cognitive Analysis of Object-denoting Nouns: Semantics and Co-occurrence] Moscow: Russkie slovari, 72–85.

A new semantic map approach to account for the development of polysemy in affixes

MARION SCHULTE
Bielefeld University
marion.schulte@uni-bielefeld.de

This paper presents a new approach to account for the polyfunctionality of derivational affixes, namely a significantly adapted version of semantic maps. Semantic maps have until now mainly been used to account for the semantics of inflectional categories across a variety of languages (e.g. Haspelmath 2003). In this paper, an adapted version of the semantic map approach will be introduced. The advantages of this adapted approach will be illustrated by a diachronic study of the semantics of two English derivational suffixes, *-age* and *-ery*.

Semantic maps represent the different functions of grammatical categories and the connections between these on a two-dimensional map. This approach has a number of advantages over other ways of representing affix semantics: it is straightforward to compare the semantics of different grammatical categories, and it is easy to account for language change. Another important advantage is the fact that semantic maps do not make a priori assumptions regarding the structure of a morphological category. For example, if a category has a single core sense and sense extensions of this primary sense, a semantic map can depict this easily. But if a category shows a different structure this can also be represented on a semantic map. Spatial prepositions, for example, are often expected to exhibit a prototypical structure (e.g. Tyler & Evans 2001). Semantic maps do not make such a priori assumptions as they utilise a bottom-up approach, and are thus very open and flexible. However, traditional maps such as Haspelmath's have been developed as a means for cross-linguistic comparison, and are mainly used for inflectional categories. The adapted approach presented here can account for the semantics of derivational affixes, and makes it possible to compare different affixes within a single language as well as across different languages. It is also a particularly suitable tool for representing the semantics of a single affix in different time periods of a language, and can thus provide valuable information on the nature of language change regarding the semantics of derivational affixes. As a major improvement over traditional semantic maps, this approach incorporates frequency information in a new way (see also Cysouw 2007 for an argument regarding the importance of this issue).

The study that provides the data to illustrate the adapted semantic map approach takes both dictionary and corpus data into account. The Oxford English Dictionary (OED), an extremely large historical dictionary, enables a large-scale diachronic investigation, which

compares the semantics of -*age* and -*ery* neologisms in Middle English (1100-1499) to Present Day English (since 1900). The results of this investigation are then compared to -*age* and -*ery* derivatives in the British National Corpus (BNC). This corpus represents language use in the late 20th century.

Semantic maps are then created for each of the suffixes, periods, and data sources investigated. Apart from being a good visual representation of the morphological categories under investigation, the analysis of these maps provides a number of valuable insights into the nature of affix semantics in general. The analysis of -*age* derivatives, for example, reveals that this category has more than one core sense. In Middle English, the readings ACTION and CHARGE are core interpretations, but in Present Day English, this has changed to ACTION and AMOUNT. And not only the core senses have shifted since Middle English, some readings that were commonly used then have disappeared completely since. There has clearly been a significant amount of change in the semantics of this suffix in the last 500 years. All of this is illustrated on the semantic maps.

Another important question the maps help to answer regards the development of polysemy in affixes. The connecting lines make polysemy visible, and change in the amount or kind of polysemous derivatives can be represented in a straightforward manner. The maps show some already well-documented cases of sense extension, e.g. ACTION → OBJECT, but also reveal unexpectedly close links between the core senses and between other readings as well. The diachronic analysis shows how the connections between these readings are strengthened over time, which leads to a highly interconnected polysemous category with many regular semantic overlaps and sense extensions. There are also significant differences in the kind of polysemy shown by -*age* and -*ery* derivatives. Although the semantic contribution of these two suffixes is often said to be very similar (e.g. Lieber 2004), the semantic maps reveal significant differences in the connections between the various readings of the derivatives. This shows that the adapted semantic map method is a good tool for bringing subtle aspects of affix semantics to light.

References

CYSOUW, M. (2007). Building semantic maps: The case of person marking. In M. Miestamo & B. Wälchli (Eds.), *New challenges in typology*. Berlin: Mouton de Gruyter, 225-247. • HASPELMATH, M. (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In M. Tomasello (Ed.), *The new psychology of language* (Vol. 2). Mahwah, NJ: Erlbaum, 211-242. • LIEBER, R. (2004). *Morphology and lexical semantics*. Cambridge: Cambridge UP. • TYLER, A. & EVANS, V. (2001). Reconsidering polysemy networks: The case of over. *Language*, 77(4), 724-765.

Functions and typology of the compounding stem:
meaning-independent elements in compounds

MASAHARU SHIMADA
University of Tsukuba
shimada.masaharu.fu@u.tsukuba.ac.jp

AKIKO NAGANO
Tohoku University
tmg0002as@gmail.com

TATSUHIRO OKUBO
University of Tsukuba
okubo.tatsuhiro@gmail.com

MASANAO ASANAO
University of Tsukuba
nao0921.s@gmail.com

It has long been observed that compounds may exhibit semantically-unmotivated morphophonological elements (Bloomfield 1933), representative examples of which are linking elements found in inflectional languages, e.g., (1a) from languages with stem-based inflection and (1b) from languages with word-based inflection. Ralli (2008) argues that the primary function of linking elements is to mark the process of compounding, i.e. they are **compound markers (CM henceforth)**. Using the framework of lexeme-based morphology, we attempt to extend the CM analysis to various meaning-independent morphophonological variations within compounds observed in non-inflectional languages or categories. Briefly, we propose:

- Compounding employs special stems or **compounding stems** (Aronoff 1994) in order to mark the wordhood of the entire construction.
- Compounding stems are **morphomic**, being formed by formal means available in each language in the **stem-formation component** (Spencer 2012). Inflectional languages (or categories) use inflectional stems or inflected forms as compounding stems, while non-inflectional languages (or categories) use other formal marking options available in their analytic grammars.

While Ralli mainly focuses on morphological CMs from inflectional languages like (1), we take the view that CMs can be syntactic or phonological formatives because irrespective of its morphological typology, any language needs to mark its compound unit in some way. In our view, morphological typology concerns not whether but how a particular language marks the compound.

First, the necessity of such a view is suggested by French data in (2). Given Ralli's (2008) description of French as inflectionally-rich in verbs but poor in nominal inflection, the CM status she ascribes to P in the [N-P-N] compound in (2b) suggests that non-inflectional categories use non-inflectional formatives to mark the compoundhood. Thus, in (1) and (2), the non-head constituents including the underlined formative constitute compounding stems. (1) and (2a) illustrate stem formation dependent on inflectional paradigms, while (2b) is a case of stem formation by a syntactic functional morpheme.

Second, the stem analysis is preferable because the compound status is marked not only syntagmatically but also via morphophonological alternation. Thus, the compound stress exhibited by many English nominal compounds can be seen as (a component of) the compounding stem that an analytic language like English can make use of in order to make the word-phrase distinction. On the other hand, so-called neoclassical compounds (e.g. *anthropology*) instantiate cases when the compounding stem is suppletive. The bound form *anthrop* is a suppletive stem of the lexeme MAN (Amiot & Dal 2008). The use of the suppletive bound stem can be seen as a marker of compounds with the stylistic feature [+learned]. Hence, the stem analysis can encompass both major and minor types of English compounds. Additionally, Nübling & Szczepaniak (2008: 3) note the existence of "subtractive linking elements" in German (e.g. *Wolle* "wool" > *Woll-kleid* "woolen dress"), which is most naturally treated in the lexeme-based stem analysis.

Like French, Japanese has paradigmatic inflection for verbs but not for nouns, so Ralli's

(2008) parameters predict the existence of CM for Japanese verbs but not for nouns. Our analysis, on the other hand, predicts that Japanese verbs and nouns show different types of compounding stems. (3) supports the latter prediction. As in (3a), Japanese verbs always take the so-called adverbial word-form in the compound non-head, independently of semantics. In constrast, Japanese nouns exhibit syntactically or phonologically derived compounding stems; (3b) shows that the particle *no* can appear in N-N compounds as a CM (Mukai 2008), while (3c) shows that the second noun of N-N compounds can take an initially voiced allomorph (Itô & Mester 1986). In addition, (3d) illustrates the suppletive compounding stem.

Chinese is a language extremely poor in overt inflection but very rich in compounding. If the compound status is to be marked in grammar, this language should possess some formal device to do so. (4) indicates that the word status of the whole combination is marked by the bimorphemic morphophonological template. First of all, Modern Chinese vocabulary consists mainly of bimorphemic words, which developed to compensate for the increased homophony among (originally) monomorphemic words (Sampson 2013). The evidence for the bimorphemic structure itself now functioning as a CM comes from the fact that when compounded, bimorphemic input words can undergo a sort of morpheme reduction to produce still bimorphemic output words. The data in (4) taken from Packard (2000) and Ceccagno & Basciano (2009) may look like blending, but crucially differ from it in that the outputs in (i) consist of two morphemes and also in that the reduction process is obligatory, hence the ungrammaticality in (ii). Admittedly, it is difficult to view the bimorphemic word template itself as a compounding stem. Under the stem analysis, the proposal above can be translated as follows: each lexeme produces a monomorphemic stem for compounding. Bimorphemic lexemes choose one of the constituent morphemes as their compounding stem based on the informativity (Shaw et al. 2013).

(1) a. Greek: *kukl-o-souto* “doll house,”
Polish: *ostr-o-słup* sharp-LE-pillar “pyramid” (Szymanek 2009)
b. German: *Wirt-s-haus* host-LE-house “inn,” *Jahr-es-zeit* year-LE-time “season”

(2) a. French: *essuie-mains* “hand towel,” *porte-drapeau* “standard bearer” (Fradin 2009)
b. French: *moulin à vent* “wind mill,” *étoile de mer* “sea star”

(3) a. TATSU “stand” > *tachi-saru* stand-leave “go off,”
tachi-sugata stand-figure “standing posture”
b. *kitsune-no tebukuro* fox-of glove “foxglove, digitalis,”
ama-no gawa heaven-of river “Milky Way”
c. KITSUNE “fox” > *no-gitsune* field-fox “wild fox”
d. TATSU “stand” > *ritsu-an* stand-idea “planning,”
KITSUNE “fox” > *byak-ko* white-fox “white fox”

(4) a. *cǎo* “straw” + *mào-zi* hat-AFFIX “hat”
> (i) *cǎo-mào* “straw hat” vs. (ii) **cǎo-màozi*
b. *yì* “righteous” + *pāi-mài* “auction”
> (i) *yì-pāi* “charitable auction” vs. (ii) **yì-pāimài*
c. *wèi-xīng* “satellite” + *diàn-shì* “television”
> (i) *wèi-shì* “satellite TV” vs. (ii) **wèixīng-diànshì*

Selected References

ARONOFF, M. (1994) *Morphology by Itself*, MIT Press. /MUKAI, M. (2008) “Recursive compounds,” *Word Structure* 1. /NÜBLING, D. & R. SZCZEPANIAK (2008) “On the way from morphology to phonology: German linking elements and the role of the phonological word,” *Morphology* 18. /RALLI, A. (2008) “Compounding markers and parametric variation,” *Language Typology and Universals* 61. /SAMPSON, G. (2013) “A counterexample of homonymy avoidance,” *Diachronica* 30. /SHAW, J, C. HAN & Y. MA (2013) “Surviving truncation: Informativity at the interface of morphology and phonology,” a paper presented at the symposium Morphology and Its Interfaces (Université Lille 3). / SPENCER, A. (2012) “Identifying stems,” *Word Structure* 5.

Paradigm leveling in non-standard Russian

NATALIA SLIOUSSAR
HSE Moscow & St. Petersburg State University
slioussar@gmail.com

VARVARA MAGOMEDOVA
St. Petersburg State University
varya.magomedova@gmail.com

In this paper, we look at a paradigm leveling process currently taking place in Russian that affects historic consonant alternations. In standard Russian, these alternations are present in some verb forms, in comparatives and after certain derivational suffixes (e.g. *su**x**oj* 'dry' / *su**x**o* 'dryly' – *su**š**e* 'drier, more dryly', *ljub**bl**ju* 'I love', *no**g**a* 'leg' – *no**ž**ka* 'small leg, furniture leg'). But many non-standard forms lack alternations or have 'incorrect' alternations unattested in standard Russian. We will look at verb forms also analyzed in our previous work (Slioussar & Kholodilova 2013) and, primarily, at comparatives.

Unfortunately, Russian corpora contain almost no non-standard forms that we were interested in, so we had to look for them on the Internet. Estimating relative frequencies of different forms found there is a challenge because the counts provided by search engines are extremely unreliable. To circumvent this problem, we came up with several strategies. E.g. we established what variants of a particular form are attested (see below), included all variants in one search, i.e. asked the search engine to look for them simultaneously, sorted the results by date, counted frequencies of different variants and then did the relevant statistical tests. We developed a collection of Perl scripts *Lingui-Pingui* that greatly facilitates this task. This program can automatically form queries from the list of morphs, send them to the search engine, download the results, sort them according to specified criteria and do various counts.

Let us look at the non-standard innovation *zafrendit'* 'to include in one's friend list' as an example. The following 1SG future forms of this verb can be found on the Internet: *zafren**ž**u* with the standard *d // ž* alternation, *zafren**žd**u* with a *d // žd* alternation originally coming from Old Church Slavonic, *zafren**dj**u* lacking alternation and several variants with alternations unattested in standard Russian. Among them are *zafren**dl**ju, zafren**dž**u, zafren**čč**u, zafren**dč**u, zafren**dš**u* and *zafren**šč**u*. These alternations result from the inappropriate use of epenthetic *l* (*d // dl*), adding the alternating consonant to the stem rather than replacing the final consonant (*d // dž*), choosing a wrong alternating consonant (*d // č, d // šč*) or the combination of two last strategies (*d // dč, d // dš*). Three of these forms allow the speaker to kill two birds with one stone: to have an alternation and to keep the stem constant. Not to miss any variants, we looked for all possible combinations of alternating consonants, stem-final *d* and epenthetic *l*.

1. We can show that very often, leveling simultaneously goes in two opposite directions. However, some innovations are more frequent than the others. In particular, underapplication of consonant alternations is more widespread than overapplication. We will discuss this problem in detail for the verbs from the G-K verb class like *žeč'* 'to burn', where numerous non-standard examples of both types can be found. Although the present/future paradigm of G-K verbs includes four forms with alternations and two forms without them, innovative present/future forms lacking alternations prevail dramatically.

2. Looking at various verbs from the I class (like the non-standard *zafrendit'* above) we can demonstrate that leveling is influenced by the following factors. Firstly, less frequent words lack alternations more often. Secondly, non-standard words have more forms without alternations. Thirdly, the proportion of forms lacking alternations is the highest for stems ending in obstruent clusters and the lowest for the stems ending in labials (where the epenthetic *l* is standardly used).

3. For the study of alternations in comparatives two groups of adjectives were selected: ones that have normative synthetic comparatives with alternations (this group is not productive) and ones that do not, but native speakers still tend to generate such forms. In the first group,

the process of alternation loss is more sporadic than in the case of verbs: some adjectives like *ubogij* ‘poky’ or *uprugij* ‘resilient’ have up to 30% of comparatives without alternations, but no significant correlation between the number of non-standard forms and lemma frequency or the last consonant of the stem could be found. The second group consisted primarily of compound adjectives ending in *-gij, -kij, -xij*. Analogous simplex adjectives have synthetic comparatives with alternations, while these adjectives have only analytic standard forms. If speakers nevertheless try to form synthetic comparatives, they lack alternations significantly more often than in the first group. The most important factor is whether the second part of the compound is used as an independent adjective. If it is not, as in the case of *dlinnorukij* ‘long-armed’ or *dlinnonogij* ‘long-legged', the majority of comparatives lack alternations. This is notable because the relevant stems with alternations can be found in many highly frequent words, e.g. *ručka* ‘small hand, handle’, *nožka* ‘small leg, furniture leg’. So it seems to be crucial whether a particular form is listed in the mental lexicon, not whether the model is available. This is similar to our results with the verbs: in the I class, the model is productive, but it plays the most important role whether a standard form with alternations from a particular verb is stored in the lexicon. The tendencies observed for comparatives were replicated in a psycholinguistic experiment where participants (25 speakers of Russian, age 18-41) were asked to produce comparatives from various real and nonce adjectives (embedded in short sentences they were asked to complete).

Many competing approaches to paradigm leveling exist. But we are cautious to interpret our results in favor of any theory. For example, the fact that underapplication of alternations is preferred to overapplication is problematic for McCarthy’s (2005) framework, being more readily compatible with accounts like (Albright 2002, 2010). However, these and other theories were primarily designed to work with different data, explaining why some groups of words developed particular established forms rather than predicting different frequencies of various non-standard innovations. We can gain access to such data only now, with the development of Internet communication. We strongly believe that the general principles underlying these data should be the same in both cases, but some adaptation is still needed.

To give another example, alternations that are unattested in standard Russian can be taken as evidence that speakers, at least in some cases, rely on conditions on outputs (form X should contain consonant A) rather than on input-output relations (stem-final consonant B becomes A in form X) (e.g. Bybee 1995). However, this does not readily predict the vast diversity of such alternations, as well as their very low frequency compared not only to “standard” alternations, but also to the cases where alternations are missing.

References

Albright, A. 2002. The identification of bases in morphological paradigms. Doctoral dissertation, University of California. • Albright, A. 2010. Base-driven leveling in Yiddish verb paradigms. *Natural Language and Linguistic Theory* 28: 475–537. • Bybee, J. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425–455. • McCarthy, J. 2005. Optimal paradigms. In *Paradigms in Phonological Theory*, ed. L.J. Downing, T.A. Hall & R. Raffelsiefen, 170–210. Oxford: Oxford University Press. • Slioussar, N., & Kholodilova, M. 2013. Paradigm leveling in non-standard Russian. In *Formal Approaches to Slavic Linguistics: The Second MIT Meeting 2011*, ed. A. Podobryaev, 243–258. Ann Arbor, MI: Michigan Slavic Publications.

Agreement attraction: a novel view from Russian

NATALIA SLIOUSSAR
HSE Moscow & St. Petersburg State University
slioussar@gmail.com

NATALIA CHEREPOVSKAIA
St. Petersburg State University
cherepovskaia@gmail.com

We discuss two experiments on number agreement attraction (AA) and on a potentially similar phenomenon in Russian and their implications for psycholinguistic approaches to agreement attraction and for the views on the grammar architecture.

1. Background. Attraction errors received a lot of attention in the last two decades (e.g. Bock & Miller 1991; Eberhard et al. 2005; Franck et al. 2002, 2006; Vigliocco et al. 1995). An example of an attraction error in subject-predicate agreement is given in (1). The verb *are* agrees not with the head of the subject DP *key*, but with a dependent DP *cabinets* (attractor). Corpus studies and production and comprehension experiments looked at such errors in various languages, and, among many things, noted that: (ia). Only Plural DPs cause significant attraction: errors like (1) are produced more often and cause smaller delay in comprehension than errors like (2). (ib). The formal resemblance factor. Errors appear more often in production and cause smaller delay in comprehension if the form of the attractor coincides with the Nom.Pl form, like in German (3a) as opposed to (3b) (Hartsuiker et al. 2003).

(1) *The key to the cabinets are rusty.*
(2) *The keys to the cabinet is rusty.*
(3) a. *die Stellungnahme gegen die Demonstrationen* 'the$_{NOM.SG}$ position against the$_{ACC.PL=NOM.PL}$ demonstrations'
b. *die Stellungnahme zu den Demonstrationen* 'the$_{NOM.SG}$ position on the$_{DAT.PL\neq NOM.PL}$ demonstrations'

Two types of accounts have been proposed: (iia). Subject DPs get their number specification from a dependent DP due to erroneous feature percolation or similar mechanisms, i.e. the syntactic structure is represented incorrectly (e.g. Franck et al. 2002; Eberhard et al. 2005). So we produce an error or miss it in comprehension. (iib). The error arises not when we build syntactic structure, but when we access information in it: when we encounter a wrong verb form in comprehension or generate it in production, we recheck the subject DP and an attractor can interfere with this process (e.g. Solomon & Pearlmutter 2004; Wagers et al. 2009).

2. Our study. Exp. 1. In Russian, Nom.Pl forms of some nouns coincide not only with Acc.Pl, but also with Gen.Sg forms: e.g. *vecherinki* from *vecherinka* 'party'. We conducted an experiment to compare the number of errors with subject DPs like (4a-c). In every trial, participants saw a predicate and then a subject and were asked to pronounce a complete sentence. Half of the predicates did not agree with the subjects in number (a modified version of the method originally proposed by (Vigliocco et al. 1995)). In target stimuli, subjects consisted of a head noun, a preposition selecting Acc or Gen and a dependent noun in all possible number combinations (inanimate nouns for which the relevant forms coincide were used).

(4) a. *bilet na koncerty* 'ticket$_{NOM.SG}$ for concert$_{ACC.PL=NOM.PL}$'
b. *komnata dlja vecherinki* 'room$_{NOM.SG}$ for party$_{GEN.SG=NOM.PL}$'
c. *komnata dlja vecherinok* 'room$_{NOM.SG}$ for party$_{GEN.PL\neq NOM.PL}$'

Results. Out of all possible number combinations in subject DPs of our target stimuli (four in Acc sets and four in Gen sets), number agreement errors occurred only in three conditions containing DPs like (4a-c) and were distributed as follows: (4a)>(4b)>(4c). I.e. there were significantly more errors with Gen.Sg=Nom.Pl attractors than with Gen.Pl≠Nom.Pl ones.

Exp. 2. Exp.1 is compared with Exp.2 looking at another example of morphological ambiguity. In Russian, some adjective forms are ambiguous between different cases, in

particular, Gen.Pl and Prep.Pl coincide (*Prep* stands for *Prepositional case*). Rusakova (2009) who studied naturally occurring errors in Russian noted several examples like (5). We studied these errors in a comprehension experiment using self-paced reading methodology (unlike with subject-predicate agreement errors, there is no obvious way to induce them in production). We had 36 sets of target sentences: 12 sets with prepositions selecting Gen, Prep and Dat. Gen and Prep sets contained sentences in the three conditions shown below (Dat sets were used as controls and contained only the errors analogous to (6c), i.e. had two WICs):

(5) *v tex razmerov* 'in those$_{\text{PREP.PL(=GEN.PL)}}$ size$_{\text{GEN.PL}}$

(6) a. *Neudachi v proshlyh sezonah zastavili komandu potrudit'sja.* 'Correct' condition, CC
'failure$_{\text{NOM.PL}}$ in previous$_{\text{PREP.PL}}$ season$_{\text{PREP.PL}}$ make$_{\text{PST.PL}}$ team$_{\text{ACC.SG}}$ work$_{\text{INF}}$'

b. *Neudachi v proshlyh sezonov…* 'Wrong compatible' condition, WCC
'failure$_{\text{NOM.PL}}$ in previous$_{\text{PREP.PL(=GEN.PL)}}$ season$_{\text{PREP.PL}}$'

c. *Neudachi v proshlyh sezonam…* 'Wrong incompatible' condition, WIC
'failure$_{\text{NOM.PL}}$ in previous$_{\text{PREP.PL(≠DAT.PL)}}$ season$_{\text{DAT.PL}}$'

Results. The violation was detected significantly later and its effect was less pronounced in WCC conditions than in WIC conditions.

3. Conclusions. Previously, the formal resemblance factor in (ib) was assumed to boost the Plural markedness effect in (ia). Our results from Exp.1 indicate that (ib) is more important than (ia), AA effects are driven more by the form overlap than by featural overlap. Results from Exp.2 provide an interesting parallel (such constructions have not been experimentally studied before). We argue that in this case we deal with a phenomenon analogous to AA. The implications are manifold:

- The arguments used in the discussion of two approaches to AA (iia-b) have been inconclusive so far. Our results are incompatible with (iia) and thus provide a strong argument in favor of (iib). In cases like (4b), there are no Pl features that could percolate, yet AA errors are attested and are even more frequent than in cases like (4c). In (5) or (6b), there is also no feature that could percolate from the adjective to the noun. Approaches in (iib) need to be modified to account for our data. Wagers et al. (2009) assume that when we have a wrong verb form and rechecking is prompted, an attractor can provoke a mis-take if it contains a Pl feature. Our data show that the surface form of the attractor is important.
- Our results are hardly compatible with grammar architectures where full-fledged morphological forms appear only at late stages (Late Insertion). All previous studies of number AA discussed only cases like (1), (3a-b), (4a) or (4c). Both types of approaches above assumed that errors are triggered by the featural specifications of the attractor. Our results show that the actual form of the attractor is also (and even more) important. This form should be available and visible at early stages of syntactic derivation to be able to influence agreement.

References

BOCK, J. K., & MILLER, C. A. (1991). Broken agreement. *Cognitive Psychology, 23,* 45–93. • EBERHARD, K. M., CUTTING, J. C., & BOCK, J. K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review, 112,* 531-559. • FRANCK, J., LASSI, G., FRAUENFELDER, U., & RIZZI, L. (2006). Agreement and movement: A syntactic analysis of attraction. *Cognition, 101,* 173-216. • FRANCK, J., VIGLIOCCO, G., & NICOL, J. (2002). Attraction in sentence production: The role of syntactic structure. *Language and Cognitive Processes, 17,* 371-404. • HARTSUIKER, R., SCHRIEFERS, H., BOCK, K., & KIKSTRA, G. (2003). Morphophonological influences on the construction of subject–verb agreement. *Memory and Cognition, 31,* 1316–1326. • RUSAKOVA, M. (2009). *Rechevaja realizacija grammaticheskix elementov russkogo jazyka* [Speech realization of some grammatical features of Russian]. Habilitation dissertation, St.Petersburg State University. • SOLOMON, E. S., & PEARLMUTTER, N. J. (2004). Semantic integration and syntactic planning in language production. Cognitive Psychology, 49, 1-46. • VIGLIOCCO, G., BUTTERWORTH, B., & SEMENZA, C. (1995). Constructing subject–verb agreement in Speech: The role of semantic and morphological factors. *Journal of Memory and Language, 34,* 186-215. • WAGERS, M.W., E.F. LAU & C. PHILLIPS (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language* 61: 206-223.

Turkish synthetic compounds and the categorial status of the Root

DENİZ TAT
Yeditepe University
deniztat@u.arizona.edu

Within the recent constructionalist approaches to language, particularly in the grammar architecture assumed in Distributed Morphology (Halle and Marantz 1993), syntax not only shoulders phrase/sentence formation, but also word formation, including both simplex and complex words. In these approaches, Roots, which are acategorial, nongrammatical morphemes with semantic content, are manipulated by functional categories to derive words, which receive phonological expression only after all syntactic and postsyntactic operations are completed. For instance, the Root √DESTR, when selected by a little *n*, creates the context for the insertion of the Vocabulary Item *destruction* whereas in the context of a little *v*, it is realized as *destroy*.

There are two main views about the categorial status of Roots. According to the first one, Roots are considered syntactic categories that project phrases and specifiers and take complements. In this view, the complement position of the √P is where the internal argument can be hosted (Marantz 1997; Harley 2009, 2011). According to the second view, Roots are non-projecting categories, such that they do not project √Ps or specifiers, take complements, or impose any position for arguments (Borer 2009, 2011; de Belder 2011; de Belder and Craenenbroeck 2011). In this view, a verbal projection, such as a *v*P is needed to host the internal argument. In this study, I consider a range of seemingly synthetic compounds in Turkish, and show that nominals that are derived directly from Roots can never form true synthetic compounds insofar as our definition of these is such that the first element in the compound is an argument of the second. This shows that Roots cannot take internal arguments as complements, supporting the view that they are non-projecting syntactic elements.

A compound that appears to be synthetic in nature at first glance can be parsed at least in two ways. For instance *truck-driver* in English can be parsed as a true synthetic compound as in (1), in which the first element initially incorporates into the verbal stem before the whole construction gets nominalized by *-er*, or as a primary compound as in (2) analogous to, let's say, *truck-garage* where the semantic relation between the first and the second element is determined by world knowledge rather than strict syntactic means.

(1) [*truck* [*drive*]-er]
(2) [*truck* [*driver*]]

In Turkish, the difference between synthetic and primary compounds is less subtle. This is because the primary compounds would always necessitate the presence of a peculiar compound marker *-(s)İ(n),* typically analyzed as agreement morphology. For instance, the Turkish equivalent of *truck-driver* and *truck-garage* would both have this marker, suggesting that they are both primary compounds.

(3) *kamyon* *sür-ücü-sü*
truck drive-NOM-SIN
'truck driver'
(4) *kamyon* *garaj-ı*

truck garage-SIN
'truck garage'

True synthetic compounds in Turkish never take the compound marker *-(s)İ(n)*. Furthermore, these allow by-phrases, agent-oriented modifiers, adverbial and aspectual modification, which all contrast with primary compounds, showing the presence of an event structure in the former and the lack of it in the latter. These contrasts are illustrated in examples (5-6).

Primary compound:

(5) a. **öğrenci tarafından kadın çiz-im-i*
student by woman √draw-NOM-SİN
Intended: 'drawing of a woman by the student'

b. **gönüllü ol-arak kadın çiz-im-i*
voluntary be-ADV woman √draw-NOM-SİN
Intended: 'drawing of a woman voluntarily'

c. *(*bir saat-te) (*nihayet) kadın çiz-im-i*
one hour-LOC eventually woman √draw-NOM-SİN
Intended: 'eventually drawing of a woman in one hour'

True synthetic compound:

(6) a. *öğrenci-ler tarafından bilgi pay-laş-ım*
student-PL by information √share-VER-NOM
'sharing of information by the students'

b. *gönüllü ol-arak bilgi pay-laş-ım*
voluntary be-ADV information √share-VER-NOM
'sharing of information voluntarily'

c. *iki hafta-dır güçlük-le bilgi pay-laş-ım*
two week-EP hardship-COM information √share-VER-NOM
'sharing of information with difficulty for two weeks'

In this paper, I conclude that (i) Roots are non-projecting categories, (ii) a verbal stem is required to host a theme argument, (iii) and, therefore, only nomimalizations with verbal stems can act as the head of true synthetic compounds. I further show that the presence of an overt verbalizing or causativizing morpheme in certain nominalizations does not always entail an event structure; nominals of this type always denote entitites. I refer to compounds headed by these as 'pseduo-synthetic compounds' which I claim are just a sub-set of primary compounds.

References

BORER, H. 2009. Roots and categories. Paper presented at the 19th Colloquium on Generative Grammar. University of the Basque Country. April 1-3, 2009.

BORER, H. 2011. In the event of a nominal. In Everaert, M., Maelj, M. and T. Siloni (eds). *The theta system: argument structure at the interface.* Oxford University Press.

DE BELDER, M. 2011. *Roots and affixes: eliminating lexical categories from syntax.* Brussels: HU Brussel / Utrecht University.

DE BELDER, M. and VAN CRAENENBROECK, J. 2011. How to merge a root. M.s. HU Brussel & Utrecht University.

HALLE, M. and MARANTZ, A. 1993. Distributed morphology and the pieces of inflection. In *A view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. K.

HALE, K. and S. J. KEYSER, 111-176. Cambridge, MA: MIT Press.

HARLEY, H. 2009. Compounding in Distributed Morphology. In *The Oxford handbook of compounding*, R. Lieber and P. Stekauer (eds). Oxford: OUP. p. 129-143.

HARLEY, H. 2011. On the identity of Roots. M.s. University of Arizona.

MARANTZ, A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *UPenn Working Papers in Linguistics* 4: 201-225.

Verbal morphemes need not entail verbal structure

DENİZ TAT
Yeditepe University
deniztat@u.arizona.edu

GREG KEY
University of Arizona
gkey@email.arizona.edu

Derivation from category-neutral Roots has been observed to differ in important respects from derivation from syntactic categories (Marantz 1997, Arad 2003). In terms of semantics, the merger of the first functional head with an acategorial Root is a domain of special meanings. A verbalizing head that merges directly with a Root may have an idiosyncratic meaning; further derivation from this syntactic object cannot result in a new, renegotiated meaning, but must instead be an extension of the semantics produced at the first merger with the Root. Marantz (2013) addresses apparently counter-examples to this generalization. Among these are cases in Japanese where a nominalizer interacts with a Root to produce idiosyncratic semantics despite being separated from the Root by a causative or inchoative verbalizing morpheme.

(1)

Root	Verb-$_{\text{INTRANS}}$	Verb-$_{\text{TRANS}}$	Nominalization
√*nag*	*nag-e-(ru)* 'to flow'	*nas-as-(u)* 'to make flow'	*nag-ashi* 'a sink'

(from Volpe (2005), cited in Marantz (2013))

Marantz proposes that the verbalizing morpheme in the nominalization is semantically null insofar as it does not entail an event. This is what allows the nominalizing head to negotiate a special meaning with the Root: the intervening morpheme is not semantically visible, such that the nominalizer is semantically adjacent to the Root. Marantz's proposal suggests that the intervening morpheme plays no role in deriving the idiosyncratic semantics. However, a number of Turkish derivations call this into question. There are several transitive verbs in Turkish that optionally have single or double morphological exponence of the causative with no consistent difference in meaning.

(2)

	Root	Single Exponence	Double exponence
a.	√*çık*	*çık-ar-* 'bring up, out'	*çık-ar-t-* 'bring up, out'
b.	√*kop*	*kop-ar-* 'detach (tr.)'	*kop-ar-t-* 'detach'
c.	√*kap*	*kap-a-* 'close (tr.)'	*kap-a-t-* 'close (tr.)'
d.	√*uç*	*uç-ur-* 'make fly'	*uç-ur-t-* 'make fly'

Nominalizations based on these forms have special meanings, much like the Japanese examples. What is striking is that these semantics vary sharply and idiosyncratically according to whether the causative verbalizer has single or double exponence.

(3)

	Single Exponence	Double Exponence
a.	*çık-ar-ma* 'naval landing'	*çık-ar-t-ma* 'sticker, decal'
b.	*kop-ar-ma* 'clean and jerk'	*kop-ar-t-ma* '(a kind of) pastry'
c.	*kap-a-ma* 'lamb stew'	*kap-a-t-ma* 'kept woman'
d.	*uç-ur-ma* (no special meaning)	*uç-ur-t-ma* 'kite (the toy)'

In the face of these facts, it cannot be maintained that the morphemes intervening between the Root and the nominalizer are invisible to the semantics. We propose instead that the nominalizations are not derived from the corresponding verbal structures. Building on Marantz's insight that these nominalizations do not entail an event, we claim that the nouns in (3) do not have an argument structure, unlike the corresponding verbal forms in (2). The

meanings of these nouns, like those of their Japanese counterparts, indicate that they are R-nominals (entitydenoting nominals) in the sense of Borer (2003, 2011), and thus do not allow agent-oriented modifiers, adverbial or aspectual modification, and cannot participate in the formation of true synthetic compounds. We therefore conclude that the different interpretations of nominalizations such as the ones in (3) do not occur at LF because their nominalizing head is semantically Root-adjacent but rather occur at the Encyclopedia component of the grammar architecture generally assumed in the Distributed Morphology framework (Harley and Noyer 1999), which contains the list of idioms in a language that need to be memorized by the language learner.

References

ARAD, M., 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language and Linguistic Theory* 22: 737-778.

BORER, H. 2003. Exo-skeletal vs. Endo-skeletal explanations. In Moore, J. and M. Polinsky (eds). *The nature of explanation in linguistic theory*, Chapter 2. CSLI.

BORER, H. 2011. In the event of a nominal. In Everaert, M., Maelj, M. and T. Siloni (eds). *The theta system: argument structure at the interface.* Oxford University Press.

HARLEY, H. and R. NOYER. 1999. State-of-the-article: Distributed Morphology. *GLOT 4.4*: 3-9.

MARANTZ, A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *UPenn Working Papers in Linguistics* 4: 201-225.

MARANTZ, A. 2013. Locality domains for contextual allomorphy across the interfaces. In Matuhshansky, O. and A. Marantz (eds). *Distributed Morphology today.* MIT Press.

Two types of *i*-defectivity in Spanish verbal paradigms

JESSE TSENG
CLLE-ERSS UMR 5263, CNRS & Université Toulouse Jean Jaurès
Jesse.Tseng@univ-tlse2.fr

A curiosity of the Spanish *-ir* conjugational class is the existence of verbs like *abolir* "abolish" that are said to have a defective paradigm. According to grammatical tradition, the verbs in this class can only be conjugated when the ending begins with the letter ‹i› (Bello §587ff), hence the name "*i*-defective". The other forms — pres. ind. 1/2/3/6, pres. subj., and sing. imperative — are said to be missing. So the pertinent parts of the paradigm of *abolir* have the following shape, with the defective cells highlighted:

infinitive	*abolir*	pres.part.	*aboliendo*	past part.	*abolido*	
pres.ind.	*abolo*	*aboles*	*abole*	*abolimos*	*abolís*	*abolen*
pres.subj.	*abola*	*abolas*	*abola*	***abolamos***	***aboláis***	*abolan*
preterit	*abolí*	*aboliste*	*abolió*	*abolimos*	*abolisteis*	*abolieron*
imp.subj.	*aboliera*	*abolieras*	*aboliera*	*aboliéramos*	*abolierais*	*abolieran*
imperat.		*abole*			*abolid*	

According to some sources, however, *abolir* is no longer an *i*-defective verb and now has a complete conjugation, with fully regular forms as indicated above. Normative authorities in fact disagree about the precise membership of the class of *i*-defective verbs, and limited empirical studies (notably Albright 2003) suggest that *i*-defectivity is a gradient and variable phenomenon. The verbs identified are always low frequency if not completely archaic, but if pressed, speakers will produce and accept supposedly defective forms, albeit with a low degree of confidence. The psychological reality and robustness of *i*-defectivity are therefore less than fully established, and the distribution shown above may be best understood as an idealization, possibly corresponding more to the situation in earlier stages of modern Spanish than to that of current standard varieties.

The proper characterization of the *i*-defective pattern also remains open to debate. It is clear that the semantic features of the missing forms provide no satisfactory explanation. Albright (2003) proposes an "anti-stress" constraint, meaning that root-stressed forms are avoided. This accounts for almost all of the paradigm except for persons 4/5 of the present subjunctive (underlined in the paradigm above), which have an unstressed root but are nevertheless considered to be defective. Maiden & O'Neill (2010) see in the *i*-defective distribution a combination of the morphomic L- and N-patterns that are recurrent in Romance verbal morphology, primarily in connection with stem allomorphy. There are indeed verbs whose stems exhibit precisely this distribution in the present tenses (e.g. *ped-* vs. *pid-* for *pedir*), but note that the parallelism with *abolir* does not extend to the complete paradigm:

infinitive	*pedir*	pres.part.	*pidiendo*	past part.	*pedido*	
pres.ind.	*pido*	*pides*	*pide*	*pedimos*	*pedís*	*piden*
pres.subj	*pida*	*pidas*	*pida*	*pidamos*	*pidáis*	*pidan*

preterit	*pedí*	*pediste*	*pidió*	*pedimos*	*pedisteis*	*pidieron*
imp.subj.	*pidiera*	*pidieras*	*pidiera*	*pidiéramos*	*pidierais*	*pidieran*
imperat.		*pide*			*pedid*	

In the "stem space" approach of Boyé & Cabredo-Hofherr (2006, 2010), Spanish verbs have 11 interrelated stems, each associated with a region of the paradigm. The defective cells in the conjugation of *abolir* correspond exactly to the regions for stems "S1", "S2", and "S5", so the defective paradigm can be modeled by stipulating that those three stems are suppletively specified as Ø. But none of the previous analyses captures the traditional idea that *i*-defectivity is correlated with the segmental content of the verb ending: the endings of the existing forms all begin with [i] or [jV]. It is hard to see how the absence of such an ending could be a trigger for defectivity, but no other generalization appears to give the desired result, and we conclude that the analysis of the *abolir*-type conjugation must make direct reference to the surface phonological form of the ending (or to an abstract thematic vowel /i/).

Bello (§591f) further describes a variant of *i*-defectivity that was already obsolete by the mid-19th century. Verbs like *escarnir* "mock" could only be conjugated with endings starting with [i], so that in addition to the *abolir*-type defectivity, forms in [jV] were also excluded: **escarnió*, **escarnieron*, **escarniendo*. These missing forms were supplied by the doublet verb *escarnecer*; these are indicated in green in the following paradigm.

infinitive	*escarnir*	pres.part.	*escarneciendo*	past part.	*escarnido*	
pres.ind.	*escarnezco*	*escarneces*	*escarnece*	*escarnimos*	*escarnís*	*escarnecen*
pres.subj	*escarnezca*	*escarnezcas*	*escarnezca*	*escarnezcamos*	*escarnezcáis*	*escarnezcan*
preterit	*escarní*	*escarniste*	*escarneció*	*escarnimos*	*escarnisteis*	*escarnecieron*
imp.subj.	*escarneciera*	*escarnecieras*	*escarneciera*	*escarneciéramos*	*escarnecierais*	*escarnecieran*
imperat.		*escarnece*			*escarnid*	

It is not widely recognized that modern Spanish once had heteroclitic conjugations involving the inchoative suffix (cf. Italian *-isc-*, French *-iss-*), and what is most striking is that the distribution of stems was exactly the same as in the case of *pedir* shown above. The [e]/[i] alternation for verbs like *pedir* is relatively well studied, with analyses typically invoking phonological notions (dissimilatory lowering of [i] before *í*, metaphonic raising of [e] before [j]). We propose that the *pedir*-type distribution of stems became a morphologized paradigmatic pattern that gave rise to the obsolete *escarnir*/*escarnecer*-type mixed conjugation, for which no phonological motivation can be identified.

Selected references

ALBRIGHT, A. 2003. "A Quantitative Study of Spanish Paradigm Gaps". In G. Garding & M. Tsujimura (eds.), *WCCFL 22 Proceedings*. Somerville, MA: Cascadilla Press, pp. 1–14. • BELLO, A. 1847. *Gramática de la Lengua Castellana destinada al uso de los Americanos*. Santiago: El Progreso. • BOYÉ, G. & P. CABREDO HOFHERR. 2006. "The structure of allomorphy in Spanish verbal inflection". *Cuadernos de Lingüística Ortega y Gasset* 13:9–24. • BOYÉ, G. & P. CABREDO HOFHERR. 2010. "Defectiveness as Stem Suppletion in French and Spanish Verbs". In M. Baerman et al. (eds) *Defective Paradigms*. London: British Academy, pp. 35–52. • MAIDEN, M. & P. O'NEILL. 2010. "On morphomic defectiveness: evidence from the Romance languages of the Iberian peninsula". In M. Baerman et al. (eds.) *Defective Paradigms*. London: British Academy, pp. 103–124.

The transcategorial irrealis suffix in some Mayan languages

IGOR VINOGRADOV
The Institute of Linguistics, Russian Academy of Sciences
igor.vinogradov@iling-ran.ru

Though the term "irrealis" (or "irreal status") is widely used in linguistic studies and grammatical descriptions, there are some doubts about its typological validity (cf. Bybee et al. 1994, Bybee 1998, and the recent study de Haan 2012). In essence, these doubts arise because of the vagueness and uncertainty of possible definitions of irreality. As a marker of the set of unreal events, irrealis is usually considered a part of verbal grammar, and consequently irreality is regarded as a verbal semantic domain. As I argue below, even such a circumscription of the applicable domain of irreality is incorrect for some languages. The study is mainly based on the analysis of original texts.

The modern Mayan languages of the Cholan-Tzeltalan and Q'anjobalan branches have a suffix of irrealis that is notable for its transcategorial nature. It can combine not only with verbs, but also with a wide range of other parts of speech. Etymologically, this suffix originates in the Protomayan modal suffix denoting dependent status for intransitive verbs *-*oq* (Robertson 1992).

There is a range of possible meanings of this suffix and these meanings are highly dependent on the part of speech of the stem. Most frequently it is added to finite verb forms. In this case it has the widest realm of uses that consists of counterfactual conditions, expressions of goal or intention and various desiderative/optative constructions (1). Generally, the list of verbal meanings of the Mayan irrealis does not significantly differ from the typological standard obtained in various comparative studies (Elliott 2000, for instance).

When the irrealis suffix is used with another part of speech it reduces the range of possible meanings. Furthermore the semantics of this suffix depend on the syntactic role of the word form, specifically on its predicative status. Non-verbal predicates are formed without any auxiliary component in Mayan languages, so in some cases the irrealis suffix added to a non-verbal stem can be regarded as syntactic marker of predicativity. With non-verbal predicates, the realm of uses of the irrealis suffix is limited to negative (2) or, less often, interrogative contexts.

When the nominal, adverbial or numeral stem that does not act as a predicate carries the irrealis suffix, it usually denotes approximateness or indefiniteness (3). The combination of the numeral 'one' and the irrealis suffix (4) is to a certain extent similar to the indefinite article in European languages, an element that is absent in Mayan languages of the Cholan-Tzeltalan and Q'anjobalan branches.

The irrealis suffix can sometimes perform the word-formative function as well. Indefinite pronouns are derived from interrogative ones by means of this suffix in some Mayan languages.

There are some differences in the use of irrealis suffix in a particular language. For example, the usage of irrealis in order to form a question was noted with certainty only in Tzeltal, while the derivation of an indefinite pronoun by means of the irrealis suffix is possible only in Tzotzil and Ch'orti'. In general, however, the two branches of the Mayan family under consideration exhibit significant similarities.

The transcategorial nature of the Mayan irrealis suffix, as a special morphological feature of a particular morpheme, reveals at least two significant semantic features of the category of irrealis. Firstly, the Mayan irrealis provides clear evidence that irreality does not correspond with modality, being a wider semantic domain. It exceeds the limits of verbal grammatical semantics, spreading to nominal and numeral semantic domains, such as, for instance, approximateness and indefiniteness. Secondly, as regard to the possible definitions of

irreality, the Mayan languages from the Cholan-Tzeltalan and Q'anjobalan branches prove the importance of referentiality as the determining factor when considering an event or an object to be real or unreal. The non-referential status of an object can trigger the marker of irrealis, at least in some languages. This idea was first suggested in Givón (1994), but is as yet still not generally accepted as irreality is considered an exclusively verbal category.

(1) Chol (Cholan-Tzeltalan)

uts'at	che'_jini	la'-to	kol-**ik**
well	then	let-still	grow-**IRR**

'*Well, let him still grow*' (Warkentin & Whittaker 1965: 16)

(2) Akatek (Q'anjobalan)

man	y-et-**oj**	naj	y-ee	tumin
NEG	3.POSS-ownership-**IRR**	CLF(metal)	3.POSS-EXIST	money

'*The money is not for him*' (Andrés & Dakin 1989: 294)

(3) Tzotzil (Cholan-Tzeltalan)

i-ch'ay	nan	xemana-**uk**	mi	mas
3.ERG.PFV-waste	maybe	week-**IRR**	or	more

'*He wasted maybe a week or more*' (Laughlin 1977: 55)

(4) Tzeltal (Cholan-Tzeltalan)

la	s-jok'oy	te	me	ay	y-ich'-oj	
PFV	3.ERG-ask	REL	PART	EXIST	3.POSS-take-PRF.PTCP	
j-ch'ix-**uk**		s-may	y-u'un	ya	s-nuk'	
one-CLF(cane)-**IRR**		3.POSS-cigarette	3.POSS-cause	IPFV	3.ERG-smoke	

'*He asked if he had taken a cigarette to smoke*' (Pérez López et al. 1994: 282)

References

ANDRÉS, J. J. & K. DAKIN. 1989. El wakax kan y otros textos acatecos. *Tlalocan* XI: 285-300.

BYBEE, J. 1998. "Irrealis" as a grammatical category. *Anthropological linguistics* 40/2: 257-271.

BYBEE, J., R. PERKINS & W. PAGLIUCA. 1994. *The evolution of grammar: tense, aspect, and modality in the languages of the world*. Univ. of Chicago Press.

DE HAAN, F. 2012. Irrealis: fact or fiction? *Language sciences* 34: 107-130.

ELLIOTT, J. 2000. Realis and irrealis: forms and concepts of the grammaticalisation of reality. *Linguistic typology* 4/1: 55-90.

GIVÓN, T. 1994. Irrealis and the subjunctive. *Studies in language* 18/2: 265-337.

LAUGHLIN, R. 1977. *Of cabbages and kings: tales from Zinacantán*. Washington DC, Smithsonian Institution Press.

PÉREZ LÓPEZ, E., M. HIDALGO PÉREZ & A. GÓMEZ GÓMEZ (eds.). 1994. *Cuentos y relatos indígenas, vol. V*. México DF, UNAM.

ROBERTSON, J. 1992. *The history of tense/aspect/mood/voice in the Mayan verbal complex*. Austin, UT Press.

WHITTAKER, A. & V. WARKENTIN. 1965. *Chol texts on the supernatural*. Oklahoma, SIL.

Defining direct/inverse systems: a canonical approach

GÉRALDINE WALTHER[1,2], ANTON ANTONOV[3], GUILLAUME JACQUES[3]
1. DDL (CNRS, U. Lyon 2) 2. LLF (CNRS, U. Paris 7) 3. CRLAO (CNRS, INALCO)
geraldine.walther@univ-lyon2.fr a.antonov@gmail.com rgyalrongskad@gmail.com

The concept of CANONICAL INFLECTION (Corbett, 2007) has been used to characterise a broad variety of inflectional systems (Thornton, 2010; Baerman *et al.*, 2010; Stump, 2012; Brown and Hippisley, 2012; Palancar, 2012). However, none of the existing work on canonical inflection seems to consider the specificity of DIRECT/INVERSE SYSTEMS (Zúñiga, 2006) and their relationship to the criteria defining canonical inflection.

This paper constitutes a first attempt towards the proposition of a set of criteria for defining the concept of direct/inverse systems. Our definition is set within the CANONICAL TYPOLOGICAL FRAMEWORK (Corbett, 2003) and focuses more specifically on the canonicity of direct/inverse marking. We also show how regular direct/inverse systems compare with the more general standard of CANONICAL INFLECTION as defined in (Corbett, 2007).

Direct/inverse systems are traditionally thought of as systems displaying hierarchical agreement. Transitive verbs within direct/inverse systems typically index two arguments and are marked so that, e.g., a 1SG agent and a 3SG patient (henceforth 1SG>3SG) are marked in the same way as a 3SG agent and a 1SG patient (3SG>1SG), except for a distinctive INVERSE MARKER. Direct *vs.* inverse forms are then distinguished through their coherence with an independent agentivity/empathy hierarchy, which typically stipulates that arguments corresponding to speech act participants (SAP) are higher in the hierarchy than third person animate arguments, which are again higher than inanimate arguments. A typical hierarchy is one by (DeLancey, 1981, 644): SAP > 3rd pers. pronoun > human > animate > natural forces > inanimate.

In our example, 1SG>3SG complies with the hierarchy and would be considered a direct form, marked by a dedicated DIRECT MARKER (DM), while 3SG>1SG does not comply with the hierarchy and would be typically marked as inverse by the contrasting INVERSE MARKER (IM). Moreover, in traditional typological approaches, the hierarchy among SAP is considered to be unstable across languages (Jacques and Antonov, 2013): 1>2 and 2>1 can be marked with special forms (Algonquian); 2>1 may receive the IM (Situ Rgyalrong); both 2>1 and 1>2 may receive the IM (Mapuche or Kiranti); the presence of DM or IM independently follows from the context (Japhug Rgyalrong). This variation is often seen as the image of the languages' specific hierarchy and typological properties.

Within this broad range of variation, it makes sense to resort to a canonical approach. It allows for defining an independent, purely formal (and typically non-existing) standard from which deviation may be measured. The standard we define for canonical direct/inverse (DI-canon) systems is based on three converging criteria: (i) compliance with a language independent, maximally regular hierarchy (any verbal argument should be higher or lower w.r.t. the other), as sketched hereafter: 1>2>3HIGHEST>...>3LOWEST, (ii) perfect symmetry across the cells of the inflectional paradigm involving the same verbal arguments (see Table 1), (iii) autonomy with respect to linguistic features external to the hierarchy.

	1	2	3H	3L
1		1>2	1>3H	1>3L
2	2>1		2>3H	2>3L
3H	3H>1	3H>2		3H>3L
3L	3L>1	3L>2	3L>3H	

Table 1. Schematic symmetrical structure of a canonical direct-inverse system (dark cells mark inverse, white cells mark direct; 3H represents a 3rd person argument higher in the hierarchy, and 3L a 3rd person lower argument)

We illustrate possible types of deviation from the direct/inverse canon based on data from Algonquian, Kiranti, Mapuche, Rgyalrong, and Sahaptian. For example, the Kiranti data in the Khaling sub-paradigm in Table 2 shows deviation from the DI-canon in two of above mentioned

dimensions: (i) inverse marking does not strictly follow the directness hierarchy: third person forms are all marked as direct forms; (ii) the IM *ʔi-* is not symmetrically distributed across the paradigm: (a) the empathy hierarchy does not strictly structure 3rd person arguments; (b) the IM also appears in supposedly direct forms such as 2>3, 1DE>2 and 1PE>2, and (c) 1>3 and 3>1, resp. 1>2 and 2>1, are not symmetrical. The Japhug data, where inverse marking is reliant on genericity, will illustrate deviation from the DI-canon with respect to criterion (iii), namely the autonomy of the hierarchical marking w.r.t. other linguistic features.

	1S	1DI	1DE	1PI	1PE	2S	2D	2P	3S	3D	3P
1S						lɔ̂m-nɛ	lɔ̂m-su	lɔ̂m-nu	lob-u	lob-u-su	lob-u-nu
1DI									ləp-i		
1DE						ʔi-lɔɔp	ʔi-ləp-i	ʔi-lɔ̂m-ni	ləp-u		
1PI									lɔɔp-ki		
1PE						ʔi-lɔɔp	ʔi-ləp-i	ʔi-lɔ̂m-ni	lɔɔp-kʌ		
2S	ʔi-lɔ̂m-ŋʌ								ʔi-lɔ̄:b-ʁ	ʔi-lɔ̂:p-su	ʔi-lɔ̂:p-nu
2D	ʔi-lɔ̂m-ŋʌ-su								ʔi-ləp-i		
2P	ʔi-lɔ̂m-ŋʌ-nu								ʔi-lɔ̂m-ni		
3S	ʔi-lɔ̂m-ŋʌ	ʔi-ləp-i	ʔi-ləp-u	ʔi-lɔɔp-ki	ʔi-lɔɔp-kʌ	ʔi-lɔɔp	ʔi-ləp-i	ʔi-lɔ̂m-ni	lɔ̄:b-ʁ		
3D	ʔi-lɔ̂m-ŋʌ-su								lɔ̂:p-su		
3P	ʔi-lɔ̂m-ŋʌ-nu								lɔ̂:p-nu		

Table 2. Non-canonical direct-inverse marking in Khaling (Kiranti), data from (Jacques *et al.*, 2012)

Finally, we characterise the canonical direct/inverse system w.r.t. the notion of canonical inflection (CI). We show that the DI-canon only partly complies with the criteria for CI. CI can be seen as an instance of maximal regularity and minumum ambiguity in the paradigmatic realisation of individual morphosyntactic property sets (henceforth MPS). CI also features COMPLETENESS (each MPS has a corresponding form realisation) and UNIQUENESS (each MPS is realised by one form only). In the maximally regular case, the paradigm of a direct/inverse system (i.e., the DI-canon) seems to fully satisfy the set of criteria drawn for defining CI. However, when one considers the individual feature realisation, hierarchical direct/inverse marking does deviate from the canonical definition of inflection. Corbett (2007) indicates that CI expresses each value of the overall set of morphosyntactic features through an unambiguous single affix. In the case of direct/inverse based hierarchical agreement marking, affixes are typically ambiguous in the sense that they underspecify the status of the argument they are supposed to encode. Only the DM and IM allow for disambiguation. Moreover, the structure of the paradigm is not completely unmotivated (as expected for CI), but conditioned by a canonical empathy hierarchy. The DI-canon thus appears as a maximally regular and minimaly ambiguous system, also complying with the completeness and uniqueness criteria, that deviates from the general canon for inflection in the individual encoding of the argument verbal features and the (at least partial) motivation of the paradigm's symmetric shape through a specific feature hierarchy.

Acknowledgments This work was partly supported by the LabEx ASLAN (ANR-10-LABX-0081) operated by the French National Research Agency (ANR).

References

M. BAERMAN, G. G. CORBETT AND D. BROWN, Eds. 2010. *Defective Paradigms: Missing Forms and What They Tell Us*. Oxford University Press.

BROWN, D. AND A. HIPPISLEY. 2012. *Network Morphology: A Defaults-based Theory of Word Structure*. Cambridge University Press.

CORBETT, G. G.. 2003. Agreement: the range of the phenomenon and the principles of the surrey database of agreement. *Transactions of the philological society*, **101**, 155–202.

CORBETT, G. G.. 2007. Canonical typology, suppletion and possible words. *Language*, **83**, 8–42.

DELANCEY, S.. 1981. An interpretation of split ergativity. *Language*, **57**(3), 626–57.

JACQUES, G. AND A. ANTONOV. 2013. Direct-inverse systems. *Language and Linguistics Compass*. In press.

JACQUES, G., A. LAHAUSSOIS, B. MICHAILOVSKY, AND D. B. RAI. 2012. An overview of Khaling verbal morphology. *Language and linguistics*, **13.6**, 1095–1170.

PALANCAR, E.. 2012. The conjugations classes of tilapa otomi: An approach from canonical typology. *Linguistics*, **50**(4), 782–832.

STUMP, G. T.. 2012. The formal and functional architecture of inflectional morphology. In A. RALLI, G. BOOIJ, S. SCALISE AND A. KARASIMOS, Eds., *Morphology and the Architecture of Grammar: On-line Proceedings of the 8th Mediterranean Morphology Meeting (MMM8), Cagliari, Italy, 14-17 Sept. 2011*, p. 255–271.

THORNTON, A. M.. 2010. Towards a typology of overabundance. Communication au colloque des Décembrettes 7.

ZÚÑIGA, F.. 2006. *Deixis and Alignment - Inverse systems in indigenous languages of the Americas*. Amsterdam: Benjamins.

The Opacity-Compactness Tradeoff: Morphomic Features for an Economical Account of Khaling Verbal Inflection

GÉRALDINE WALTHER[1,2] GUILLAUME JACQUES[3] BENOÎT SAGOT[4]
1. DDL (CNRS, U. Lyon Lumière) 2. LLF (CNRS, U. Paris Diderot)
3. CRLAO (CNRS, Inalco) 4. ALPAGE (Inria, U. Paris Diderot)
geraldine.walther@univ-lyon2.fr rgyalrongskad@gmail.com benoit.sagot@inria.fr

Morphomic features (Aronoff, 1994) have been argued to allow for more elegant and more economical descriptions of inflectional systems (Bonami and Boyé, 2010). However, morphomic accounts are opaque from the morphosyntactic point of view and thus require an additional mapping between the morphomic features and the actual morphosyntactic property sets (MPS) expressed by the inflected forms (Stump, 2006) — something which is not needed if the MPSs are realised directly. Yet, no measures have been proposed so far to quantitatively assess the tradeoff between the lower complexity achieved by relying on opaque morphomic features and the additional need for an explicit feature mapping.

In this paper, we show that the descriptive economy achieved by capturing the inherent morphomic structure of inflectional systems can effectively be assessed through dedicated compactness measures.

We present two competing analyses of Khaling verbal inflection based on original data from Jacques *et al.* (2012). These two analyses will be called descriptions A and B hereafter. Khaling is a highly inflectional language of the Kiranti subbranch of the Sino-Tibetan language family, spoken by about 15,000 people in Eastern Nepal. Verbs realise up to 572 verb forms, among which 300 realis forms. Due to numerous syncretisms, these 300 realis forms reduce to only 100 distinct forms.

Among our two descriptions, DESCRIPTION A uses realisation rules based on standard MPS (such as A1S.P2S). DESCRIPTION B comprises rules realising morphomic features (such as B, C, D, etc. in Table 2) and and their corresponding feature mapping functions (such as B: 1DU.INCL). We show that the paradigm structure of Khaling verbs is best described using morphomic features of the type proposed for Nepali by Bonami and Boyé (2010) together with another morphomic feature related to the direct/inverse marking system based on the language's empathy hierarchy, as commonly used for describing Algonquian languages (Silverstein, 1976; Zúñiga, 2006).

We show that if one ignores the direct/inverse marking (i.e., the presence of the inverse prefix ʔi-), the resulting additional syncretisms decrease the number of distinct forms from above mentioned 100 to only 68. These additional syncretisms are illustrated by the coloring in Table 1 for the non-past positive subparadigm. Additional morphomic feature mappings, such as posing a unique feature F for all forms expressing a P=2SG, then allows for further syncretic reduction. Table 2 shows how for the non-past positive subparadigm alone, the number of 75 distinct MPSs can be collapsed into 22 morphomic feature bundles such as single feature F or feature bundle X>1.2SG. Among these 22 morphomic feature bundles, 12 are complex features, but 10 are simple features, thus greatly simplifying the the expression of form realisation rules within the grammar (*lêːp-nu* becomes simply the realisation of feature K instead of that of features 3PL>3, 3DU>3DU, 3DU>3PL, 2SG>3PL plus the inverse marked 2SG>3PL).

↓A | P→	1SG	1DU.INCL	1DU.EXCL	1PL.INCL	1PL.EXCL	2.SG	2DU	2PL	3SG	3DU	3PL
1SG						loɔ̂m-nɛ	loɔ̂m-su	loɔ̂m-nu	lob-u	lob-u-su	lob-u-nu
1DU.INCL									lɔp-i		
1DU.EXCL						ʔi-lɔɔp	ʔi-lɔp-i	ʔi-loɔ̂m-ni	lɔp-u		
1PL.INCL									lɔɔp-ki		
1PL.EXCL						ʔi-lɔɔp	ʔi-lɔp-i	ʔi-loɔ̂m-ni	lɔɔp-kʌ		
2SG	ʔi-loɔ̂m-ŋʌ								ʔi-lɔ̄ːb-u	ʔi-lɔ̂ːp-su	ʔi-lɔ̂ːp-nu
2DU	ʔi-loɔ̂m-ŋʌ-su									ʔi-lɔp-i	
2PL	ʔi-loɔ̂m-ŋʌ-nu									ʔi-loɔ̂m-ni	
3SG	ʔi-loɔ̂m-ŋʌ	ʔi-lɔp-i	ʔi-lɔp-u	ʔi-lɔɔp-ki	ʔi-lɔɔp-kʌ	ʔi-lɔɔp	ʔi-lɔp-i	ʔi-loɔ̂m-ni	lɔ̄ːb-u		
3DU	ʔi-loɔ̂m-ŋʌ-su									lɔ̂ːp-su	
3PL	ʔi-loɔ̂m-ŋʌ-nu										lɔ̂ːp-nu

Table 1. Positive non-past paradigm for the Khaling verb LOP 'catch', represented with standard features

Our two descriptions have been implemented on 167 verbs, producing 50,100 verb forms. We have measured the descriptions in terms of descriptive economy following Sagot and Walther (2011). This measure relies on the information theoretical concept of Minimum Description Length (MDL) (Rissanen, 1984). The measure allows in particular for showing the amount of information within different parts of the description, such as regular morphonological rules at morph boundaries ("phono+morphono"), the types of

X	X>1	1>X	B	C	D	E	F	G	H	I	J	K
2SG	loɔ̂m-ŋʌ	loɔ̂m-nɛ	lɵp-i	lɵp-u	lɔɔp-ki	lɔɔp-kʌ	lɔɔp	lɵp-i	loɔ̂m-ni	lɑ̄:b-ʉ	lô:p-su	lô:p-nu
2DU	loɔ̂m-ŋʌ-su	loɔ̂m-su										
2PL	loɔ̂m-ŋʌ-nu	loɔ̂m-nu										
3S	loɔ̂m-ŋʌ	lob-u										
3DU	loɔ̂m-ŋʌ-su	lob-u-su										
3PL	loɔ̂m-ŋʌ-nu	lob-u-nu										

Table 2. Positive non-past paradigm for the Khaling verb LOP 'catch', reorganised after having introduced morphomic feature mappings
E.g., B: 1DU.INCL, F: P=2SG, I: 2SG>3SG or 3SG>3SG, ...

phonological "operations" applied by the realisation rules, the information comprised within the "feature" specification and the feature mapping rules, and the general "structure" of the description, including the complexity of the realisation rules or the specified inflection classes.

In our experiment, the measure shows that the length of standard description A (with realisation rules realising transparent fully specified MPSs) is 6.2 Kbits, while the length of morphomic description B (realising above mentioned morphomic features in the rule sets coupled with separate inverse marking and morphomic mapping functions) is 5.4 Kbits.

These results show that using morphomic features makes for a description that is by 12% more compact than the non-morphomic description. In particular, the additional cost of feature mapping functions (here formalised with a mechanism similar to rules of *Paradigm Linkage* in PFM (Stump, 2006)) does not lead to a higher overall description length (due to the simplification of the realisation rules themselves). As illustrated by Figure 1, the additional information needed to stipulate the mapping between the morphomic features and the actual realised morphosyntactic property sets appears to be a more than reasonable tradeoff compared to the reduced length of the general description's structure.

Combined with supplementary morphomic features, direct/inverse marking as traditionally used as a distinctive feature within descriptive linguistics thus proves to be quantitatively relevant for describing some morphological systems, as it can lead to formulating maximally compact descriptions.

Acknowledgments This work was performed within the framework of the LabEx EFL (ANR-10-LABX-0083), Strand 6 "Language Resources", operation LR4.11. It was partly supported by the LabEx ASLAN (ANR-10-LABX-0081). Both LabEx's are operated by the French National Research Agency (ANR).

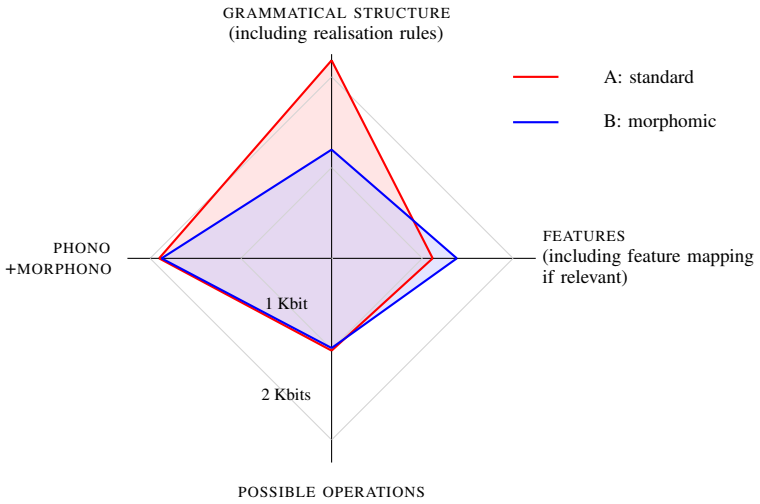


Fig. 1. Comparative description lenghts for two accounts of Khaling verbal inflection, breakdown on grammar components

References

ARONOFF, M.. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Linguistic Inquiry Monographs. MIT Press.

BONAMI, OLIVIER AND GILLES BOYÉ. 2010. Opaque paradigms, transparent forms in nepali conjugation. Workshop On Theoretical Morphology 5.

JACQUES, GUILLAUME, AIMÉE LAHAUSSOIS, BOYD MICHAILOVSKY, AND DHAN BAHADUR RAI. 2012. An overview of Khaling verbal morphology. *Language and linguistics*, **13.6**, 1095–1170.

RISSANEN, J.. 1984. Universal coding, information, prediction, and estimation. *IEEE Transactions on Information Theory*, **30**(4), 629–636.

SAGOT, BENOÎT AND GÉRALDINE WALTHER. 2011. Non-canonical inflection : data, formalisation and complexity measures. In C. MAHLOW AND M. PIOTROWSKI, Eds., *Systems and Frameworks in Computational Morphology*, volume 100, p. 23–45, Zurich, Switzerland: Springer.

SILVERSTEIN, MICHAEL. 1976. Hierarchy of Features and Ergativity. In R. M. DIXON, Ed., *Grammatical Categories in Australian Languages*.

STUMP, GREGORY T.. 2006. Heteroclisis and paradigm linkage. *Language*, **82**, 279–322.

ZÚÑIGA, FERNANDO. 2006. *Deixis and Alignment - Inverse systems in indigenous languages of the Americas*. Amsterdam: Benjamins.

POSTER SESSION

The nominal morphology of Lovari from an analogical perspective

András Márton Baló
Doctoral School in Linguistics, Theoretical Linguistics Programme, ELTE
baloam@gmail.com

The available information regarding the nominal inflectional paradigms of the Lovari dialect of Romani spoken in Hungary is contradictory, especially concerning the oblique stem. The data provided by descriptive grammars and textbooks do not always and fully coincide with the information provided by native informants. The relevant literature generally approaches the phenomena from a rule-based aspect, synchronic or diachronic, mentioning analogical effects only as a "last resort". (Analogy at work in Lovari is very conspicuous; the analogical processes and effects in the Lovari verbal paradigms are discussed by Baló 2012 for instance.) This paper attempts to provide evidence that analogy-based approaches make language change, as well as unstable and variegated forms and inflectional classes easier to understand and grasp than they would be in a traditional synchronic framework or through a diachronic analysis when the processes and the data are otherwise inexplicable.

Romani is a dialectally most diverse Indo-European language that is often exposed to contact-related influences (cf. Matras 1995). One important result of an intensive contact period of Early Romani with Greek is a cross-linguistically unique feature (Elšík 2000), the strict split between the morphology of inherited and borrowed vocabulary (also termed as the thematic-athematic dichotomy in Romani linguistics; for further information cf. Matras 2002), which is seen in a new light if we examine the possible processes behind the apparent erosion of this age-old system.

The difference between the inherited and borrowed layers of the lexicon manifests itself in the oblique stem. What we can see on a synchronic level is that the thematic oblique ending for the masculine is *-es-/-en-* (singular/plural), as opposed to athematic *-os-/-on-*. Elšík 2000 considers this from a diachronic perspective and claims that the reconstructed Proto-Romani ending of all oblique stems is *-en-/-es-*. In Early Romani, the lexical items borrowed newly from Greek took on different oblique endings, such as *-os-*, *-us-*, *-is-*, *-as-*, through intra-paradigmatic levelling. He also adds that the oblique endings were re-analysed into an oblique marker *-s-* and a "classification" vowel, or, we might as well say following a traditional term, a thematic vowel, which he considers to belong to the stem in some dialects. These additional oblique endings, if they existed in Lovari at some point, have completely disappeared. Not even descriptive grammars list different athematic classes based on a thematic vowel (cf. e.g. Cech—Heinschink 1999, Hutterer—Mészáros 1967 or Choli-Daróczi—Feyér 1989). No matter what the nominative ending of a noun is, it will fall into one of the two classes mentioned above (the oblique ending being either *-es-/-en-* or *-os-/-on-*). Therefore, it is probably more useful for us to consider the surface forms and their relation, and interpret words in a more holistic manner instead of breaking them down into even smaller morphemes (cf. Ackerman et al. 2009). The *-os-/-on-* pattern, possibly due to its frequency, simply spread onto the whole athematic layer.

Let us consider two examples: Early Romani **sapuni* 'soap' (present-day Lovari *sapuj*) and Early Romani *?*doktori* 'doctor'. The Early Romani oblique form *sapunis-* has been replaced by *sapujos-*, and similarly "*doktoros-*, i.e. not the original **doktoris-*" (Elšík 2000: 23). The second example, however, is problematic, because there is no reference to the word in earlier sources, and the oblique appears as *doktoros-* in the above-mentioned, recent Lovari sources. Other dialects may behave differently, but *doktori* could easily be a fairly new loanword, which has never had an oblique in *-is-* in Lovari.

Another well-known example is the case of the Greek-derived word *fōro* 'town'. We learn from the sources and informants that the forms of both the singular and the plural oblique

stems are ambiguous: they may be *fōres-/fōros-* and *fōren-/fōron-*, respectively. As Elšík 2000 states, diachronically, *fōros-* replaced *fōres-*, so that the oblique form could resemble the nominative singular. This process, however, goes against the basic layout of the "thematic" inflection, where the oblique singular stem ends in *-es-*, no matter what the nominative ending is (for example nominative singular *bālo* 'pig' and oblique singular stem *bālés-*). It might as well be the case that the loss of the word-final consonant resulted in a form similar to many thematic nouns, and the oblique form is slowly taking on the thematic pattern, too – or, if the historical aspect is considered, re-acquiring it. Using the Saussurean proportion, we can represent it as follows.

(1) *bāló : bālés- = fōró : x*
x = fōrés-

As for the oblique plural stem, Matras 2002 claims that *fōren-* became *fōron-*, possibly based on an analogy to the nominative singular. On the other hand, Hutterer—Mészáros 1967 mentions that the original form of the oblique plural stem of *sokro* 'father-in-law', from Romanian *socru*, is *sokron-*, but it appears more and more frequently in the form *sokren-*, and even the oblique singular stem can be *sokres-* instead of *sokros-*, as attested by informants. A similar process to the singular might have taken place in the plural as well.

If we disregard the thematic dichotomy for a moment, we find that basically there are two declension classes in Lovari: one masculine and one feminine. The thematic dichotomy, as Elšík 2000 already notes, is decaying. But is it actually decaying, or do we just see a balance in variation maintained by constant language change?

In the case of Romani, it is difficult to determine what the original forms of a certain word were exactly. The fact that the forms *fōren-*, *fōres-*, *sokren-* and *sokres-* are in use – and they are spreading, weakening the role of the forms in *-os-/-on-* – suggests that the thematic classes seem to exert an analogical force on the athematic borrowings, at least as far as the masculine is concerned.

This can be related to the fact that many borrowings become obscure; for a bilingual speaker of Hungarian and Lovari, the words *tudōšo* and *juhāsi* might be transparent borrowings; older borrowings, like *fōro* and *socro* might have become more integrated into the system. Generalisations may be made based on surface patterns that are "stronger" in some aspect; this might be the case for the historically deeper-rooted thematic pattern which is in constant opposition with the athematic pattern, which is not as old but has become well-established due to the high degree of contact Romani has been exposed to.

References

Ackerman, Farrell, James P. Blevins & Robert Malouf. 2009. "Parts and Wholes: Implicative Patterns in Inflectional Paradigms". *Analogy in Grammar: Form and Acquisition* ed. by James P. Blevins & Juliette Blevins, 54-82. Oxford: Oxford University Press. • Baló, András Márton. 2012. "Arguments from Lovari loan-verb adaptation for an analogy-based analysis of verbal systems." *Current Issues in Morphological Theory: (Ir)Regularity, Analogy and Frequency* (Current Issues in Linguistic Theory 322.) ed. by Ferenc Kiefer, Péter Siptár P. & Mária Ladányi, 3-22. Amsterdam: John Benjamins. • Cech, Petra & Mozes F. Heinschink. 1999. *Basisgrammatik. Arbeitsbereicht 1a des Projekts Kodifizierung der Romanes-Variante der österreichischen Lovara* (ed. by Dieter W. Halwachs). Vienna: Verein Romano Centro. • Choli-Daróczi, József & Levente Fejér. 1988. *Cigány nyelvkönyv* [Romani Textbook]. Budapest: Magyarországi Cigányok Kulturális Szövetsége. • Elšík, Viktor. 2000. "Romani nominal paradigms: their structure, diversity, and development". *Grammatical relations in Romani: The noun phrase.* (Current Issues in Linguistic Theory, 211.) ed. by Viktor Elšík & Yaron Matras, 9-30. Amsterdam: John Benjamins. • Hutterer, Miklós & György Mészáros. 1967. *A lovári cigány dialektus leíró nyelvtana* [Descriptive Grammar of the Lovari Dialect of Romani]. Budapest: Magyar Nyelvtudományi Társaság. • Matras, Yaron ed. 1995. *Romani in Contact. The History, Structure and Sociology of a Language.* Amsterdam: John Benjamins • Matras, Yaron. 2002. *Romani: A linguistic introduction.* Cambridge: Cambridge University Press. • Saussure, Ferdinand de. 1966. *Course in General Linguistics.* New York: McGraw-Hill Book Company. • Vekerdi, József. 2000. *A Comparative Dictionary of Gypsy Dialects in Hungary*. Budapest: Terebess.

Italian VN compounds are not exocentric

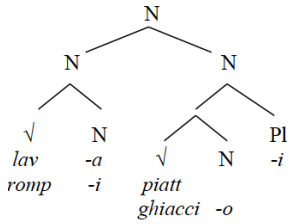
LUDOVICO FRANCO*^ – M. RITA MANZINI^ – LEONARDO SAVOIA^
*Universidade Nova de Lisboa; ^Università di Firenze
franco.ludovico@gmail.com - rmanzini@unifi.it - lsavoia@unifi.it

What is the nature of the verb in Italian 'exocentric' Verb-Noun (VN) compounds, cf. (1)?

(1) a. *lavapiatti* wash-dishes ('dishwasher')
b. *rompighiaccio* break-ice ('icebreaker')

There are many different approaches in the literature to this issue and such point is still a matter of contemporary debate (Floricic, 2008; von Heusinger & Schwarze, 2013). The relevant V item has been considered: an imperative form (Rohlfs, 1968; Progovac & Locke, 2009); an indicative third person singular (Di Sciullo, 1992); a deverbal nominal form (Zuffi, 1981; Bisetto, 2004); a root plus a thematic vowel (Scalise, 1992; Vogel & Napoli, 1995; cf. also Ralli, 2008, who argues that the thematic vowel acts as a linker). All these hypotheses have weak points (see Ferrari Bridges, 2005; Floricic, 2008). Let us assume that the first constituent of Italian VN compounds consists of a root form, i.e. a predicative lexical base, followed by a 'thematic vowel'. Following an insight of Manzini & Savoia (2005: 487-489), we propose that the thematic vowels of verbs and nouns have the same categorial signature N(ominal class). Thus *rompi* (break) in (1b) or *lava* (wash) in (1a) can be analysed as in (2) where the lexical base, indicated as √, expressing predicative content, combines with an N inflection (*-i*, *-a*). In the same way, ghiaccio consists of the predicative base ghiacci- 'ice' and the N inflection *–o* (we take the *-i* of piatti in (1a) to be plural morphology). Stems of the √+N forms are not individuated as verbs or noun. A √+N structure is read as a Noun as a default – only if embedded under TMA (Tense/Mood/Aspect structures) is it read as a verb. Under these assumptions, structures like (2) need not be exocentric. Rather they are regular left-headed compounds where the category neutral stem *lava*, *rompi* provides both the semantic head and the syntactic head. In particular, in the absence of other specifications, the *lava*, *rompi* stems correctly project N by default.

(2)



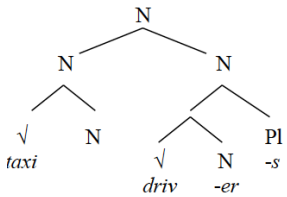
One obvious question concerns the exact nature of N. We assume that predicative bases √ – not only 'wash' or 'break', but also 'ice' or 'dish' – are associated with at least an argumental slot. The internal argument slot of the latter is generally referred to in the literature as the R-Role (Higginbotham, 1985; Williams, 1994). We take it that the N class vowel is the lowest level saturation of the R-role, eventually to be closed by the determiner D. We assume much the same for 'wash' or 'break', eventually to be closed by referential arguments. Another question concerns the non-existence of 'nouns' such as **rompi, *lava* in isolation; this has proved in the past a powerful objection to assuming that they head the N compound in (2). In other words, why would the stem on the left in (2) be an N only in compounds? The natural assumption is that the internal theta-role of the lexical base must be closed by more than the N inflection, specifically by an N level constituent, such as *piatti* or *ghiaccio* in (2) – whence the compounding. A third question concerns the fact that *rompi* or *lava* are perfectly good as verbal forms in isolation. We naturally assume that this is so only to the extent that they are embedded TMA structures. One of the proposals alluded to at the beginning takes the left-

hand member of the compound in (2) to be a 2nd sg imperative. On the contrary, we argue that such forms can become verbal only to the extent that they are embedded in particular under the modal position C characteristic of imperatives (for Romance, see Rivero, 1988). Manzini & Savoia (2005, cf. also Rohlfs, 1969; Giannelli, 1976; Stefanini, 1970) further note that in the Italian variety of Firenze (Florence) forms morphologically corresponding to the 2nd sg imperative can appear as complements of some verbal classes which usually require infinitival complements in Romance, as in (3a-b). (3a') provides comparison with the prosodically different morphological infinitive, (3b') with the morphologically identical imperative. Again modal value accrues to the *spiega* 'explain' and *rompi* 'lit: break' in (3) from the high modal position of the verb (associated with enclisis, cf. Kayne, 1991).

(3) a. *va* *a* *s'pjɛga-ʎʎi* *he* …
go.IMP.2SG to explain√-him COMP ('go to explain to him that…')
a'. *va* *a* *spjɛ'ga-ʎʎi* *he*…
go.IMP.2SG to explain.INF-him COMP ('go to explain to him that…')
b. *'basta* *un* *'rompi-ʎʎi* *'l* *anima*
be.enough.PRS.3SG NEG break√-him the soul ('it is better not to bother him')
b'. *un* *'rompi-ʎʎi* *θanto* *'l* *anima!*
NEG break.IMP.2SG-him so.much the soul ('don't bother him so much')

Finally, we shall provide some cross-linguistic discussion of the ways in which different languages realize agentive/instrumental synthetic compounds (henceforth: A/ISC). Languages seem to realize the A/ISC configuration, using inflectional morphology or derivational morphology. Thus an architecture containing inflectional and derivational processes in a single grammatical component (Williams, 1981) seems to be on the right track. Germanic languages, like English, tend to use 'derivational' morphemes (e.g. the morpheme –er) to express forms of the dishwasher or icebreaker type (forms like scarecrow would have a zero morpheme instead of the overt -er suffix, according to e.g. Marchand, 1969; Scalise & Bisetto, 2005; see also Bauer, 2008). The idea pursued here is that English uses a so-called derivational tool to do what in Romance is done with inflection, that is to join a nominal class slot, N, to a (verbal) root, √.

(4)



Interestingly, Old English in (5) seems to adopt the inflectional strategy for A/ISCs, since according to the analysis given in Kastovsky (1985: 246ff, cf. also Gast, 2009) the endings *-a, -e* have as their primarily function to determine the inflectional class (i.e. case/number) of the lexical item to which they attach. In fact, the Old English examples closely resemble Latin NV items with analogous agentive/instrumental function such as the ones in (6).

(5) a. *æw-brec-a* b. *mere-far-a*
law-break-infl ('law-breaker') sea-travel-infl ('sea-traveller, sailor')
(6) a. *agr-i-col-a* b. *part-i-cep-s*
field-gen-farm-infl ('farmer') part-gen-take-infl ('part-taker')

Selected References: BAUER, L. 2008. Exocentric compounds. *Morphology* 18, 51-74. – FLORICIC, F. 2008. The Italian verb–noun anthroponymic compounds. *Morphology* 18, 167-193. – HIGGINBOTHAM, J. 1985. On Semantics. *Linguistic Inquiry* 16, 547-621. – MANZINI, M.R. & SAVOIA, L.M. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*, vol. III, Alessandria, dell'Orso. – SCALISE, S. 1992. Compounding in Italian. *Rivista di Linguistica* 4, 175-199. – WILLIAMS, E. 1994. *Thematic Structure in Syntax*. Cambridge, MA., MIT Press.

What makes me sleepy? Morphological polyfunctionality of causatives and reflexives in Mari and Udmurt

NIKOLETT F. GULYÁS
Eötvös Loránd University
nikolett.fgulyas@gmail.com

Impersonal constructions of Finno-Ugric languages have principally been studied from the viewpoint of verbal morphosyntax (cf. Stipa 1962, Schiefer 1981). Functional investigations have emerged recently in linguistic typology (Malchukov & Siewierska 2011) and this kind of approach has been applied for the major Finno-Ugric languages (Posio & Vilkuna 2013). In my presentation, I will focus on two lesser-studied FU languages, these are Mari and Udmurt. The goal of this study is to describe some particular impersonals within the terms of morphological polyfunctionality.

In both languages there are very rich verbal derivational and case marking systems (cf. Bartens 2000). The present study will describe two verbal suffixes demonstrating morphological polyfunctionality; causatives and reflexives.

Causation is a valency (transitivity) increasing strategy (Song 2001). Morphological causation is expressed by the derivational affix *-kt* and *-alt* in Mari and *-t* in Udmurt. In canonical cases, the causer is marked with the nominative followed by a verb bearing the causative suffix, while the causee is marked with the accusative in Udmurt. In Mari, the causee is sometimes marked with the dative:

(1) Udmurt

Mon	*so-je*	*gožtet*	*gožtį-t-iśko.* (Winkler 2001: 56)
I	(s)he-ACC	letter[ACC]	write-CAUS-PRS.1SG

'I have the boy write the letter.'

(2) Mari

Pet'a	*Lena-lan/Lena-m*	*šür-əm*	*šolt-əkt-en-Ø.* (Kalinina et al. 2006: 446)
Peter	Helen-DAT/Helen-ACC	soup-ACC	boil-CAUS-PST-3SG.II

'Peter made Helen cook (lit. boil) soup.'

In constructions (3) and (4), the causative verb has only one obligatory argument standing in the accusative:

(3) Ud. *Ataj-ez* *berįk-t-e.* (Y. S.)

(4) Ma. *Ačа-m* *vakə̑-kt-a.* (T. E.)
father-ACC vomit-CAUS-PRS.3SG
'The father is nauseous.'

According to functional considerations, these constructions demonstrate a slightly different usage of causative verbs that I reckon as impersonal. Following Siewierska (2008: 116), I consider given constructions to be impersonal if they lack a subject bearing canonical properties. Subject is determined here in functional terms; a canonical subject is an argument that is referential, topical, agentive, animate and definite (cf. Malchukov & Ogawa 2011: 23). As the translations of the latter two examples show, the accusative-marked argument is the logical subject of the causative impersonal construction, that is, it bears the thematic role of the Patient or the Experiencer.

A same functional differentiation is presented in the usage of reflexive suffixes in both languages. The dative-marked argument of the construction displays functional properties similar to the former type but can be considered less salient:

(5) Ud. (*Mon*) *pinal-lį* *śoti-śk-i-z.* (Y. S.)
(I) child-DAT give-REFL-PST-3SG

‘I gave (something) to the child unintentionally.’/‘I was forced to give (something) to the child.’

The data examined in this study were elicited from native speakers using two questionnaires which focused on pragmatic neutrality and acceptability factors. Examples were provided by three informants for each language. To outline the characteristics of the use of these special kinds of impersonals, I will present data of a corpus based survey. The corpus consists of the parallel translations of a Russian novel (PM 2010), which contains 12500 words and texts collected from reference grammars (Bereczki 1990, Csúcs 1990).

References

BARTENS, RAIJA 2000: Permilaisten kielten rakenne ja kehitys. *MSFOu* 238.

BERECZKI, GÁBOR 1990: *Chrestomathia Ceremissica*. Budapest: Tankönyvkiadó.

CSÚCS, SÁNDOR 1990: *Chrestomathia Votiacica.* Budapest: Tankönyvkiadó.

KALININA, ELENA, DMITRIY KOLOMATSKY, ALEXANDRA SUDOBINA 2006: Transitivity increase markers interacting with verb semantics. Evidence from Finno-Ugric languages. In: Kulikov, L., A. Malchukov, P. de Swart (eds.): *Case, Valency and Transitivity*, 441–464. Amsterdam: John Benjamins.

MALCHUKOV, ANDREJ & AKIO OGAWA 2011: Towards a typology of impersonal constructions. A semantic map approach. In: Malchukov, A. & A. Siewierska (eds.): *Impersonal constructions. A cross-linguistic perspective,* 19–56. Amsterdam: John Benjamins.

MALCHUKOV, ANDREJ & ANNA SIEWIERSKA 2011: Introduction. In: Malchukov, A. & A. Siewierska (eds.): *Impersonal constructions. A cross-linguistic perspective,* 1–15. Amsterdam: John Benjamins.

PM 2010: *Pavlik Morzov -paralleelitekstikorpus*. Turku: Research Unit for Volgaic Languages, University of Turku. (Manuscript)

POSIO, PEKKA & MARIA VILKUNA 2013: Referential dimensions of human impersonals in dialectal European Portuguese and Finnish. *Linguistics* 51 (1): 177–229.

SCHIEFER, ERHARD 1981: Das anonyme Subjekt in den finnisch-ugrischen Sprachen. In: Bereczki Gábor–Molnár József (eds.): *Lakó-emlékkönyv*, 140–157. Budapest.

SIEWIERSKA, ANNA 2008: Introduction: Impersonalization from a subject-centred vs. agent-centred perspective. *Transactions of the Philological Society*, 106 (2): 115–137.

SONG, JAE JUNG 2001: Linguistic Typology. Morphology and Syntax. Harlow: Longman.

STIPA, GÜNTER 1962: Impersonale Ausdrucksformen. In: Commentationes Fenno-Ugricae in honorem Paavo Ravila. *MSFOu* 125: 579–592.

WINKLER, EBERHARD 2001: *Udmurt*. Lincom Europa.

The Grammatical Category of Case in Early Development of Slovak-speaking Children

JANA KESSELOVÁ
Faculty of Arts University of Presov in Presov
jana.kesselova@unipo.sk

The aim of the research is to describe and to explain the development of grammatical category of case in early development of Slovak-speaking children (up to 36 months). The main research question is as follows: What is the development of understanding and production of case in early development of Slovak-speaking children? In order to find an answer to this question, we must decide the following sub-questions:
(i) What is the relation between case semantics and lexical semantics?
(ii) Which cases represent preferentially acquired core of case subsystem in Slovak-speaking children?
(iii) What is the relation between case semantics and pragmatics of child utterance?
We are using two methods: qualitative and quantitative. Qualitative method presents longitudinal studies of 5 Slovak-speaking children communicating with their parents and care givers in natural conditions and in standard situations (bathing, feeding and playing). Video records were made by children's parents (one hour per month). We are using transcription system CHAT of CHILDES (CHIld Language Data Exchange System; http://www.laboratorium.detskarec.sk/ukazka.php). Quantitative method presents Slovak adaptation of MacArthur-Bates Communicative Development Inventories – Test of Communicative Behaviour II (TEKOS II http://www.laboratorium.detskarec.sk/tekos2.php). The research sample consists of 1065 children (539 girls and 526 boys) from 17 to 36 months. Information about child's speech development was provided by parents.

(i) The first case forms in early development of Slovak-speaking children appear in 4 semantic categories. The first category is ***objectness*** within which names of things and people in nominative and accusative are formed. The second category called ***locationality*** includes formation of the grammatical case in Slovak-speaking children. The substantive cases are formed as three opposites: up/down, close/far, inside/outside. We notice these opposites in non-verbal gestures first, followed by development of verbal means (lexis and grammar). All case forms are formed within this category:

(1) the opposite ***inside/outside***
(1a) beginning of contact with interior of an object: the preposition *do/into* + the genitive case
do domčeka – into the house
(1b) duration of contact with interior of an object: the preposition *v/in* + the locative case
v lese – in the forest
(2a) beginning of contact with surface of an object: *na/on* + the locative case
na deku – on the blanket
(2b) duration of contact with surface of an object: *na/on* + the accusative case
na deke – on the blanket

(3) the opposites ***close/far*** (spatial and equally social contact as a means of showing the need to be with somebody)
(3a) beginning of contact with a person: *k/to* + the dative case
k mame – to mum
(3b) duration of contact with a person: s/with + the instrumental case
s mamou – with mum

(4) the opposite ***down/up***
(4a) down: *pod*/*under* + the instrumental case
pod stolom – under the table
(4b) up: *nad*/*above* + the instrumental case
nad fľaškou – above the bottle

The third semantic category is ***benefitiality*.** A child uses dative or accusative case with preposition *for* to name a beneficiary (*mam**e**, pre mam**u** – for mum).* The ***sociativity*** is the fourth semantic category which reflects a child's need to be with somebody, be close to somebody: *with* + the instrumental case (*s mamou – with mum*). This category coincides with locationality category.

(ii) The case formation in early development of Slovak-speaking children can be divided into two periods: successive a simultaneous. The grammatical cases (nominative and accusative) appear in a child's speech successively. In the development of grammatical case, first the opposition of nominative – accusative is formed. The quantitative research shows that from 21st month accusative is used by more than a half of Slovak-speaking children. Semantic cases (dative, locative and instrumental) appear together (simultaneously) in a different order, age and with different frequency in each child (in 5 longitudinal studies from 19 to 25 months). However, semantic concretization of cases and pragmatic functions of utterances with case forms develop in all children identically. An example of instrumental: While in 17th month an instrumental is used by 5 % of children, in 36th month it is used by almost 90 % of children. In early development of Slovak-speaking children the first to be formed is sociative instrumental (*with father*) and instrumental of device (*by car*). Instrumental of place with prepositions appear less frequently (in order *under – in front of – above – behind*). The development of the case is gradual not in stages. The preferential core of the case system (non-prepositional accusative, dative of benefit, sociative instrumental, accusative case of plural, direction genitive, direction dative, nominative of plural, locative of place, instrumental of means, direction accusative) is acquired by Slovak-speaking children in this order: in 17th month of age from 1 to 10 % of children uses the cases. In 24th month approximately a half of children do so and in 36th month almost all (90%) of children in our sample (1065 children).
(iii) In children acquiring Slovak the preferential core of the case system can be seen in utterances with three pragmatic functions: con-situational information (a child comments an activity, that he or she is currently doing), information about an intention (a child wants to do something with a tool), challenge or a wish (a child wants to do an activity with a close person). Development of grammatical category of case in early development of Slovak-speaking children demonstrates how cognition (transition to structures of relatedness, cognitive coordination of an aim and means to achieve it), social experience, experience with item manipulation and with movement in space reflect into preferentially acquired cases.

References

EISENBEISS, S. – NARASIMHAN, B. – VOEIKOVA, M. 2009. The acquisition of case. In: Malchukov, A. and Spencer, A. (eds.) *The Oxford Handbook of Case*. Oxford OUP, 369–383. • GVOZDEV, A. N. 1961. *Voprosy izučenija detskoj reči,* 273–358. Moskva: Izdateľstvo APN RSFSR • PAČESOVÁ, J. 1968. *The Development of vocabulary in the child,* 162–251. Brno: Univerzita J. E. Purkyně. • VOEIKOVA, M. – DRESSLER, W. 2002: Introduction. In: Maria D. Voeikova and Wolfgang Dressler (eds.) *Pre- and Protomorphology: Early Phases of Morphological Development in Nouns and Verbs*. Lincom Studies in Theoretical Linguistics 29. Munich: Lincom Europa, 3–5.

Modular affixation
Form-related operations in an Item-and-Process model of word formation

ANDREAS NOLDA
University of Szeged
andreas@nolda.org

In this poster, I shall argue for a modular conception of affixation processes and other form-related operations that allows for flexible and general solutions of problems in word-formation description. This conception is part of a formal theory of word formation in the Item-and-Process tradition, called "Pattern-and-Restriction Theory" (PR; Nolda 2012), which makes crucial use of functional, multidimensional word-formation patterns. In this respect, PR resembles theories such as Lexeme Morpheme Base Morphology (Beard 1995) or Lexical Relatedness Morphology (Bochner 1993), although it surpasses them considerably in terms of formal explicitness and flexibility.

While PR has been fully formalised in axiomatic manner, the poster will present the basic theoretical tenets and the empirical advantages of the conception as informally as possible, by applying it to selected examples from German. The analyses of the examples discussed in the poster will, in addition, be made available as a computer implementation of the theory for a hands-on experience.

According to PR's basic idea, the word-formation component of a linguistic system provides a set of *word-formation patterns*. A pattern is viewed as a multidimensional combination of functional word-formation means: a *formal means*, certain *categorial means*, and a *semantic means*. Those means are modelled as mathematical operations on forms, categories, or concepts, respectively. One and the same formal means can be combined with different categorial and semantic means of the same arity, resulting in formally related, but categorially and/or semantically different patterns. Given that formal means also include affixation operations, apparent polysemy or homonymy of affixes is not treated as a lexical matter, but as a grammatical one (for arguments against a lexical conception of affixes, cf. Wurzel 1989; Beard 1995: chap. 4).

Patterns not only differ in terms of the means they involve, but also in terms of the restrictions they impose on possible bases. This is taken into account by assigning each pattern of the system a corresponding *base restriction*.

Formal means capture all sorts of form-related morphological – and, possibly, morphophonological – manipulation, including additive, subtractive, and modificational operations. Instead of taking basic operations such as concatenation, subtraction, or substitution as primitives of the theory, formal means are typically conceived in PR as specialisations or combinations thereof. By doing so, complex formal means can be defined in a modular way from simpler operations.

Simple affixation operations are defined as one-place operations concatenating a variable base form and a constant affix form in a given order. For example, the formal means of *ung*-suffixation in German can be defined as the concatenation of a base form with the affix form *ung* (in that order). From simple affixation operations, complex affixation operations can be derived in various ways. I shall indicate three of them here.

In a first type of complex affixation operations, a simple affixation operation is combined with a suprasegmental operation. An example at hand is the *ˈei*-suffixation used in a pattern for the derivation of location nouns from profession nouns in German: *ˌMetzger ˈei* 'butcher's shop' < *ˈMetzger* 'butcher', *Konˌditor ˈei* 'confectionery' < *Konˈditor* 'confectioner', *ˌAbt ˈei*

'abbey' < *ˈAbt* 'abbot', etc. ("ˈ" and "ˌ" denote primary and secondary lexical accents, respectively). This suffixation operation combines simple *ˈei*-suffixation by functional composition with a previously applied deaccentuation operation, the latter substituting primary accents in base forms by secondary ones. Provided that base and affix forms have inherent suprasegmental phonological properties, complex affixation operations of this type make distinctions between 'class I' and 'class II' affixes (or 'level I' and 'level II' affixes) dispensable – at least insofar, as such distinctions are to account for suprasegmental differences in the product forms (for discussion, cf. Booij 1987, Bochner 1993: chap. 6).

A second type of complex affixation is *conditional affixation*, which is particularly useful for modelling allomorphic variation with complementary distribution of formally related affix forms. Such affixation operations are defined by distinguishing several cases, involving different affix forms, with one of the cases possibly functioning as a default. As an example consider a variant of the *ˈei*-suffixation mentioned above, figuring in a pattern for the derivation of iterative process nouns from verbs in German, e.g. *ˌTanz er ˈei* 'pointless dancing' < *ˈtanz* (*en*) 'to dance', *ˌKletter ˈei* 'pointless climbing' < *ˈkletter* (*n*) 'to climb', *ˌWackel ˈei* 'pointless wobbling' < *ˈwackel* (*n*) 'to wobble' (orthographic case is, of course, phonologically insignificant). Here, *ˈei*-suffixation, with preceding base deaccentuation, is only applied to base forms ending in *er* or *el*; in the default case, however, deaccentuated base forms are suffixed first by ('morphomic') *er* and then by *ˈei*. (This affixation operation appears also to be involved in a pattern deriving location nouns from verbs; for an overview cf. Motsch 2004: 334 f., 351 f.)

A third type of complex affixation operations are *substitutive affixations*, which substitute an affix form for an element of the base forms. Such an affixation operation may be assumed for a pattern for the derivation of German profession nouns from competence domain nouns, such as *ˌPoˈlit ik er* 'politician' < *ˌPolit ˈik* 'politics' or *ˈChem ik er* 'chemist' < *ˌChem ˈie* 'chemistry' (a pattern which presumably is restricted to certain neoclassical base nouns). The formation of those nouns can be analysed in a uniform manner by reference to a formal means which first deletes the final morphological segment of the base form and accents the final syllable of the remainder, before suffixing both *ik* and *er*. As a result, *ˈie* in *ˌChem ˈie* is replaced by *ik* in *ˈChem ik er*, while the *ˈik* in *ˌPolit ˈik* is only changed suprasegmentally in *ˌPoˈlit ik er*. Note that this analysis neither resorts to adjustment rules in the sense of Aronoff (1976) nor necessarily has to assume bound stem variants like *Chem* (as Aronoff 1994 does for *nomin ee* < *nomin ate*).

References

ARONOFF, M. 1976. *Word Formation in Generative Grammar*. Linguistic Inquiry Monographs 1. Cambridge, Mass.: MIT Press. • ARONOFF, M. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Linguistic Inquiry Monographs 22. Cambridge, Mass.: MIT Press. • BEARD, R. 1995. *Lexeme Morpheme Base Morphology: A General Theory of Inflection and Word Formation*. Albany: State University of New York Press. • BOCHNER, H. 1993. *Simplicity in Generative Morphology*. Berlin: Mouton de Gruyter. • BOOIJ, G. 1987. Lexical Phonology and the organization of the morphological component. In: Gussmann, E. (ed.), *Rules and the Lexicon: Studies in Word-formation*, Lublin: Katolicki Uniwersytet Lubelskiego, 43–65. • MOTSCH, W. 2004. *Deutsche Wortbildung in Grundzügen*. 2nd ed. Schriften des Instituts für deutsche Sprache 8. Berlin: de Gruyter. • NOLDA, A. 2012. Konversion im Deutschen – Muster und Beschränkungen: Mit einem Grundriss einer allgemeinen Theorie der Wortbildung. Habilitation thesis, Humboldt-Universität zu Berlin. • WURZEL, W. U. 1989. Von der Inadäquatheit einer Affixmorphologie: Weshalb morphologische Kategorienmarker nicht als eigene Einheiten im Lexikon repräsentiert sein können. *Linguistische Studien* A 194: 277–298.

Morphological priming in Turkish nominal compound processing

SİBEL ÖZER, ANNETTE HOHENBERGER

Middle East Technical University, Ankara, TURKEY

sibelozer.ege@gmail.com,
hohenber@metu.edu.tr

Compounding, constructing new words out of previously known lexemes by means of simple concatenation mostly, can be counted as one of the major word production mechanisms in the majority of languages including Turkish.[1,10,11 and 4] Their importance in the history of human languages warrants a detailed study with respect to the language faculty and related cognitive aspects.[12] In the last decade, compound production as well as comprehension have become highly debated and investigated areas of research. Morphological priming is one frequently employed paradigm for the investigation of compounding. Whether morphologically complex words undergo a decomposition-composition process, respectively, during comprehension and production or whether they are all listed in full form in the lexicon is one key question hitherto addressed using picture naming task in several studies related to English, German and Dutch nominal compound words [3, 6, 14, and 15].

The present study is concerned with compound production in Turkish. Various types of Turkish compounds were investigated [2, 5, 7, 8, 9, and 13]:

(1) bare JCs (‘akbalık’, ‘dace’),
(2) indefinite (‘dil balığı’, ‘flounder’)
(3) definite (‘gölün balığı’, ‘fish of the lake’) izafet constructions

in a morphological priming paradigm by means of a picture naming task. In the general implementation of this task, subjects name black-white line drawings of simple objects in a limited and pre-specified time-interval while at the same time, they have to ignore distractor words which are presented visually(or auditorily). The locus of interest in this paradigm is the evaluation of possible linguistic effects of the distractor word presentation on picture naming performance. In this study, distractor words were Turkish nominal compounds and picture names(e.g., ‘balık’, ‘fish’) were morphologically related (depicted either first or second part of the compound) or completely unrelated to these distractor words. Results of the experiment revealed equal amounts of morphological priming effect in all compound types investigated, that is, morphologically related compounds led to shorter naming latencies compared to unrelated distractors, a result which is in line with the decompositional (also possibly opportunistic framework proposed by Libben[12]) view of compound processing. Furthermore, significant animacy effect found on naming latencies irrespective of the compound type, underlines another possible essential factor in compound processing. Finally, distractor-wise analysis revealed marginally significant reaction time advantages for the head part of the compound as compared to the modifier part, a finding which suggests a possible special role for the head constituent during lexical access.

References

[1] BAUER, Laurie. Introducing linguistic morphology. 2003. [2]DEDE, Müşerref Ağan. A syntactic and semantic

analysis of Turkish nominal compounds. 1978. PhD Thesis. University of Michigan. [3]DOHMES, Petra; ZWITSERLOOD, Pienie; BÖLTE, Jens. The impact of semantic transparency of morphologically complex words on picture naming. Brain and Language, 2004, 90.1: 203-212. [4]GÖKSEL, Aslı. Compounds in turkish. Lingue e linguaggio, 2009, 8.2: 213-236. [5]GÖKSEL, Asli; KERSLAKE, Celia. Turkish: A comprehensive grammar. Psychology Press, 2005. [6]KOESTER, Dirk; SCHILLER, Niels O. Morphological priming in overt language production: Electrophysiological evidence from Dutch. Neuroimage, 2008, 42.4: 1622-1630. [7]KÖNIG, Wolf Dietrich. Nominalkomposita im Türkischen. Zeitschrift für Sprachwissenschaft, 1987, 6.2: 165-185. [8]KORNFILT, Jaklin. Turkish. descriptive grammars. London and New York, 1997. [9]LEWIS, Geoffrey L. Turkish grammar. Clarendon Press, 1985, 2001. [10][LIEBER, Rochelle; STEKAUER, Pavol (ed.). The Oxford handbook of compounding. Oxford University Press, 2009. [11]LIBBEN, Gary; JAREMA, Gonia (ed.). The representation and processing of compound words. Oxford University Press, 2006. [12] LIBBEN, Gary. The nature of compounds: A psychocentric perspective. Cognitive neuropsychology, 2014, ahead-of-print: 1-18. [13]UNDERHILL, Robert. Turkish grammar. MIT Press, 1976. [14]ZWITSERLOOD, Pienie; BÖLTE, Jens; DOHMES, Petra. Morphological effects on speech production: Evidence from picture naming. Language and Cognitive Processes, 2000, 15.4-5: 563-591. [15]ZWITSERLOOD, Pienie; BÖLTE, Jens; DOHMES, Petra. Where and how morphologically complex words interplay with naming pictures. Brain and Language, 2002, 81.1: 358-367.

Irony processing and contextual effects in schizophrenia – A neurolinguistic perspective

ZSUZSANNA SCHNELLa#, ESZTER VARGAb*, TAMÁS TÉNYIb*, MÁRIA SIMONb*, A. HAJNALb*, TAMÁS DÓCZYd,f*, SÁMUEL KOMOLYe*, J. JANSZKYc,e,f*, EDINA HAMVASb*, S. FEKETEb* and RÓBERT HEROLDb*

a Department of Linguistics and Institute of Psychology, University of Pécs, Hungary
b Department of Psychiatry and Psychotherapy, Medical Faculty, University of Pécs, Hungary
c Neuro CT Diagnostic Center, Hungary
d Department of Neurosurgery, Medical Faculty, University of Pécs, Hungary
e Department of Neurology, Medical Faculty, University of Pécs, Hungary
f MTA-PTE Clinical Neuroscience MR Research Group, Hungary
*Supported by the *National Brain Research Program (NAP) KTIA_NAP_13_1_2013_001 Grant*
#Supperted by *TÁMOP 4.2.4. A/2-11-1-2012-0001 „National Excellence Program and SROP-4.2.2.C-11/1/KONV-2012-0005 Well-Being in the Information Society project*

zsuzsischnell@gmail.com

Schizophrenic patients are known to be diagnosed with atypical Theory of Mind (ToM) mechanisms even during remission, which is believed to be responsible, or at least, to contribute to their deficit in pragmatic competence, within that, to the processing of non-compositional constructions. Several studies reveal a connection between irony comprehension and ToM capacities (Frith & Frith, 2003; Happé, 1993), given that schizophrenic patients demonstrate difficulty in irony processing, which has been associated with their deficient mentalization skills. We examined the neural correlates of irony understanding in schizophrenic patients, as an indicator of ToM capacity, and evaluated how surface cues (i.e., linguistic help) could affect irony comprehension. Schizophrenic patients in remission and healthy controls were submitted to event-related functional MRI scanning while performing three tasks: 1. irony, 2. irony with linguistic help and 3. control tasks.

Mental background of ToM and schizophrenia

Schizophrenia is a pervasive mental disorder often resulting in incapacitation and in social marginalization. The basis of this neuropsychiatric illness is based on a failure of social participation, since patients generally have problems with the perception of reality. Their mental illness generates unreal sensations that they perceive as part of reality. This mental disturbance is named as mentalization deficit, also known as Theory of Mind (ToM) deficit in literature, stating that schizophrenic patients have difficulty representing their own and others' mental states (will, desire, intention, feelings, beliefs and knowledge) realistically (Dennett, 1989; Frith & Frith, 1999; Woodruff & Premack, 1978), making Theory of Mind a core component of social cognition. For this reason the close examination of mentalization ability and its interaction with language in communicative settings based on social cognitive skills can eventually shed light on the yet unknown cognitive background mechanism responsible for this complex neuropsychiatric disorder resulting in social marginalization and often isolation, and may result in a training paradigm that can contribute to the reintegration of such patients into society.

Linguistic background

Irony is a type of polysemy, a non-compositional construction, where intended meaning does not correspond to the meaning of the constituents. That is, ironic meaning cannot be interpreted on the basis of sheer linguistic input (semantically), only and strictly integrating

contextual cues, i.e., pragmatically. Thus, to be able to decode the intended (that is, ironic) meaning of the utterance, a holistic meaning construction is necessary that breaks away from the semantic analysis. For the successful deciphering of ironic meaning the listener relies on cognitive strategies, which we define as mentalization, based on a change of perspectives, for which cues are drawn from the situational and linguistic context. This process of meaning construction involves the coordination of an implicit communicative intent, where the implicit meaning is the opposite of what is explicitly expressed (Giora 2001). Irony comprehension thus represents a form of non-compositional language use that requires not only semantic and syntactic decoding, but also the deciphering of the speaker's non-linguistic implicature (Sperber & Wilson, 1995; 2002, Alberti 2011). Therefore, the integration of intended meaning and social context is essential for irony comprehension in order to be able to represent the speaker's mental state, and decipher their intention, and eventually the intended meaning of the utterance heard. It is important for the listener to recognize that the actual intention expressed by the speaker is in fact, the opposite of the utterance heard. Thus, irony understanding requires not only the correct interpretation of communicative intentions, but also the ability to construct a coherent narrative based on contradictory information. Up to now, only a few brain imaging studies have examined the neural basis of irony comprehension. The present study would like to contribute to this line of research and reveal potential correlations between language and cognition which can eventually yield a communicative strategy-based treatment to facilitate integration and reintegration of those affected by the disorder.

The study

Our aim was to investigate irony comprehension and the underlying brain activity in patients with schizophrenia. We examined various phases of irony processing separately (such as context phase, statement phase, as well as question on comprehension and answer phase). We also evaluated whether providing more explicit contextual information modifies the patients' performance in irony comprehension and the related brain activities. We hypothesized that patients with schizophrenia would perform worse and exhibit an abnormal brain activation pattern during irony comprehension (Varga et al. 2013, 2014). We assumed that linguistic cues depicting the speaker's emotional state improve patients' performance and modify the network activation. To model complex, real life discourse settings as they take place in natural languages, we used short social scenarios presented as auditory stimuli in an event related design, and examined the differences in irony processing at various stages of the paradigm in patients and controls.

Results

The results indicate that patients handled irony significantly worse than healthy controls. The patients showed stronger brain activity in the parietal and frontal areas in the early phase of irony task, however the healthy controls exhibited higher activation in frontal, temporal and parietal regions in the latter phase of the irony task. Interestingly the linguistic help not only improved the patients' ToM performance, but it also evoked similar activation pattern to healthy controls.

Selected references

ALBERTI, G. 2011. ℜeA*LIS: Interpreters in the World, Worlds in the Interpreter.* Budapest. Akadémiai Kiadó • DENNETT D. 1989. *The intentional stance*. A Bradford Book • FRITH, C. D., & FRITH, U. 1999. Interacting minds – A biological basis. *Science*. 286(5445), 1692–1695 • FRITH, U., & FRITH, C. 2003. Development and

neurophysiology of mentalizing Philosophical Transactions: Biological Sciences, 358, 459–473 • GIORA, R. 2001. *Irony and its discontent*. Amsterdam. John Benjamins. 163-182 • HAPPÉ, F. G. 1993. Communicative competence and theory of mind in autism: A test of relevance theory. Cognition, 48(2), 101–119 • SPERBER, D., & WILSON, D. (1995). *Relevance*: Communication and cognition (2nd ed.) Oxford: Blackwell • SPERBER, D., & WILSON, D. (2002). Pragmatics, modularity and mind-reading. *Mind and Language*, 17(1&2), 3–23 • VARGA et al. 2013. Irony comprehension and context processing in schizophrenia during remission – A functional MRI study. *Brain and Language* 126. 231-242 • VARGA et al. 2014. Compensatory effect of general cognitive skills on non-literal language processing in schizophrenia. *Journal of Neurolinguistics* • WOODRUFF, G., & PREMACK, D. (1978). Does the chimpanzee have a theory of mind. *Behavioral and Brain Sciences* 1(4), 515–526.

Unusual and double plural markers in a representative sample of German pre-school children and adults

EUGEN ZARETSKY[1], BENJAMIN P. LANGE[2], KATRIN NEUMANN[3], HARALD A. EULER[3]
[1]Department of Phoniatrics and Pediatric Audiology, Goethe-University, Frankfurt/Main, Germany
[2]Institute of Medical Psychology and Medical Sociology, Center of Psychosocial Medicine, Faculty of Medicine, University of Göttingen, Germany
[3]Department of Phoniatrics and Pediatric Audiology, Ruhr-University, Bochum, Germany
yevgen.zaretsky@kgu.de, kontakt@benjaminplange.de, katrin.neumann@rub.de, euler@uni-kassel.de

Numerous studies on plural acquisition in German examined the distribution of plural allomorphs of the target language in the speech of children, mostly of pre-school age, and in rare cases in adult samples. However, there is hardly any information on infrequent phenomena of double plural markers and pseudosuffixes used as plural allomorphs. In this study, both children and adults were examined in this respect. 810 German children (age range 3;8-6;0 years, median=4;2) and 582 adults (18-96, median=24;0) were asked to produce plural forms of 8 nonce words taken from the validated language test *SETK 3-5* (Grimm 2001): *eine Ribane > Ribanen*, *ein Tulo > Tulos*, *eine Plarte > Plarten*, *ein Biwo > Biwos*, *ein Dolling > Dollinge*, *eine Kland > Klände*, *ein Ropf > Röpfe*, *eine Tapsel > Tapseln*. Indefinite articles allowed for consideration of dichotomized (feminine vs. non-feminine) grammatical gender. Children were examined by means of color pictures, adults had to fill out questionnaires.

Apart from 9,713 usual (single) plural markers, 175 double (and sometimes even triple) plural markers were documented: 70 in the answers of children and 105 in the answers of adults: *-se* (38), *-ner* (7), *-sen* (29), *-nen* (53), *-ser* (1), *-ers* (3), *-ne* (14), *-ren* (3), *-ns* (13), *-re* (5), *-res* (2), *-ses* (3), *-nse* (2), *-rne* (2). All plural markers with a double-digit frequency contained either *-(e)n* or *-s*, that is two most frequent plural allomorphs in our sample: *-s* (2,964 occurrences=30.0% of all plural markers), *-e* plus umlaut (864=8.7%), *-(e)n* (4,432=44.8%), umlaut (98=1.0%), *-e* (1,289=13.0%), *-er* (246=2.4%).

A categorical regression was conducted with the classification of double plural markers (*-ns*, *-re*, *-res* etc.) as the dependent variable and the following independent variables: classification of test subjects as children or adults, number of syllables in the singular forms of *SETK 3-5* items, expected (correct) plural allomorphs, dichotomized gender, application or non-application of the schwa deletion rule, and word final sounds (e.g., /n/, schwa, /f/): adjusted R^2=.65, the only important variable being expected plural allomorphs: β=1.37***. However, it should be mentioned that the extension of the same model to other *SETK 3-5* items (real words *Fisch*, *Vogel*, *Apfel*, *Schiff*, *Gabel*, *Buch*, *Stuhl*, *Bild*, *Hand*, *Glas*) explained much more variance (94%) with significant results for expected plural allomorphs (β=.73***), number of syllables (β=-1.32***), and word final sounds (β=.40***).

There was a highly significant difference in the distribution of double plural markers in the answers of children and adults: $\chi^2_{(13)}$=86.40***. The most obvious difference was a clear preference of adults for the plural marker *-nen* (47% of all double plural markers vs. 6% in the answers of children) and a clear preference of children for the plural marker *-se* (41% vs. 9%). Other tendencies were a high frequency of *-ne* in the answers of adults compared to children (12% vs. 1%) and a high frequency of *-ns* in the answers of children compared to adults (16% vs. 2%). Children often (in≥10% of all double plural markers) used markers *-se* (*Ropfse*), *-sen* (*Ropfsen*, *Plartesen*), and *-ns* (*Tulons*, *Plartens*, *Tapselns*), adults often used *-nen* (*Biwonen*, *Tulonen*), *-sen* (*Tulosen*, *Ribanesen*, *Biwosen*), and *-ne* (*Tulone*, *Biwone*). In the next step an

attempt was made to explain the preference for certain double plural markers in these most frequent double plural forms given in brackets.

In order to control which word final sounds are associated with which plural markers in the target language, a word frequency list was built based on a German language corpus (Leipzig University 2010). The list contains the 1,000 most frequent nouns of each of the three genders: masculine, feminine, and neuter (tokens). The schwa (*Plarte*, *Ribane*) occurs most frequently either with *-(e)n* or *-s*, /f/ (*Ropf*) either with *-e* (with or without umlaut) or *-s*, full vowels (*Tulo*, *Biwo*) either with *-s* or *-n*, /l/ (*Tapsel*) either with *-e* or *-(e)n* (however, in cases where the schwa deletion rule is applicable, *-e* is substituted by *-s*). Hence, the combinations of the two most expected plural markers (according to phonotactic rules) can account for the forms *Ropfse*, *Plartesen*, *Tulons*, *Plartens*, *Tapselns*, *Tulosen*, *Ribanesen*, *Biwosen*. According to the same frequency list, the sounds /s/ and /n/ in the word final position can be combined with either *-e* or *-(e)n*, hence the high frequency of forms *Ropfse(n)*, *Biwone(n)*, *Tulone(n)*.

The tendency of both children and adults to overgeneralize *-s* and *-(e)n* might be linked to the almost universal applicability of *-s* and a high frequency of *-(e)n*. For instance, in the same frequency list, *-s* is compatible with 21 word final sounds, *-e* with 16, *-(e)n* with 17, *-er* with 11, umlaut with 3, and *-e* plus umlaut with 12. Nouns requiring *-(e)n* account for 46% of the frequency list, followed by *-e* (29%), *-s* (9%), *-e* plus umlaut (9%), *-er* (6%), and umlaut (1%). Noteworthy is a low frequency of *-s* in comparison with our sample.

14 children and 2 adults used one plural marker which exists neither in the Modern High German nor in dialects: *Klande**l***, *Röfpe**l***. Furthermore, 52 children and one adult tended to delete *-el* and thus to reduce nouns to what they probably believed to be plural forms: *Tapsel* > *Taps**e***, *Taps**en***, *Taps**s***, where the word final sounds look like plural markers. This uncertainty about the link between the pseudosuffix *-el* and pluralization might be traced back to the high frequency of nouns ending in *-el* and receiving umlaut – the most seldom and the least iconic German plural allomorph – as plural marker, which might have led some children to misinterpret *-el* as the plural marker. Furthermore, because innumerable German nouns sometimes receive *-el* in its diminutive meaning and lose it again depending on context, some children might have internalized schemata making this element removable (cf. 6 occurrences of another diminutive suffix *-i(e)/y* only with the item *Tapsel* > *Tapsi*).

Hence, the preference of both children and adults for double plural markers containing *-s* and *-(e)n* can be attributed to an almost universal compatibility of *-s* with word final sounds and a very high frequency of *-(e)n*. Another obvious regularity was found in combinations of two most expected plural markers allowed by the rules of phonotactics. Although both *-s* and *-n* can function as linking elements in compounds like *Arbeit**s**recht* and *Karte**n**haus*, it can hardly be assumed that children or adults misused them as linking elements between stems and second plural markers in the words like *Tulonen*. Rather, once again, the rules of phonotactics prompted a higher compatibility of /n/ in the word final position with the plural marker *-(e)n*, which was demonstrated on the basis of a frequency list. Interestingly, both children and adults tended to add double plural markers to the four items following simplest German pluralization rules (*-s* after full vowels, *-n* after schwa), probably as an expression of creativity in linguistically advanced groups. There is some uncertainty as far as the link between the pseudosuffix *-el* and pluralization is concerned. The most important factor predicting the distribution of double plural markers in nonce words is the correct plural allomorph which occurs more often in such combinations than other allomorphs. If real words are also taken into account, number of syllables and word final sounds are also of importance.

References

GRIMM, H. 2001. SETK 3-5. Sprachentwicklungstest für drei- bis fünfjährige Kinder. Göttingen, Germany: Hogrefe.

LEIPZIG UNIVERSITY, DEPARTMENT FOR NATURAL LANGUAGE PROCESSING. 2010. Wortschatz, http://corpora.uni-leipzig.de/download.html.

WORKSHOP 1:

Modelling compound properties - new approaches and new explanations

Organizers:

Sabine Arndt-Lappe (Düsseldorf)
Melanie J. Bell (Cambridge, UK)
Martin Schäfer (Jena)
Barbara Schlücker (Berlin / Konstanz)

From *bread baskets* to *muffin tops*: What can we learn about the mental lexicon from lexical creativity?

RÉKA BENCZES
Eötvös Loránd University & Indiana University
rbenczes@gmail.com

One of the most remarkable features of linguistic creativity in English is the predominance of metaphorical and metonymical compounds in everyday language. This is immediately apparent from the vast number of examples that can be found in dictionaries. One such example is *belly button*, denoting the 'navel', coined in 1934 (*Oxford English Dictionary*; henceforth *OED*). *Belly*, however, fell from grace in the Victorian era and became a taboo word; the wonderfully apt *bread basket* and *pudding house* were coined instead for the same concept (Burridge 2004: 42). Needless to say, both expressions are metaphor- and metonymy-based: 1) an image schematic metaphor establishes the image-based resemblance between the shape of a bread basket or pudding house (the latter of which is a round container for puddings) and the belly; and 2) a CAUSE FOR EFFECT metonymy provides further motivation by creating a mapping between the thing that is eaten (bread and pudding, respectively) and the result (a bulging belly).

The ongoing appeal of metaphor- and metonymy-based compounds is also demonstrated by the more recently formed *muffin top*, which can be considered as one of the success stories of English word-formation. It was originally coined in 2003 by two Australian TV presenters to denote the spare flesh that overhangs loose-fitting jeans. In 2006, it was named as "Word of the Year" by the Australian *Macquarie Dictionary* and was also elected among the "most creative" terms the same year by the American Dialect Society. It eventually entered the online edition of the *OED* in March 2011.

Nevertheless, despite the preponderance of metaphorical and metonymical compounds in English, not much has been said about them in morphological literature on the grounds that they are not based on productive word-formation processes due to their "exocentric" nature. Apparently, in exocentric (or headless) compounds the head "falls outside" of the expression (hence the term "exocentric"), and the compound expression is not a hyponym of the head element. Such coinages are distinguished from endocentric ones, where the compound represents a subcategorization of the entity expressed by the head element (e.g., *apple tree* is a type of tree).

The endocentric–exocentric distinction has had a profound – and rather negative – effect on the scope of morphological research into compounding in the sense that linguistic literature has a strong tendency to focus on exocentric combinations only peripherally (if they are mentioned at all). This fact is also underlined by Scalise and Guevara (2006: 185): "In fact, while there is an extensive literature on endocentric compounding, the references to the theoretical and/or typological treatment of exocentric compounds are *very rare*" (emphasis added). Although descriptivist approaches do make reference to exocentric combinations, these are typically more superficial than the detailed classifications of endocentric compounds (see, for example, Jespersen 1954; Adams 1973). An exception to this trend is Marchand (1960), who devotes a whole chapter to the classification of exocentric compounds. As regards to the transformational generativist account, it left the issue of metaphorical and metonymical compounds untouched, probably for the simple reason that the theoretical framework was unable to accommodate such combinations.

Cognitive linguistics, however, has demonstrated that these "exocentric" compounds are indeed analysable with the application of conceptual metaphor and metonymy and blending among others. As emphasized by Langacker (1987), Talmy (1988) and Croft and Cruse (2004), both metaphor and metonymy can be considered as a type of construal operation, and

as such, a certain way of interpreting/conceptualizing the world around us. What this implies, therefore, is that the use of and reliance on conceptual metaphors and metonymies in word formation must also be an absolutely natural process. Benczes (2006) has extensively shown that such expressions can be analyzed remarkably well within a cognitive linguistic framework. The use of metaphors and metonymies in novel compound formation opens up a limitless supply of innovation and creativity in novel word-formation, as such expressions make use of the creative associations that exist between concepts; associations based on similarity, analogy or contiguity.

However, how can all this be applied to a better understanding of the mental lexicon? In line within the cognitive linguistic framework, and by analyzing numerous examples, the talk will attempt to bring together the semantic network approach of Langacker (1987) with more recent psycholinguistic results on compound representation (El-Bialy, Gagné & Spalding 2013; Gagné 2009: 265; Ji, Gagné & Spalding, 2011; Libben 2006). As it will be demonstrated throughout the presentation, creative language use involves ad hoc, contextual and analogical information that is "rooted in long-term memory or the immediate physical or linguistic context" (Gerrig & Gibbs 1988: 14), and any account of the mental lexicon must necessarily be able to explain and accommodate such examples of everyday creativity.

References

ADAMS, V. 1973. *An Introduction to Modern English Word Formation*. London: Longman.

BENCZES, R. 2006. *Creative Compounding in English: The Semantics of Metaphorical and Metonymical Noun–Noun Combinations*. Amsterdam & Philadelphia: John Benjamins.

BURRIDGE, K. 2004. *Blooming English: Observations on the roots, cultivation and hybrids of the English language*. Cambridge: CUP.

CROFT, W. & D.A. CRUSE. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

EL-BIALY, R., Ch.L. GAGNÉ & Th.L. SPALDING. 2013. Processing of English compounds is sensitive to the constituents' semantic transparency. *Mental Lexicon* 8/1: 75–95.

GAGNÉ, Ch.L. 2009. Psycholinguistic perspectives. In: R. Lieber & P. Štekauer (eds.), *The Oxford Handbook of Compounding*. Oxford: OUP, 253–71.

GERRIG, R.J & R.W. GIBBS. 1988. Beyond the lexicon: Creativity in language production. *Metaphor and Symbolic Activity* 3/1: 1–19.

JESPERSEN, O. 1954. *A Modern English Grammar on Historical Principles. Part VI: Morphology*. London: Bradford and Dickens.

JI, H., Ch.L. GAGNÉ & Th.L. SPALDING. 2011. Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language* 65: 406–30.

LANGACKER, R.W. 1987. *Foundations of Cognitive Grammar, Vol I*. Stanford, CA: Stanford UP.

LIBBEN, G. 2006. Why Study Compound Processing? An overview of the issues. In: G. Libben & G. Jarema (eds.), *The Representation and Processing of Compound Words*. Oxford: OUP, 1–22.

MARCHAND, H. 1960. *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach*. Wiesbaden: Otto Harrassowitz.

Oxford English Dictionary, The. 1989. 2nd ed., online version. Oxford: Oxford University Press.

SCALISE, S. & E. GUEVARA. 2006. Exocentric compounding in a typological framework. *Lingue e Linguaggio* 2: 185–206.

TALMY, L. 1988. The relation of grammar to cognition. In: B. Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*. Amsterdam & Philadelphia: Benjamins, 165–205.

Compounds in German – Influences on prosodic constituency

PIA BERGMANN
University of Freiburg – German department
pia.bergmann@germanistik.uni-freiburg.de

In prosodic morphology, compounds in German are traditionally mapped onto two prosodic words that together recursively form another prosodic word (henceforth: pword). Accordingly, the compound *Zahnarzt* 'dentist' (lit. 'tooth doctor') is analyzed as $((zahn)_{\omega}(arzt)_{\omega})_{\omega}$. Evidence for this analysis stems from the fact that the syllable related law of onset maximization is blocked by the intervening pword boundary: Instead of [tsaː.naɐtst], which we would expect according to onset maximization, we yield [tsaːn.ʔaɐtst]. Syllable boundaries thus coincide with pword-boundaries. Additionally, glottal stop insertion strengthens the initial boundary of the second pword. Although this view on prosodic constituency of morphological compounds is rather uncontroversial (cf. Hall 1999, Raffelsiefen 2000, Wiese 2000), two points of criticism may be raised. The first one refers to the fact that the analyses are mostly based on introspectional data only. There are hardly any accounts of how compounds are actually realized in spoken German, and whether the phonological and phonetic evidence for pword status – like blocking of resyllabification and glottal stop insertion – is really existent, and if so, whether there are gradient differences in boundary strength. The second point of criticism relates to the first one by focussing on factors that may systematically influence the phonetic realization of compounds. While it is hardly ever denied that factors like speech rate or speech style may have an influence on acoustic reductions at the pword boundary, other factors like frequency or morphological informativeness are only beginning to being looked upon (but see e.g. Pluymaekers et al. 2005, 2010 for Dutch and Ernestus & Hanique 2012 for a critical overview).

Against this background, the purpose of the present talk is twofold: Firstly, it will shed some light on the acoustic-phonetic realization of the pword boundary in German compounds, and secondly, it will discuss the effect of word frequency and other independent variables on pword boundary strength. The results to be presented stem from two types of data:

(1) In a carefully designed acoustic production study, twelve north-western German speakers produced target items in a laboratory setting. All targets were embedded into carrier sentences. The boundary phenomena of interest are glottal stop insertion / glottalization and degemination across the pword boundary. The independent variables are word frequency, accentuation, vowel quantity and segmental context. For instance, glottal stop and/or glottalization are investigated in word pairs like high-frequency *Zahnarzt* ('dentist') vs. low-frequency *Bahnarzt* ('doctor of the railway') (n = 756); degemination is investigated in word pairs like high-frequency *Stadttour* ('city tour') vs. low-frequency *Watttour* ('tour of the tidelands') (n = 896). The results indicate that prosodic boundary strength in compounds is indeed systematically influenced by frequency, but structural factors like vowel quantity and segmental context cannot be dismissed either.

(2) An ongoing corpus study of spontaneous speech investigates durational reduction at the pword boundary in compounds with /l/-initial second constituents (e.g. *Urlaubslohn* 'holiday pay') which are then compared to derivations with *–los* (e.g. *arbeitslos* 'unemployed') (n = 414). From a prosodic point of view, both types of complex words are similar in that they consist of two pwords: $((arbeits)_{\omega}(los)_{\omega})_{\omega}$. Despite this similarity, however, it may be hypothesized that durational reduction at the pword boundary varies systematically with other factors. The independent variables of primary interest are morphological structure

(compounding vs. derivation), semantic transparency and absolute word frequency. Preliminary results suggest that semantic transparency does not play a role for boundary realization in complex words, whereas morphological structure and word frequency lead do systematically different realizations.

References

HALL, A.T. 1999. The Phonological Word: A Review. In: A. T. Hall and U. Kleinhenz (eds.), *Studies on the Phonological Word*. Amsterdam: John Benjamins, 1-22.

HANIQUE, I., ERNESTUS, M. 2012. The role of morphology in acoustic reduction. *Lingue e linguaggio* 2, 147-164.

PLUYMAEKERS, M., ERNESTUS, M. AND BAAYEN, R.H. 2005. Lexical Frequency and Acoustic Reduction in Spoken Dutch. *Acoustical Society of America Journal* 118, 2561-2569.

PLUYMAEKERS, M., ERNESTUS, M., BAAYEN, H.R. AND BOOIJ, G. 2010. Morphological effects on fine phonetic detail: The case of Dutch -igheid. In: C. Fougeron, B. Kühnert, M. D'Imperio and N. Vallée (eds.): *Laboratory Phonology*. Berlin / New York:de Gruyter, 511-531.

RAFFELSIEFEN, R. 2000. Evidence for word-internal phonological words in German. In: R. Thieroff, M. Tamrat, N. Fuhrhop and O. Teuber (eds.): *Deutsche Grammatik in Theorie und Praxis*. Tübingen:Max Niemeyer Verlag, 43-56.

WIESE, R. 2000. *The phonology of German.* Oxford: Oxford University Press.

Relational competition and meaning construction in compound interpretation and production

CHRISTINA L. GAGNÉ & THOMAS L. SPALDING
Department of Psychology, University of Alberta
cgagne@ualberta.ca spalding@ualberta.ca

The available evidence on the processing of morphologically complex words is consistent with a lexical system that is highly productive and actives both whole-word and constituent representations (Ji, Gagné, & Spalding, 2011; Kuperman, Bertram, & Baayen 2010; Libben 2006). Such a system requires that compound processing involve the integration of constituent representations. Thus, it is critical to understand what this integration entails. Several studies indicate that integration involves meaning construction (e.g., Fiorentino & Poeppel 2007) and the use of semantic relations to link the constituents (Gagné & Spalding 2009; Ji, et al. 2011). Of course, compounds vary in terms of the extent to which the constituents' meanings contribute to the meaning of the compound, and there is ongoing debate about whether the semantic representations of the constituents of opaque compounds are available during compound processing. Even so, there is some recent evidence that even opaque compounds are affected by relational competition (Gagné & Spalding, 2014; Spalding & Gagné, 2014), suggesting that meaning construction is attempted even for opaque compounds.

In this talk, we present studies that examine whether semantically priming the first constituent of a compound influences the processing of that compound, and whether any such priming effect depends on the semantic transparency of the compound's constituents. We investigate this question using a lexical decision task to evaluate interpretation, where the primary evidence for meaning construction exists. We also use a typing task to evaluate written production, where one might believe that meaning construction would be unnecessary. We find that semantic priming affects processing in both the interpretation and production tasks, but that such effects are complicated. Semantic transparency of the constituents affects whether semantic priming results in changes to processing, but it is not just the semantic transparency of the primed constituent that exerts an influence—for example, the semantic transparency of the head affects whether semantically priming the modifier results in a change in lexical decision latencies or typing times. In fact, the semantic transparency of the primed constituent seems not to matter in the typing task, and matters only in conjunction with the semantic transparency of the head in the lexical decision task.

Overall, the results are consistent with a meaning construction process that gives rise to competing meanings during both interpretation and production of compound words. If meaning construction is occurring, then competition among the interpretations might offset benefits due to increased availability of lexical and morphemic information (see, for example, El-Bialy et al. 2013; Ji et al. 2011). That is, because constructed meanings compete, manipulations that increase the availability of a constituent's semantic, lexical, or morphological representations may either facilitate or inhibit processing, depending on the other properties of the item being processed and on the properties of the processing task. A better understanding of the integration process might help resolve some of the conflicting results in the literature, especially those concerning the availability of morphemic and semantic representations of the constituents.

References

EL-BIALY, R., GAGNÉ, C. L. & SPALDING, T. L. (2013). Processing of English compounds is sensitive to the constituents' semantic transparency. *The Mental Lexicon, 8,* 75-95.

FIORENTINO, R. & POEPPEL, D. (2007). Compound words and structure in the lexicon. *Language and Cognitive Processes*, *22*, 953-1000.

GAGNÉ, C. L. & SPALDING, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language*, *60*, 20-35.

GAGNÉ, C. L., & SPALDING, T. L. (2014). Relation diversity and ease of processing for opaque and transparent English compounds. In F. Rainer, Dressler, W., Gardini, F, Luschützky, HC (Eds.). *Morphology and Meaning: Current Issues in Linguistic Theory.* Amsterdam: John Benjamins.

JI, H., GAGNÉ, C. L., & SPALDING, T. L. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language, 65*, 406-430.

KUPERMAN, V., BERTRAM, R., & BAAYEN, R. H. (2010). Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language*, *62*, 83-97.

LIBBEN, G. (2006). Why study compound processing: An overview of the issues. In G. Libben & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 1-22). Oxford, New York.

SPALDING, T. L., & GAGNÉ, C. L. (2014). Relational diversity affects ease of processing even for opaque English compounds. *The Mental Lexicon, 9,* 48-66.

From genitive suffix to linking element in German

KRISTIN KOPF
Johannes Gutenberg University Mainz
kristin.kopf@uni-mainz.de

Nominal compounds in German can have one of two structures: In the first one, similar to English, the two nouns are simply combined (*Tisch+decke* 'tablecloth'). The second makes use of intervening linking elements (*Geburt+**s**+ort* 'birth-LE-place', *Akte-**n**-ordner* 'file-LE-folder'), a phenomenon that can also be observed in several other Germanic languages (e.g. Dutch, Luxemburgish).

The origin of most German linking elements is generally considered to be syntactic in nature (cf. Pavlov 1983, Nitta 1987, Demske 2001). Due to the frequent preposing of genitive attributes in Early New High German (1350–1600), reanalysis of a genitive attribute and its head lead to a new, compound-like structure:

(1)	*des*	*Leib-s*	*Erbe-n*	>	*die*	*Leib-*	*s-*	*erbe-n*	'legitimate heirs'
	DEF.GEN.SG	body-GEN	heir-PL		DEF.NOM.PL	body	LE	heir-PL	

By analogy, these fossilized genitive suffixes spread to new formations or existing compounds in which a former syntactic origin is no longer indicated due to specific morphological or semantic properties. This resulted in a new pattern of compounding that quickly spread.

German compounds with linking elements are a widely studied aspect of word formation. However, thorough quantitative corpus studies that focus on the genesis and expansion of this pattern are still lacking. The present study aims to fill this gap by addressing the following questions:

1. At which point can we assume a new pattern of compound formation as opposed to singular, lexeme specific instances of reanalysis?
2. Can the ambiguous structures leading to reanalysis be quantified?
3. Was the expansion and current distribution of German linking elements guided by morphological, phonological or semantic factors and how did these interact?

The study is based on data from a corpus of Early New High German and New High German with texts ranging from 1500 up 1710. It is a strongly modified version of the Bergmann & Nerius (1996) corpus. The corpus is divided in 8 periods, 5 regions and two text types, making a detailed analysis of a changing system possible.

This diachronic data is essential for the interpretation of today's system. About 39% of today's newly formed compounds make use of linking elements (see Figure 1). A number of theories have been proposed to explain the persistance of this seemingly useless phonological material (Aronoff & Fuhrhop 2002, Wegener 2003, Krott et al. 2004, Berg 2006, Nübling & Szczepaniak 2008). The present study seeks to shed light on the role the proposed morphological, phonological or semantic functions have played in the emergence of linking elements.

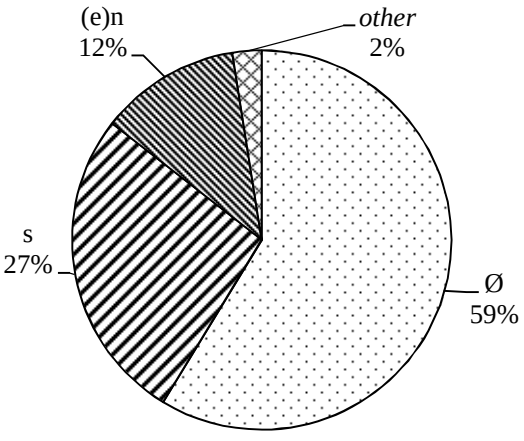


Figure 1: Distribution of linking elements in recent German neologisms (2009–2010, n=1,754; own study).

References

Aronoff, Mark & Fuhrhop Nanna (2002): *Restricting suffix combinations in German and English: Closing suffixes and the monosuffix constraint*. In: Natural Language & Linguistic Theory 20, 451–490. ● Berg, Thomas (2006): *The internal structure of four-noun compounds in English and German*. In: Corpus Linguistics and Linguistic Theory 2, 197–231. ● Bergmann, Rolf/Nerius, Dieter (eds., 1996): *Die Entwicklung der Großschreibung im Deutschen von 1500 bis 1700*. Heidelberg. ● Demske, Ulrike (2001): *Merkmale und Relationen. Diachrone Studien zur Nominalphrase des Deutschen*. Berlin, New York. ● Krott, Andrea & Hagoort, Peter & Baayen, Harald (2004): *Sublexical Units and Supralexical Combinatorics in the Processing of Interfixed Dutch Compounds*. In: Language and Cognitive Processes 19.3, 453–471. ● Nitta, Haruo (1987) : *Zur Erforschung der 'uneigentlichen' Zusammensetzungen im Frühneuhochdeutschen*. In: ZfdPh 106. 400–416. ● Nübling, Damaris & Szczepaniak, Renata (2008): *On the way from morphology to phonology: German linking elements and the role of the phonological word*. In: Morphology 18, 1–25. ● Pavlov, Vladimir M. (1983): *Von der Wortgruppe zur substantivischen Zusammensetzung*. Zur Ausbildung der Norm der deutschen Literatursprache (1470–1730), Vol. 4. Berlin. ● Wegener, Heide (2003): *Entstehung und Funktion der Fugenelemente im Deutschen, oder: warum wir keine *Autosbahn haben*. In: Linguistische Berichte 196, 425–457.

Branching direction and phonetic reduction in triconstituent compounds

GERO KUNTER & INGO PLAG
Heinrich-Heine-Universität Düsseldorf
gero.kunter@uni-duesseldorf.de, ingo.plag@uni-duesseldorf.de

Studies investigating the order of derivational affixes in English (e.g. Plag & Baayen 2009) have provided evidence that affix boundaries differ in boundary strength: in a word of the form [[base-X]-Y], the outer boundary between [base-X] and affix Y is stronger than the inner boundary between the base and affix X. Hay (2003) and Plag & Baayen (2009) argue that these strength differences affect the degree not only of decomposability, but also of phonological integration: affixes at weaker boundaries will show a higher degree of phonological integration than affixes at stronger boundaries.

Similar boundary effects are reported in Sproat (1993) and Sproat & Fujimura (1993), who investigate the phonetic implementation of the same segment string at different types of boundary (no boundary, affix, compound, phrasal, utterance). Their findings are compatible with the conclusion that weaker boundaries show more phonological integration: for instance, the acoustic duration of the same segments at an affix boundary is shorter than at a compound boundary. Likewise, Hay (2007) finds a difference in duration between the prefix *un-* and non-morphemic word-initial *un-*, as well as differences in vowel reduction.

Taken together, the results from the phonetic and the affix order studies predict that the phonetic implementation of a word with more than two morphological constituents will reflect its morphological structure, i.e. its internal bracketing. We hypothesize that the embedded form shows more phonetic reduction than forms at higher derivational levels ('Embedded Reduction Hypothesis').

To test this prediction, we analyzed triconstituent compounds. Traditionally, the branching direction of triconstituent compounds (e.g. [*child care*] *center* vs. *university* [*textbook*]) has been extensively discussed in the context of stress assignment. Recent studies (for example, Kösling 2013 and Kösling et al. 2014) have shown, that, contrary to what has been assumed in much of the earlier literature (in the wake of Liberman and Prince 1977), the branching direction of noun-noun-noun compounds cannot be read off from the stress pattern.

In the present study, we use the data from Kösling et al.'s experiment to explore the prediction of the Embedded Reduction Hypothesis that the embedded constituents show more reduction than the third, non-embedded, constituent. The statistical analysis confirms that prediction. The constituents of the embedded compound were significantly shorter than the single constituent at the higher compositional level. This effect holds for left-branching and right-branching compounds. The results present clear evidence that the branching direction, i.e. morphological structure, is phonetically encoded.

This poses a challenge for most formulations of Lexical Phonology (e.g. Kiparsky 1982), which argue that the internal structure of morphologically complex forms is not accessible anymore at the post-lexical stage, with which phonetic variation like sub-phonemic durational differences or phonetic reduction are usually associated. If there is reliable evidence that the acoustic signal contains phonetic detail signaling the internal structure of morphologically complex words, and if this detail affects the processing of the acoustic signal by listeners, the strict division between lexical and post-lexical components needs to be revised.

References

HAY, JENNIFER B. 2003. *Causes and consequences of word structure*. London: Routledge.

HAY, JENNIFER B. 2007. The phonetics of 'un'. In Judith Munat (ed.), *Lexical creativity, texts and contexts*. Amsterdam: John Benjamins. 39-59.

KÖSLING, KRISTINA. 2013. *Prominence assignment in English triconstituent compounds*. PhD thesis. Universität Siegen. Available at http://dokumentix.ub.uni-siegen.de/opus/volltexte/2013/674/.

KÖSLING, KRISTINA, GERO KUNTER, R. HARALD BAAYEN & INGO PLAG. 2013. Modelling pitch contours in English tripartite compounds. *Language and Speech*. DOI: 10.1177/0023830913478914

KIPARSKY, PAUL. 1982. From cyclic phonology to lexical phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations*. Dortrecht: Foris Publications. 131-175.

LIBERMAN, MARK Y. & ALAN PRINCE. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8(2): 249-336.

PLAG, INGO & R. HARALD BAAYEN. 2009. Suffix Ordering and Morphological Processing. *Language* 85(1): 109-152.

SPROAT, RICHARD. 1993. Looking into words. In Sharon Hargus & Ellen M. Kaisse (eds.), *Studies in lexical phonology*. San Diego, CA: San Diego Academic Press. 173-195.

SPROAT, RICHARD & OSAMU FUJIMURA. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21: 291-311.

Individual differences in compound processing

GARY LIBBEN
Brock University
gary.libben@gmail.com

In the psycholinguistic literature, a great deal of progress has been made in the understanding of how compound constituents contribute to the overall meaning of compound words and the understanding of the conditions under which compound constituents are activated in lexical processing (e.g., Spalding, Gagné, Mullaly & Ji. 2010; Marelli & Luzzatti, 2012). Traditionally, however, matters of individual variation have received relatively less attention. In this paper, a psycholinguistic model of the mental representation of compounds is presented that highlights the extent to which individuals and populations can differ in terms of how compounds are represented and processed. The model has two key components:

(1) Morphological Transcendence: This is the claim that the representation of compound words in the mind changes through the lifespan as a result of the experience that an individual language user has with processing specific compounds and morphological families.

(2) Morphological Superstates: This is the claim that compound constituents exist psychologically in a morphological superstate up to the point at which they are measurable through acts of language production or comprehension.

The postulation of morphological transcendence and morphological superstates predicts that the constituent activation effects that can be measured through psycholinguistic experiments will depend on the compound processing experience of the individual language user and the specific performance demands of the experimental paradigm. These predictions were tested through a series of experiments with native speakers of English.

Experiment 1 employed the P3 experimental paradigm (Libben, Weber, & Miwa, 2012) in which participants are visually presented with compound stimuli in a Primed Progressive Demasking paradigm. The task of the participants is to say the stimulus word aloud as quickly as possible and then to type it. It was found that individual participant characteristics including education, linguistic background and metalinguistic knowledge affected performance on both tasks. Moreover, the typing production data revealed a substantial effect of task performance characteristics as predicted by the morphological superstate hypothesis.

In Experiment 2, the P3 paradigm was employed in dyadic word recognition and production paradigm in which two participants are tested at the same time. One member of the dyad sees progressively demasked stimuli and says them aloud. The other participant types the compound stimuli (as one would in a classic dictation task). Results show that performance on the task is not only influenced by the characteristics of individual participants and the disparity between the two participants in the dyad. A critical difference between these two variables is that the former is plainly evident to participants in the experiment at the time of testing, whereas the latter is not. We introduce a new data visualization technique, Participant Profiles, that facilitates the understanding of the relation between individual and dyadic participant characteristics as well as their relation to patterns of compound processing.

References

LIBBEN, G., WEBER, S., & MIWA, K. (2012). P3: A technique for the study of perception, production, and participant properties. *The Mental Lexicon,* 7:2, 237-248.

MARELLI, M., & LUZZATTI, C.G. (2012). Frequency Effects in the Processing of Italian Nominal Compounds: Modulation of Headedness and Semantic Transparency. Journal of Memory and Language, 66, 644-664.

SPALDING, T., GAGNÉ, C. MULLALY A.,& JI, H. 2010. Relation-based interpretations of noun–noun phrases: A new theoretical approach. In Susan Olsen (ed.), New impulses in word-formation (Linguistische Berichte Sonderheft 17), 283–315. Hamburg: Buske.

A model of present-day English compound spelling

CHRISTINA SANCHEZ-STOCKHAMMER
University of Erlangen-Nuremberg
christina.sanchez@fau.de

A question many language users are frequently faced with when spelling English compounds is whether to use open spelling (*drinking fountain*), hyphenation (*far-off*) or solid spelling (*airport*). In the literature, two views dominate:

1. The spelling of English compounds is chaotic (e.g. Fowler 1926: 243; Merriam-Webster 2001: 99).
2. The spelling of English compounds develops from an open via a hyphenated stage towards solid spelling (e.g. Quirk et al. 1985: 1537).

These two views are not necessarily mutually exclusive, since a different speed of development of individual compounds may result in what looks like superficial chaos. A large-scale empirical study was conducted in order to show that neither of these general views is entirely correct and that there are determinants of present-day English compound spelling.

In my paper, I am going to present the central points of that study, which goes beyond previous research on the spelling of English compounds (e.g. Sepp 2006; Rakić 2009; Kuperman and Bertram 2013), since is not limited to noun+noun compounds and uses a wider compound concept. Compounds are defined as complex lexemes

a) which consist of at least two constituents occurring as free synchronically recognisable and semantically relevant lexemes each and
b) which contain no affixation on the highest structural level.

The study uses various types of linguistic evidence to test over 60 hypotheses concerning the spelling of English compounds. Among other things, it analyses

- lemma lists from six dictionaries, one of which served for the compilation of the Master Compound List comprising more than 10,000 compounds.
- corpus data from analogously structured corpora for different periods and varieties of English (BLOB-1931, LOB, FLOB, BE2006 for British English; Brown and FROWN for American English).
- information on the first spelling recorded in the *Oxford English Dictionary*.

In contrast to previous research, which has mainly focused on variation in compound spelling, the study that I am going to present in my paper follows an approach that is based on the observation that the spelling of a number of English compounds is largely standardised and does not seem to vary. My study establishes empirically which out of a large number of features (such as word length, stress pattern etc.) are responsible for this special behaviour. The variables with a statistically significant effect are then used as predictors for the spelling of other compounds. Not only do the results that I am going to present in my paper call into question the general views of English compound spelling outlined above: in addition, they were used to develop various spelling algorithms with different degrees of complexity. The

predictive accuracy of the decision trees was tested on compounds belonging to different samples:

1. compounds with exclusively open, hyphenated or solid spelling in 5-6 dictionaries.
2. compounds with spelling variation across dictionaries – but which exhibit a preference for one particular spelling variant.
3. compounds from CompText, an 8,864-word corpus of present-day British English compiled specifically for the study.

I am going to show in my paper that even a minimal algorithm performs better than educated English native speakers in a decision test on the most prototypical English compounds. This seems to suggest that while it is indeed possible to recognise principles underlying the present-day spelling of English compounds, these are not immediately obvious.

The theoretical analysis of the data discusses the application of different theoretical models to the phenomenon of English compound spelling. My paper presents an overview of the most important aspects and reaches the conclusion that both a modified prototype model (cf. e.g. Rosch 1973) and a constructionist account (cf. e.g. Goldberg 2006; Stefanowitsch and Gries 2003) are suitable for the modelling of the data. In addition, analogy (cf. e.g. Bybee 2003; Krott, Baayen and Schreuder 2001; Plag, Kunter and Lappe 2007), which is frequently contrasted with rule-based principles, can be reconciled with my feature-based approach by taking into account that the features used have actually emerged from a large set of exemplars (cf. also Bell and Plag 2012: 517).

References

BELL, MELANIE, and INGO PLAG (2012): "Informativeness is a determinant of compound stress in English." *Journal of Linguistics* 48/03, 485-520.

BYBEE, JOAN (2003): *Phonology and Language Use*. 2nd ed. Cambridge: CUP.

FOWLER, HENRY W. (1926): *A Dictionary of Modern English Usage*. Oxford: Clarendon.

GOLDBERG, ADELE (2006): *Constructions at Work: The Nature of Generalization in Language.* Oxford: OUP.

KROTT, ANDREA, HARALD BAAYEN and ROBERT SCHREUDER (2001): "Analogy in morphology: Modeling the choice of linking morphemes in Dutch." In: *Linguistics* 39/1, 51-93.

KUPERMAN, VICTOR and RAYMOND BERTRAM (2013): "Moving spaces: Spelling alternation in English noun-noun compounds." In: *Language and Cognitive Processes* 28/07, 939-966.

MERRIAM-WEBSTER (2001): *Merriam-Webster's Guide to Punctuation and Style*. 2nd ed. Springfield, Massachusetts: Merriam-Webster.

Oxford English Dictionary (2009). 2nd ed. on CD-Rom (v. 4.0.0.3). Oxford: OUP.

PLAG, INGO, GERO KUNTER and SABINE LAPPE (2007): "Testing hypotheses about compound stress assignment in English: A corpus-based investigation." In: *Corpus Linguistics and Linguistic Theory* 3/2, 199-233.

QUIRK, RANDOLPH et al. (1985): *A Comprehensive Grammar of the English Language*. London: Longman.

RAKIĆ, STANIMIR (2009): "Some observations on the structure, type frequencies and spelling of English compounds." In: *SKASE Journal of Theoretical Linguistics 6*. http://www.skase.sk/Volumes/JTL13/pdf_doc/04.pdf.

ROSCH, ELEANOR (1973): "On the internal structure of perceptual and semantic categories." In: Moore, Timothy E. (ed.): *Cognitive Development and the Acquisition of Language.* New York: Academic Press, 111-144.

SEPP, MARY (2006). *Phonological Constraints and Free Variation in Compounding: A Corpus Study of English and Estonian Noun Compounds.* American Dissertation.

STEFANOWITSCH, ANATOL and STEFAN TH. GRIES (2003): "Collostructions: Investigating the interactions of words and constructions." In: *International Journal of Corpus Linguistics* 8, 209-243.

The polysemy of Italian VN compounds in a diachronic perspective

PAVEL ŠTICHAUER
Charles University in Prague
pavel.stichauer@ff.cuni.cz

The VN compounding (of the type *portalettere*) in Italian (and, in general, in the Romance languages) is a well-studied and well-described phenomenon (cf., e.g., Tollemache 1945; Scalise 1994: 134-136; Bisetto 2004: 45-47; Scalise - Bisetto 2008: 133-135; Dardano 2009: 190-198; Ricca 2010, among others) – to the extent that one may question the necessity to deal with it in another study. The structural patterns (e.g., the nature of verb form) and also the semantic outputs (agentive, instrumental, eventive, etc.) are effectively very well known.

However, recent research has shown that at least two questions are worth being further developed. First, it is the quantitative aspect of the productivity and of the overall distribution of the VN compounds, which is a question revisited especially thanks to the development of large language corpora (cf. Ricca 2010). Second, it is the diachronic development of the pattern in question. Even though it is generally known that VN compounds represent a morphological innovation which is common to virtually all Romance languages (cf. Bauer 2010), a precise quantitative – diachronically based description – is still missing (but cf. Moyna 2011 for a diachronic treatment of Spanish compounding). Thus, for example, the question of the limited number of verb bases which enter the pattern (*porta-, salva-, mangia-*, ...), as well as the question of the historical coverage of the meanings (e.g., is the agentive meaning the primary one?), calls for a thorough elaboration.

In fact, beside the empirical results which are obviously very important, a theoretical question arises as well, namely the diachronic variability of word-formations rules (whatever definition is adopted – Aronoff 1976 vs. Corbin 1987 or some other new formulations...) and the associated semantic instructions of these WFRs (e.g., the status of the unitary output base hypothesis, cf. Ricca 2010).

Both questions – empirical and theoretical – are precisely what I wish to address in this paper, and I wish to analyse them along the following lines. As far as the empirical basis is concerned, I will present the diachronic data drawn from the corpus LIZ 4.0 within the time span that goes from the 16th to the 19th century (the 20th century being the period on which all the available synchronic studies are based and will thus be excluded from the survey), for each century separately; I shall also present the complementary data from the major diachronic dictionaries (from the 1612 *Vocabolario degli Accademici della Crusca* up to the 1859 Tommaseo's dictionary, along with some specialised dictionaries, cf. for a general lexicographic overview Marazzini 2009).

This empirical basis will reveal that the quantitative characteristics are well different for the 16th century as opposed to the present-day situation in that even within the class of the most frequent „leading bases" (Ricca 2010: 247; Gather 2001), the number of different types (and their token frequencies) is rather low (for *porta-* the 16th century corpus offers just *portaflagello, portalettere, portanovelle, portaseggiette*). This indicates that even if the VN compounding has always been *available* (in Corbin's sense of *disponibilité*), the pattern has become fully *profitable* (again in the sense of *rentabilité*) only recently. At the same time, also the range of possible meanings is rather limited being there – in the time span in question – especially the agentive VN compounds.

This empirical situation thus leads directly to the theoretical question whether the semantic instruction associated with the WFR in question has always been identical regardless of any diachronic variability. In fact, as shown by Ricca (2005; 2010; cf. also Magni 2010), the VN compounds may not only denote agents, instruments, events, locations (though the quantitative representation of each type is not equal), but they also function – productively –

as adjectival modifiers (e.g., *cannone sparaneve*, *giro spaccagambe*). Ricca raises the theoretical question whether this adjectival output is part of the WFR; and he also alludes to the marginality of the above mentioned meanings in the history of Italian (claiming that event VN compounds are rare also in diachrony).

I maintain that both questions are closely connected, i.e. the diachronic empirical research could provide us with some concrete answers also to the theoretical problems regarding the development of the WFRs in general. I intend to put forward the hypothesis according to which there is a core semantic instruction tied up with the readily *available* pattern which is, at least as far as the VN compounds are concerned, diachronically constant (cf. for Spanish similar observations made by Moyna 2011: 210-211), and the increasing *profitability* could effectively exploit it or also leave it unexploited; at the same time, the modifier function is simply a contextual enrichment which is not part of the WFR (along the lines of Ricca 2010: 254).

The first claim is supported by the diachronic evidence (e.g., *salvadanaio* is an old instrument noun showing thus that also in diachrony the VN compounds were not exclusively agents; and much new evidence comes also from specialised historical dictionaries); the second claim is a little harder to support diachronically given the data scarcity in the period under investigation. Nevertheless, we shall present a couple of examples which go in this direction.

References

ARONOFF, M. 1976. *Word Formation in Generative Grammar*, Cambridge (MA): The MIT Press.

BISETTO, A. 2004. Composizione con elementi italiani. In: Grossmann, Maria e Rainer, Franz (a cura di), *La formazione delle parole in italiano*. Tübingen: Niemeyer.

BAUER, B. 2010. Word formation. In: Maiden, Martin and Smith, John Charles, Ledgeway, Adam (eds.), *The Cambridge History of the Romance Languages. Volume I. Structures.* Cambridge: Cambridge University Press, 2010, pp. 532–563.

CORBIN, D. 1987. *Morphologie dérivationnelle et structuration du lexique*, 2 voll., Tübingen: Niemeyer.

DARDANO, M. 2009. *Costruire parole. La morfologia derivativa dell'italiano*. Bologna: il Mulino.

GATHER, A. 2001. *Romanische Verb-Nomen Komposita. Wortbildung zwischen Lexikon, Morphologie und Syntax*. Tübingen: Narr.

IL TOMMASEO. 2004. *Dizionario della lingua italiana di Niccolò Tommaseo e Bernardo Bellini (1857-1879).* Testo integrale degli otto tomi originali e interrogazione per lemmi e a tutto testo in CD-ROM. Bologna: Zanichelli.

LIZ 4.0. 2001. *Letteratura italiana Zanichelli*. CD-ROM dei testi della letteratura italiana, a cura di Pasquale Stoppelli ed Eugenio Picchi. Bologna: Zanichelli.

MAGNI, E. 2010. From the periphery to the core of Romance [VN] compounds. *Lingue e Linguaggio*, 9, 1, 2010, pp. 3–39.

MARAZZINI, C. 2009. *L'ordine delle parole. Storia di vocabolari italiani.* Bologna: il Mulino.

MOYNA, M. I. 2011. *Compound Words in Spanish. Theory and History*. Amsterdam: John Benjamins.

RICCA, D. 2005. Al limite tra sintassi e morfologia: i composti aggettivali V-N nell'italiano contemporaneo. In: Grossmann, M. – Thornton, A. M. (a cura di). *La formazione delle parole. Atti del XXVII Congresso Internazionale di Studi della Società di Linguistica Italiana. L'Aquila, 25-27 settembre 2003*. Roma: Bulzoni, pp. 465–486.

RICCA, D. 2010. Corpus data and theoretical implications. With special reference to Italian V-N compounds. In: Scalise, Sergio and Vogel, Irene (eds.), *Cross-Disciplinary Issues in Compounding*. Amsterdam: John Benjamins, 2010, pp. 237–254.

SCALISE, S. 1994. *Morfologia*. Bologna: il Mulino.

SCALISE, S. & A. BISETTO 2008. *La struttura delle parole*. Bologna: il Mulino.

TOLLEMACHE, F. 1945. *Le parole composte nella lingua italiana*. Roma: Rores.

WORKSHOP 2:

Information-theoretic approaches to morphology

Organizer:

James P. Blevins (Cambridge)

Uncertainty and discriminability in information-theoretic WP

Jim Blevins
University of Cambridge
jpb39@cam.ac.uk

Michael Ramscar
Eberhard Karls University, Tübingen
ramscar@gmail.com

Classical WP models depart from prevailing symbol-processing paradigms in two fundamental respects. The first is that classical WP models aim to minimize the UNCERTAINTY that arises in associating units with grammatical properties, not to minimize the size of units, or the 'redundancy' of inventories. Associated with this aim is a conception of languages as 'complex systems' in which dimensions of variation do not function as individually meaningful 'signs' or even as 'realizations' but instead serve to DISCRIMINATE larger units that are meaningful within the system. This talk suggests how formalization in terms of information theory and discriminative learning theories clarifies these properties of classical WP models.

Uncertainty

The relevance of information theory to morphology generally is enhanced by the fact that the communicative function of morphological variation conforms well to the idealization of communication as "selecting from a set of possible messages" (Shannon 1948: 22). WP models are particularly amenable to formalization in information-theoretic terms given their emphasis on minimizing indeterminacy in the mapping between features and forms. The goal of uncertainty reduction dictates the choice of units and structures; neither words nor paradigms have a privileged status *a priori* in WP models. The abstraction of word-sized units and their assignment to larger paradigmatic structures is justified by the reduction in uncertainty that this achieves in a grammatical analysis. The word is the basic unit of morphological analysis because it is, on the one hand, a maximally discriminable "perceptual gestalt" Hockett (1987: 52), and, on the other, "a more stable and solid focus of grammatical relations than the component morpheme by itself" (Robins 1959: 128). The association between words and paradigm cells 'anchors' a form within a system of grammatical contrasts, facilitating implicational deductions that further constrain uncertainty within the system. Pushing analyses below the word level increases indeterminacy, not only by isolating units that are individually more ambiguous than the forms they are extracted from, but also by disrupting patterns that locate a form within an inflectional paradigm or derivational family.

The roles that words and paradigms play in reducing uncertainty are summarized in (1).

(1) a. The uncertainty that arises in relating word forms and grammatical properties is less than (or equal to, in the case of simple forms) the sum of the uncertainty that arises in associating their component morphs with properties.

b. The uncertainty associated with a paradigm is less than (or equal to, in the case of single-cell paradigms) the sum of the uncertainty of their component cells.

(1a) reformulates the claim that "grammatical statements … are more profitably abstracted from words as wholes than from individual morphemes" (Robins 1959: 128), whereas (1b) ex-

presses a version of the Low Conditional Entropy Conjecture of Ackerman & Malouf (2013). Both claims can be formalized in terms of standard entropy measures.

Discriminability

A discriminative perspective further clarifies the structure of WP models, and offers a principled account of morphological patterns that are unexpected or problematic on decompositional models, whether morphemic or realizational. Although classical WP models are often faulted for failing to recognize morphemic structure, the classical position reflects a discriminative rather than an associative interpretation of form variation.

> As with the rest of western Antiquity, Priscian's grammatical model is word and paradigm and he expressly denied any linguistic significance to divisions in what would now be called morphemic analysis below the word. (Robins 1997: 70)

The locus of form contrasts is clearly identified in analyses that specify what Matthews (1991: 191) terms "different formal changes by which oblique Cases were derived from the 'upright' Case". Yet this variation does not signify or even realize grammatical meaning but discriminates between larger meaningful units. Hence discriminative learning approaches (Ramscar et al. 2010) provide the most appropriate models for the formalization of the exponence patterns of classical WP accounts. A discriminative perspective also offers natural explanations for the fact that combinations of elements are often distinctive, that elements in isolation may differ from the 'same' element when it occurs within a larger form, as well as for the observation that no division or multiple divisions of a form may be motivated.

References

Ackerman, F. & Malouf, R. (2013). Morphological organization: The Low Conditional Entropy Conjecture. *Language* **89**, 429–464.

Hockett, C. F. (1987). *Refurbishing our Foundations: Elementary Linguistics from an Advanced Point of View*. Amsterdam: John Benjamins.

Matthews, P. H. (1991). *Morphology*. Cambridge: Cambridge University Press.

Ramscar, M., Yarlett, D., Dye, M., Denny, K. & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognititive Science* **34**, 909–957.

Robins, R. H. (1959). In defence of WP. *Transactions of the Philological Society* **58**, 116–144. Reprinted in *Transactions of the Philological Society* **99**, 2001, 116–144.

Robins, R. H. (1997). *A Short History of Linguistics*. London: Longman.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal* **27**, 379–423, 623–656.

Implicative structure and inflectional classification

Olivier Bonami
Université Paris-Sorbonne, Institut Universitaire de France, Laboratoire de Linguistique Formelle
olivier.bonami@paris-sorbonne.fr

Descriptions of inflection class systems take many forms, depending on the level of documentation of the language, the theoretical preferences of the author, and the goals of the classification. Pedagogical grammars are often content with giving a broad classification in major classes. At the other end of the spectrum, various attempts at making sense of the structure of inflection systems presuppose a classification into fine-grained micro-classes that exhaustively partition the set of lexemes (see e.g. Stump and Finkel's (2013) notion of a *plat*). The two types of classifications can be linked by assuming a hierarchically-organized system of classes, where macro-classes are seen as groupings of micro-classes (Dressler and Thornton, 1996). Although there are various ways of designing such hierarchies, we limit ourselves here to tree-shaped hierarchies with monotonous inheritance of inflectional properties.

Inflectional classifications are often used to reason about the typology of inflection systems. For such reasoning to be meaningful, it is crucial that the classifications defined for different languages be commensurable. Linguists use combinations of heuristics to decide on a classification, which are hardly ever defined precisely. Thus existing classifications for different languages may not be commensurable. Of particular concern is the fact that inflectional classifications are often dependent on disputed decisions on segmentation: whether two lexemes should be taken as belonging to the same class is deduced from the fact that they use the same affixal exponents in the same paradigm cells. However in the absence of a consensus as to the exact boundary between stem allomorphy and affixal exponence, this results in the possibility of nontrivially different classifications.

One obvious way of addressing this issue is to implement systematized heuristics and apply these to large inflected lexica (Brown and Evans, 2012; Lee and Goldsmith, 2013). While the heuristics might not capture equally well the structure of different systems, at least criticizing the heuristics will be made easier, and the classifications will be commensurate. In this presentation we will exemplify a formally simple-minded way of doing this, where infection classes are deduced directly from implicative relations between surface forms. For concreteness I will show examples from French conjugation. All data is taken from the `Flexique` lexicon (Bonami et al., 2014).

The general strategy is very similar to that of Brown and Evans (2012). We start by establishing a distance matrix between all lexemes in the lexicon, and then use a clustering algorithm to deduce a tree-shaped classification. To evaluate distance between lexemes, we start by inferring a classification of patterns of relatedness between pairs of paradigm cells. Notice that this is the same initial step used by Ackerman et al. (2009) and subsequent literature. The inflection of each lexeme can then be represented by a tuple of patterns relating each pair of cells, and the distance between two lexemes can be defined as the Hamming distance between tuples. This is exemplified in Tables 1 and 2, where a sample of French verbs are reduced to three of their indicative present cells. In this subparadigm, LAVER and RASER inflect in exactly the same way: the distance between them is 0. Both LEVER and FINIR differ from LAVER in two positions, and are thus at distance 2 from it, but they differ from each other in all three positions. BOIRE differs from LAVER in all three dimensions and is thus more distant from LAVER than either FINIR or LEVER.

From the distance matrix one may deduce a hierarchy using some hierarchical clustering algorithm. For concreteness we use here agglomerative average linkage clustering (Sokal and Michener, 1958): items that are minimally distant form a cluster; distance between clusters is evaluated by taking the average of the pairwise distances between the items they contain. The resulting clustering for the current dataset is shown in Figure 1.

lexeme		cells			Tuple of patterns		
		3SG	2PL	3PL	⟨ 3SG ⇌ 2PL	, 3SG ⇌ 3PL , 2PL ⇌ 3PL	⟩
LAVER	'wash'	lav	lave	lav	⟨ $X \rightleftharpoons X$e	, $X \rightleftharpoons X$, Xe ⇌ X	⟩
RASER	'shave'	ʁaz	ʁaze	ʁaz	⟨ $X \rightleftharpoons X$e	, $X \rightleftharpoons X$, Xe ⇌ X	⟩
LEVER	'raise'	lɛv	ləve	lɛv	⟨ Xɛ$C \rightleftharpoons X$əCe	, $X \rightleftharpoons X$, XəCe ⇌ XɛC	⟩
FINIR	'finish'	fini	finise	finis	⟨ $X \rightleftharpoons X$se	, $X \rightleftharpoons X$s , Xe ⇌ X	⟩
BOIRE	'drink'	bwa	byve	bwav	⟨ Xwa ⇌ Xyve	, $X \rightleftharpoons X$v , XyCe ⇌ Xwa$C$	⟩

Table 1: A toy inflection system based on French conjugation

	LAVER	RASER	LEVER	FINIR	BOIRE
LAVER	—	0	2	2	3
RASER	0	—	2	2	3
LEVER	2	2	—	3	3
FINIR	2	2	3	—	3
BOIRE	3	3	3	3	—

Table 2: Hamming distances deduced from Table 1

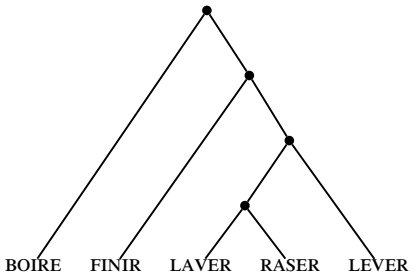


Figure 1: Dendrogram deduced by hierarchical clustering from the distance matrix in Table 2

The present method strikes a balance between unsupervised learning and the practice of descriptive linguists. This makes it useful as a tool for empirical exploration. Importantly, we implement a clear notion of inflectional microclass: two lexemes that intuitively have the exact same inflection have a distance of 0. By contrast, Brown and Evans's method always finds a non-zero distance between two nonidentical paradigms. The price to pay for this is that some linguistic knowledge is built in the classification of patterns of alternation. This is done however without relying on a constant segmentation scheme, avoiding the methodological quandaries that plague traditional discussions of inflection classes.

In the talk we will show the results of applying the present method to the full paradigm of the 4952 verbs in the `Flexique` dataset, and compare the resulting classification to the hand-designed classification of Kilani-Schoch and Dressler (2005). We will also discuss the heuristics implicit in the choice of Hamming distances and particular clustering algorithms, and argue that these decisions, although they must be discussed and justified, are direct analogues of the decisions taken by a descriptive linguist building a classification. Thus if anything, inference of inflection classes from patterns of alternation should help assessing the reliability of classification schemes. Finally, we will discuss how paradigm entropy in the sense of Ackerman et al. (2009) can be used to assess the cohesion of clusters.

References

• Ackerman, F., Blevins, J. P., and Malouf, R. 2009. Parts and wholes: implicative patterns in inflectional paradigms, pp. 54–82. *In* J. P. Blevins and J. Blevins (eds.), Analogy in Grammar. Oxford University Press, Oxford. • Bonami, O., Caron, G., and Plancq, C. 2014. Construction d'un lexique flexionnel phonétisé libre du français. *In* Actes du quatrième Congrès mondial de linguistique française, Berlin. • Brown, D. and Evans, R. 2012. Morphological complexity and unsupervised learning: validating Russian inflectional classes using high frequency data, pp. 135–162. *In* F. Kiefer, M. Ladányi, and P. Siptár (eds.), Current Issues in Morphological Theory: (Ir)regularity, analogy and frequency. John Benjamins, Amsterdam. • Dressler, W. U. and Thornton, A. M. 1996. Italian nominal inflection. *Wiener Linguistische Gazette* 55-57:1–26. • Kilani-Schoch, M. and Dressler, W. 2005. Morphologie naturelle et flexion du verbe français. Gunter Narr Verlag, Tübingen. • Lee, J. and Goldsmith, J. A. 2013. Automatic morphological alignment and clustering. Presented at the 2nt AIMM. • Sokal, R. R. and Michener, C. D. 1958. A statistical method for evaluating systematic relationships. *The University of Kansas Scientific Bulletin* 38. • Stump, G. T. and Finkel, R. 2013. Morphological Typology: From Word to Paradigm. Cambridge University Press, Cambridge.

Reading skill and early inflectional processing

LAURIE BETH FELDMAN
The University at Albany, SUNY &
Haskins Laboratories

&

PETAR MILIN
Eberhard Karls University Tubingen, Germany
University of Novi Sad, Serbia

KIT CHO
The University at Albany, SUNY

A now classic debate in the domain of morphological processing is whether native speakers process regular (talked-talk) and irregular (spoke-speak) verb forms in the same manner. Those who advocate a single processing account characterize effects of regularity in terms of graded form similarity along with consistent meaning similarity and argue that processing of all words benefits from similarity between prime and target (Basnight-Brown, Chen, Shu, Kostić, & Feldman, 2007; McClelland & Elman, 1986; Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997; Rueckl & Raveh, 1999; Rumelhart & McClelland, 1986; Seidenberg & Elman, 1999). As a result, regularity effects emerge because many irregular past and present forms have less similarity between their orthographic and/or phonological form (e.g., letter overlap is higher in TALKED-TALK than in SPOKE-SPEAK) or because regular stems tend to differ from irregular stems in terms of their similarity to other words along semantic dimensions (Baayen & del Prado Martín, 2005). Alternatively, advocates of a dual route account claim that recognition of regular past-tense verbs (talk–talked) is a rule-based process where an -ed past tense marker is affixed to the stem of the verb. For the many (about 180) irregular verbs in the English language characterized by a past tense that does not include an -ed ending, there is no rule. Further, when past tense forms of irregular verbs cannot be formed by rule, those forms must be stored in rote memory (Pinker & Ullman, 2002; Prado & Ullman, 2009). Because many assert that frequency effects reflect storage (but see Baayen, 2009; Baayen, Wurm, Aycock, 2007), "dual routists" also assert that frequency effects should be more robust for irregular than for regular forms.

In the visual lexical decision task, we compare forward masked morphological facilitation after identity, inflected ED and ING and unrelated primes for regular and irregular verbs. Within regularity, we manipulate surface frequency.

Results replicate the study of Milin et al. (2014): an effect of spelling proficiency that fails to interact with either target type or prime type, and a tensor product of two item-based principal components of form (PC1) and frequency (PC2). In addition, effects of morphological versus unrelated prime type were significant and it did not interact with the target inflection (simple, ED and ING). Crucially, the contrast between regular vs. irregular past tense primes was weak for simple targets (talked-talk marginally faster than spoke-speak) and absent for inflected forms of those same targets (talked-talking and spoke-speaking fail to differ).

Results challenge the decomposition account of facilitation that is at the core of the dual route account of inflectional processing.

References

Baayen, R. H. (2009). Demythologizing the word frequency effect: A discriminative learning perspective The Mental Lexicon. 5(3):436-461.

Baayen, R. H., Hendrix, P., & Ramscar, M. (2013). Sidestepping the combinatorial explosion: Towards a processing model based on discriminative learning. Language and Speech, page in press.

Baayen, R. H., & del Prado Martín, F. M (2005). Semantic density and past-tense formation in three Germanic languages Language, 666-698

Baayen, R. H., Milin, P., Filipović Durdević, D., Hendrix, P., and Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. Psychological Review, 118(3):438{481.

Baayen, R. H., Wurm, L. H., Aycock, J. (2007). Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities The Mental Lexicon 11/2007; 2(3):419-463.

Basnight-Brown, D., Chen, H., Shu, H., Kostić, A. and Feldman, L. B. (2007). Monolingual and Bilingual Recognition of Regular and Irregular English Verbs: Does Sensitivity to Word Form Vary with Language Experience? *Journal of Memory and Language.57*, 65-80. doi:10.1016/j.jml.2007.03.001

Milin, P., Ramscar, M., Cho, K., Baayen, R.H. & Feldman, L. B. (2014). Processing partially and exhaustively decomposable words: An amorphous approach based on discriminative learning

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences, 6,* 456–463.

Prado, E. & Ullman, M., (2009). Can Imageability Help Us Draw the Line Between Storage and Composition? Journal of Experimental Psychology: Learning, Memory, and Cognition 2009, Vol. 35, No. 4, 849–866

Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C. S., & Mars, F. (1997). Morphological priming, connectionist networks, and masked fragment completion. Journal of Memory and Language, 36, 382–405.

Rueckl, J. G., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. Brain and Language, 68, 110–117.

Rumelhart, D., & McClelland, J. (1986). On learning the past tense of English verbs: implicit rules or parallel distributed processing? In J. McClelland, D. Rumelhart and the PDP Research Group (Eds.), Parallel distributed processing: Explorations in the microstructure of cognition. Cambridge, MA: MIT Press.

Seidenberg, M. S., & Elman, L. M. (1999). Networks are not hidden rules. Trends in Cognitive Sciences, 3, 353–365.

Morphemes in context and in isolation: A recursive view on measures of frequency and complexity.

EMMANUEL KEULEERS[1], PETAR MILIN[2], AND JAMES P. BLEVINS[3]
[1]Ghent University, [2]University of Novi Sad, [2]Eberhard Karls University of Tübingen, [3]Cambridge University
emmanuel.keuleers@ugent.be

Measures of frequency and complexity are used pervasively in the study of word processing. For instance, in research on the recognition of complex words, surface frequency, lemma frequency, family size, and derivational entropy, have all been shown to have an influence on morphological processing (e.g., Amenta & Crepaldi, 2012).

In this talk, we will show that nearly all measures of frequency and complexity are variants of three fundamental measures: element frequency, element dispersion, and element entropy, which can be derived recursively by assuming that the contexts in which elements are seen become themselves the elements of interest in the next recursion.

For instance, let us consider unbound morphemes to be elements of interest, and let us consider context to be the material surrounding the element up to the word boundaries. The token count of the contexts in which the element (morpheme) occurs is the measure of element frequency (in this case morpheme frequency). The type count of the contexts is the measure of element (morpheme) dispersion and the entropy of the probability distribution of contexts is the element entropy. Interestingly, the final two measures are also known as family size and derivational entropy (e.g., Baayen, Feldman, Schreuder, 2006).

If we now consider the elements in context to be the elements of interest, we are operating over words in the context of surrounding words. The measure of element frequency now corresponds the classic measure of word frequency. Next, the measure of element dispersion, which could be called word family size, simply indicates in how many word trigrams a word of interest occurs centrally. Finally, the probability distribution of word trigrams is the basis for the computation of the measure of element entropy at this level.

It should now be clear that in the following recursion, measures of word trigram frequency, word trigram dispersion, and word trigram entropy are obtained, and so on.

While we have so far considered context as elements directly surrounding the elements of interest, the definition of context is variable. For instance, starting with morphemes as elements of interest, but looking only at material directly to the left or to the right give us measures closely related to prefixation and suffixation at the first step of recursion. At a higher level, considering the edges of a document in which a word occurs as boundaries, gives us contextual diversit (Adelman, Brown, & Quesada, 2006) as the measure of element dispersion.

This systematic approach yields several interesting results. In our talk, we will focus on three findings in particular.

First, to our knowledge, word dispersion (the family size measure at the word level) has not been investigated. However, this measure seems to be superior to word frequency in predicting lexical decision times on words presented in isolation. This finding by itself is quite surprising, as word frequency is the most important and most robust predictor of isolated

word processing. This shows that our approach fills gaps: it allows us to identify important measures which are simply absent from the literature.

Second, at a lower level of granularity, considering phonemes as elements of interest generates the measure of phoneme entropy. In a phoneme-trigram context, this measure provides a simple and elegant insight in the concept of sonority, with elements with higher sonority having higher entropies.

Finally, the same finding occurs at the level of morphemes, where measures of morpheme entropy give an indication of the freedom with which morphemes can combine.

References

ADELMAN, J. S., BROWN, G. D. A., & QUESADA, J. F. (2006). Contextual Diversity, Not Word Frequency, Determines Word-Naming and Lexical Decision Times. *Psychological Science*, *17*(9), 814–823.

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*.

BAAYEN, R. H., FELDMAN, L. B., & SCHREUDER, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. Journal of Memory and Language, 55(2), 290–313.

The effect of grammatical context on processing Serbian inflected nouns and verbs

ALEKSANDAR KOSTIĆ
Laboratory for Experimental Psychology
Faculty of Philosophy, University of Belgrade
akostic@f.bg.ac.rs

Inflected languages are characterized by grammatical congruency among different word classes. Thus, for example, Serbian nouns are congruent with adjectives and possessive pronouns in case, grammatical number and gender, verbs are congruent with personal pronouns in person, grammatical number and sometimes gender etc. Any violation of this congruency will produce an ungrammatical sentence. The congruency is in most cases realized via inflectional suffixes attached to the root morpheme of an open class word.

In four experiments using the lexical decision task it was demonstrated that Serbian inflected noun forms are primed by adjectival and pronominal context. When preceded by congruent adjectival and pronominal primes, noun forms were processed faster than when preceded by incongruent context. (Gurjanov, Lukatela, Moskovljević, Savić & Turvey, 1985a; Gurjanov, Lukatela, Savić & Turvey, 1985b).

The information-theoretic approach to processing inflected morphology is based on the assumption that the cognitive system is sensitive to the uncertainty of inflected word forms, specified in terms of the amount of information (bits). The general formalism from which the amount of information is derived consists of two distinct terms: the probability of an inflected form and the grammatical component. The grammatical component in case of nouns is the number of syntactic functions/meanings (R) of an inflected noun form, while in case of verbs it is the number of congruent personal pronouns (Q) /Equ. 1 and 2/.

$$I_m = \left[-\log_2 \left(\frac{\frac{F_m}{R_m}}{\sum_{j=1}^{k} \frac{F_{m_j}}{R_{m_j}}} \right) \right] \qquad I_m = \left[-\log_2 \left(\frac{\frac{F_m}{Q_m}}{\sum_{j=1}^{k} \frac{F_{m_j}}{Q_{m_j}}} \right) \right]$$

In a series of experiments it was demonstrated that the amount of information derived from the above equations accounts for almost all processing variability of Serbian inflected noun and verb forms presented in isolation (Kostić, 1995; Kostić & Havelka, 2002; Kostić, Marković & Baucal 2003). The question is how to account for priming effects observed with Serbian inflected nouns. The approach adopted in this study is the following: facilitation and inhibition effects are *co-implicative* and take place *simultaneously* within a defined morphological paradigm. For example, Serbian regular feminine nouns appear in six distinct inflected forms which constitute the paradigm of feminine nouns. Congruent primes will decrease the amount of information carried by the target stimulus and proportionally increase the amount of information carried by other paradigm constituents.

This approach was evaluated in three experiments using the lexical decision task. In the first experiment two noun forms (suffixes *–i* and *–u*) were preceded by neutral context (***) and congruent and incongruent adjectival primes. In the second experiment the same noun forms were preceded by neutral context and congruent and incongruent possessive pronouns. In the third experiment the first and second person verb in the present tense were preceded by neutral context and congruent and incongruent personal pronouns.

In case of adjectival and pronominal priming, modification of the amount of information carried by a noun form should be applied at the level of probability of a noun

form because adjective forms, personal pronoun forms and noun forms share the same syntactic functions/meanings. Likewise, the amount of information carried by verb person should also be modified at the level of its probability (Equ. 1 and 2).

There are two ways to carry out these modifications. The most obvious way is to apply the formalism for conditional probabilities. The equation for conditional probabilities is defined as the ratio of the probability of the joint of events *A* and *B*, and the probability of *B*. An alternative way is to assume that the modified probability of the target in congruent contexts is the sum of the probabilities of the prime and the target. Consequently, the probability of the remaining forms within a paradigm is proportionally decreased.

In regression analyses it was demonstrated that the amount of information derived from the equation for conditional probabilities does not account for significant proportion of processing latency variability in the three experiments. The informational values derived from the alternative approach (i.e. summing up prime and target probabilities) accounted for almost all processing variability in the experiments (see Figures 1 and 2).

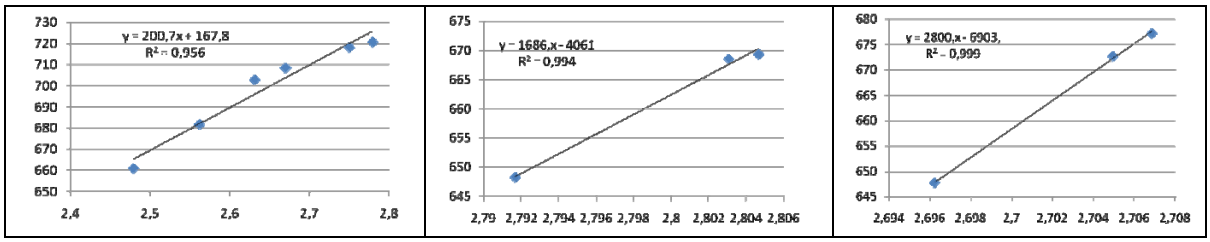


Figure 1. Relation between the amount of information and processing latency in Experiments 1 and 2. Left: noun targets with suffixes *–i* and *–u* in Exp. 1; Middle: suffix *–i* in Exp. 2, right: suffix *–u* in Exp. 2

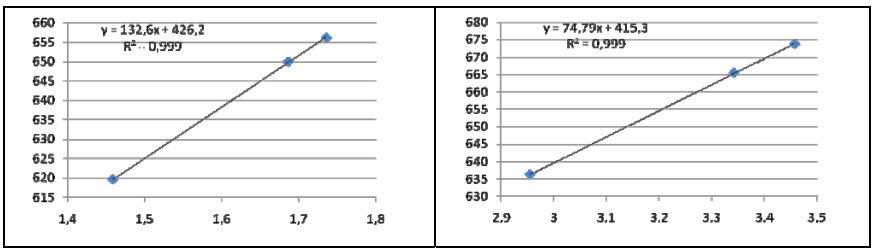


Figure 2. Relation between the amount of information and processing latency in Experiment 3. Left scattergram: first person present tense; Right scattergram: second person present tense

There is, however, a serious conceptual problem with this approach. Summing prime and target probabilities implies that the two events are independent (mutualy exclusive). This, in turn, requires revision of the general information-theoretic framework that describes the processing consequences of grammatical congruency among different word classes.

References

GURJANOV, M., LUKATELA, G., MOSKOVLJEVIĆ, J., SAVIĆ, M., &: TURVEY, M. T. 1985a. Grammatical priming of inflected nouns by inflected adjectives. *Cognition*, 11, 55-71.

GURJANOV, M., LUKATELA, G., LUKATELA, K., SAVIĆ, M., &: TURVEY, M. T. 1985b. Grammatical pr iming of inflected nouns by the gender of possessive adjectives. *Journal of Experimental Psychology: Learning, memory, and Cognition,* Vol 11, No 4,. 692-701.

KOSTIĆ, A. 1995. Informational load constraints on processing inflected morphology. In L. B. Feldman (Ed.), *Morphological Aspects of Language Processing*. New Jersey. Lawrence Erlbaum, Inc., Publishers.

KOSTIĆ, A. & HAVELKA, J. 2002. Processing of a verb tense. *Psihologija*, 35, 3-4, str. 299-316.

KOSTIĆ, A., MARKOVIĆ, T. & BAUCAL, A. 2003. Inflectional morphology and word meaning: orthogonal or co-implicative cognitive domains? u H. Baayen & R. Schreuder (Eds.): *Morphological Structure in Language Processing*. Mouton de Gruyter. Berlin. str.1- 45.

Why is *Maria* read slower than *Μαρία*?

SOPHIA MANIKA
UiL OTS, Utrecht University
s.manika@uu.nl

Introduction Consider the following:

(i) John *praises*, very often, *Maria* and not his other friends - low inflectional entropy (LH)
(ii) John *helps*, very often, *Maria* and not his other friends - high inflectional entropy (HH)

Processing of the Proper Name (PN) *Maria* is influenced by the ease of processing of the verb with which it forms a predicate, namely, *Maria* in (i) is processed harder than *Maria* in (ii). What differentiates *praises* from *helps* is their accessibility in long term memory (LTM) as measured by their value of the ***inflectional entropy of their verbal paradigm*** (H), an information-theoretic measure that quantifies the representational similarity of an inflectional paradigm in the lexicon and is a function of each verb type's frequency and syntactic functions it serves [1]. High inflectional entropy (HH) describes a more "uniform" distribution of the verb types in the inflectional paradigm that increases the type's accessibility. Low inflectional entropy (LH) describes a more "distinct" distribution of the verb types in the inflectional paradigm that lowers the type's accessibility. Consequently, response times of inflected verb types that belong to an HH paradigm (in ***comprehension***) are accelerated in comparison to inflected verb types that belong to a LH paradigm when the words are in ***isolation*** ([2], [3] a.o.). The effect pertains also ***within*** a sentence and since inflectional entropy is associated with the amount of resources consumed during first lexical activation it influences, accordingly, the processing time of the object of the sentence as well [4]. So, in the case of an object like *Maria* verbs with HH require fewer resources during first activation and their object is processed faster than verbs with LH, the objects of which are delayed due to the costly first activation of the verb.

Inflectional entropy is a continuous measure and can also be defined at a between language level according to the language's morphological wealth. In fact, rich inflection allows for distinct forms per syntactic function resulting in a more uniform distribution of the paradigm so languages with rich morphology (e.g. Greek) have verbs with higher values of inflectional entropy and lie on the right side of the entropy continuum. In contrast, poor inflection forces few forms to be used in more functions than others increasing their frequency and, consequently, languages with poorer morphology (e.g. Dutch) have verbs with lower values of inflectional entropy and are positioned on the lower end of the entropy continuum This means that, all other factors kept constant, in a sentence like (i), the verb *praises* would be read faster in Greek ('epeno') than in Dutch ('prijzen'), because Greek is an HH language and Dutch is a LH language.

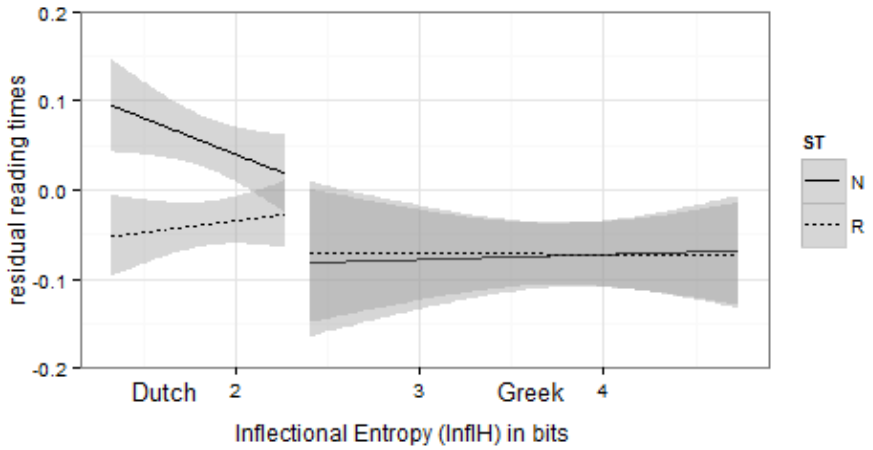
The present paper studies the hypothesis that the effects of inflectional entropy can be generalized between languages. More precisely, it will be demonstrated that, indeed, Greek verbs are retrieved faster than Dutch verbs as a result of different positioning of the language on the entropy continuum and that crucially, *Μαρία* is processed faster than *Maria* as a result of the higher amount of available resources in Greek.

Experiments Forty-two native Dutch and forty-one native Greek students read sentences like (i) and (ii) in two self-paced reading experiments. The words between the verb and the PN were kept constant across items. The PNs in subject and object position were controlled for frequency and number of syllables. The inflectional entropy of the matrix verb was varied and the two entropy ranges between languages did not overlap (the highest value of entropy of a Dutch verb was lower than the lowest value of entropy of a Greek verb).

Reading times in each word were measured and analyzed with linear mixed effects regression analyses using a 2-stage model, as described in [5]. In the first stage logarithmic transformed RTs of each word were regressed against the number of the word's letters, position of the word in the sentence and position of the item in the task. The residuals of that model were used as the dependent variable for the second stage and were crossed with language (Dutch vs Greek) and object type (PN vs reflexive, which served as a baseline). Analyses with entropy values were also conducted and yielded the same results, because entropy values correspond (at least for the present experiments) to a language and vice versa. For the sake of simplicity, the analyses reported are the one with language.

Results Verb retrieval: The main effect of entropy on RTs was significant ($t=-2.758$, $p<.05$) at the region of the verb. Verbs with higher inflectional entropy were read faster than verbs with lower inflectional entropy.

Object integration: The main effect of object type on RTs was significant at the region of the object ($t=-2.458$, $p<.05$). Sentences with PNs were read significantly slower than those with reflexives. There was a significant interaction between object type and language ($t=2.476$, $p<.05$). Simple slopes reveal that high entropy language (Greek) accelerates RTs for sentences with PN ($t=-3.342$, $p<.01$).



Discussion Inflectional entropy is tightly connected to the morphological richness of a language that in turn affects the speed of sentence processing. Processing of a PN requires introduction of a new entity in the discourse, which is a costly operation. In a morphologically poor language, like Dutch, that consumes a lot of resources already at the time of the verb's retrieval, it is even costlier. The fact that the reflexive condition was processed equally fast in both languages provides us with a baseline and enables us to conclude that this delay, in Dutch PN *Maria*, is a result of the shortage in available resources, as predicted by the position of Dutch in the inflectional entropy's continuum.

References

[1] KOSTIĆ, A. (1995). Information load constraints on processing inflected morphology. In: Morphological Aspects of Language Processing. Hillsdale, NJ: Erlbaum.

[2] MILIN, P., FILIPOVIĆ ĐURĐEVIĆ, D. & MOSCOSO DEL PRADO MARTÍN, F. (2009). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from serbian. Journal of Memory and Language, 60(1), pp. 50-64.

[3] MOSCOSO DEL PRADO, F., KOSTIĆ, A. & BAAYEN, R. (2004). Putting the bits together: An information theoretical perspective on morphological processing. Cognition, 94, pp. 1-18.

[4] MANIKA, S., AVRUTIN, S. & REULAND, E. (2014). The effect of verb's inflectional entropy on the processing of reflexive objects. In *Selected Papers from UK-CLC4*, http://www.uk-cla.org.uk/proceedings/volume_2_36.

[5] FINE, A. B., JAEGER, T. F., FARMER, T. A, & QIAN, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*

Inflectional paradigms and classes in sentence reading

Petar Milin*† Peter Hendrix* R. Harald Baayen*‡
*Eberhard Karls University, Tübingen, Germany
†University of Novi Sad, Serbia ‡University of Alberta, Canada
{petar.milin, petrus.hendrix, harald.baayen}@uni-tuebingen.de

Milin, Filipović Đurđević, & Moscoso del Prado Martín, 2009 showed that the relative entropy of Serbian nouns correlates positively with response latencies in lexical decision, where the relative entropy operationalized divergence between the distribution of the inflected variants of a given word from the distribution of case exponents across all words in that word's inflectional class. By running two experiments using masked priming paradigm: lexical decision and self-paced sentence reading, Baayen, Milin, Filipović Đurđević, Hendrix, & Marelli, 2011 replicated the same effect. To account for the influence of the prime, the authors used a weighted relative entropy.

The present research question is whether the same effect of relative entropy occurs in normal reading. Two eye-tracking experiments were conducted. Experimental items were single-line sentences containing two targeted nouns. In the first experiment the two nouns belonged to the same inflectional class, while in the second experiment inflectional classes were different.

Results replicated previous findings but also shed new light on the effect of relative entropy in normal reading. On the one hand, the weighted relative entropy measure was predictive when nouns belonged to the same inflectional class. On the other hand, for nouns belonging to different inflectional classes a simple, unweighted relative entropy measure sufficed.

The pattern of results suggests that the second noun is discriminated against the first noun. Additionally, the necessity of a weight for nouns belonging to the same inflectional class suggests that in this case more cognitive effort is required to discriminate between the two words.

References

Baayen, R., Milin, P., Filipović Đurđević, D., 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.

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

A functional take on Estonian inflectional paradigms

FERMÍN MOSCOSO DEL PRADO MARTIN
University of California, Santa Barbara
fmoscoso@linguistics.ucsb.edu

Inflectional richness is often taken as an indication of linguistic "complexity" (e.g., McWhorter, 2011). However, from an informational-theoretical standpoint, it has been noted that the paradigms of morphologically rich languages are perhaps not all that complex (e.g., Blevins, Malouf & Ackerman, 2019; Malouf & Ackerman, 2013). Whereas the view of inflectional morphology assumes that it is difficult to know how to fill the cells in a paradigm in an inflectionally rich language, the latter approches show that a few forms in the paradigm, the classical principal parts, are strongly informative about the rest, and therefore substatially reduce the actual difficulty or complexity that is entailed by the paradigm. This has led some researchers to postulate a "low entropy conjecture" (e.g., Malouf & Ackerman, 2013): In terms of their morphological complexity languages differ less than one might expect.

In this study, I follow this "low entropy conjecture" one step further. One needs to ask the question of what is the functional role inflectional classes in the first place. Although it is relatively accepted that noun classes, such as gender, can have a semantic, and thus functional component to them, it is generally thought that the inflectional classes themselves are of little semantic relevance. It seems however rather illogical that a system subject to such evolutionary pressure as is language would indulge in gratuitous levels of complexity in a systematic fashion. Rather, I show that the variation observed in the forms filling the same paradigmatic cell has itself a functional justification in reducing semantic uncertainty.

I present a detailed, corpus-based, information theoretical analysis of the Estonian nominal system. Firstly, following Malouf and Ackerman (2013), I show that the principal parts of Estonian paradigms can be recovered from corpus data (and I provide a new algorithm to achieve this automatically). I follow by investigating how the choice of inflected variant to fill in a paradigm cell is in itself informative about the noun's semantics. This is in slight contradiction with Blevins (2008)'s statement that "classes are cued by variation in form, without evident syntactic or semantic correlates". Finally, using word co-occurrence statistics, I provide a detailed account of mutual implication relations between word meaning, inflectional class, and inflectional cell in all possible pairings. With this I show that, rather than orthogonal, these three factors, together with the stem itself, are strongly mutually informative. The consequence of the mutual informations described above is that, at the end, the choice of form to fill a paradigm cell, the choice of meaning associated with an inflected form, or the choice of case for a particular word, displays surprisingly little uncertainty (i.e., entropy), in line with the predictions of the low entropy conjecture.

Selected References

- JAMES P. BLEVINS (2008) Declension classes in Estonian. *Linguistica Uralica* XLIV, 241–267.
- JAMES P. BLEVINS, FARRELL ACKERMAN and ROBERT MALOUF (2009) Parts and wholes: Implicative patterns in inflectional paradigms. In JAMES P. BLEVINS and JULIETTE BLEVINS (eds.), *Analogy in Grammar: Form and Acquisition*. Oxford University Press, 54–81.
- FARRELL ACKERMAN and ROBERT MALOUF. 2013. Morphological organization: The low conditional entropy conjecture. *Language*. 89:429-464.

The possible role of Entropy in processing argument dependencies in Hungarian

CSABA PLÉH KORNÉL NÉMETH DÁNIEL VARGA
Eszterházy College, Eger Budapest U of Technology and Economics
pleh.csaba@ektf.hu cornel.nemeth@gmail.com daniel@mokk.bme.hu

It is a commonplace of research in sentence processing for at least 40 years now, since the beginnings of the use of Thematic Roles and different versions of Frame Theories in processing studies, to find facilitative effects between verbs and their arguments (Tanenhaus et al, 1995, Kintsch, 1998). Thus reading CUT would facilitate the reading of instrumental arguments. In a less trivial way, reading MET or PLAYED also would facilitate reading instrumentals. This is a rather crucial issue in sentence processing, to the extent that Bornkessel & Schlesewsky (2006) even developed a full-fledged crosslinguistic theory of the temporal activation of the verbal argument frames and the insertion of noun phrases into the slots based on neuronal processing evidence.

In our research we concentrate on oblique arguments of Hungarian verbs exploiting the fact that the argument relations are coded by case markers in Hungarian. Therefore, argument processing is closely tied with morphological processing. As a matter of fact most of the arguments with abstract relations are coded by locative case markers as illustrated in the table.

Type	Example	Gloss
unambigouous	Emlékszik a fiú**ra** Haragszik a tanító**ra**. Készül a verseny**re**. Találkozott a lán**nyal.** Csókolózott a rendőr**rel.**	remembers the boy.ON angers the teacher.ON Prepares the race.ON Met the girls.INST Kissed the cop.INST
ambiguous in marking	Gondol a lány**ra.** Gondol valami**t.**	thinks girl.ON thinks something.ACC
ambiguous argument /adjunct	Gondolkodik a lány**on.** Gondolkodik a haj**ón.**	thinks girl.ON thinks boat.ON

Earlier studies on the processing of argument structures in Hungarian have shown that these assumed interactions between morphology and sentence processing do indeed hold. Gervain and Pléh (2003), for example, showed that prenominal verbs facilitate the processing of constructions like 'Anna thought of the boat', and postverbal nouns that are ambiguous between a locative-adjunct and an abstract-argument reading are read slower that arguments. Compare 'Anna RUMINATED on the boat' versus 'Anna RUMINATED on the problem', where in the later case the locative meaning is excluded. Fekete and Pléh (2011) had shown the relevance of the argument-based approach to psycholinguistic processing (experimenting in Hungarian with bidirectional comitative constructions like John was kissing with Mary), compared with unidirectional comitative such as 'John was messing with Mary'. They observed that there was a difference in the processing of anaphors referring to the Subject. Anaphors in THE singular (He was…) were processed slower

after bidirectional antecedents, showing a 'deep anaphora effect' in the sense introduced by Hankamer. & Sag (1976).

All of these studies did try to control for frequency effects. Case ending and semantic interpretation ambiguities, however, were treated by them in a categorical manner. In the studies to be presented at the workshop we shall concentrate on the statistical structure of the possible ambiguities.

1. We would combine the structures illustrated in the table with a consideration of entropy measures. From the MOKK (2006) and the MAZSOLA corpus (Sas, 2008) entropy computations would be made for the relations between a given verb and the noun endings in a 'plus/minus two content words' frame. These distributions will be compared to the theoretical propositions.
2. These entropy estimates will be used as predictors in self-paced word by word reading time experiments for the reading of nouns in sentence contexts. In this we way we would like to learn whether the entropy relations between verbs and case endings do have an explanatory power in processing and verb-noun attachments.

References

Bornkessel, I. & Schlesewsky, M. (2006) The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review, Vol 113*(4), 787-821 • Fekete, I. & Pléh, Cs. (2011) Bidirectional and unidirectional comitative constructions in Hungarian: A psycholinguistic investigation at the interface of argument structure and semantics. *Acta Linguistica Hungarica, Vol. 58 (1–2), pp. 3–23* • Gervain, J. & Pléh, C. (2004). [Anna is thinking on the ship, or argument expectancies in Hungarian sentence understanding]. In: Judit Gervain, J. & Pléh, C. (eds): *A láthatatlan megismerés,* 112–25. Gondolat, Budapest. • Hankamer, J. & Sag. I. 1976. Deep and surface anaphora. *Linguistic Inquiry 7*: 391–428. • Kintsch,W. 1998. *Comprehension: a paradigm for cognition.* Cambridge University Press, Cambridge. • MOKK (2006): A WEB based Hungarian frequency dictionary: http://szotar.mokk.bme.hu/szoszablya/searchq.php • Sass, B. (2008): The Verb Argument Browser. In: Sojka, P. et al. (eds.): Proceedings of TSD 2008, Brno, Czech Republic, , LNCS 5246, 187-192,http://corpus.nytud.hu/vab • Tanenhaus,M.K.,Spivey-Knowlton,M. J.,Eberhard,K., & Sedivy,J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 632–634.

Implicative organization facilitates morphological learning

SCOTT SEYFARTH, FARRELL ACKERMAN, & ROBERT MALOUF
University of California, San Diego and San Diego State University
sseyfarth@ucsd.edu, fackerman@ucsd.edu, rmalouf@mail.sdsu.edu

In word-based morphology, implicative relationships among related wordforms are used to facilitate the learning of complex morphology (see Blevins, to appear, ch. 7; Ackerman & Malouf 2013). For example, an Italian singular suffix *-a* (*persona)* typically implies a plural suffix *-e* (*persone*), while singular suffix *-o* (*gatto*) implies plural suffix *-i* (*gatti*), and vice versa. On this basis, a learner might correctly predict that a novel singular form *rosa* has the plural form *rose*. However, this perspective on morphological organization is challenged by experimental results in artificial language learning, which find that learners are unable to acquire these relationships unless they are supported by additional phonological cues in the stem (Gerken et al. 2009; Brooks et al. 1993, Frigo & MacDonald 1998). We argue that these results were biased by ecological factors, and that implicative relations alone are sufficient to enhance paradigm learnability.

We first identify four methodological features of prior experiments that restrict general learnability: (1) passive exposure and/or rote repetition, rather than active trial-and-error learning; (2) a requirement to learn both novel lexemes and inflections in a short time; (3) a requirement to infer relations among abstract linguistic labels rather than referents; (4) presentation of labels before their referents, which is known to delay association learning (Ramscar et al. 2010).

We present three novel artificial grammar experiments in which subjects demonstrate knowledge of formal paradigmatic relationships, despite experimental randomization designed to avoid redundant phonological or semantic information that would signal class membership. In Experiment 1, subjects attempted to learn how to inflect familiar nouns for number (singular, dual, plural) in an alien language. Six nouns were randomly selected from a pool of 30 household objects and divided into two classes which had suffix inflections as shown in Figure 1. In each training trial, subjects saw one, two, or many pictures of a noun and tried to guess how the aliens would refer to that set. Subjects were then given feedback and shown the correct answer.

After 90 training trials, subjects were tested on their knowledge. In each testing trial, subjects first saw one, two, or many pictures of a new, previously-unseen object, and were given its label. Subjects were then asked to label a different number of the same noun. In critical trials, the new noun was first presented in a form that reliably predicted the test form (e.g., singular *-taf* implies dual *-guk*). The results showed that subjects successfully took advantage of implicative relationships: when subjects could use two compatible suffixes for the second inflection, performance was significantly better when they were first presented with a predictive inflection than when they were not ($p < 0.001$). Experiments 2-3 tested more complex paradigms with 9 nouns and 3 classes. In these paradigms, performance was better only for bidirectional relationships, where the given and unseen suffixes both implied each other ($p < 0.01$).

Our results indicate that subjects acquired paradigmatic relations without redundant cues. This finding suggests that ecological factors drove previous results, which helps resolve

	class 1	class 2
SG	chair-taf	bed-yez
DU	chair-guk	bed-cav
PL	chair-lem	bed-lem

Experiment 1: (n=33)

class 1	class 2	class 3
chair-taf	bed-yez	table-yez
chair-guk	bed-cav	table-cav
chair-lem	bed-lem	table-nup

Experiment 2: (n=28)

class 1	class 2	class 3
chair-taf	bed-yez	table-seb
chair-guk	bed-cav	table-cav
chair-lem	bed-lem	table-nup

Experiment 3: (n=32)

Figure 1: Five (in Exp. 1), six (Exp. 2), or seven suffixes (Exp. 3) were randomized to avoid phonological patterns.

the apparent mismatch between prior experimental and typological data. Some researchers have argued that enriching word classes with additional syntactic or semantic redundancies could permit acquisition if phonological cues are inadequate (Ouyang, Boroditsky, & Frank 2012; Mintz 2002; Braine 1987). However, our results indicate that learners can acquire useful paradigmatic relations without any additional supporting cues if the potential influence of domain-general learning factors is properly taken into account (Frank & Gibson 2011).

References

ACKERMAN, F., & MALOUF, R. (2013). Morphological Organization: The Low Conditional Entropy Conjecture. *Language*, *89*(3), 429–464. doi:10.1353/lan.2013.0054

BLEVINS, J. (to appear). Word and Paradigm Morphology. Oxford University Press.

BRAINE, M. (1987). What is learned in acquiring word classes: a step toward an acquisition theory. In *Mechanisms of Language Acquisition* (pp. 65–87). Hillsdale, NJ: Lawrence Erlbaum.

BROOKS, P., BRAINE, M., CATALANO, L., & BRODY, R. (1993). Acquisition of gender-like noun subclasses in an artificial language: the contribution of phonological markers to learning. *Journal of Memory and Language*, *32*, 76–95.

FRANK, M. C. & GIBSON, E. (2011). Overcoming memory limitations in rule learning. *Language Learning and Development*, *7*(2), 130–148

FRIGO, L., & MACDONALD, J. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language, 39,* 218–245.

GERKEN, L., WILSON, R., GÓMEZ, R. L., & NURMSOO, E. (2009). The relation between linguistic analogies and lexical categories. In *Analogy in Grammar: Form and Acquisition*. Oxford University Press.

MINTZ, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition*, *30*(5), 678–686.

OUYANG, L., BORODITSKY, L., & FRANK, M. C. (2012). Semantic Coherence Facilitates Distributional Learning of Word Meanings. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society.*

RAMSCAR, M., YARLETT, D., DYE, M., DENNY, K., & THORPE, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science*, *34*(6), 909–957. doi:10.1111/j.1551-6709.2009.01092.x

Structural attraction in Croatian:
Effects of inflectional paradigmatic structure in morphosyntactic processing

ANDREA D. SIMS
The Ohio State University
sims.120@osu.edu

Structural attraction is when a target agrees with a local controller, rather than the proper but more distant one: *How much correction*$_{SG}$ *of syntactic errors*$_{PL}$ *are*$_{PL}$ *there, anyway?* It is widely considered to be a locality effect in morphosyntactic processing. Attraction effects are relatively rare in case-marking languages (Nicol and Wilson 2000), but it is unclear whether this derives from shallow processing of the noun's (non-nominative) form, or reflects deeper processing of morphological feature structure. Croatian offers the opportunity to tease apart these issues.

Attraction effects in Croatian are sensitive to syncretism: attraction is found (rarely) when non-nominative local controllers are homophonous with nominative; see (1), in which the verb *tjerala su* seems to agree with the local accusative NP *ova plemena*, rather the conjoined structural subjects. Importantly, *ova plemena* exhibits systematic nominative-accusative syncretism, so in a local context it could be construed as a viable controller for the verb.

(1) *Vječiti ratovi i opasnosti od uvijek mogućih napada…*
perpetual wars(M).NOMPL and dangers(F).NOMPL from always possible attacks

tjerala su ova plemena u saveze.
drove.N.PL AUX.3PL these.N.ACCPL tribes(N).ACCPL into alliances

'Perpetual wars$_{(M).NOMPL}$ and dangers$_{(F)NOM.PL}$ from ever-possible attacks …, drove$_{N.3PL}$ these tribes$_{(N).ACCPL}$ into alliances.' (http://hr.wikipedia.org/wiki/Indijanci, March 2008)

Similar facilitation has been elicited through forced production in Slovak (Badecker and Kuminiak 2007) and German and Dutch (Hartsuiker et al. 2003).

Studies of this syncretism effect in structural attraction tend to assume that it reflects shallow processing of the morphological form of the local controller, but they fail to distinguish between superficial homonymy and systematic syncretism, the latter being a deep fact of morphological structure. This leaves open the question: Does homophony that reflects morphological organization (syncretism) facilitate structural attraction more than accidental homophony? We might expect this given evidence that the processing of inflected forms is sensitive to the amount of information they convey (Milin et al. 2009; Moscoso del Prado Martín et al. 2004). We might also predict this based on resolution of morphosyntactic feature conflict by syncretism in which the systematicity of the syncretism matters, e.g., German double agreement (Groos and van Riemsdijk 1981).

This paper presents a way to quantify accidental vs. systematic syncretism using information-theoretic measures, and uses it to predict structural attraction effects in Croatian. I show experimentally that some patterns of syncretism in Croatian (e.g, nom pl - acc pl) induce attraction effects more than other patterns do (e.g., nom pl - gen sg), and I hypothesize that this difference is explained in terms of the information that a syncretic form conveys about the morphosyntactic values it expresses. I quantify the information a given Croatian inflected form conveys in terms of Shannon entropy. (Since syncretic forms are morphosyntactically ambiguous, there is on average more surprisal associated with the morphosyntactic values of a syncretic form than a non-syncretic form.) I also ask the extent to which the distribution of syncretic forms in a given paradigm is shared across inflection classes and word classes, and

quantify this in terms of Hamming distance calculated over nearest neighbor distributions. In combination, the two measures offer a way to quantify the extent of syncretism within a paradigm, and more importantly, the degree of systematicity of a pattern of syncretism.

The results suggest that morphosyntactic processing is sensitive to the difference between accidental and systematic inflectional homophony in the morphological system. Morphosyntactic processing and agreement patterns in Croatian thus shed light on the intricate structuring of inflectional paradigms and the cognitive processing mechanisms underlying structural attraction effects.

References

BADECKER, W., and F. KUMINIAK. 2007. "Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak." *Journal of memory and language* 56(1): 65-85.

GROOS, A., and H. VAN RIEMSDIJK. 1981. "Matching effect in free relatives: A parameter of core grammar." In *The theory of markedness in generative grammar: Proceedings of the 1979 GLOW conference*, edited by A. BELLETTI, L. BRANDI and L. RIZZI, 171-216. Pisa: Scuola Normale Superiore di Pisa.

HARTSUIKER, R., H. SCHRIEFERS, K. BOCK, and GERDIEN M. KIKSTRA. 2003. "Morphophonological influences on the construction of subject-verb agreement." *Memory and cognition* 31(8): 1316-1326.

MILIN, P., V. KUPERMAN, A. KOSTIĆ, and R. H. BAAYEN. 2009. "Words and paradigms bit by bit: An information-theoretic approach to the processing of inflection and derivation." In *Analogy in grammar: Form and acquisition*, edited by J. P. BLEVINS and J. BLEVINS. Oxford: Oxford University Press.

MOSCOSO DEL PRADO MARTÍN, F., A. KOSTIĆ, and R. H. BAAYEN. 2004. "Putting the bits together: An information theoretical perspective on morphological processing." *Cognition* 94: 1-18.

NICOL, J., and R. WILSON. 2000. "Agreement and case-marking in Russian: A psycholinguistic investigation of agreement errors in production." In *Proceedings of the Annual Workshop on Formal Approaches to Slavic Linguistics: The Philadelphia Meeting 1999*, edited by T. HOLLOWAY KING and I. SEKERINA, 314-327. Ann Arbor: Michigan Slavic Publications.

Diversity and relative entropy as indices of relational competition in the processing of English compounds

THOMAS L. SPALDING & CHRISTINA L. GAGNÉ
Department of Psychology, University of Alberta
spalding@ualberta.ca cgagne@ualberta.ca

Recent research has indicated that processing compound words involves an attempt at semantic composition of the constituent words (Fiorentino & Poeppel 2007; Kounios et al. 2003; Koester et al. 2007), and that this meaning construction process involves an attempt to identify a semantic relation linking the constituents (Gagné & Spalding 2009; Ji, et al. 2011). The current project investigates relational competition during the processing of transparent and opaque English compounds, using an information theoretic approach. In particular, we examine the impact of relational diversity, which reflects the extent to which the possible interpretations concentrate on a small set of relations, and relational relative entropy, which measures the relational information gain associated with the particular compound (i.e., relational relative entropy measures how much the relational distribution for a specific item differs from the distribution based on all items). These measures are derived from the possible relations task (Gagné & Spalding 2014), which generates, for each item, a distribution of possible relational interpretations based on participant responses to a question in which they are asked to indicate the most likely literal meaning for a pair of words (e.g., *hog wash*).

We found that although both semantically opaque and semantically transparent compounds have morphemic structure, the response latencies in a lexical decision task for these items are differentially affected by diversity and entropy. Diversity interacted with compound frequency for transparent compounds but not for opaque compounds. Also, when information from the compound leads to a particular meaning (i.e., when there is low diversity and high entropy), this convergence resulted in faster processing for transparent compounds, but slower processing for opaque compounds.

These results suggest that relation-based interpretations are computed whenever morphological constituents (and their corresponding semantic/conceptual representations) become available, even in situations where these meanings are incompatible with the established meaning and, thus, must be rejected.

References

FIORENTINO, R., & POEPPEL, D. (2007). Compound words and structure in the lexicon. *Language and Cognitive Processes*, *22,* 953-1000.

GAGNÉ, C. L., & SPALDING, T. L. (2014). Relation diversity and ease of processing for opaque and transparent English compounds. In F. Rainer, Dressler, W., Gardini, F., Luschützky, H.C. (Eds.). *Morphology and Meaning: Current Issues in Linguistic Theory* (pp. 153-162). Amsterdam: John Benjamins.

GAGNÉ, C. L., & SPALDING, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language, 60*, 20-35.

JI, H., GAGNÉ, C. L., & SPALDING, T. L. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language, 65*, 406-430.

KOESTER, D., GUNTER, T. C., & WAGNER, S. (2007). The morphosyntactic decomposition and semantic composition of german compound words investigated by ERPs. *Brain and Language, 102*, 64-79.

KOUNIOS, J., BACHMAN, P., CASASANTO, D., GROSSMAN, M., SMITH, R. W., & YANG, W. (2003). Novel concepts mediate word retrieval from human episodic associative memory: Evidence from event-related potentials. *Neuroscience Letters*, *345*, 157-60.

WORKSHOP 3:

Synchrony and Diachrony of Inflectional Classes: Theoretical and Empirical Considerations

Organizers:

Petros Karatsareas (Bristol/Nicosia)
Enrique L. Palancar (Paris/Surrey)
Timothy Feist (Surrey)

The inflectional classes of the Rtau verb

Anton Antonov Guillaume Jacques
CRLAO, EHESS, INALCO, CNRS
anton.antonov@inalco.fr rgyalrongskad@gmail.com

This paper deals with the verbal inflectional classes of Rtau (locally known as *rəsɲəske*), a Rgyalrongic language spoken in Rtau country, Sichuan province, China. The data presented here is based on ongoing fieldwork by the authors.

In Rtau, there are two main verb classes that we can call intransitive and transitive, respectively. Verbs belonging to the intransitive class have at most two different forms, one with a first person (singular or plural) reference, the other with a non-first (second or third) person reference. The first person form has only a limited array of possible rhymes: we never find front or central vowels, but only open syllable nasal rhymes *–ã* and *–õ*, velarized vowels *–oˠ* and *–aˠ* or the back rounded *–u*. In the non-first person (henceforth 2/3) form however, almost all possible rhymes are attested, including open and closed syllables (except for non-velarized *–o*). Vowel alternations (or the lack thereof) allow us to posit the existence of six inflectional classes for verbs ending in open syllables, with class 6 including verbs exhibiting no alternations, whose rhyme can be any of *–u*, *–oˠ*, *–aˠ*, *–õ* and *–ã* (Table 1). The alternations can be stated in a straightforward way: centralized vowels *–ə* and *–ɵ* change to *–õ*, and front and open (unrounded and non-velarized) vowels change to *–ã*. In the case of verb stems ending in*–r* or *–v*, the first person is always derived by replacing the entire rhyme by *–ã* or *–õ* depending on the main vowel of the rhyme. Stems ending in *–m* (the only other final consonant available) are always Tibetan loanwords and do not present any alternation.

Verbs belonging to the transitive class have at most six different forms, illustrated by the perfective (where the verb is preceded by a directional prefix, which can be neglected for the purposes of this presentation) paradigm of the verb 'to kill' (cf. Table 2). If we disregard the inverse prefix *v–/f–*, there are four distinct forms: 1SG→3; 2→3; 1PL→3 and 2/3→1 (with the inverse prefix); and 1→2 and 3→2/3 (with the inverse prefix). As with intransitive verbs, vowel alternations allow us to posit six inflectional classes, depending on the final vowel of the verb stem. Table 3 presents all six classes (verb forms are shown without the inverse prefix *v–/f–*.) Class 6 includes all verbs with stems ending in *–u*, *–oˠ*, *–aˠ*, *–õ* and *–ã*. In the case of verb stems ending in a closed syllable, final *–v* drops in first person forms; in the 2SG→3 form it stays but there is vowel fronting as in *–zgriv* 'you accomplished'; final *–r* drops in the first and second person but is preserved in the third person and in 1→2 forms while final *–m* is immune to any changes and verbs ending in this consonant present no stem alternations.

A closer look reveals that in the case of intransitive verbs the first person form is generally predictable from the 2/3 form except in a few isolated counterexamples. It is therefore possible to analyze the 2/3 form as the base form and the first person form as being derived from it by a set of morpho-phonological rules involving fusion with a suffix *–ã*, which is realized as *–õ* when the rhyme is centralized. In the case of transitive verbs, it is possible to consider the third person form as the base form; the 1PL→3 and 2/3→1 stems can be analyzed as resulting from fusion with the first person *–ã* suffix. The 1SG→3 form presents rounding of the vowels with an additional *–w* glide in the case of mid-low and low vowels. These alternations can be accounted for by assuming the existence of a suffix whose underlying form is *–w*. The 2→3 form has vowel fronting with an additional *–j* glide for mid-low and low vowels. Here the underlying form *–j* can be posited. All these morphophonological rules are summarized in Table 5. A further support for positing these vowel fusion rules is the fact that they are not restricted to the verbal system, but also apply to the ergative *–w* and genitive *–j* case markers. Table 6 illustrates some examples of vowel fusion in nouns.

Table 1: Vowel alternations in open-syllable intransitive verbs in Rtau

	1	2	3	4	5	6
meaning	look at	move	like	be full	be ill	be hot
1	*scəqã*	*mbəɕã*	*rgã*	*fkõ*	*ŋõ*	*c^{h}u*
2/3	*scəqi*	*mbəɕe*	*rga*	*fkə*	*ŋʊ*	*c^{h}u*

Table 2: *fse* 'to kill'

	1s 1p	2	3
1s		*nə-se*	*nə-sow*
1p			*nə-sã*
2	*nə-fsã*		*nə-sej*
3		*nə-fse*	

Table 3: Vowel alternations in open-syllable transitive verbs in Rtau

	1	2	3	4	5	6
meaning	drink	kill	dig	dress up	give	cut
1SG→3	*–t^{h}u*	*–sow*	*–ɴq^{h}ərow*	*–zgu*	*–k^{h}ow*	*–tsu*
1PL→3, 2/3→1	*–t^{h}ã*	*–sã*	*–ɴq^{h}ərã*	*–zgõ*	*–k^{h}õ*	*–tsu*
2→3	*–t^{h}i*	*–sej*	*–ɴq^{h}ərej*	*–zgi*	*–k^{h}ej*	*–tsu*
3→3, 1→2	*–t^{h}i*	*–se*	*–ɴq^{h}əra*	*–zgə*	*–k^{h}ʊ*	*–tsu*

Table 4: Vowel alternations in closed syllable transitive verbs in Rtau

meaning	accomplish	give back	close	rob
1SG→3	*–zgru*	*–xsow*	*–zdəm*	*–stow*
1PL→3, 2/3→1	*–zgrõ*	*–xsõ*	*–zdəm*	*–stõ*
2→3	*–zgriv*	*–xsev*	*–zdəm*	*–stej*
3→3,1→2	*–zgrəv*	*–xsev*	*–zdəm*	*–stər*

Table 5: Vowel fusion in Rtau verbs

Stem \ Suffix	1SG→3 *–w*	1 *–ã*	2→3 *–j*
i	*u*	*ã*	*i*
e	*ow*	*ã*	*ej*
a	*ow*	*ã*	*ej*
ə	*u*	*õ*	*i*
ʊ	*ow*	*õ*	*ej*

Table 6: Vowel fusion in Rtau nouns

base form	meaning	ergative	genitive
kəta	dog	*kətow*	*kətej*
vdzi	man	*vdzu*	*vdzi*
xə	hybrid of yak and cow	*xu*	*xi*

Reaching out beyond the wordform: Periphrasis and inflectional classes

MATTHEW BAERMAN
Surrey Morphology Group, University of Surrey
m.baerman@surrey.ac.uk

ENRIQUE L. PALANCAR
SeDyL-CELIA (UMR8202), CNRS
Surrey Morphology Group, University of Surrey
e.palancar@surrey.ac.uk, palancar@vjf.cnrs.fr

It is common knowledge that inflectional classes are built on inflectional distinctions based on affixal allomorphy. This understanding is in accord with a view of inflectional morphology relying on the concept of a morphological word that consists of a phonological word with morphological internal structure. But across grammatical systems, it is also common to find periphrastic forms conveying the same kind of inflectional information realized in wordforms, suggesting that such periphrastic forms could be treated as inflectional and be thus part of the paradigm (e.g. gradient forms of adjectives in English (*more/most interesting* vs. *bigger/biggest*); the perfect of the passive in Latin (*amatus est* 's/he was loved' vs. *amavit* 's/he loved'). Against the traditional view that periphrasis is a purely syntactic phenomenon, approaches to inflection from autonomous morphology or from inferential-realizational models consider periphrasis as a morphological phenomenon (e.g. Ackerman and Stump 2004, Vicent 2011, Chumakina & Corbett 2012, etc.).

In this paper, we want to extend this view of periphrasis one step further to cover inflectional classes too. We claim that the morphological information stored in a lexeme for the selection of inflectional allomorphs can have scope beyond the phonological word and reach over to the non-lexical component of a periphrastic expression. We illustrate this with two case studies on verb inflection in sufficient detail.

One case is the inflectional system of verbs in Otomi; a family of languages spoken in Mexico that belongs to the Oto-Pamean branch of Oto-Manguean. Verbs in Otomi inflect for TAM values by means of markers that we call 'inflectional formatives' and that often realize person of subject too. Example (1) illustrates the formative *tú* of the 1st p. completive realis in Tilapa Otomi. Otomi inflectional formatives are clitics, e.g. they are obligatorily hosted as enclitics on certain functional heads, like the negative in (2), or more facultatively to any preceding phrase, allowing hesitation pauses like in (3).

(1) **tú**=nde **tú**=^{h}teni='mbe
1.CPL.REALIS=want 1.CPL.REALIS=follow[3OBJ]=PL.EXCL
'we wanted to follow them.' [txt]

(2) pe hín=**dú** nde='mbe
but NEG=1.CPL.REALIS want=PL.EXCL
'but we didn't want.' [txt]

(3) kha^{h}ka'mba=**tú...** ^{h}ko='mbe
then=1.CPL.REALIS return=PL.EXCL
'Then we...returned.' [txt]

Examples like (2-3) indicate that the entire inflection of Otomi verbs is periphrastic. This system probably emerged from the grammaticalization of old auxiliary verbs which are beyond reconstruction at this stage and no longer have a synchronic correlate, making this a non-canonical type of periphrasis. What is interesting for a theory of morphology is that verbs of some Otomi languages fall into different inflectional classes because they select different inflectional formatives for the realization of a set of TAM values in the paradigm. Examples (4) show this with the subparadigm of completive irrealis, and they serve to show how allomorphic selection by the verb reaches out beyond the wordform.

(4)	CPL, IRR	Class I	Class II	Class III
		'hit'	'bathe'	'fix'
	1st	**gu** ho	**gutu** hi	**giti** hó^{h}ki
	2nd	**gi** syo	**gugu** hi	**giti** hó^{h}ki
	3rd	**ta** syo	**ti** hi	**ti** hó^{h}ki

Otomi naturally brings us to better-known instances of auxiliary selection that have, nonetheless, been treated in the literature as motivated by the syntax-semantics interface. For example, French verbs can be said to fall into two inflectional classes attending to whether they select the auxiliaries *etre* 'be' or *avoir* 'have' for their past tense. Both classes are open-ended, but only the former has intransitive verbs -known as 'inaccusatives'. The class includes all reflexive verbs together with a small subset of (mainly motion) verbs and their lexical derivatives. While this class has a semantic core (cf. Mithun 1991), the semantics no longer helps predicting the verbs of the subset, and they must be learned (i.e. *mourir* 'die' selects *etre* but *expirer* 'expire, die' or *périr* 'perish' select *avoir*, similarly *monter* 'go up' selects *etre* but *grossir* 'grow' or *pouser* 'grow (plants)' select *avoir*, etc.).

References:

ACKERMAN, FARRELL AND STUMP, GREGORY. 2004. Paradigms and periphrastic expression: a study in realization-based lexicalism, in Louisa Sadler and Andrew Spencer (eds.), *Projecting Morphology*, 111-58. Stanford, CA: Center for the Study of Language and Information.

CHUMAKINA, MARINA AND GREVILLE G. CORBETT (eds.). 2012. *Periphrasis: The Role of Syntax and Morphology in Paradigms* [Proceedings of the British Academy 180. Oxford: OUP/British Academy.

LOPORCARO, MICHELE. 2007. On triple auxiliation in Romance. *Linguistics* 45.1. 173-222

MITHUN, MARIANNE.1991. Active / agentive case marking and its motivation. *Language* 67:3, 510-546.

VINCENT, NIGEL. 2011. Non-finite forms, periphrases, and autonomous morphology in Latin and Romance, in Maria Goldbach, Marc-Olivier Hinzelin, Martin Maiden and John Charles Smith (eds.), *Morphological Autonomy: Perspectives from Romance Inflectional Morphology*, 421-439. Oxford: OUP.

Potential and actual variation in morphosyntax

NEIL BERMEL, LUDĚK KNITTL AND JEAN RUSSELL
University of Sheffield
n.bermel@sheffield.ac.uk

This talk will explore variation within two morphosyntactic 'slots' of a traditional inflectional class, and question what that says about the description of this class and its historical evolution. Although it comes from an emergentist perspective on language, emergentist enquiry into the structure of language has tended to focus on syntactic variation (ditransitivity, transitive-causative constructions, adjective suitability, phrasal verbs, see e.g. Brooks & Tomasello 1999, Ambridge *et al.* 2008, Boyd & Goldberg 2011). Such data naturally come to the fore in English, but data from other, morphologically rich languages present an opportunity to examine choice and variation in slots where 'overdifferentiation' has taken place (as in Brown 2007), but the number of variants is very limited and their form arbitrary.

Our starting point is new data from Czech, a language which offers an interesting view into the nature of paradigmatic relations. Inflection in Czech generally falls into the pattern described in the workshop call, where sets of morphosyntactic/morphosemantic features are expressed by different formants across classes of lexical items whose membership is in part arbitrary. In addition, however, within these inflectional classes, variation or competition between formants exists in multiple functional slots in paradigms, and the appearance of this variation also seems unpredictable. Štícha (2009), an extensive survey of one of these places, notes that while some general phonological and derivational tendencies in its usage can be observed, hard-and-fast rules dictating the use of one or another formant are absent, and diachronic developments also show the expansion and contraction of formant usage moving in mutually contradictory directions.

Our data on variation in the genitive and locative singular of the so-called 'masculine hard inanimate' class of nouns have been collected through trawls in 100-million-token corpora of written Czech and through large-scale (N=550) experiments probing native speakers' acceptability judgments of particular forms and their production of them in written forced-choice tasks. In each of our two cases, there is a historically innovative formant {u} that is highly productive and forms a marker of this declension class, which competes with a 'conservative' formant {a} or {ě} that is far less productive, although not unproductive, and also serves as a marker of this case in other classes. Some lexemes appear in the corpus exclusively with one of the two formants (either is possible), while a limited number of high-frequency nouns appear with both formants.

For example, in the SYN2005 corpus, we find lexemes with no (or almost no) variation:

1) *hrad* 'castle': gen. *hradu* (n=2602), not **hrada*;
2) *les* 'forest': gen. *lesa* (n=4336), not **lesu* (80 lexemes with this pattern);
3) *zámek* 'stately home': loc. *zámku* (n=2078), not **zámce*;
4) *prales* 'primeval forest': loc. *pralese* (n=192), not **pralesu* (isolated lexemes with this pattern)

In the SYN2005 corpus, variation in the genitive singular slot of this paradigm is found with 112 relatively common lexemes and in the locative singular slot we find variation with 392 lexemes, e.g.:

5) *rok* 'year': gen. *roku* (n=71041) or *roka* (n=905);
6) *betlém* 'creche': gen. *betlému* (n=33) or *betléma* (n=106)
7) *hrad* 'castle': loc. *hradu* (n=122) or *hradě* (n=1518);
8) *les* 'forest': loc. *lesu* (n=5) or *lese* (n=2845)

Initial statistical analyses of our experimental data suggest that native-speaker judgments and choices are sensitive to observed frequencies (Bader and Häussler 2009, Divjak 2008, Kempen and Harbusch 2008), and the degree of this sensitivity is for forced-choice tasks quite notable: a regression analysis using primarily these observed frequencies as factors shows significantly improved predictions vis-à-vis the null model (R^2 = 0.30 to 0.68) However, in contrast to what might be expected from the literature on language acquisition, items with low (but not zero) relative frequency continue to be produced at low (but not zero) rates and to enjoy at least middling acceptability, cp. Arppe and Järvikivi 2007. There is some further evidence for the existence of 'sub-cases' distinct to this inflectional class, which have a morphosyntactic or morphosemantic basis and seem to play a role in maintenance of the less-common formant. This is shown in the fact that the data from the two cases under investigation have different profiles, with the genitive more sensitive to frequency-based effects and regional variation, while the locative shows more influence of the form's syntactic context.

Our findings raise some questions about the way in which concepts from language acquisition have been used to model language change, i.e. the developments within the language system shared by a community of (primarily adult native) language users. Some concepts, such as entrenchment, appear to transfer well from one domain to another, while others, such as pre-emption (*inter alia* Boyd & Goldberg 2011), may need further modification to adequately describe what is happening in morphosyntactic systems.

The data seem to point not to the rationalisation of an inflectional class system as a dead weight, but to its potential for on-going reanalysis and even elaboration. This might explain why, in contrast to Brown's (2007) description of the atrophying of a once live instance of overdifferentiation, Czech presents a relatively enduring picture of morphological variation in a similar paradigm.

Touches on the following issues from the CFP: synchronic motivation and function of inflectional class systems; possible relationships of inflectional class systems with other grammatical features (phonological, syntactic, semantic); position of inflectional class features in the grammar; diachronic emergence and decay of inflectional class systems; acquisition of inflectional class systems.

References:

AMBRIDGE, B. *et al.* 2008. The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition* 106: 87–129.

ARPPE, A. & J. JÄRVIKIVI. 2007. Take empiricism seriously! In support of methodological diversity in linguistics. *Corpus linguistics and linguistic theory* 3 (1). 99–109.

BADER, M. & J. HÄUSSLER. 2009. Toward a model of grammaticality judgments. *Journal of Linguistics* 45. 1–58.

BOYD, J. & A. GOLDBERG. 2011. Learning what not to say: The role of statistical preemption and categorization in a-adjective production. *Language* 87: 55–83.

BROOKS, P. & M. TOMASELLO. 1999. How children constrain their argument structure constructions. *Language* 75: 720–738.

BROWN, D.. 2007. Peripheral functions and overdifferentiation: The Russian second locative. *Russian Linguistics* 31: 61–76.

CZECH NATIONAL CORPUS. Ústav Českého národního korpusu FF UK, Prague 2000–2010. Available on the web: <http://www.korpus.cz>.

DIVJAK, D.. 2008. On (in)frequency and (un)acceptability. In Barbara Lewandowska-Tomaszczyk (ed.), *Corpus linguistics, computer tools and applications – State of the art*, 213–233. Frankfurt: Peter Lang.

KEMPEN, G. & K. HARBUSCH. 2008. Comparing linguistic judgments and corpus frequencies as windows on grammatical competence: A study of argument linearization in German clauses. In Anita Steube (ed.), *The discourse potential of underspecified structures*, 179–192. Berlin: Walter de Gruyter.

ŠTÍCHA, F. 2009. Lokál singuláru tvrdých neživotných maskulin (*ve vlaku* vs. *v potoce*): úzus a gramatičnost. *Slovo a slovesnost* 70: 193–220,

Third-person realis morphology and verb classes in the Zamucoan languages

Luca Ciucci
Scuola Normale Superiore di Pisa
luca.ciucci@sns.it

The Zamucoan family only includes two living languages (Ayoreo and Chamacoco) spoken in the Northern Chaco, between Bolivia and Paraguay. Both languages are endangered and are spoken by approximately 4500 and 1600 people, respectively. The linguistic family also includes an extinct language, Ancient Zamuco, described in the XVIII century by the Jesuit Father Ignace Chomé.

In the Zamucoan languages the verb expresses neither tense nor aspect, but it still presents a realis / irrealis opposition, although this only shows a defective paradigm in the currently spoken languages. The verb has prefixes indicating the person category and suffixes indicating the number category (specifically, plural). Verb morphology allows the identification of different paradigms. The most conspicuous identifying feature for the various paradigms consists in the morphology of the 3-person realis (3-realis), which is characterized by remarkable morphological variation (Ciucci 2013).

The aim of this paper is precisely to show how, in these tenseless languages, the 3-realis polymorphism has led to the emergence of inflectional classes. Table 1 illustrates the verb classification of the Zamucoan languages, as based on the 3-realis marker. One can immediately note that Chamacoco is the Zamucoan language with the most complex verb morphology. From a merely descriptive point of view, one can detect the person prefix (*PP*), the thematic vowel (*TV*) and the root: *PP-TV-Root*. The first two elements may be absent. In practice, depending on the shape of the 3-realis, verbs can be classified as follows: (1) 'radical verbs', *ø-ø-Root*; (2) 'thematic verbs', *ø-TV-Root*; (3) 'prefixal verbs', *PP-TV-Root*. In addition, one can identify further subgroups depending on the type of 3-realis *PP*. In Chamacoco, one can even identify subgroups depending on the *TV*. These groups and subgroups are not related, at least sinchronically, to any semantic or valency property.

<table>
<tr><td></td><td colspan="4">prefixal verbs</td><td>thematic verbs</td><td>radical verbs</td></tr>
<tr><td rowspan="2">Ayoreo (AY)</td><td colspan="2">First class</td><td colspan="2">Second class</td><td>Third class</td><td>Fourth class</td></tr>
<tr><td colspan="2">/tɕ/-verbs</td><td colspan="2">/t/-verbs</td><td>thematic verbs</td><td>radical verbs</td></tr>
<tr><td rowspan="2">Ancient Zamuco (AZ)</td><td colspan="2">First class</td><td colspan="2">Second class</td><td>Third class</td><td>Fourth class</td></tr>
<tr><td colspan="2">/tɕ/-verbs</td><td colspan="2">/t/-verbs</td><td>thematic verbs</td><td>radical verbs</td></tr>
<tr><td rowspan="3">Chamacoco (CH)</td><td colspan="4">First macro-class</td><td colspan="2">Second macro-class</td></tr>
<tr><td colspan="2">First class</td><td colspan="2">Second class</td><td>Third class</td><td>Fourth class</td></tr>
<tr><td>/tɕ/-verbs</td><td>/ts/-verbs</td><td>/t/-verbs</td><td>/d/-verbs (/d/-, /j/-)</td><td>thematic verbs (/i/-, /u/-, /ɨ/-)</td><td>radical verbs</td></tr>
</table>

Table 1. Verb paradigms as based on the 3-realis marker.

This paper will analyse the causes that might have originated the 3-realis polymorphism and will discuss the properties of each inflectional class, which may concern, e.g., the morphological expression of: (1) the 1-person; (2) the 3-person irrealis; (3) the presence of irregularities or subregularities in the paradigm.

Comparing the verb paradigms of the Zamucoan languages, one can show that some verbs have changed their inflectional class. Specifically for Chamacoco, one can observe that new

verb groups have arisen (Ciucci 2013). For instance, consider the /tɕ/-verbs paradigm, which is the most regular and the largest group in all Zamucoan languages. As one can see from the table, this paradigm has split in two pardigms in Chamacoco (/tɕ/- vs /ts/-verbs):

(1) AZ. *tɕ-ise* (3-realis) ‘to reach’, AY. *tɕ-ise* (3-realis) ‘to reach’, CH. *tɕ-iɛ* (3-realis) ‘to reach’

(2) AZ. *tɕ-iasore* (3-realis) ‘to help’, AY. *tɕ-osõre* (3-realis) ‘to help’, CH. *ts-osɨr* (3-realis) ‘to help’

This is clearly a recent innovation, as shown by the fact that in many Chamacoco verbs /tɕ/- and /ts/- can alternate. During the talk, the person markers of the Zamucoan languages will be compared with the person marking system of other geographically adjacent languages.

Finally, the 3-realis will be compared with the 3-person of possessable nouns. There is an interesting parallelism between verb and noun morphology: as it happens, in the possessive inflection, one can distinguish ‘prefixal nouns’, ‘thematic nouns’ and ‘radical nouns’. There is however a remarkable complementarity in the Zamucoan 3-person morphology: while most verbs are ‘prefixal’ (see again the Table), most nouns are ‘thematic’.

The only 3-person prefixes common to both verbs and nouns are /d/- and /j/-, to be observed in a tiny group of Chamacoco verbs. In Ayoreo and Chamacoco the vast majority of nouns characterized by /d/- are kinship terms: this is the only semantic criterion related to 3-person morphology which one can detected in the whole Zamucoan family.

References

Bertinetto, Pier Marco 2009. Ayoreo (Zamuco). A grammatical sketch. *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore* 8 n.s.

Bertinetto, Pier Marco 2011. How the Zamuco languages dealt with verb affixes. *Word structure* 4/2. 215-220.

Bertinetto, Pier Marco (to appear). Ayoreo (Zamuco) as a radical tenseless language. *Miscellanea in onore di Alberto Mioni*.

Chomé, Ignace 1958 [ante 1745]. Arte de la lengua Zamuca. Présentation de Suzanne Lussagnet. *Journal de la Société des Américanistes de Paris* 47. 121-178.

Ciucci, Luca 2007/08. Indagini sulla morfologia verbale nella lingua ayoreo. *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore* 7 n.s.

Ciucci, Luca 2009. Elementi di morfologia verbale del chamacoco. *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore* 8 n.s.

Ciucci, Luca 2010a. La flessione possessiva dell’ayoreo. *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore* 9,2 n.s.

Ciucci, Luca 2010b. La flessione possessiva del chamacoco. *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore* 9,2 n.s.

Ciucci, Luca 2013. *Inflectional Morphology in the Zamucoan languages*. Ph.D. thesis. Pisa: Scuola Normale Superiore.

Higham, Alice, Maxine Morarie & Greta Paul 2000. *Ayoré-English dictionary.* Sanford, FL.: New Tribes Mission. 3 vols.

Kelm, Heinz 1964. Das Zamuco: eine lebende Sprache. *Anthropos* 59. 457-516 & 770-842.

Ulrich, Matthew & Rosemary Ulrich 2000. *Diccionario Ishiro (Chamacoco) – Español / Español – Ishiro (Chamacoco).* Misión Nuevas Tribus Paraguay: Asunción.

Inflection classes and orthogonal conditions

GREVILLE G. CORBETT
University of Surrey
g.corbett@surrey.ac.uk

Introduction: dimensions of generalization

Inflectional systems demonstrate the need for generalizations in the lexical 'dimension'. I review such generalizations (inflectional classes) with data from a less familiar language, namely Burmeso. I then show that we also require orthogonal (cross-cutting) generalizations in the other dimension, first for Burmeso, and then more generally. I provide a typology of these generalizations or 'conditions'; this typology is based on the antecedent of the condition and the type of paradigm affected.

Dimension 1: inflection classes

Consider the following verbal forms of Burmeso (Donohue 2001: 100, 102). They establish that we need generalizations over groups of lexemes, that is, inflection classes.

		'see' (class I)		'bite' (class II)	
	assignment	SG	PL	SG	PL
I	male	j-ihi-	s-ihi-	b-akwa-	t-akwa-
II	female, animate	g-ihi-	s-ihi-	n-akwa-	t-akwa-
III	miscellaneous	g-ihi-	j-ihi-	n-akwa-	b-akwa-
IV	mass nouns	j-ihi-	j-ihi-	b-akwa-	b-akwa-
V	banana, sago tree	j-ihi-	g-ihi-	b-akwa-	n-akwa-
VI	arrows, coconuts	g-ihi-	g-ihi-	n-akwa-	n-akwa-

The prefixal gender-number markers mark agreement with the absolutive argument. The two examples given each represent a substantial number of verbs. Donohue (2001: 101) states explicitly that inflection class membership is not predictable from obvious semantic correlations, and that the classes are approximately equal in size. We need the notion inflection class to describe Burmeso since (i) the distribution of the affixes can be described only by reference to the particular lexeme; and (ii) each affix is sufficient to predict every other within the paradigm. In fact, the inflection classes of Burmeso are close to canonical (Corbett 2009).

Dimension 2: conditions on inflection

Since we need inflection classes we might try to eliminate generalizations in the other dimension. However, Burmeso again shows that this is not possible. We need to specify gender-number syncretisms (such as: [V SG] = [IV SG] = [IV PL] = [III PL]), since they apply equally to the two inflection classes, but they involve different phonological forms. Once the need for these orthogonal conditions on inflection is accepted, they deserve full scrutiny. Arguably, they have received somewhat patchy attention. One type which *has* justifiably aroused great interest and detailed study is the morphomic pattern. And that is precisely the type found in Burmeso, since the syncretism noted has no extra-morphological justification. But this is just one type of condition, as I demonstrate.

Types of conditions on inflection

We begin with the clearest examples, and then descend a slippery slope. In Russian, a fully explicit account of noun inflection might have eight different inflection classes. However, we usually represent four inflectional classes, and specify that each contains both count and non-count nouns. The condition, then, is this: if a noun is low on the Animacy Hierarchy (antecedent) it lacks a plural sub-paradigm (consequent). This avoids two types of duplication: (i) on the formal side, there are four major inflectional classes and each of them includes count nouns and singularia tantum; (ii) on the semantic side, similarly, the distinction cross-cuts the four classes.

This condition is convincing because it is fully orthogonal to the inflectional classes. Its antecedent rests on the lexical semantics of the noun. Its consequent involves the abstract content of the paradigm (the number of cells, not their realizations). This suggests a typology, built on possible antecedents of the condition and possible consequents. Going further, the animacy condition of Russian (as in ACC=GEN *student-a* 'student' vs ACC=NOM *zakon* 'law') also has a semantic antecedent, but its consequent is different: it determines patterns of syncretism, hence the form paradigm. Our Burmeso condition has part of speech as its antecedent (all verbs) and its consequent again involves the form paradigm. Then there is the (not exceptionless) syllable-counting condition on plural formation in Serbo-Croat, giving *grad-ov-i* 'cities' versus *prozor-i* 'windows'. Here the antecedent is phonological (single vs multi-syllable stem) and the consequent (augment or none) involves the realization paradigm. The antecedent of a condition on inflection may invoke four types of information: semantic, part of speech, morphological, or phonological. The consequent may involve the content paradigm (the most abstract type of paradigm), the form paradigm, or the realization paradigm (Stump 2012). This implies a twelve-member typology of conditions (above), and it proves remarkably well attested, as demonstrated by further data from a wide range of typologically diverse languages.

	consequent		
antecedent	content paradigm	form paradigm	realization paradigm
semantic	Russian number	Russ. animacy	
part of speech		Burmeso sync.	
morphological			
phonological			SC plural

Conclusions

Morphological theory requires the availability of generalizations in two dimensions: inflection classes and orthogonal conditions on inflection. The latter deserve more consistent study. While I have sketched a typology of these conditions, we still need to investigate their distribution and relative frequency. Their diachrony is intriguing.

References

CORBETT, G. G. 2009. Canonical inflectional classes. In: Fabio Montermini, Gilles Boyé and Jesse Tseng(eds) *Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux*, 1-11. Somerville, MA: Cascadilla Proceedings Project. • DONOHUE. M. 2001. Animacy, class and gender in Burmeso. In: Andrew Pawley, Malcolm Ross & Darrell Tryon (eds) *The Boy from Bundaberg: Studies in Melanesian Linguistics in Honour of Tom Dutton*, 97–115. Canberra: Pacific Linguistics. • STUMP, G. 2012. The formal and functional architecture of inflectional morphology. In Angela Ralli, Geert Booij, Sergio Scalise & Athanasios Karasimos (eds.), *Morphology and the Architecture of Grammar: On-line Proceedings of the Eighth Mediterranean Morphology Meeting (MMM8), Cagliari, Italy, 2011*, 255-271. • WURZEL, W. 1984. *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie. • ZALIZNJAK, A. 1967. *Russkoe imennoe slovoizmenenie* Moscow: Nauka.

The singular/plural split and the making of verb inflectional classes in San Pedro Amuzgo

TIMOTHY FEIST
Surrey Morphology Group, University of Surrey
t.feist@surrey.ac.uk

ENRIQUE L. PALANCAR
SeDyL-CELIA (UMR8202), CNRS
Surrey Morphology Group, University of Surrey
e.palancar@surrey.ac.uk, palancar@vjf.cnrs.fr

Amuzgan is a small family of languages spoken in Mexico which forms one of the branches of the Oto-Manguean phylum. The Amuzgan family consists of two languages: Xochistlahuaca Amuzgo and San Pedro Amuzgo (SPA). Oto-Manguean languages are renowned for having complex inflectional systems (e.g. Baerman, forthcomming; Finkel & Stump, 2009; etc.). Amuzgan is no exception. The overall structure of the inflection of the two languages in the family is very similar, but in this paper we focus on SPA, and we base our analysis on the data in the large dictionary by Stewart & Stewart (2000) and on the grammatical description by Buck (2000).

In the inflection of many verbs in SPA, there is a paradigmatic split (i.e. a systematic differentiation) of singular subject forms and plural subject forms. While the split is a unique property of Amuzgan (i.e. it is not found elsewhere in Oto-Manguean), what makes it interesting for a theory of inflection and inflection classes is its complex array of both formal and distributional properties across the verbal lexicon.

Verbs in SPA inflect for TAM and person of the subject and they fall into two large inflection classes we treat for convenience as A and B. In (1) and (2) we illustrate each class with the incompletive (INCPL) and the irrealis (IRR) of the verbs 'stay' and 'break' (Tone is represented by numbers, 1 being the highest. Ballistic syllables are indicated by an acute accent). A-verbs have only two inflected forms for person: one for the 3rd person and another for anything else (person of subject is encoded by means of enclitics). B-verbs are more complex. Person and number are realized by a combination of affixal and prosodic formatives (differences in tone and ballistic stress). Note that the INCPL prefix *co*^{2}- is used in B-verbs only in the plural forms. This we take as a first instance of the existence of the singular/plural split in verbs.

(1) 'stay' (A)

	INCPL	IRR
1SG	co^{2}-ntho^{3}=ja^{2}	n-ntho^{3}=ja^{2}
2SG	co^{2}-ntho^{3}=ʔuʔ^{2}	n-ntho^{3}=ʔuʔ^{2}
3SG	co^{2}-ntho^{3}-in^{3}	n-ntho^{3}-in^{3}
1PL.EXCL	co^{2}-ntho^{3}=já^{13}	n-ntho^{3}=já^{13}
1PL.INCL	co^{2}-ntho^{3}=jaa^{1}	n-ntho^{3}=jaa^{1}
2PL	co^{2}-ntho^{3}=oʔ^{3}	n-ntho^{3}=ʔoʔ^{2}
3PL	co^{2}-ntho^{3}-in^{3}	n-ntho^{3}-in^{3}

Furthermore, B-verbs fall into a range of other inflectional classes attending to how the different TAM distinctions are encoded. For example, the B-verbs 'break' in (2) and 'suck' in

(3) belong to different classes because they select a different prefix for the IRR, i.e. the former takes *n-*, the latter *ngi-*.

(2)	‘break’ (B)	
	INCPL	IRR
1SG	ma^{2}-t‹ʔ›án^{12}	n-t‹ʔ›án^{12}
2SG	ma^{2}-tán(ʔ)^{13}-ʔ	n-tán(ʔ)-ʔ^{13}
3SG	i^{3}-tánʔ^{13}	n-tánʔ^{13}
1PL.EXCL	co^{2}-t‹ʔ›án^{12}	n-t‹ʔ›án^{12}
1PL.INCL	co^{2}-t‹ʔ›áan^{2}	n-t‹ʔ›áan^{2}
2PL	co^{2}-tánʔ^{2}=oʔ^{3}	n-tánʔ^{2}=oʔ^{3}
3PL	co^{2}-tánʔ^{2}	n-tánʔ^{2}

(3)	‘suck’ (B)	
	INCPL	IRR
1SG	ma^{2}-tí	ngi^{2}-tí
2SG	ma^{2}-tí-ʔ	ngi^{2}-tí-ʔ
3SG	i^{3}-tí^{3}	ngi^{2}-tí^{3}
1PL.EXCL	co^{2}-tí	ngi^{2}-tí
1PL.INCL	co^{2}-tíi	ngi^{2}-tíi
2PL	co^{2}-tí=oʔ^{3}	ngi^{2}-tí=oʔ^{3}
3PL	co^{2}-tí^{3}	ngi^{2}-tí^{3}

In addition to the distribution of the INCPL prefix *co^{2}-*, it is often the case that a verb has a distinct stem only used to build the plural forms. This can be seen in the verb for ‘twist, split’ in (4). What makes these verbs remarkable is that they inflect for the plural as if they belonged to a different inflectional class. This can be seen if the forms of the singular of the IRR are compared with the IRR of the verb for ‘eat’ in (5), and those of the plural with the verb ‘suck’ in (3). This phenomenon suggests that the singular/plural split is embedded with the lexeme itself, as if we had two different verbs in one.

(4)	‘twist, split’ (B)	
	INCPL	IRR
1SG	ma^{2}-kǫ́n^{12}	nt-kǫ́n^{12}
2SG	ma^{2}-kon^{2}-ʔ	nt-kon^{2}-ʔ
3SG	i^{3}-kon^{3}	nt-kon^{3}
1PL.EXCL	co^{2}-tǫ́n	ngi^{2}-tǫ́n
1PL.INCL	co^{2}-toon	ngi^{2}-toon
2PL	co^{2}-ton=oʔ^{3}	ngi^{2}-ton=oʔ^{3}
3PL	co^{2}-tón^{2}	ngi^{2}-tón^{2}

(5)	‘eat’ (B)	
	INCPL	IRR
1SG	ma^{2}-k‹ʔ›íia	nt-k‹ʔ›íia
2SG	ma^{2}-kiia(ʔ)-ʔ	nt-kiia(ʔ)-ʔ
3SG	i^{3}-kiiaʔ^{2}	nt-kiiaʔ^{2}
1PL.EXCL	co^{2}-k‹ʔ›íia	nt-k‹ʔ›íia
1PL.INCL	co^{2}-k‹ʔ›iia	nt-k‹ʔ›iia
2PL	co^{2}-kiiaʔ =oʔ^{3}	nt-kiiaʔ =oʔ^{3}
3PL	co^{2}-kíiaʔ^{3}	nt-kíiaʔ^{3}

With a phenomenon such as this, a number of questions arise: What is the distribution of the classes A and B in the lexicon? How many TAM classes are there? How representative or frequent is the pattern shown by the verb in (4)? What is the status of the plural stem? Is it basic (given by the lexicon)? Or is it built on the singular? In a given verb whose singular stem has certain phonological properties, is the class membership of the plural subparadigm predictable? In our paper, we will address these and other similar questions in the quest to understand where the cues for inflectional class structure can reside.

References:

BAERMAN, M. Forthcoming. “Inflectional class interactions in Oto-Manguean”, *in* J-L. Léonard and A. Kihm (eds.), *Issues in Meso-American morphology*. Paris: Michel Houdiard. ● BUCK, MARJORIE J. 2000. "Gramática Amuzga", *in* Cloyd Stewart and Ruth D. Stewart. 2000, *Diccionario Amuzgo de San Pedro Amuzgos, Oaxaca*. [Serie de vocabularios y diccionarios indígenas Mariano Silva y Aceves, 44], 363-473. Mexico City: Instituto Lingüístico de Verano. ● FINKEL, R AND G STUMP. 2009. “Principal parts and degrees of paradigmatic transparency”, *in* J. P. Blevins and J. Blevins (eds), *Analogy in Grammar: Form and Acquisition*. Oxford: Oxford University Press, 13-53. ● STEWART, CLOYD AND RUTH D. STEWART (eds.). 2000. *Diccionario Amuzgo de San Pedro Amuzgos, Oaxaca*. [Serie de vocabularios y diccionarios indígenas Mariano Silva y Aceves, 44]. Mexico City: Instituto Lingüístico de Verano.

Emergence and decay of inflectional class systems—an evolutionary perspective

FRANCESCO GARDANI
WU Vienna
francesco.gardani@wu.ac.at

Productivity is a central property of inflectional classes and, as such, decisive in order to understand both the synchronic organization of inflectional systems and their evolution. Based on an investigation of the nominal inflectional classes of Latin and Old Italian (Gardani 2013), this paper investigates the role that productivity plays in the dynamics of emergence and decay of inflectional class systems. It thus offers insights into long-term morphological change, covering a time of approximately 2,000 years.

On a wider perspective, the paper follows the line of research on the evolution of morphology recently propelled forwards by Carstairs-McCarthy (2010). More specifically, the investigation is theoretically couched in the functionalist framework of Natural Morphology (Dressler 2003), from which also the methodological approach is derived. In the model proposed, productivity is conceived of as a force of attraction peculiar to inflectional classes which is reflected both in the morphological integration of loanwords, in the inflectional class assignment of indigenous neologisms arisen via conversion, and in the behavior of class shifters. The measurement of productivity is operated on semi-synchronic cuts, and the diachronic trajectory is outlined by connecting the results of the single cuts. The data on loanword integration are drawn from the contact languages, Ancient Greek, Germanic varieties, Arabic, Byzantine Greek and Old French, thus reflecting different chronological depths in the analysis. The corpora are constructed on a variety of sources, including the *Thesaurus Linguae Latinae,* for Latin, and the *OVI* corpus, for Old Italian.

This paper shows that there exists a link between the lifecycle of inflectional classes and a universal preference for biunique form-meaning relationships, with respect to the morphological realization of certain morphosyntactic features. For example, provided that a given language is sensitive to the feature of number, an inflectional class which realizes the values of number by means of formatives which are biunique (or rank high on a scale ranging from biuniqueness to ambiguity), is more likely to undergo an increase in its grade of productivity than a class whose formatives are ambiguous (or rank low on a scale ranging from biuniqueness to ambiguity) with respect to the realization of the same feature value(s). Also, the paper demonstrates that the need for a (progressively more) biunique realization of number values can promote a process of morphogenesis. In Old Italian, starting from 1200, three new nominal inflectional classes emerged, viz. *poeta poeti* ‘poet’, *promessa promessora* ‘promise’, and *nome nomora* ‘name’, and were quite productive. As the data show, these classes arose from extant classes whose formatives were ambiguous with respect to the realization of both the singular and the plural, via replacement of those formatives by formatives that realized plural in a unique way.

References

CARSTAIRS-MCCARTHY, A. 2010. *The Evolution of Morphology*. Oxford: Oxford University Press.

DRESSLER, W. U. 2003. Degrees of grammatical productivity in inflectional morphology. *Italian Journal of Linguistics* 15(1). 31–62.

GARDANI, F. 2013. *Dynamics of Morphological Productivity. The Evolution of Noun Classes from Latin to Italian*. Leiden & Boston: Brill.

OVI = *Corpus OVI dell’italiano antico*. 2005–. CNR. Available at http://gattoweb.ovi.cnr.it/.

TLL = 1900–. *Thesaurus Linguae Latinae*. Editus auctoritate et consilio Academiarum quinque Germanicarum, Berolinensis, Gottingensis, Lipsiensis, Monacenis, Vindobonensis. Lipsiae: B.G. Teubner.

Reflexes and remnants of inflectional classes in Maay nominal morphology

MARY PASTER
Department of Linguistics and Cognitive Science, Pomona College
mary.paster@pomona.edu

This paper describes aspects of nominal morphology in Maay, a language of southern Somalia. Dialects of closely related Somali are described as having inflectional classes, but in Maay these classes have broken down or been rearranged. In this paper, I focus on two aspects of nominal morphology – gender and plural marking – that exhibit inflectional class behavior in other East Cushitic languages. I discuss inflectional class patterns observed in those languages, then describe the remnants of the corresponding systems in Maay. I discuss how the inflectional class systems in the other languages, likely inherited, have changed in Maay and propose some possible mechanisms and explanations for the changes.

The Ethnologue classifies Maay (Paster 2007, Comfort & Paster 2009, Paster 2010) as East Cushitic. Tosco (2012: 278) advocates for a smaller Omo-Tana group within East Cushitic, and within this an Eastern ('Somaloid') subgroup including Maay, Somali, Rendille, and Tunni, among others. Eastern Omo-Tana (EOT) languages have gender class systems wherein every noun is masculine or feminine. Masculine nouns are identified (in the singular) by *k*-initial variants of suffixes including the definite marker, demonstratives, and possessive markers; feminine nouns take *t*-initial versions of these suffixes. Nominal gender appears to be relatively stable within the group, except for some interesting changes that appear to be in progress currently in some Maay nouns, since speakers disagree on their gender. As is widely known, however, the apparent gender of nouns in these languages varies in the plural in EOT languages. Rendille (Oomen 1981), for example, has a gender polarity pattern wherein nouns change to the opposite gender in the plural. In Standard Somali, most nouns switch gender in the plural (Saeed 1987). In Central Somali (Saeed 1982), plurals in *-o* show complete polarity, while plurals in *-yal* have masculine gender regardless of their gender in the singular (see below for more on the *-o* and *-yal* plural suffixes). The unidentified dialect of Somali discussed by Lecarme (2002) exhibits the reverse of the Central Somali pattern: plurals in *-yal* exhibit polarity, while plurals in *-o* are masculine. Maay has yet another variation on this pattern: all plural nouns are masculine, regardless of which plural suffix is used (1). This is apparently true of Tunni as well (Tosco 1997: 43). It seems likely that true polarity was present in Proto-EOT and that the masculine plural patterns were innovated, since polarity also exists elsewhere in Cushitic outside this group, e.g., in Oromo (Andrzejewski 1960).

Languages in this group, especially Somali, are also documented as having declension classes for plural formation, although as Diriye (2000) notes, there is no consensus on the number of classes even in a given dialect; cf. Andrzejewski (1964, 1979), Hyman (1981), Banti (1988), Morin (1991), Saeed (1999). In Maay, the declension classes have been completely reorganized: there are only two productive plural markers, and their distribution for different speakers ranges from somewhat to completely phonologically determined. The distribution of plural markers in Maay appears chaotic when data from multiple speakers are aggregated, but individual speakers have clearly describable systems of plural formation that suggest a change in progress to a completely phonologically based system wherein the *-o* suffix occurs with consonant-final stems while the *-yal* suffix occurs with vowel-final stems or optionally with consonant-final stems in place of *-o* (Paster 2013). An interesting pattern is documented where both plural suffixes are used simultaneously on a single noun, with no consistent meaning change relative to singly-marked plurals (Paster 2010); a further development seems to be underway where multiple plural marking indicates a larger number (2). I suggest that the

phonologically based pattern in Maay is emergent in response to the breakdown of an earlier declension class system – perhaps due to the introduction of non-native lexical items that lacked assignment to a declension class, since Maay has many apparently borrowed nouns that are not cognate with Somali.

Examples

(1) Masculine plurals in Maay (examples simplified to phonemic notation)

a. Feminine nouns in singular — Feminine → masculine in plural

d͡ʒeer-tey	'my hippo'	d͡ʒeer-o-key	'my hippos'
gewer-tey	'my daughter'	gewer-o-key	'my daughters'
mindi-tey	'my knife'	mindi-yal-key	'my knives'

b. Masculine nouns in singular — Masculine stays masculine in plural

awo-key	'my grandfather'	awo-yal-key	'my grandfathers'
bakaile-key	'my rabbit'	bakaile-yal-key	'my rabbits'
ad͡ʒir-key	'my thigh'	ad͡ʒir-o-key	'my thighs'

(2) Plural marking in Maay

a. Vowel-final stems take only *-yal*

buundo-yal	'bridges'	muata-yal	'ducks'	mindi-yal	'knives'

b. Pattern I: C-final stems take *-o ~ -yal ~ -o-yal*

mukulal-yal	~ mukulal-o	~ mukulal-o-yal	'cats'
eey-yal	~ eey-o	~ eey-o-yal	'dogs'

c. Pattern II: C-final stems take *-o ~ -yal* vs. *-o-yal*, distinguishing the number of items

luɣ-o	~ luk-yal	'feet'	luɣ-o-yal	'many, many feet'
geeð-o	~ geet-yal	'trees'	geeð-o-yal	'many, many trees'

References

ANDRZEJEWSKI, B. 1960. The categories of number in noun forms in the Borana dialect of Galla. *Africa* 30: 62-75. • ANDRZEJEWSKI, B. 1964. *The Declensions of Somali Nouns*. London: School of Oriental and African Studies. • ANDRZEJEWSKI, B. 1979. The Case System in Somali. London: School of Oriental and African Studies. • BANTI, G. 1988. Two Cushitic systems: Somali and Oromo nouns. Pp. 11-49 in H. van der Hulst and N. Smith, eds. *Autosegmental Studies on Pitch Accent*. Dordrecht: Foris. • COMFORT, J. & M. PASTER. 2009. Notes on Lower Jubba Maay. Pp. 204-216 in M. Matondo, F. McLaughlin and E. Potsdam, eds. *Selected Proceedings of the 38th Annual Conference on African Linguistics*. Somerville, Massachusetts: Cascadilla Proceedings Project. • DIRIYE ABDULLAHI, M. 2000. *Le Somali, dialectes et histoire*. PhD thesis, University of Montreal. • HYMAN, L. M. 1981. Tonal accent in Somali. *Studies in African Linguistics* 12.2: 169-201. • LAMBERTI, M. 1985. *Die Somali-Dialekte.* Hamburg: Buske. • LECARME, J. 2002. Gender 'polarity': Theoretical aspects of Somali nominal morphology. Pp. 109-141 in P. Boucher and M. Plénat, eds. *Many Morphologies*. Massachusetts: Cascadilla Press. • MORIN, D. 1991. Marques et relations en Somali: le problème des déclinaisons. *Linguistique africaine* 6: 75-102. • OOMEN, A. 1981. Gender and plurality in Rendille. *Afroasiatic Linguistics* 8.1: 35-75. • PASTER, M. 2007. Aspects of Maay phonology and morphology. *Studies in African Linguistics* 35.1: 73-120. • PASTER, M. 2010. Optional multiple plural marking in Maay. Pp. 177-192 in F. Rainer, W.U. Dressler, D. Kastovsky, and H. C. Luschützky, eds. *Current Issues in Linguistic Theory* 310*: Variation and Change in Morphology*. Amsterdam: Benjamins. • PASTER, M. 2013. An I-language approach to inter-speaker variation in Maay. Ms., Pomona College. • SAEED, J.I. 1982. *Central Somali: A Grammatical Outline. Monographic Journals of the Near East: Afroasiatic Linguistics* 8.2. Malibu, California: Undena Publications. • SAEED, J.I. 1987. *Somali Reference Grammar*. Wheaton, Maryland: Dunwoody Press. • SAEED, J.I. 1999. *Somali*. Amsterdam/Philadelphia: John Benjamins Publishing Company. • TOSCO, M. 1997. *Af Tunni*. Köln: Rüdiger Köppe. • TOSCO, M. 2012. The unity and diversity of Somali dialectal variants. Pp. 263-280 in N.O. Ogechi, J.A.N. Oduor, and P. Iribemwangi, eds. *The Harmonization and Standardization of Kenyan Languages: Orthography and Other Aspects*. Cape Town: Centre for Advanced Studies of African Society.

Verb morphology and conjugation classes in Dunan (Yonaguni)

THOMAS PELLARD
CRLAO, EHESS, INALCO, CNRS
thomas.pellard@gmail.com

MASAHIRO YAMADA
Kyoto University
yamada.masahiro.4r@kyoto-u.ac.jp

Japanese and other Japonic languages have a relatively simple and transparent verb morphology, with few non-canonical phenomena. They generally exhibit a highly agglutinative structure with little morphophonology, few conjugation classes, limited stem allomorphy, and very few irregular verbs. This explains why morpheme-based constructive approaches to morphology have been popular for the description of these languages. However, the verb morphology of Dunan (a.k.a Yonaguni, ISO 639-3: yoi), a highly endangered Ryukyuan language spoken by around 500 people on Yonaguni Island (Okinawa pref., Japan), is much more complex and departs in several interesting ways from the simpler system exhibited by its relatives. In particular, the existence in Dunan of a rich system of conjugation classes, of a rather high degree of allomorphy, and of several non-canonical phenomena, is noteworthy. The analysis presented here is based on original data gathered during fieldwork that fill important gaps in previous descriptions.

The same morphosyntactic features are not marked uniformly for all verbs in Dunan, and several different patterns partition verbs into conjugation classes. The exact number of these (more than a dozen) depends upon the amount of morphophonology one is ready to posit, but in any case it is a function of three mostly independent factors: *a*) stem allomorphy, *b*) suffix allomorphy, and *c*) tone alternation (metatony). Class membership is historically correlated to segmental shape, but this has been largely obscured by historical changes, and a synchronic account needs to refer to several principal parts. A word-and-paradigm (W&P) analysis based on principal parts has more merits than a morpheme-based approach, which requires recourse to abstract underlying forms and a set of complicated rules of limited scope. The morpheme-based approach, far from illuminating the structure of Dunan's verb morphology, rather obscures its alternation patterns and its network of implicative relationships between verb forms. A word-based approach is anyway required for irregular and fully suppletive patterns (e.g. 'be, exist' *an*, negative stem *minu*), which escape a cogent morpheme-and-rule treatment. A W&P approach has thus the merit of allowing a unified treatment of all verbs.

Unlike the typical case of Indo-European languages, in Dunan verbs can be partitioned and grouped together primarily according to their pattern of stem variation rather than suffix allomorphy, which is less prominent. A verb can have up to three different stem forms, which often exhibit a reduction/augmentation pattern, e.g. 'remember' *ubw* ~ *ubui* ~ *ubuir*, 'drop' *ut* ~ *utu* ~ *utus*. Such alternations are phonologically unmotivated but are the result of historical change, and they are hardly amenable to a purely phonological treatment. It does not seem desirable for example to posit a [DORSAL] → [CORONAL] rule which applies in front of the medial verb suffix *-i* but not in front of the imperative *-i* (e.g. 'pull' *sunt-i* vs. *sunk-i*).

Verb stems in Dunan underlie sets of forms that do not realize any coherent set of morphosyntactic features. Thus no common feature can be found between the imperative and circumstantial forms, even though they systematically share the same stem in all conjugation classes. Stems are thus best viewed as purely morphological objects.

Though each verb has only a restricted set of stem forms, the total number of stems needed to be posited in order to account for all stem alternations across the different classes is far greater. This is because not only stem alternations depend on conjugation class, but the distribution of stems within a paradigm is also class-specific. Thus, while the shortest stem *ubw* of 'remember' is used in its perfect forms (e.g. *ubwan*), for 'drop', it is the longest one *utus* that is found (e.g. *utusyan*). The situation is reversed for negative forms (e.g. *ubuir-anun* vs. *ut-anun*), and there is

globally no uniform distribution of stems within paradigms. For example the same stem *utu* is found in the imperative (*utu-i*) and the prohibitive (*utu-nna*) for 'drop', but two different stems *ubuir* and *ubui* are found for 'remember' (*ubuir-i* vs. *ubui-nna*).

Suffix allomorphy is also observed, though it is less developed than stem allomorphy, and it is not predictable from stem alternation patterns. This concerns the simple present *-u* ~ *-∅*, the circumstantial *-uba* ~ *-iba*, or the perfect *-ya* ~ *-a* ~ *-yu* ~ *-u*. Such variation is not fully correlated with patterns of stem alternation but further subclassifies Dunan verbs into conjugation classes.

Dunan has three word-tones (H, L, and F) which are lexically determined, but some verbs undergo tonal alternations in some inflectional forms (Tab. 1). Six patterns of paradigmatic metatony can be distinguished, which can be reduced to four: a single non-alternating class and three L ~ F alternating classes. These tones are melodies which apply to whole word forms, and they cannot thus be segmented into a stem part and an affixal part. Metatonic forms do not always share a common segmental stem-shape, and putting the burden of metatony on stems would only lead to multiply them. On the other hand, metatonic forms do not share a common base of morphosyntactic features, and therefore metatony itself cannot be considered to be an exponent but should recognized as a purely morphological feature.

Table 1: Metatony in Dunan conjugations

	I	II	III	IV	V	VI
Imperative					L	
Prohibitive				L		
Hortative			L			
Circumstantial						
Simple Present	H	L				F?
Simple Past						
Negative					F	
Perfect				F		
Medial			F			
Sequential						

Interestingly, several conjugation classes are non-canonical in the sense that they are barely distinguished from others. The perfect is sometimes the only form that discriminates between two classes, with some extreme examples such as 'boil' and 'get boiled', which are completely homophonous except for their perfect forms, i.e. *nyan* vs. *nyun*. Perfect forms exhibit however some variation, which threatens the distinction between several classes, as some speakers seem to be leveling the -(*y*)*a*/-(*y*)*u* allomorphy to -(*y*)*a*. This can be straightforwardly explained by the fact that perfect forms are less common and that the -(*y*)*u* forms are in minority. Contrary to an approach based on principal parts and implicative relationships, it does not seem that a diacritic approach to conjugation classes can offer an insightful account of this change.

Many "impostors" are also found, and many forms are ambiguous about class membership when taken isolatedly. Even the perfect is ambiguous by itself, and a *Xsyan* form could belong to two completely different classes. Only reference to other forms, such as the simple present (e.g. *Xn* or *Xsirun*) can resolve the ambiguity. A systematic homophony is also found between etymologically related transitivity pairs for some forms (e.g. *utun* present perfect of 'fall' or simple present of 'drop'), between the medial and imperative forms for most classes (e.g. 'get up' *that-i*), and some perfect forms are ambiguously marked with a suffix *-u*, which with a different stem is an exponent of the simple present. The existence of ambiguous forms shows that simply listing the presence of a stem or a word form for a verb is insufficient: the key information is to know what paradigm cell such a form realizes.

Selected references

BLEVINS, J. P. 2006. Word-based morphology. *Journal of Linguistics* 42: 531–573. • CORBETT, G. G. 2009. Canonical inflectional classes. In MONTERMINI, F., BOYÉ, G. & JESSE, T. (eds.) *Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux*. Somerville: Cascadilla Proceedings. • CORBETT, G. G. & BAERMAN, M. 2006. Prolegomena to a typology of morphological features. *Morphology* 16(2): 231–246. • STUMP, G. & FINKEL, R. A. 2013. *Morphological typology: From word to paradigm.* Cambridge: Cambridge University Press. • UWANO, Z. 2011. Yonaguni hōgen dōshi katsuyō no akusento shiryō (2). *Kokuritsu Kokugo kenkyūjo ronshū* 34: 135–164. • YAMADA, M. PELLARD, T. & SHIMOJI, M., 2013. Dunan (Yonaguni)-go no kan'eki bunpō to shizen danwa shiryō. In TAKUBO, Y. (ed.) *Ryūkyū rettō no gengo to bunka*, Tokyo: Kuroshio shuppan, 291–324.

Morphomic stems or inflectional classes?

Noura Ramli and Andrew Spencer
Department of Language and Linguistics, University of Essex, UK
{nramli, spena}@essex.ac.uk

Inflectional classes (ICs) are typically defined in terms of affix allomorphy and are generally taken to require arbitrary (morphomic) labels: [CLASS:{I, II, III, …}]. But paradigm complexity can also result from stem allomorphy and this is not normally described in terms of ICs. Bonami & Boyé (2003) have argued that French lacks conjugation classes and that instead, complexity is defined in terms of stem sets. Additionally, it has been argued that stem alternations can serve as realizations of morphosyntactic property sets (MPSs), without necessarily resorting to non-concatenative inflectional morphology (Baerman & Corbett 2012). However, Spencer (2012) has argued that stem must be morphomic in PFM, defined by stem formation rules and selected by stem selection rules ('Block 0'), not by realization rules. Stems/root alternants can systematically realize inflections only through non-concatenative or prosodic exponents operating as morphophemically defined realization rules ('Block I, II, ...'), e.g. the passive *u-i* vocalism in Classical Arabic verbs of the f-ɛ-l class (*kataba* ⇒ *kutiba* 'it was written') or the phenomenally complex prosodic 'inflectional series' found in Chinantecan conjugation (Palancar, to appear) (and, arguably Nuer, Baerman 2012, which we will briefly re-analyse).

Traditional descriptions of Arabic lack ICs but modern dialects have developed stem alternations which partition the lexicon into classes (Camilleri 2012). We illustrate with Transitional Libyan Arabic. Verbs inflect for PFV/IPFV series. Inflections are identical for all verbs for a given series. In 'weak' verbs, stem allomorphy distinguishes 3rd person PFV forms from 1/2 (Table 1). This might be analysable as multiple exponence of person/number MPSs (Baerman & Corbett 2102). However, in 'sound' verbs the stem allomorphy distinguishes 3PL/3SG.F from 1/2 and the 3SG.M forms, a morphomic distribution (Table 2). The complexity of the system in terms of conditional entropy (Ackerman et al. 2009) is similar to that of an IC system defined by affixal homophony. We provide two analyses within PFM (Stump 2001), one in terms of ICs, the other in terms of hierarchically organized morphomic stems, lexically defined stem classes and stem selection rules in both cases treating the stems as morphomic. We (tentatively) conclude that these give the same results, but since all accounts need stem formation/selection anyway we should favour the stem-based analysis and reject ICs.

Recently Stump (2012) has argued from Sanskrit conjugation for four types of IC including a 'metaconjugational' class, in which for some lexemes a given IC determines one set of MPSs (e.g. imperfect) and for others it determines another (e.g. aorist). But the endings are the same for all lexemes: the variation is entirely defined by stem allomorphy. Stump requires these stem patterns to be 'conjugational' , however, because his rules of paradigm linkage crucially (it seems) appeal to arbitrary IC labels. We reanalyse his data in terms of stem selection.

We tentatively conclude that stem-based complexity is distinct from exponence-based complexity, offering further support for the autonomy of morphology. Future research will focus on the search for a clear characterization of the two types.

Table 1: Weak verbs perfective series

	hollow		defective
	ɣa:b 'be absent'	ʒa:b 'bring'	ʕadda 'depart' (Form II)
1SG	ɣub-(i-)t	ʒib-(i-)t	ʕaddee-t
1Pl	ɣub-na	ʒib-na	ʕaddee-na
2SG. M	ɣub-(i-)t	ʒib-(i-)t	ʕaddee-t
2SG.F	ɣub-ti	ʒib-ti	ʕaddee-ti
2PL.M	ɣub-tu	ʒib-tu	ʕaddee-tu
2PL.F	ɣub-tin	ʒib-tin	ʕaddee-tin
3SG.M	**ɣa:b**	**ʒa:b**	**ʕadda**
3SG.F	**ɣa:b-it**	**ʒa:b-it**	**ʕadd-it**
3PL.M	**ɣa:b-u**	**ʒa:b-u**	**ʕadd-u**
3PL.F	**ɣa:b-in**	**ʒa:b-in**	**ʕadd-in**

Table 2: 'Sound' verbs perfective series

	laʕab 'play'	ʕaraf 'know'	nḍarab 'be hit' (Form VI)
1SG	*laʕab-t*	*ʕaraf-t*	*nḍarab-t*
1Pl	*laʕab-na*	*ʕaraf-na*	*nḍarab-na*
2SG. M	*laʕab-t*	*ʕaraf-t*	*nḍarab-t*
2SG.F	*laʕab-ti*	*ʕaraf-ti*	*nḍarab-ti*
2PL.M	*laʕab-tu*	*ʕaraf-tu*	*nḍarab-tu*
2PL.F	*laʕab-tin*	*ʕaraf-tin*	*nḍarab-tin*
3SG.M	*laʕab*	*ʕaraf*	*nḍarab*
3SG.F	**liʕb-it**	**ʕurf-it**	**nḍurb-it**
3PL.M	**liʕb-u**	**ʕurf-u**	**nḍurb-u**
3PL.F	**liʕb-in**	**ʕurf-in**	**nḍurb-in**

Selected references •Ackerman, F et al. 2009. 'Parts and wholes'. In J. Blevins and J. Blevins (eds) *Analogy in Grammar*, 54-82, OUP •Baerman M 2012. 'Paradigmatic chaos in Nuer'. *Language* 88:467-94 •Baerman M & G Corbett 2012. 'Stem alternations and multiple exponence' *Word Structure* 5:52-68 •Bonami, O & G Boyé 2003. 'Supplétion et classes flexionnelles dans la conjugaison du français' *Langages* 152:102-26 •Camilleri, M 'Morphological Complexity in Maltese' 2012. *Proc MMM8*, 92-113 •Palancar E (to appear) 'Revisiting the complexity of the Chinantecan verb conjugation classes' in Léonard & Kihm (eds) *Issues in Meso-American Morphology*. Michel Houdiard •Spencer A 2012. 'Identifying stems' *Word Structure* 5:88-108 •Stump, G 2012. 'The formal and functional architecture of inflectional morphology' *Proc MMM8*, 254-68.

Interrelation of grammatical gender and inflectional class:
A case study of Russian 'common gender' nouns

OLGA STERIOPOLO
ZAS, Berlin
olgasteriopolo@hotmail.com

There are two conflicting claims concerning correlation between grammatical gender and inflectional class of Russian nouns. Some claim that grammatical gender can be predicted form inflectional class (Corbett 1982, 1991; Corbett & Fraser 2000). In contrast, others claim that inflectional class can be predicted from grammatical gender (Crockett 1976, Thelin 1975). However, there is a class of nouns in Russian — the so-called "common gender" nouns — which cannot be accounted for in any of these proposals.

Common gender nouns denote individuals, like *s'irot-á* 'orphan' and *sudj-á* 'judge' that can trigger either masculine, or feminine agreement (1a–b). Compare with other Russian nouns that can trigger only masculine or only feminine agreement (2–3).

(1) a. bol'š-ój s'irot-á
big-**MASC**.N.SG orphan-N.SG (**MASC; CLASS II**)
'big orphan'

b. bol'š-ája s'irot-á
big-**FEM**.N.SG orphan-N.SG (**FEM; CLASS II**)
'big orphan'

(2) a. bol'š-ój j'únoš-a
big-**MASC**.N.SG young-N.SG (**MASC; CLASS II**)
'big young person (male)'

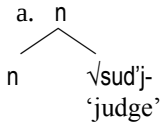
b. *****bol'š-ája j'únoš-a
big-**FEM**.N.SG young-N.SG (**FEM; CLASS II**)
'big young person (male)'

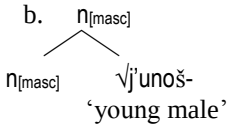
(3) a. bol'š-ája n'án'-a
big-**FEM**.N.SG nurse-N.SG (**FEM; CLASS II**)
'big nurse'

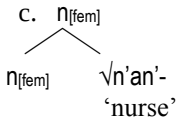
b. *****bol'š-ój n'án'-a
big-**MASC**.N.SG nurse-N.SG (**MASC; CLASS II**)
'big nurse'

Common gender nouns differ from other Russian nouns, because their grammatical gender cannot be predicted from inflectional class and vice versa, their inflectional class cannot be predicted from grammatical gender. For example, in (1a–b), the common gender noun *s'irot-á* 'orphan' belongs to the inflectional class II, which is evident from the inflectional suffix *–a* (only class II nouns have the inflectional suffix *–a* in Russian). Nonetheless, it can be either masculine or feminine, which is evident from either masculine (1a) or feminine (1b) agreement. With this respect, the question arises: What is so special about common gender nouns, which sets them apart from all other nouns?

I propose that the difference between common gender nouns and other Russian nouns is that the former are unmarked for grammatical gender, while the latter are marked for either [MASC], or [FEM] (4a–c).

(4) a. 

b. 

c. 

This proposal makes the following predictions. First, in case of common gender nouns, either masculine or feminine agreement can be used when they refer to an individual

whose sex is unknown. However, when the sex of an individual is known, one or the other agreement will be used. This prediction is borne out (5a–b).

(5) a. bol'š-ój s'irot-á (referring to a male individual)
big-**MASC**.N.SG orphan-N.SG (**MASC; CLASS II**)
'big orphan'

b. bol'š-ája s'irot-á (referring to a female individual)
big-**FEM**.N.SG orphan-N.SG (**FEM; CLASS II**)
'big orphan'

Second, in case of other Russian nouns, we predict that it would not matter whether the sex of an individual they refer to is known or not. If they are marked for [MASC], they will only trigger masculine agreement, regardless of the sex of an individual they refer to. Similarly, if they are marked for [FEM], they will only trigger feminine agreement. This prediction is borne out (6a–b).

(6) a. bol'š-ája n'án'-a (referring to a male individual)
big-**FEM**.N.SG nurse-N.SG (**FEM; CLASS II**)
'big nurse'

b. bol'š-ája n'án'-a (referring to a female individual)
big-**FEM**.N.SG nurse-N.SG (**FEM; CLASS II**)
'big nurse'

I show in detail how the current proposal works with respect to Russian expressive suffixes. First, I argue that expressive suffixes cannot be marked for grammatical gender in Russian. Instead, they are marked for the inflectional [CLASS II]. Second, I argue that the grammatical gender of a derived word can be predicted from the inflectional class of an expressive suffix; which gives additional evidence for the first claim discussed above (grammatical gender can be predicted from inflectional class). Third, I illustrate how this works with respect to common gender nouns. I argue that since common gender nouns are unmarked for grammatical gender, their gender cannot be predicted from the inflectional class of an expressive suffix. Thus, a derived common gender noun with an expressive suffix will always trigger either masculine or feminine agreement (7), just as it does without an expressive suffix (1).

(7) a. bol'š-ój s'irot'-**ín**-a
big-**MASC**.N.SG orphan-**EXPR**-N.SG (**MASC; CLASS II**)
'big orphan (expressive)'

b. bol'š-ája s'irot'-**ín**-a
big-**FEM**.N.SG orphan-**EXPR**-N.SG (**FEM; CLASS II**)
'big orphan (expressive)'

References

Corbett, G. (1982). Gender in Russian: An account of gender specification and relationship to declension. *Russian Linguistics, 6*(2), 197–232. Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press. Corbett, G., & Fraser, N. (2000). Default genders. In B. Unterbeck, & M. Rissanen (Eds.), *Gender in grammar and cognition* (pp. 55-98). Berlin; New York: Mouton de Gruyter. Crockett, D. (1976). *Agreement in Contemporary Standard Russian*. Slavica: Cambridge University Press. Thelin, N. (1975). Notes on general and Russian morphology. *Studia Slavica Upsaliensia, 15*, 30-32.