Contents

Invited Speakers

Volker Gast
Towards an Empirical Study of Sentence Semantics: Examples from the Domains of Scalar Operators and Concessives

Ray Jackendoff and Jenny Audring
Morphology in the Mental Lexicon

David Pesetsky
Exfoliation: Towards a Derivational Theory of Clause Size

Presentations
(in alphabetical order according to the last name of the first author)

Gábor Alberti and Judit Farkas
Operator Projections Moving as Remnants in Hungarian

Jwahir Alzamil
Examining the Fluctuation Hypothesis at Early Stages of English L2 Article Acquisition: The Role of Instruction and Metalinguistic Feedback

András Bárány
Global Case Splits and Cyclic Agree: Arguments for Dissociating Case and Agree

Pilar Barbosa, Cristina Flores, and Ana Bastos-Gee
Variable Use of Strong Preterites in European Portuguese: A Quantitative and Theoretical Approach

Laura Becker and Matías Guzmán Naranjo
Psych Predicates from a Cross-Linguistic Perspective

Anne Beshears
Licensing the Correlative Construction via the Semantics of the Demonstrative

Anna Bondaruk, Bożena Rozwadowska and Wojciech Witkowski
Constraints on Psych Verb Passivisation in Polish

Pavel Caha
Bobaljik’s Root Suppletion Generalization as an Instance of the Adjacency Condition

Anna Čermáková, Zuzana Komrsková, Marie Kopřivová, and Petra Poukarová
Causal Connective protože across Registers
Bożena Cetnarowska
Polish Derived Nominals with Thematic Adjectives as Argument-Supporting Nominals

Lucie Chlumská
The Myth of Comparability? Issues in the Design and Analysis of a Monolingual Comparable Corpus

Anna Chromá and Filip Smolík
Person Reference and Social Understanding: Laboratory Data from Czech Toddlers

Jan Chromý, Štěpán Matějka, and Jakub Dotlačil
Do Bi-Aspectual Verbs in Czech Express Grammatical Aspect?

Alexandra Cornilescu
On Dative Clitics and Nominalizations in Romanian

José Deulofeu
On the Border between Relative and “Consecutive Clauses” Constructions: The Case of [NP Det [-def] N que X] Construct in Spoken French: A Corpus Based Analysis

Mojmír Dočekal
Constituent Negation, Subjunctive and Alternatives

Tomáš Duběda
Living on the Edge: Integration vs. Modularity in the Phonology of Czech Anglicisms

Bożena Duda
Nocturnal and Vizard as Late Modern English Metonymy-Based Synonyms of Prostitute

Katalin É. Kiss and Erika Asztalos
Discourse-Motivated Word Order Variation in Udmurt

Patrick Elliott
Selection without Syntax: Explaining the Difference between Clauses and Nominal Phrases

Patrick Elliott and Andrew Murphy
Non-Intersective Readings in the Slavic Noun Phrase: Diagnosing Two Varieties of Left Branch Extraction

Joseph E. Emonds
How Formal Lexical Entries Produce Paradigms: The French Clitic System

Lenka Fárová
Reporting Verb Said and Its Czech Translations
Abdelkader Fassi Fehri
New Roles and Architecture for Gender: Evidence from Semitic, Berber, and Romance

Ludovico Franco, Giulia Bellucci, Lena Dal Pozzo, and Rita Manzini
Locatives, Part and Whole in Uralic

Rafał L. Górska and Maciej Eder
Piotrowski’s Law and Four Cases in the History of Polish

Wojciech Guz
Resumptive Pronouns in Polish co Relative Clauses

Matías Guzmán Naranjo and Elena Pyatigorskaya
Comparing Naive Discriminative Learning, Sublexicon Phonotactics, and Analogical Learning

Kateřina Havranová
A New Syntactic Analysis of Dutch Nominal Infinitives

Anders Holmberg and On-Usa Phimsawat
Minimal Pronouns

Klára Jágrová, Irina Stenger, Roland Marti, and Tania Avgustinova
Lexical and Orthographic Distances between Czech, Polish, Russian, and Bulgarian – a Comparative Analysis of the Most Frequent Nouns

Ángel L. Jiménez-Fernández
Subextraction in Dative Clitic and Differential Object Marking Constructions

Tamás Káldi, Anna Babarczy and Ágnes Bende-Farkas
The Hungarian Focus: Presuppositional Content and Exhaustivity Revisited

Jiri Kaspar
Information Status: An Experimental Study at the Syntax-Semantics Interface

Paul Kiparsky
Counterscopal Morphology

Miroslav Kubát and Radek Čech
Quantitative Analysis of US Presidential Inaugural Addresses

Ivona Kučarová
φ-Feature Resolution in a DP Coordination:
A Case Study at the Syntax-Semantics Interface

Yuriy Kushnir
The (Not So) Inverse Linking Pattern in Plains Cree
Michał Kuźmicki
Nasal Place Neutralization in Classical Nahuatl: A Derivational OT Account

James E. Lavine
The Structure of v-Heads and the Status of the Initiator Role

Michal Láznička
The Influence of Sequential Iconicity on the Ordering of Czech Adverbial Clauses

Eva Lehečková and Jan Křivan
Czech Habitual Morpheme -va- as an Expression for Coding Gaps in Time Reference

Natalia Levshina
A Pilot Study of T/V-Pronouns in European Languages Based on a Parallel Corpus of Online Film Subtitles

Wei-lun Lu
Use of Parallel Texts as a Methodology in Cultural-Cognitive Linguistics: A Case Study on English-Chinese and English-Czech Parallel Texts

Dagmar Machová
Polyfunctionality and German Modals

M. Rita Manzini and Leonardo M. Savoia
N Morphology and Its Interpretation: Romance Feminine/Plural –a

Michaela Martinková
Boundary Crossing Motion Events in Czech Source and Target Texts: Evidence from a Parallel Corpus

Vladimir Matlach and Diego Gabriel Krivochen
Measuring String ‘Randomness’: Applications of Quantitative Methods to Natural and Formal Languages

Petra Mišmaš
Multiple and Short Wh-Movement as Wh-Movement to the Peripheries

Urszula Modrzyk
A Corpus-Based Study of Gradience of a Complex Preposition Category: The Case of Polish na rzecz, do spraw and w celu

Andrew Murphy
Gang Effects with Left-Branch Extraction: Arguments for Weighted Constraints

Elliot Murphy and Andrea Santi
Processing and Acceptability Properties of Predicate Ordering Effects in Copredication
Olga Nádvorníková
Parallel Corpus in Translation Studies: Analysis of the Shifts in Segmentation in Phrases in the Czech-English-French Part of the Parallel Corpus InterCorp

Mark Newson and Márton Kucsera
Multiple Wh-Structures in Hungarian: A Late Insertion Analysis

Marcin Orszulak
Long-Distance Wh-Extraction in Polish. Problems of Structure and Derivation

Jamal Ouhalla
The Morphosyntax of Pluractionality

Jeffrey Keith Parrott
Gender and the ‘Y’ Question

Elena Rudnitskaya
The Analysis of the Evenki Analytic Negation

Denisa Šebestová
English Translation Counterparts of the Czech Particles copak, jestlipak, kdepak

Chao-Kai Shih
A V Raising Account of Verb-Initiality in Isbubukun Bunun

Masaharu Shimada and Akiko Nagano
Mirativity and Focus in DP

Jana Šindlerová and Aleš Tamchyna
Emotions Translated: Enhancing a Subjectivity Lexicon Using a Parallel Valency Lexicon

Filip Smolík
Imageability and Past Tense: Imageability Effects Are Not Limited to Irregular Verbs

Filip Smolík and Veronika Bláhová
Grammar or Social Understanding? Both Mental State Language and Grammatical Skills Predict Children’s Use of Personal Pronouns and Verb Inflections

Réka Sólyom
Metaphors, Metonymies, and Blends in English–Hungarian Equivalents of an Online ICT Glossary

Eva Švejdarová
A Sociolinguistic Study of Hybrid Language Practices, Attitudes towards Language and the Role of Language in Identity Formation of Latina Students at a Public University in Los Angeles
Magdalena Szczyrbak
Stance and “Saying” Routines in Expert-Lay Communication

Krisztina Szécsényi and Tibor Szécsényi
Definiteness Agreement in Hungarian Multiple Infinitival Constructions

Jen Ting
On the New Expression Bucuo V in Taiwan Mandarin and Its Implications for Rule Borrowing

Enikő Tóth and Péter Csatár
Preverbal Focus and Syntactically Unmarked Focus in Hungarian: A Comparison

Gergő Turi and Balázs Surányi
Information Structure and Inverse Quantifier Scope Interpretation

Wout Van Praet
Subject-Complement Asymmetry in Copular Clauses: Investigating the Realization of the ‘General’ NP across Different Copular Types

Jorge Vega Vilanova
Specificity and Past Participle Agreement in Catalan: A Diachronic Approach

Kateřina Veselovská
Using Universal Dependencies for Cross-Linguistic Comparison of Syntactic Features in Evaluative Structures

Ludmila Veselovská
Economy of Czech Group Nouns and Quantifiers

Marcin Wągiel
Fake Guns, Privative Adjectives and Taxonomies of Concepts

Susi Wurmbrand
How to Do Your Best—Feature Sharing and Rule H

Susi Wurmbrand
Semantic Gender Features of Nouns—Visible and Active in Syntax

Joanna Zaleska
Testing Phonotactic Constraints: A Language Game Experiment

Hedde Zeijlstra
Contentives or the Alleged Universal Noun-Verb Distinction: An Asymmetric Perspective

Magdalena Zíková
What Can a Discourse Behavior of a Demonstrative Tell Us about Its Stage of Grammaticalization?
Invited Speakers
Towards an Empirical Study of Sentence Semantics:  
Examples from the Domains of Scalar Operators and Concessives  
in English and German  

Volker Gast  
Friedrich Schiller University Jena, Germany  
volker.gast@uni-jena.de

The study of sentence semantics in a model-theoretic framework (formal semantics), which 
grew out of logic and mathematics (e.g. Frege 1879, Russell 1905, Montague 1973, Cresswell 1973) and which abstracts away from any extra-sentential or non-linguistic information, has 
increasingly taken the factor of ‘context’ into account, thus leading to the establishment of a 
‘formal’ branch of pragmatics (e.g. Kadmon 2001, Potts 2010). The consideration of the 
context is of particular relevance to the study of elements that make reference to alternatives, 
such as focus particles (Rooth 1985), or that refer to contextually given background 
assumptions, such as concessives (König 1991). Given that contextual embedding is arguably 
much less intuitively accessible than truth conditions, and given that information relating to 
the context of speech is often implicit, the study of language in context requires a 
methodology that relies on naturalistic data, and that goes beyond categorical judgements of 
truth or falsity, or (in)felicity.

While in descriptive linguistics a canon of methods has been developed that is based on 
natural language data (as represented in linguistic corpora, e.g. Biber et al. 1998), and that 
allows for the integration of probabilistic (in addition to categorical) generalizations (cf. 
Baayen 2008), studies in formal semantics or pragmatics rarely rely on naturalistic, 
contextually embedded data. This is probably at least partly due to the lack of corpora with 
annotations as are required for the analysis of elements such as focus particles or concessives. 
For example, for a proper analysis of focus operators information about the scope and the 
focus of a particle is indispensable, and information about the focus alternatives (for any 
given example) is at least useful. What is required, thus, is corpora with ‘rich’ annotations, i.e. 
annotations at various levels of linguistic analysis (minimally syntax, semantics and 
pragmatics) that are theoretically informed and provide the type of information that is needed 
for a fine-tuning of the formal analyses.

In this talk I will aim to show how the formal study of sentence meaning can be enriched 
by using a corpus-based, quantitative methodology, and by creating ad hoc corpora annotated 
at various levels. Such annotations are highly specific and require a combination of automatic 
annotations – for instance, syntactic parsing (Klein & Manning 2003) and word sense 
disambiguation (Pedersen & Kolhatkar 2009) – with manual annotations (Gast et al. 2015) – 
such as indications of the scope and focus of a particle, and the (implicit or explicit) focus 
alternatives available in the discourse environment.

I will present two case studies, one on scalar additive operators (e.g. Engl. even) and one 
on concessives (e.g. Engl. although). I will compare the relevant English expressions to their 
German equivalents. Specifically, I will show what distinguishes German scalar operators 
such as selbst and sogar from each other – and how they relate to Engl. even – and what
determines the distribution of English concessive subordinators such as although and while, in comparison to the corresponding German expressions.

References
Morphology in the Mental Lexicon

Ray Jackendoff\textsuperscript{a} and Jenny Audring\textsuperscript{b}

\textsuperscript{a}Tufts University, Medford, MA, USA; \textsuperscript{b}University of Leiden, Netherlands
Ray.Jackendoff@tufts.edu; j.audring@hum.leidenuniv.nl

We explore a theory of morphology grounded in the outlook of the Parallel Architecture (PA, Jackendoff 2002), drawing in large part on Construction Morphology (Booij 2010). The fundamental goal is to describe what a speaker stores and in what form, and to describe how this knowledge is put to use in constructing novel utterances. A basic tenet of PA is that linguistic structure is built out of independent phonological, syntactic, and semantic/conceptual structures, plus explicit interfaces that relate the three structures, often in many-to-many fashion.

Within this outlook, morphology emerges as the grammar of word-sized pieces of structure and their constituents, comprising morphosyntax and its interfaces to word phonology, lexical semantics, and phrasal syntax. Canonical morphology features a straightforward mapping among these components; irregular morphology is predominantly a matter of noncanonical mapping between constituents of morphosyntax and phonology.

As in Construction Grammar, PA encodes rules of grammar as schemas: pieces of linguistic structure that contain variables, but which are otherwise in the same format as words – in other words, the grammar is part of the lexicon. Novel utterances are constructed by instantiating variables in schemas through Unification. A compatible morphological theory must likewise state morphological patterns in terms of declarative schemas rather than procedural or realizational rules.

Non-productive morphological patterns can be described in terms of schemas that are formally parallel to those for productive patterns. They do not encode affordances for building new structures online; rather, they motivate relations among items stored in the lexicon. Productive schemas too can be used in this way, in addition to their standard use in building novel structures; hence they can be thought of as schemas that have “gone viral.” Interestingly, this classification proves useful also for extending syntactic schemas to idioms and other fixed expressions.

This raises the question of how lexical relations are to be expressed. Beginning with the well-known mechanism of inheritance, we show that inheritance should be cashed out, not in terms of minimizing the number of symbols in the lexicon, but in terms of increased redundancy (or lower entropy). We propose a generalization of inheritance to include lexical relations that are nondirectional and symmetrical, and we develop a notation that pinpoints the regions of commonality between pairs of words, between words and schemas, and between pairs of schemas.

We conclude that linguistic theory should be concerned with relations among lexical items, from productive to marginal, at least as much as with the online construction of novel forms. We further conclude that the lexicon is richly textured, in a fashion that invites comparison with other domains of human knowledge.
Exfoliation: Towards a Derivational Theory of Clause Size

David Pesetsky
Massachusetts Institute of Technology
pesetsk@mit.edu

We too easily become used to facts about language that should strike us as strange. One of these is the menagerie of clause-types and clause-sizes in the world's languages categorized with ill-understood labels such as finite, non-finite, full, reduced, defective, and worse. For almost a half-century, the standard approach to these distinctions has treated them as a consequence of lexical choice — a legacy of arguments by Kiparsky & Kiparsky (1970) and Bresnan (1972), who showed (1) that verbs that select a clausal complement select for the complementizer and finiteness of that complement, and (2) that finiteness and complementizer choice have semantic implications. In an early-1970s model of grammar in which selection and semantic interpretation were properties of Deep Structure, these discoveries directly entailed the lexicalist view of clause type that is still the standard view today.

So compelling was this argument at the time, that its 1960s predecessor (Rosenbaum 1967) was all but forgotten — the idea that distinctions are derivationally derived as the by-product of derivational processes such as Raising. As a consequence, it has gone unnoticed that in a modern model of grammar, where structure is built by Merge (and both selection and semantic interpretation are interspersed with syntactic operations), the arguments against the derivational theory no longer go through.

In this talk, I present a series of arguments for a modernized return to a derivational theory. I argue that a reduced clause is the response to specific situation: a clause-external probe that has located a goal such as the subject in the upper phase of its CP-complement, when that goal does not occupy the edge of its CP. Since anti-locality prevents that goal from moving to the clausal edge (Erlewine 2015 and predecessors), a last-resort operation called Exfoliation deletes outer layers of the clause until the goal occupies the edge without ever having moved there. Evidence comes from an array of English, Romance and Icelandic puzzles usually taken to bear on case and licensing, which in this theory bear instead on whether or not Exfoliation happened legally, as well as from complementizer-trace effects, the typology of anti-Agreement (Baier 2016), and Zulu hyper-Raising (Halpert 2015).
Presentations
Operator Projections Moving as Remnants in Hungarian

Gábor Alberti and Judit Farkas

Department of Linguistics, University of Pécs, / Research Institute for Linguistics, Hungarian Academy of Sciences
alberti.gabor@pte.hu, juttasusi@gmail.com

Not only "argument-size" δP constituents (APs, AdvPs, or most typically DPs) can appear in the specifier position Spec,α of an operator head α in the cartographic syntactic structure of Hungarian sentences (see É. Kiss–Kiefer 1994, Piñón 1995, Alberti–Medve 2000, É. Kiss 2002, Brody–Szabolcsi 2003) where α is a Qunatifier, (Identificational) Focus or Aspectual head (see the schematic "deep structure" in (1a)). We claim that such "sentence-size" βP constituents can also appear there (in the above-mentioned Spec,α operator position) as an FP, NegP, AspP, among others, as is shown in (1b) below. As is also shown in (1b), however, the "sentence-size" βP constituent can only move as a remnant (Koopman–Szabolcsi 2000, Alberti 2004), presumably due to its extra size. Semantically, the content of βP is interpreted in Spec,α; while phonetically, a γP part of βP should be extracted so that only an "argument-size" δP constituent appear in the Spec,α position (but it is not the content of δP itself which is interpreted as a filler of Spec,α).

(1) a. \([αP \ βP γ \ ... \ V \ ... \ δP \ ... \] \]
   b. \([αP βP δP i \ Øj] k \ α \ Øk \ ... \ Øi \ ... \] \]

Examples (2–3) below serve as illustrations of our claim, with two different "evaluations" of the quadruple of head categories (α,β,γ,δ). Note that the phonetic material of δP is interpreted, indeed, as the representative of a phrase βP, since δP in (2) cannot be interpreted as a clause-level quantifier at all (fel is 'up also') whilst δP in (3) can be interpreted as a clause-level quantifier only if the word order is as follows: Marit is felhívtam 'Mari.Acc also up.called.1Sg' ('It also holds for Mari that I called up her.').

(2) \(⟨α,β,γ,δ⟩ = ⟨Q, Asp,V,Adv⟩: [QP [AspP AdvP1 Asp+Øi Øj] k Q+Vl Øk \ Øp \ ... \ Øi \ ... \] \]
(A: '[AspP Fel akartad hívní Marit.']) B: ‘És[QP [AspP fel] is [Q hívtam Marit]’!

   up wanted.2Sg call.Inf Mari.Acc and up also called.1Sg Mari.Acc

   (A: ‘You wanted to call up Mari.’) B: ‘And I did call up her.’

(3) \(⟨α,β,γ,δ⟩ = ⟨Q,F,Asp,D⟩: [FP DP1 F+Øi Øj] k Q+Vl Øk \ Øp \ [AspP \ ... \ Øi \ ... \] \]
(A: ‘[FP Marit akartad felhívní.]’) B: ‘És[QP [FP Marit ] is [Q hívtam fel]’!

   Mari.Acc wanted.2Sg call.Inf and Mari.Acc also called.1Sg up

   (A: ‘It is Mari that you wanted to call up.’) B: ‘And it is Mari, indeed, that I called up.’

Example (5) shows that if a noun phrase is "sentence-size" in the sense that a pre-D operator zone appears in it (in Giusti’s (1996) spirit, see also Alberti et al. 2015) and Alberti and Farkas (to appear), the huge noun phrase (βP=QPo,P) occupies the clausal Spec,Foc
(Spec,α) as a remnant whose phonetic material is practically a DP (δP), with an extracted part which is also a DP (γP).

\[ (α,β,γ,δ) = (F,\text{Pos},D,D) : [\text{FP}]_{\text{QPos}DP,\text{QPos}DP} \text{F}+V_i [\text{VP}...\text{Q}...\text{Q}_k...\{\text{DP}...\text{Q}...\text{I}...\}] \\
\text{[FP Csak [\text{Pos} mindkét kollégának] [F ellenzem az elküldését!]}!} \]

The only both colleague.Dat oppose.DefObj.1Sg the away.send.Nmn.Poss.3Sg.Acc

‘It is only the option according to which both colleagues would be sent away that I am definitely against [as for me, one of them can be sent away].’

**Keywords:** cartographic syntax, operators, Hungarian

**References**


Alberti, Gábor, Judit Farkas, and Veronika Szabó. 2015. “Arguments for Arguments in the Complement Zone of the Hungarian Nominal Head.” *Approaches to Hungarian* 14: 3–36.

Amsterdam: John Benjamins Publ. Company.


**Table 1.** Cases that our whole paper discusses (besides cases (2–4), discussed above)

<table>
<thead>
<tr>
<th></th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>Q</td>
<td>Q</td>
<td>F</td>
</tr>
<tr>
<td>β</td>
<td>Asp</td>
<td>F</td>
<td>Neg</td>
</tr>
<tr>
<td>γ</td>
<td>V</td>
<td>Asp</td>
<td>F</td>
</tr>
<tr>
<td>δ</td>
<td>Adv</td>
<td>D</td>
<td>V</td>
</tr>
</tbody>
</table>
Examining the Fluctuation Hypothesis at Early Stages of English L2 Article Acquisition: The Role of Instruction and Metalinguistic Feedback

Jwahir Alzamil
Newcastle University, Newcastle upon Tyne, UK
jawahir.al-zamil@newcastle.ac.uk

The formal-linguistics-based study of second language acquisition (SLA) often ignores the implications for language teaching, including acquisition of articles. Krashen (1985) and Schwartz and Gubala-Ryzak (1992) argue that negative evidence (regardless of type) does not influence L2 acquisition. To address this situation, the present study is positioned at the interface between the SLA and corrective feedback literature. The study examines the Fluctuation Hypothesis (FH) which proposes that L2 learners whose L1 lacks articles fluctuate between specificity and definiteness (Ionin et al. 2008), two settings of the Article Choice Parameter, whereas L2 learners whose L1 has articles will not fluctuate, due to L1 transfer. Fluctuation means that L2 learners use the in [+specific] contexts; however, English links the use of the to definiteness, not specificity. Like English, Arabic distinguishes articles based on definiteness; nevertheless, it lacks a phonologically overt indefinite article (comparable with the English a). The present study examines how instruction alone and instruction accompanied by metalinguistic feedback interact with article use during the early stages of L2 acquisition.

The study involved 93 participants. Of these, 78 Arabic-speaking EFL participants were divided into a metalinguistic group (MG) that received instruction accompanied by metalinguistic feedback (n=26), an instruction group (IG) that received instruction alone (n=26) and a control group (n=26). In addition, there were 15 native English speaking controls. All were at elementary level, based on the Oxford Quick Placement Test. Testing comprised two tasks: a) a forced-choice elicitation task; and b) a cloze task. A six-week intervention followed a pre-test/post-test/delayed post-test design. During the first and second weeks, the participants completed the consent forms, background questionnaire and the proficiency test. During the week following the pre-test (the third week), the MG and IG separately received one instruction session (90 minutes) on English articles. The reason for providing the MG with instruction before giving them metalinguistic feedback is to investigate in greater detail the role and effects of metalinguistic feedback when it is accompanied by instruction. However, both groups differed in relation to what they received in the subsequent four sessions (one each day for 90 minutes’ duration), which were conducted in the fourth week. Both groups were assigned the same activity tasks; however, only the MG was corrected and provided with metalinguistic feedback. Both groups were post-tested again in the sixth week.

Results of the pre-test revealed that the three groups: a) made a high rate of omission errors in using a and the; and b) they associated the use of the to specificity and not to definiteness, as they did not transfer article semantics from Arabic, which is against the FH. The fluctuation occurred at the individual level, as all the participants fluctuated, surpassing the findings of Ionin et al. (2008) in [-article] language speakers; whereby 10 of the 19
participants (53%) fluctuated. Post-test and delayed post-test results demonstrated that instruction alone and instruction with metalinguistic feedback had various effects over omission errors and fluctuation. For omission errors, the MG’s omission errors rate decreased from 35.6% to 6.7% in the post-test and to 8.7% in the delayed post-test, whereas the IG’s rate decreased from 27.9% to 22.1% in the post-test and to 26% in the delayed post-test. For fluctuation, the IG continued to fluctuate at the individual level, but the MG did not. The results revealed that: a) L1 semantic features may not be transferred, at least in the case of L1 Arabic speakers, which indicates that L1 speakers of [+article] languages who are at early stages of L2 acquisition are similar to [-article] language speakers in terms of fluctuation; b) the omission errors in the use of the and a show that L1 Arabic speakers were not sensitive to the fact that Arabic lacks an equivalent to the English article a, since the omission error rate of a and the was similar; and c) instruction unaccompanied by metalinguistic feedback is not that effective.

**Keywords:** Fluctuation; article instruction; corrective feedback

**References**


In languages with so-called global case splits, the person of the subject and the object determine case-marking on one of these arguments together. This means that Case assignment to the object must be delayed until the verb has agreed with the object and the subject, challenging the widespread assumption that Case and Agree are aspects of a single operation. I argue that it is correct to dissociate Case and Agree and I provide an analysis that derives global case splits using cyclic Agree (Béjar and Rezac 2009) and syntax-internal impoverishment (Keine 2010).

1 Data In Kashmiri (Indo-European), the object is Dat only if the OBJ’s person is higher than the SBJ’s person on a scale such as 1 > 2 > 3. Examples are shown in (1–2) (Wali and Koul 1997, 155). When the subject’s person is higher than the object’s, the object appears as Nom.

(1) a. ı chu-s-ath I. Nom ts ı be-1sg-2sg you.Nom teaching
   ‘I am teaching you.’ 1→2: Nom

   b. ı chu-kh me parñan un ınva va ı teaching
   ‘You are teaching me.’ 2→1: Dat

(2) a. ts ı you. Nom chi-h-an su parñan un ınva va ı teaching
   ‘You are teaching him.’ 2→3: Nom

   b. su chu-y tse parñan un ınva va ı teaching
   ‘He is teaching you.’ 3→2: Dat

2 Problem Such data pose a challenge to the idea that Case and Agree happen simultaneously, since Case assignment to the object would happen too early in (1a) and (2a). Both (1) and (2) examples show that the person of the object alone does not determine its case.

3 Analysis I adopt a version of cyclic Agree (Béjar and Rezac 2009): probes can agree as long as they have unvalued features. Person features are sets of features: [speaker], [participant] and a generalised feature [π]. A scale like 1 > 2 > 3 is represented by relations among sets.

(3) \([1] = \{\text{speaker}, \text{participant}, \pi\} \supset [2] = \{\text{participant}, \pi\} \supset [3] = \{\pi\}\]

\(v\) agrees with the direct object first. If and only if the subject’s features are a proper superset of the object’s features (i.e. 1 ⊃ 2, 1 ⊃ 3, or 2 ⊃ 3), \(v\) can agree with both the object and the subject, but Agree between \(v\) and the subject fails otherwise.

I propose that languages with global case splits differ from languages without them by allowing a different order of operations on \(v\) (Heck and Müller 2007; Georgi 2014). In Kashmiri, φ-Agree precedes Case ([φ-Agree ≺ Case]). As long as \(v\) can enter Agree relations, it will not assign Case. This delays Case assignment until \(v\) can no longer agree. Languages with [Case ≺ φ-Agree] cannot have global case splits because Case assignment happens too early to be influenced by agreement: \(v\) will assign Case, independently of the sets of φ-features involved.

There are then two possible derivations: when \(v\) is valued by two arguments the object has Nom (direct configurations), cf. (1a) and (5b). When \(v\) is valued by one argument the object has Dat (inverse configurations), cf. (1b) and (5c). The lack of case-marking on the object in the other examples can be derived by the impoverishment rule shown in (4), which applies when \(v\) has two values, indicated by [α, β] in (4). This means that \(v\)’s Dat case is deleted in direct configurations like (1a) and (2a) but not otherwise, cf. (5b) and (5c).

(4) Dat → ∅ / un v: [α, β]

4 Conclusions This paper provides a formally explicit analysis of Case assignment in languages with global case splits that avoids countercyclic derivations and derives effects of person hierarchies without syntax-external hierarchies and constraint rankings.
(5) a. Steps ① and ② of (1a,b)

b. Direct configuration, cf. (1a)

c. Inverse configuration, cf. (1b)

Keywords: Comparative syntax, Global case splits, Case, Agree

Variable Use of Strong Preterites in European Portuguese:
A Quantitative and Theoretical Approach

Pilar Barbosa, Cristina Flores, and Ana Bastos-Gee
University of Minho, Portugal
pbarbosa@ilch.uminho.pt; cflores@ilch.uminho.pt; acba18@gmail.com

1. Goals
In the Minho region in Portugal, the 1st and 3rd person singular forms of strong preterites can be leveled and the form used for leveling can vary:

(1) Pois ela estive lá três anos parada [68M3A]
   *Anyhow she was 1P.SG there three years still*

(2) Eu fui para a tropa em cinquenta e quatro [38H4B]
   *I went 3P.SG to the troop in fifty and four*

The goals of this paper are: (a) to determine which linguistic and social factors favor this variation, and (c) to offer an explanatory account of this phenomenon.

2. Methodology
We examined 50 interviews from the stratified corpus ‘Sociolinguistics Profile of the Speech of Braga’. We selected speakers from three age groups (26-59 / 60 – 75 / > 75) and four different education levels (without degree / 4-9 school years /10-12 school years / university degree). A total of 3595 tokens were coded.

We ran a mixed effect binominal analysis using Rbrul (Johnson, 2009), with five independent factors: ‘age’, level of education; ‘verb’, ‘subject expression’ and ‘subject position’. The factor ‘verb’ included the forms ‘estive/ esteve’ (1P/3P of to be-stative), ‘tive/ teve’ (1P/3P of to have), ‘fiz/ fez’ (1P/3P of to do) and ‘fui/ foi’ (1P/3P of to be and to go).

3. Results
The regression analysis identified as main predictors the factors ‘speaker’ (a random effect), ‘subject expression’, ‘verb’ and ‘age’. Variable leveling is almost insignificant in the younger group, thus probably indicating a process of ongoing change. It rarely occurs with null subjects. Subject position does not have a predictive effect.

There is a clear difference between ter/estar/fazer, on one hand, and ser/ir, on the other. While in the former case leveling can be realized by either the form for the 1st or the form for 3rd person, in the case of ser/ir only one form is used, namely the 3rd person form foi. There is also an individual tendency to use only one form for leveling: a given speaker either uses the 3rd person or the 1st person, but not both forms of one verb randomly. Thus, there is inter-linguistic variation in this regard.

4. Analysis
In our analysis, we follow the “late insertion” model of Distributed Morphology (Halle & Marantz 1993). We develop an account of these agreement leveling effects that is based on the interaction between the internal syntax of the strong preterites and the late insertion of
underspecified functional Vocabulary Items. We propose a derivation of the different forms in the standard dialect and then we offer an analysis of leveling where intra-speaker variation is tied to the probabilistic application of impoverishment rules along the lines of Nevins and Parrot (2010). Inter-speaker variation is due to different choices as to which feature sets are subject to impoverishment: the features for Person or T. In the case a verb such as ter, this yields /teve/ or /tive/, respectively. In the case of ser or ir the resulting forms are /foj/ in both cases.

References

Phenomenon. Psych verbs have been observed to be a challenge for theories of argument realization due to the linking problem (Dowty 1991), and because they exhibit argument structure alternations exclusive to them (Belletti and Rizzi 1988), since the experiencer and stimulus tend not to be encoded by the canonical cases (NOM and ACC, respectively) of a transitive verb. Psych verbs can be divided into two classes according to their syntactic behaviour (e.g. Croft 1993, Pesetsky 1995, Iwata 1995), i.e. subject-experiencer (e.g. like) and object-experiencer verbs (e.g. please). There has been relatively little recent research on cross-linguistic comparison of psych predicates (e.g. Bossong 1998), and nothing in the form of quantitative corpus typology (see (Pijpops and Speelman 2015) for work on Dutch).

Aim of the study. Work on parallel corpora for typological research, although promising, is still lacking for most phenomena (e.g. Dahl 2007, Levshina Forthc, Meyer and Cysouw 2014). Firstly, we want to help fill this gap by providing a systematic analysis of psychological predicates in cross-linguistic, quantitative perspective. Secondly, we can capture more than the default coding strategy for a given semantic concept, providing a more fine-grained typology of syntactic patterns. Moreover, we can compare token and type frequencies of the psych verb constructions within languages and cross-linguistically, which sheds light on cross-linguistic preferences for constructions expressing the semantics of psych verbs.

Methodology. We chose the following languages: Spanish, Portuguese, French, German, Dutch, and Russian. We used a multi-lingual parallel corpus based on TED talks (Kulkarni 2015). We extracted all constructions for 12 semantic concepts: be happy, be surprised, love, like/please, be interested, enjoy, hate, be worried, fear, be upset, be bored, be sad. These were selected on the basis of high token frequency and semantic type (positive vs negative emotions). Both the causative (frighten, worry) and stative variants (fear, be worried) (cf. Croft 1993) were taken into consideration. We then extracted the possible syntactic constructions (and different lexemes, in case more than one is used for stative and causative meanings) to express those semantic concepts. (The analysis of a couple of languages is still pending.)

Hypotheses and expected results. (i) Semantics is reflected in syntax: if the experiencer is semantically more prominent, we expect it to be coded by a core case (NOM, ACC, DAT), the stimulus being expressed as oblique (e.g. X is surprised by Y) or left out (X is worried). Causative psych verbs focusing on the stimulus are expected to be transitives with the stimulus as subject (X interests Y). (ii) Higher frequency concepts will both show less syntactic variation, and be the more likely candidates to have different causative and stative forms. Because high frequency concepts will be more entrenched (Schmid 2016), they will be less prone to syntactic variation, and because of the high frequency, they should be able to
more easily maintain different lexemes (similarly to how suppletion is more likely with high frequency words).

**Keywords:** corpus linguistics; parallel corpora; typology, psych predicates

**References**


Licensing the Correlative Construction via the Semantics of the Demonstrative

Anne Beshears
Queen Mary University of London, London, England
a.e.beshears@qmul.ac.uk

The correlative construction is characterized by a relativizing clause (the correlative, CorrelP in (2)) which is linked with the main clause through a correlate (DemP in (2)) (Bhatt 2003; de Vries 2005; Liptak 2009), where the correlate may be either a demonstrative phrase or a pronominal (de Vries 2001).

While correlatives are often analyzed as being base-generated at the left periphery (Srivastav 1991; Dayal 1996), Bhatt (2003) proposes that the correlative construction has the following structure (1) in which the correlative and the demonstrative correlate are part of the same constituent (2). The correlative may then be fronted (3) and leave a trace.

1. \([ \text{CorrelCP} \] \[ \text{DemP} \] \[ \text{Dem NP} \] \)
2. \(\text{Ram} [ [ \text{CorrelP} \text{jo CD sel par he} ] \text{DemP us CD ko} ] \text{k'aridega (Hindi)} \)
   \(\text{Raam which CD sale on is that obl CD acc buy.fut} \)
   ‘Raam will buy [ which CD is on sale, that CD ].’
3. \([ \text{CorrelP} \text{jo CD sel par he} ] \text{ram ti} [ \text{DemP us CD ko} ] \text{k'aridega (H.)} \)
   \(\text{which CD sale on is Raam that obl CD acc buy.fut} \)
   ‘[ which CD is on sale ], Raam will buy [ t, that CD ].’

Bhatt’s analysis leaves open the question of why the correlative only enters the syntax where there is a corresponding demonstrative or pronominal, and semantically, how the correlative CP and the demonstrative compose to give the proper interpretation.

Using technology from Elbourne (2008)’s treatment of demonstratives, I propose that the correlative clause enters the syntax as an overtly pronounced index of the demonstrative. The correlative clause is therefore selected by the demonstrative as an argument.

4. \([ \text{DemP} [ [ i \text{ this }] \text{ R } \text{ NP} ] \]
5. This CD [Gesturing at a ‘Dark Side of the Moon’ poster] is one of my favorites.
6. \([\text{this CD}] = \sigma z. x. (z=\text{ poster} \ x \ \text{is a CD in s} \ x \ \text{is represented by } z' \text{proximal} (\text{poster}, w, a, t)) \)
   The unique individuals z and x, such that z is a poster, x is a CD, and z represents x, and the poster is proximal according to the speaker a, at the speech time t, in a possible world w.

   The demonstrative has an underlying structure like that in (4) (Elbourne 2008, following Nunberg 1993). The index, a lexical item i, picks out an object, and R is a contextually salient relationship. The morpheme this contributes the information above number, proximity, and definiteness.

   Given example (5), the index i picks out the semantic value ‘Dark Side of the Moon’ poster, where the morpheme this contributes the proximal, singular, and definiteness features. The R relation, in this case, ‘is represented by’, relates the index to an interpretation in the main clause. Thus, this CD has a semantic value as in (6).

   I propose that the Hindi correlative (2) is an overt pronunciation of the index of the demonstrative, with the structure in (7). The demonstrative takes the correlative jo CD sel par he ‘which CD is on sale’ as the index i. In this context, R is simply an identity relation (ie. ‘=’), and the NP is CD. Using English to represent the Hindi and reflecting the Hindi word order (2), this has a semantic value of (7).

   7. \([ [ [ \text{which CD is on sale} ] \text{ that R } \text{ CD} ] \]
8. \([\text{which CD is on sale; that CD} ]=\text{the unique } x \text{ s.t. } x \text{ is a CD and } x=\text{the CD on sale.} \)

   It is the indexical nature of the demonstrative itself (or the pronominal, in some languages) which allows the correlative clause to enter the syntax.
The semantic composition of the demonstrative phrase easily extends to cases in which the correlative phrase and the correlate phrase have distinct noun phrases (9). In this case, *jo zordar arət boḥot parefankarti he* ‘which loud woman is causing a lot of trouble’ is selected as the index of the demonstrative, which corresponds through an R relation of identity to the NP in the main clause, *moti lambi arət* ‘fat tall woman’.

The full demonstrative phrase has a semantic value as in (10).

9. *[jo zordar arət boḥot parefankarti he] vo lambi moti meri sa:s he (H.) which loud woman much trouble does pres that tall fatty.f my MIL is

‘Which loud woman is causing a lot of trouble, that large tall woman is my mother in law.’

10. $\sigma x.(z = \text{loud woman} ^ x \text{is a ‘tall fatty’ in } s ^ x = z ^ \text{distal(loud woman,w,a,t)})$

This analysis predicts that other phrase may also appear as an overt pronounciation of the index of the demonstrative, which proves to be the case. In the following example in Hindi (11), a relativized DemP appears in the fronted, correlative position. In this case, *ʋo cʰatur laɽka jo cʰor hɛ* ‘that clever boy who is a thief’ is associated with and moved from the index of the demonstrative *ʋo* ‘that’ in the main clause.

11. *[ʋo cʰatur laɽka jo cʰor he], ap jante hɛ [t, vɔ] jel gaya he (H.) that clever boy who thief is you know.impfct pres that jail went pres

[DemP That clever boy who is a thief ], you know that [ t that/he ] went to jail.’

Similarly, this analysis can be extended to constructions in languages which are not traditionally thought of as having a correlative construction. For example (12), in Mandarin Chinese (familial) possessives like the following, *Zhangsang* is being selected as an argument of the pronominal, base-generated as an overtly pronounced index.

12. Lili guyong le Zhangsan ta baba (Mandarin)

*Lili hire LE Zhangsang he father*

‘Lily hired [DemP Zhangsang he father].’

The analysis which I have presented here goes beyond previous analyses of the correlative construction by showing the relationship between the correlative clause and the demonstrative correlate situated in the main clause. Bhatt (2003), the correlative (CP) and the demonstrative are both within the same constituent, but it is clear under this account why the demonstrative licenses the correlative construction and the semantic composition of the correlative-correlate constituent follows easily from the semantics of the demonstrative phrase.

Further, this analysis predicts that other types of DPs may enter the syntax as an overt pronounciation of the index of the demonstrative.

References


Constraints on Psych Verb Passivisation in Polish

Anna Bondaruk,1 Bożena Rozwadowska2 and Wojciech Witkowski2

1John Paul II Catholic University of Lublin, Poland; 2Wrocław University, Poland
bondaruk@kul.pl; browad@uni.wroc.pl; wojciechw62@gmail.com

Object Experiencer (OE) verbs behave in a special way, and thus have been on the research agenda for almost 3 decades. Cross-linguistically, they are three-way ambiguous between agentive, eventive and stative interpretations (cf., Arad 1998, 1999, Landau 2010, Alexiadou and Iordăchioaia 2014, a.o.). Whether OE verbs are grammatically special is subject to controversy. Passivisation is one of the relevant properties in this debate. Agentive OE verbs, as expected, form verbal passives. However, the status of passives formed by OE verbs in eventive and stative contexts remains unclear. Pesetsky (1995) and Landau (2010) take only eventive OE verbs to be able to form verbal passives. Grafmiller (2013) argues that all psychological predicates, even stative ones, are capable of forming verbal passives. For Fábregas and Marín (2015), the so-called ‘formal psych predicates’, the class that includes some OE and Subject Experiencer (SE) verbs, resist the verbal passive in Spanish altogether.

The aim of this paper is to examine the passivisation options that are attested for Polish OE predicates and to check whether the type of the passive structure available for a given predicate is dependent on its interpretation. Two groups of OE verbs have been investigated, stative and non-stative (or eventive), distinguished by Biały (2005). Stative OE verbs include: interesować ‘interest’, martwić ‘worry’, trapić ‘plague’, zasmucać ‘grieve’, etc., whereas the non-stative OE verbs comprise: dręczyć ‘torment’, rozbawić ‘amuse’, straszyć ‘scare’, zaskoczyć ‘startle’, etc. Passive sentences containing 33 stative and 46 non-stative OE verbs have been extracted from the National Corpus of Polish (www.nkjp.pl). Additionally, specifically constructed passive sentences with the two types of OE verbs have been collected. The data have been analysed to check what type of passive they represent - verbal or adjectival. We assume that in Polish verbal passives are distinguished from adjectival ones by means of the auxiliary verb, i.e. zostać ‘to become’ vs. być ‘to be’, respectively. Moreover, other criteria for distinguishing adjectival from verbal passives available in the literature (cf. Wasow 1977, and Gehrke to appear) have been applied to the Polish data, such as (i) the interpretive difference (ongoing, completed event, non-dynamic state), (ii) the possibility of using degree words to modify (certain) adjectival passives, (iii) the possibility of using the adjectival prefix nie- ‘un-’, (iv) control into purpose clauses, (v) unrestricted appearance in the progressive, (vi) the modification by ostatnio ‘recently’, and (vii) the use of frame adverbials such as od ‘since’ or durative/iterative wciąż ‘all the time’.

The analysis of the data supports the split of OE verbs into two classes and shows that non-stative OE verbs more readily form verbal passives, as they co-occur with the auxiliary zostać to a greater degree. Stative, verbs, on the other hand, show a strong preference for the auxiliary być. The sporadic passive sentences with stative OEs and the auxiliary zostać imply a causing subevent, a change of state and thus should be treated as eventive, not stative. This observation allows us to conclude that the special behavior of OE verbs is related to their
polysemy and is reflected in Polish in the way they passivise, depending on the type of event they describe.

**Keywords:** Object Experiencer verbs, passivisation, verbal passives, adjectival passives, Polish.

**References**


Gehrke, Berit. Forthcoming. “Adjectival participles, event kind modification and pseudo-incorporation.” *Natural Language and Linguistic Theory*.


Bobaljik’s Root Suppletion Generalization as an Instance of the Adjacency Condition

Pavel Caha
Masaryk University, Brno, Czech Republic
pcaha@mail.muni.cz

Background. The talk discusses Bobaljik’s (2012:4) “Root Suppletion Generalization” (RSG). The generalization says that adjectival root suppletion can be triggered only by comparative/superlative markers that are in the same word as the root. Bobaljik captures RSG by a new condition on allomorphy, according to which phrasal nodes block allomorphy; in other words, allomorphy is restricted to relations within a complex head.

There is a tension between Bobaljik’s implementation and some recent work (e.g., by Koopman and Szabolcsi (2000), a.m.o.), which suggests that words do not correspond to heads, but should in fact be analyzed as phrasal objects. The question is whether and how these two independent strands of work can be reconciled.

My main claim is that Bobaljik’s account of RSG in terms of a specific locality condition is superfluous. I show that in all the relevant cases, allomorphy is independently excluded by the so-called Adjacency Condition (Siegel 1978) (which Bobaljik also adopts in addition to his new condition). The adjacency condition says that the form of the root may be sensitive only to a morpheme that is directly attached to it. The consequence is that we can have a framework where the RSG is explained by a property (“structural adjacency”) that can be attributed both to two nodes in a head or two nodes in a phrase. There is no evidence from RSG for the claim that “words are special.”

The data. The empirical strength of RSG can be best seen in languages where morphological and periphrastic comparatives may be formed side by side; see (1) from Bobaljik’s work.

(1) Greek POS ‘good’ CMPR = 1 word ‘bett-er’ CMPR = 2 words ‘lit. more good’

In (1), the CMPR affix ter triggers root suppletion; the non-affixal pjo doesn’t. According to Bobaljik, this is because pjo is not in the same complex head. But this explanation is not needed. We independently expect the adjacency condition to rule this example out: pjo cannot trigger suppletion because it is not the first morpheme added to the root: -os intervenes.

Explaining RSG by adjacency has good consequences in Bulgarian, where (according to Bobaljik) comparatives are formed by a separate (clitic) word po. Under Bobaljik’s account of RSG, po- is thus expected not to trigger suppletion. This is almost true, because Adjs in BG take an obligatory Agr suffix, which intervenes between po and the root, e.g., po-[[[guljam]-a], [CMPR-[[big-fem]]]. However, there is one case where (unexpectedly for Bobaljik) po triggers suppletion (mnogo – po-veče ‘many-more’). Crucially, po-veče is unique and shows no agreement, so the adjacency condition correctly describes the facts.

Another argument for adjacency comes from Finnish. In Finnish, the order of morphemes in the superlative is such that it is incompatible with a head movement account (we get ROOT-SUPERL-COMPR). This order can only be derived by phrasal movement, so (nothing
else said) Bobaljik expects that Finnish shows no root suppletion. But this is not the case, a fact compatible with the adjacency account. Finally, I turn to languages such as English where it appears that the adjacency condition falls short of explaining the ban on forms like *more bett, because the two morphemes are adjacent. However, I argue that the example actually decomposes roughly into [[mo-re] bett], so the first morpheme that attaches to –er is mo- (itself a suppletive version of much), and hence, –er cannot reach beyond the first merged mo to condition allomorphy on the root.

References

Causal Connective *protože* across Registers

Anna Čermáková, Zuzana Komrsková, Marie Kopřivová, and Petra Poukarová
Charles University, Prague, Czech Republic
anna.cermakova@ff.cuni.cz; zuzana.komrskova@ff.cuni.cz; marie.koprivova@ff.cuni.cz; petra.poukarova@ff.cuni.cz

Written and spoken variety of Czech differ in many respects. Any research dealing with spoken language needs to balance between linguistic features inherent to spoken language per se (cross-linguistically) and features that belong to that particular language. Comparisons between written and spoken Czech (as research into spoken Czech seems to be mostly comparative) have been so far focused on morphological, lexical and phonological features. Very few studies include syntactic features as well (cf. Bermel 2000) and none is, to our knowledge, based on large scale authentic spoken data. There is no comprehensive description of the syntax of spoken Czech and we thus “lack a benchmark from speech that can be relied on” (Bermel 2000, 63).

This study aims to contribute to research focused on comparing written and spoken Czech. It is situated on the borderline of lexis, syntax and pragmatics and will look in detail at Czech causal conjunction *protože* within four major registers (Biber 1991): spontaneous spoken language will be examined in contrast with three major written registers (media/newspaper texts, fiction and academic texts). The research is based on extensive corpus data (available through the Czech National Corpus project); both quantitative and qualitative approaches will be applied. The analysis also aims to look at the current part-of-speech annotation of the corpus data with regard to its applicability to spoken data.

Causal connectives belong to the most frequent conjunctions across all registers; *protože* being the most frequent, having, however, strikingly different frequencies of occurrence in the four compared registers. In written language *protože* is most frequent in fiction but it is nearly three times more frequent in our spoken data. Research into causal connectives suggests that some connectives specialize in certain domains and that there are various degrees of causality. Some studies suggest a diachronic variation showing a process of shifting from prototypical subordinator to discourse marker (Stenström 1998, Burridge 2014). We focus our analysis on the functions of *protože*, distinguishing between content or real world uses, epistemic uses (cf. Czech “příčina” and “důvod”, Svoboda 1954), speech act uses (Sweetser 1990) and discourse markers across the four registers.

**Keywords:** causal connectives, spoken vs. written language, corpus linguistics

**References**
Polish Derived Nominals with Thematic Adjectives
as Argument-Supporting Nominals

Bożena Cetnarowska
University of Silesia, Katowice (Poland)
cetnarow@gmail.com; bozena.cetnarowska@us.edu.pl

Thematic group adjectives (such as Italian, presidential) are claimed to be restricted cross-linguistically to nominals which are not argument-supporting ones (see Marchis 2010 for Romanian and Spanish, Alexiadou and Stavrou 2010 for Greek). Event nominals modified by group adjectives are interpreted by the above-mentioned authors as names of simple events.

The present paper investigates the occurrence of thematic group adjectives in event nominals in Polish. It is argued that, even in the presence of group adjectives, such derived nominals can exhibit the characteristics of argument-supporting nominals (i.e. complex event nominals in Grimshaw’s 1990 typology). They can take internal arguments (as in 1a, 1b), occur with aspectual modifiers (1a) and with agent-oriented modifiers (1b).

(1)

a. ciągłe policyjne kontrole autokarów
   constant police.adj checks.nom coaches.gen
   ‘constant checks of tourist coaches by the police’

b. umyślnie prezydenckie ulaskawienie gangsterów
   deliberate presidential pardoning.pfv.nom gangsters.gen
   ‘the president’s pardoning (the) gangsters deliberately’

The derived nominals in (1b) and (2) are headed by –nie/-cie verbal nouns, which are recognized as complex event nominals by Rozwadowska (1997) by virtue of a number of verb-like properties. Most of those properties can be identified also in verbal nominals accompanied by thematic group adjectives, e.g. the presence of grammatical aspect, the occurrence of the reflexive clitic się (in 2a, 2b) and the negative particle nie- (in 2b).

(2)

a. celowe studenckie spóźnianie się na zajęcia
   intentional student.adj coming_late.ipfv.nom refl on classes
   ‘students’ coming late to classes intentionally’

b. papieskie nieangażowanie się w konflikty polityczne
   papal not_getting_involved.ipfv.no refl in conflicts.acc political
   ‘the pope’s not getting involved in political conflicts’

Although thematic adjectives can modify either perfective or imperfective verbal nouns (cf. 1b and 2a), they are more common and more felicitous with imperfective nominals. The change of the grammatical aspect of the verbal noun in (2a, 2b) from the imperfective to the perfective diminishes the felicity of the resulting nominals in (3a, 3b).

(3)

a. ?celowe studenckie spóźnienie się na zajęcia
   intentional student.adj coming_late.pfv.nom refl on classes
   ‘students’ coming late to classes intentionally’

b. ?papieskie niezaangażowanie się w konflikty polityczne
   papal not_getting_involved.pfv.nom refl in conflicts.acc political
   ‘the pope’s not having become involved in political conflicts’
It will be suggested that this is due to the compatibility of the mass reading of the imperfective nominals (Alexiadou, Iordachioaia and Soare 2010) with the default plural interpretation of group adjectives (Marchis 2010, Moreno 2015).

**Keywords:** group adjectives, derived nominals, complex events

**References**
The Myth of Comparability? Issues in the Design and Analysis of a Monolingual Comparable Corpus

Lucie Chlumská
Institute of the Czech National Corpus, Prague, Czech Republic
lucie.chlumska@korpus.cz

Monolingual comparable corpora, i.e. corpora consisting of two sets of texts in the same language and selected according to similar criteria (Baker 1995, 232), have become a valuable and frequently used source of data in Corpus-based Translation Studies, especially in the translation universals research. Hypotheses concerning the so-called T-universals (Chesterman 2004, 8) refer to the way translators process the target language, i.e. they are based on the comparison between translations on one hand and non-translated texts as a reference sample on the other, requiring carefully balanced data.

Corpus design is generally a complicated task; however, when it comes to translations, a whole new set of possible issues arises. One of the major problems is the disproportion in source languages: in the Czech context, English is by far the most translated language, as it accounts for more than a half of all translations (Czech National Library statistics 2003–2012). However, to be able to study the translation universals phenomenon properly, researchers need a variety of texts translated from many different languages. The need for a large corpus thus in fact contradicts the requirement of its heterogeneity and balance; there are simply not enough varied data to be included.

Another issue is the character of translated texts in the two translation directions: from a smaller language (such as Czech) to a bigger language (such as English, German, French etc.) and vice versa. The type of texts and their cultural status are often not at all comparable, e.g. widely read classics in small languages v. virtually anything in the bigger ones (Bernardini & Zanettin 2004, 57). This imbalance is most visible in parallel corpora; however, it also has implications for a monolingual comparable corpus. And finally, even though there might be similar text types to be compared and included in the corpus, cultural norms may be different and the texts may in fact have different characteristics (e.g. originally written cookbooks may differ greatly from the translated ones).

Although these issues add a big question mark behind the claim of comparability, they certainly should not prevent researchers from attempting to compile a monolingual comparable corpus in the first place. The main objective of this paper is thus to summarize major setbacks in the comparable corpus design (with the Jerome monolingual comparable corpus as a case study), to show their implications for the data analysis and interpretation (on the concrete example of simplification and convergence in translated Czech), and to suggest possible alternatives where applicable. The discussion specifically covers the following topics: text size in comparable corpora and its impact on statistical testing, text type / genre diversity issue in lexical analysis (based on n-gram and POS-gram extraction) and possible consequences of one source language prevalence.

Keywords: language of translation, translation universals, corpus design, Czech
References
There is evidence that children do not always grasp the correct principle of 1st and 2nd person reference (I/You) at first: they use proper names, make reversal errors, or use both forms interchangeably (see Tanz, 2009: 49–60). The difficulties may rest in the linguistic as well as the social aspects of the phenomenon of person reference. For the correct use and comprehension of 1st and 2nd person reference, children need to achieve certain level in their syntactic and morphological development, and they need to understand both the difference between self and others, and the perspective of the speaker. The social abilities considered here may be related to later Theory of Mind which is the ability to understand people as subjective individuals with different emotions, desires and beliefs.

The factors of linguistic and social development have rarely been examined together (but see Marková and Smolík, 2014). In our study, we analyze children's performance in production and comprehension of person reference, as well as in tasks aimed on selected social abilities and general language development. We expect that the effect of the social performance will explain significant amount of variation in person reference above and beyond the effect of general language development, and vice versa.

Our participants were typically developing 2.5 years old Czech toddlers (age range: 28–32 months). So far, data from 51 children (from the whole sample of 66 children) were analyzed. The laboratory testing took place individually and was divided into two sessions (45+25 minutes). To measure the Theory of Mind precursors, the children were given three tasks concerned with

(1) visual perspective taking: the child and the mother sat opposite each other; a card depicting an animal was shown to the child and the child was asked to show the picture to the mother; the orientation of the card and the child's gaze direction were evaluated;

(2) pretend play: the examiner pretended to drink some water and eat some cookies and then invited the child to engage in these activities too; the adequacy of the child's engagement was evaluated; and

(3) intention understanding: the examiner failed to do certain actions (e.g. to reach for a box she have previously let fall) so that her intention was obvious but never expressed verbally; she prompted the child to help her by looking at the child and addressing her/him with her/his name; the child's understanding of the intention and her/his helping behavior was evaluated.

Loveland (1984) and Ricard et al. (1999) have shown linkage between pronouns and visual perspective in English children. The other two tasks were not yet used in relation to the person reference. We adopt five other tasks from the previous studies concerned with person reference production and comprehension and we use one lexical and one grammatical comprehension test as measures of the general language development.
Our key analyses examine the effects of pretense, intention understanding and lexical development on production and comprehension of person reference. Lexicon and pretense have significant concurrent effects on the comprehension score (see table 1), and lexicon, pretense and intention understanding have significant concurrent effects on the production score (see table 2). The models suggest that person reference in production and comprehension is related to somewhat different sets of social cognition measures. However, both the models contain a linguistic as well as a social predictor, as hypothesized. The results thus reveal that both the linguistic factors and social development measures have independent effects on the acquisition of person reference.

**Keywords:** deixis; person reference; Theory of Mind

**Table 1** The first regression model examines the effect of lexical score and pretend play understanding on the correct comprehension of person reference. We use aggregated score from the three tasks on comprehension of person reference as dependent variable.

<table>
<thead>
<tr>
<th></th>
<th>p-value</th>
<th>Δ R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical score</td>
<td>&lt; 0.001</td>
<td>0.291</td>
</tr>
<tr>
<td>Pretense</td>
<td>&lt; 0.001</td>
<td>0.156</td>
</tr>
</tbody>
</table>

**Table 2** The other model examines the effect of lexical score, pretend play and intention understanding on the correct production of person reference. We use aggregated score from the two tasks on production of person reference as dependent variable. Both, the correct name (3rd person) and pronoun (1st + 2nd person) production were counted as correct responses.

<table>
<thead>
<tr>
<th></th>
<th>p-value</th>
<th>Δ R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical score</td>
<td>&lt; 0.001</td>
<td>0.122</td>
</tr>
<tr>
<td>Pretense</td>
<td>&lt; 0.001</td>
<td>0.142</td>
</tr>
<tr>
<td>Intention</td>
<td>&lt; 0.001</td>
<td>0.210</td>
</tr>
</tbody>
</table>

**References**
Do Bi-Aspectual Verbs in Czech Express Grammatical Aspect?

Jan Chromý,1 Štěpán Matějka1, and Jakub Dotlačil2

1Charles University in Prague, Czech Republic; 2University of Groningen, Netherlands
jan.chromy@ff.cuni.cz; stepan.matejka@ff.cuni.cz; j.dotlacil@gmail.com

As other Slavic languages, Czech distinguishes both grammatical (perfective and imperfective) and lexical aspect (cf. Schmiedtová – Flecken, 2008). Aside from perfective (Pf) and imperfective (Ipf) verbs, there is a specific group of so-called bi-aspectual (BA) verbs which can be classified as both perfective and imperfective (Komárek, 2002) because they can occur in contexts that are not compatible either with Pf verbs (e.g. combination with a phasal verb), or Ipf verbs (e.g. combination with až ‘when’) (cf. Veselý, 2008).

The main hypothesis we will present in the poster is that the bi-aspectual verbs are in fact aspectless, i.e. there is no real evidence that they do have grammatical aspect. This hypothesis is supported by four self-paced reading experiments we conducted so far. In the main experiment, we examined the differences in reading times (RTs) of sentences that were distinguished only by their main clause verb. Every verb appeared in one of three forms: a BA verb (e.g. akceptovat ‘to accept’), or its Pf equivalent (e.g. přijmout ‘to accept’), or its Ipf equivalent (e.g. přijímat ‘to accept’). The verbs were followed by dokud ‘until’ which is not compatible with Pf verbs.

In the experiment, we used 12 sentences with durative BA verbs (or their equivalents) and 12 sentences with punctual/iterative BA verbs (or their equivalents) as experimental items (see Item Examples below) and 72 filler sentences. All the verbs were transitive, they were in past tense and had plural objects. In other words, our experiment tested the compatibility of 4 types of predicates with dokud: BA durative, and BA punctual/iterative, Pf verbs, Ipf verbs. Our main question was if there is a difference in RTs of different predicate types. Altogether, we tested 82 subjects. Based on linear mixed-effects model, we found significantly higher RTs on dokud in sentences with Pf verbs than Ipf or both types of BA verbs. These findings suggest that BA verbs are processed differently than Pf verbs. In the other two experiments, we found that there are differences between processing of punctual/iterative BA verbs and Ipf verbs. The fourth experiment is focused on processing of BA, Pf and Ipf verbs used in sentences with až ‘when’ clause which is not compatible with Ipf verbs.

On the basis of the experiments, it seems that BA verbs in Czech do not behave the same as Pf or Ipf verbs. BA verbs can be used in contexts in which PF and Ipf verbs cannot. It seems that from the processing point of view, it makes sense to count them as grammatically aspectless. A less strong claim would be that they are aspectually underspecified. This is in accordance also with the fact that there are no structural constraints for using BA verbs and that most BA verbs are borrowings which tend to become either imperfective, or perfective in their later evolution (Jindra, 2008).

Keywords: aspect; bi-aspectual verbs; self-paced reading
Item Examples
(1) Tolerantní podnikatel akceptoval/přijal/přijímal změnu právního předpisu, tolerant businessman accepted change legal regulation, dokud firma vytvářela vysoké zisky.
until company produced high incomes.
(2) Pražský radní koncipoval/navrhl/navrhoval plány dalšího rozvoje, Prague councillor outlined plans future development, dokud primátorka slibovala rozsáhlé změny.
until mayor promised vast changes.

References
On Dative Clitics and Nominalizations in Romanian

Alexandra Cornilescu

University of Bucharest, Bucharest, Romania
alexandracornilescu@yahoo.com

1. Introduction: the data, aim

It has recently been shown that C(litic) D(oubling) is sensitive to the vP configuration where the Dative DP merges. The difference is that CD is required in unaccusative configurations (1), while being optional in ditransitive configurations (2).

(1) Îi vin idei lui Ion. *Vin idei lui Ion.
   cl-Dat come ideas Ion(Dat) come idei Ion(Dat) ‘John is getting ideas.’

(2) Am oferit/ Le-am oferit bomboane copiilor. ‘I offered sweets to the children.’
   have- offers/cl-Dat-have offered sweets.Acc children.the.Dat

A second difference is that verbs for which CD of the Dative is obligatory do not nominalize, while Verbs which show optional CD allow nominalizations. (*venirea unei idei lui Ion (the coming of an idea to Ion), oferirea de bomboane copiilor (the offering of sweets to the children).

Aim The goal of the paper is to explain the nature of this unfailing correlation, between obligatory Dative doubling by the clitic and the absence of nominalizations. This presuppose:

a) understanding the difference between configurations which require Dative doubling and configurations where doubling by the clitic is optional; b) understanding why configurations with obligatory Dative clitics do not feed nominalizations.

2. Why unaccusative configurations require doubling of the Dative by the clitic

Assumptions: 1. Since in languages with overt Appl, Appl is always a suffix, Appl should always merge above some VP projection, so that V raises to the left of the Appl suffix.(Georgala 2012, Larson 2010). 2. Argument-introducing heads (e.g.Voice, Appl, P) also have expletive variants, where they bear no θ-feature, but are endowed with a case feature (Wood & Marantz, 2015). Thus, when the Dative is non-core, and it is thus not part of the verb’s argument structure, Appl is endowed with both a θ- and a case feature and introduces a supplementary argument. When the Dative is part of the verb’s a-structure, Appl may simply be an expletive case-licensor. 3. Clitics are functional heads which license their associates, when all their features have been valued. On their way to T, clitic DPs stop in a position above the vP (=K(ase)P(hrase)), where they value their Case feature (Sâvescu 2009). 4. Romanian Datives are uniformly licensed by Appl heads (Diconescu&Rivero, 2007, Tigău 2014).

The analysis Consider the unaccusative vP of a change of location verb, like veni (come). The event structure of such a vP is reduced to v + a PP-small clause (Irwin 2011), with the Theme as subject and a Goal, realized either as a PP or a null P and a Dative DP. If the ApplP must be sister to VP, its only position in the unaccusative vP is above vP (3). The ApplP is syncretic with the KP, and may or may not also contain θ-features. The clitic raises pied-piping the whole DP, which remains accessible to the clitic for licensing since it has exited
the vP phase. Evidence for the intermediate landing site is that CD-ed IOs may bind the subject or scope over it. (e.g. Ce i-au cumpărat Evei frații ei? (what cl-Dat have bought Eve-Dat brothers.the her).

\[(3)[\text{Appl/KP} \text{Appl/K[uPers, uCase]} [\text{vP}[\text{vin}],[\text{PP}\text{DP}\text{idei} P[\emptyset]} [\text{DPGoal} [\text{DP} \text{lui Ion}[\text{DP} \text{îi [uPers pro]]}}]

The question remains why v+Appl cannot license the IO, requiring the clitic. We propose to relate the obligatory nature of the clitic in unaccusative vPs to the “defective” nature of unaccusative verbs (Pesetşky and Torrego, 2004:512): unaccusatives Agree and may trigger movement, but do not erase the uninterpretable features of the goal, since their φ-set is defective, so that a regular head is needed for a proper application of Agree. The combination v+Appl is defective in this sense. The clitic is required to provide an alternative set of φ-features for regular Agree, allowing v+Appl+φ to value the case of the Dative DP.

In a nutshell, the obligatory nature of the clitic in unaccusative configuration follows from the defective nature of the the unaccusative verb, the only verbal lexical head in the vP.

3. Nominalizations of unaccusatives and ditransitives

When one turns to the absence of nominalizations, another property of the configuration in (3) becomes significant, namely the position of Appl above the vP, i.e. in the functional domain of the vP. The configuration in (3) cannot be input to nominalization, since the Romanian event nominalizers do not take scope over the functional domain of the verb with the possible exception of Aspect. The nominalizer precisely replaces Appl in (3), leaving the Dative unlicensed. In contrast, ditransitive configurations are correctly expected to allow nominalizations, since, under any analysis, the Appl head merges below v.

**Keywords:** applicative heads, dative clitics, nominalization, unaccusative verbs, case

**References**
Săvesco Oana. 2009. *A syntactic analysis of pronominal clitic clusters in Romance: The view from Romanian.* PhD diss., NYU.
Tigău Alina. 2014. The two-object construction in Romanian and German.In Cosma Ruxandra e. a. (eds), *Komplexe Argument-Strukturen,* Berlin: Walter de Gruyter. 85-144
Wood Jim and Alec Marantz Alec. 2015. The interpretation of external arguments, mss. lingbuzz/002487
I consider utterances like the following taken from spoken French corpora, in which an indefinite NP is associated with a large range of constructs introduced by the complementizer que:

(1) c’est des endroits que tu peux pas y aller (Coral-Rom)
these are places where you cannot go

(2) c’est des filles qu’on peut pas leur parler
these are girls that it is impossible to speak to them

(3) il y a deux sortes d’ouvriers il y en a que tu leur parles ils comprennent de suite et il y en a que bon au début ils sont pas d’accord avec toi mais il faut leur expliquer ils comprennent après quand même (Corpaix)
‘There are two kinds of workers there are some that (when) you discuss with them they understand immediately and there some (others) that well first they don’t agree with you but you need to explain the things to them they understand nevertheless at the end’

My presentation addresses the syntactic status of these utterances. The traditional analysis in French literature for (1) and (2) is to posit a non standard type of relative clause based on que as a complementizer and y, leur, ils as resumptive clitic pronouns (Guiraud 1966), which is considered for French as an alternative strategy to the standard wh-relatives (des endroits où tu peux pas aller, des filles à qui on ne peut pas parler). Abeillé & Godard (2006) observe that this analysis cannot be extended to (4) where no resumptive pronoun appears:

(4) il y a des feux qu’il faut appeler des pompiers tout de suite (Corpaix)
there are fires that it is necessary to call the firemen immediately

They propose a unified analysis for examples (1) to (4), in which que introduces a plain IP licenced by the head Noun. To this syntactic structure is associated a pragmatic interpretive rule Topic–Comment, explaining that the content of the modifying clause conveys a characteristic property of the head. However, this analysis does not fit with the examples of this pattern found by manual search from the sampled C-Oral-Rom Corpus (300000 words) of spoken French. In all the 10 occurrences found, the determiner is always indefinite, whereas no such restriction is observed in wh-relative clauses. Moreover, whereas examples with definite determiners quoted in the literature (Guiraud 1966, Gapany 2004):

(5) l’homme que je lui ai dit ça (the man that I said that to him)

are strongly sociolinguistically marked (Gadet 2003), some of our examples are from educated speakers in casual conversation. One way to take these facts in account is to analyze these constructs as instances of non standard consecutive clauses. In the standard corresponding consecutive clauses: the que-clause is governed by a N with an obligatory indefinite determiner and licenced by a quantifier adjective node (tel = such) in (4’):

(4’) il y a des feux tels qu’il faut appeler les pompiers tout de suite
I propose that the “consecutive” clause is licensed by a [+ quality] underspecified feature (grossly meaning « of such quality »). This feature is borne by the quantifier in the standard construct and by the indefinite determiner in the non-standard one. This semantic feature is independently motivated by the apparent tautological sentence: un homme est toujours un homme. The intended meaning: » a man (as exemplar of the species) shows always the characteristic qualities of a man » involves the activation of the [+ quality] feature in the second occurrence of homme. In both constructs the que-clause can be considered as « niece licensed » (licenced by a daughter of the NP), which is not the case for a relative clause. Under this analysis, the semantic interpretation of the whole construct is straightforward: the que-clause does not directly convey, as proposed by Abeillé & Godard, a characteristic property of the head, rather it brings a fact or a complex situation on the basis of which this property can be inferred. The “consecutive” clause helps in specifying the underspecified [+quality] feature. Notice that in our example (3) not only can the clitics not be considered as resumptive pronouns, but the construct following que is not a standard IP but a discourse pattern built upon a concatenation of IPs. This kind of extension of the syntactic form of the construct introduced by que is not observed in our data in relative clauses headed by definite determiners, whereas it can be found in overt consecutive clauses:

(5) et le mec il est tellement bouleversé qu’ il prend le papier il le lit pas il le met dans sa poche + puis il s’en va
and the guy he is so much upset that he takes the sheet he doesn’t read it he puts it in his pocket and then he goes away

Considering [NP Det [-def] N que X] as a variant of “consecutive adjunct clause” explains in a natural way this parallel behavior. Additionally, it dispenses with the extra Topic-comment interpretive rule to the benefit of a rule independently necessary for standard consecutive clauses and, beyond, for comparative clauses.

A further consequence of this analysis is that, contra current litterature assumptions, there are no instances of relative clauses with resumptive pronoun in contemporaneous spoken French, probably with the exception of subject position (Sportiche 2011).

**Keywords:** Syntax; French; corpus linguistics; relative clause

**References**


Constituent Negation, Subjunctive and Alternatives

Mojmir Docekal
Masaryk University, Brno, Czech republic
docekal@phil.muni.cz

Background: negation and subjunctive mood are among the most popular topics in the recent literature but formal semantics of Slavic constituent negation (CN) is still missing. CN usually requires adjacency, (1), to its focused constituent (see Borschev et al. (2006), Jasinskaja (2010)) but in case of complex NPs (2) or PPs the CN has to be non-adjacent. Similarly for German focus particles like nur or sogar – Büring and Hartmann (2001), Reis (2005). Let’s assume that even Slavic CN is a focus particle (FP). In this talk I propose a uniform FP semantics for CN and concentrate on examples like (3) where CN has a scope over a subjunctive CP (see Tomaszewicz (2009)).

(1) *ne včera ne majitele, ale ...
not yesterday not owners but ...
‘yesterday not owners but …’
(2) ne majitele *ne pivovarů
not owners not breweries-gen
‘not owners of breweries’
(3) ne že bych chtěl být kojícím otcem
not that SUBJ want-l-sg BE nursing father
‘Not that I want to be a nursing father.’

The contribution of this abstract: I provide new data on CN and subjunctive in Slavic languages. Based on a corpus study (SYN2010, 500 randomly selected sentences with CN), results of which are summarized in Figure 1/Table 1 – syntactic contexts where CN appears – I propose a uniform formal semantic for Czech (Slavic) CN in (4). (4) can be applied to all types of constituents (=A from (4)) modified by CN, as will be demonstrated in the talk. The most interesting and complex among the contexts is CN modifying CP (like in (3)), I focus on this construction and propose that: (i) the subjunctive mood transfers the alternatives from the embedded clause (in line with Villalta (2008)); (ii) the alternatives can be processed by FP or used by an operator licensing NPIs/minimizers (see Guerzoni (2004)). Application of (4) to (2): (4-a). Application of (4) to (3): (4-b), again CN is applied to a set, but not to a set of entities like in (4-a) but to a set of possible worlds (propositions), still being of ⟨α, t⟩ type.

Subjunctive mood allows the negation to access the focus alternatives. In the corpus SYN2010 CN modifying an embedded clause in 99% selects just subjunctive. This claim is supported by an independent evidence: subjunctive mood in (5-a) allows the evaluation of the alternatives (on the likelihood scale) from the embedded clause. These alternatives are needed by the root negation to act as a licensor for the minimizer sebemenší podezření ‘slightest suspicion’ in the embedded clause. Indicative mood doesn’t transfer the alternatives, ungrammaticality of (5-b) explained. Discussion and proposal: the above data give raise to a clear generalization: Slavic CN is a focus oriented particle with semantics in (4). Subjunctive mood evaluates the alternatives for a potential FP/predicate selecting or comparing the contextual alternatives (NPI licensor e.g.). (6-a) shows that verbal negation in the embedding clause licenses strict NPI in the embedded clause (c.f. Gajewski 2011 a.o.), unlike CN in (6-b), the intervention in strict NPI licensing caused by the existential presupposition of CN.
(4) If A is of type \(\langle \alpha, t \rangle\), CN A is of type \(\langle \alpha, t \rangle\) too, and \(\left[\left[ CN A \right]\right]\) is the set of all B such that B doesn't have the property \(\left[\left[ A \right]\right]\) (i.e. \(B \notin \left[\left[ A \right]\right]\)) but it has some property \(P\) that is an alternative to A (i.e. in \(\left[\left[ A \right]\right]/\text{ALT}(A)\))

\[
\begin{align*}
\text{a) } & \lambda x[\neg \text{owners_of_breweries}(x) \land \exists P[P \in \text{ALT(breweries)} \land \text{owner_of}_P(x)]] \\
\text{b) } & \lambda p[\neg \text{want}(I, p) \land \exists P[P \in \text{ALT(want)} \land P(p)]]
\end{align*}
\]

(5) a. Policisté nechtěli, aby gangsteři měli šebemenši podezření.
‘The policemen didn't want the gangsters to have the slightest suspicion.’

b. #Policisté nevěděli, že gangsteři měli šebemenši podezření.
‘Policemen didn't know that gangsters had slightest suspicion.’

(6) a. Policisté nechtěli, aby \([\text{StrictNPI} \text{ani jeden gangster}]\) unikl.
‘Policemen didn’t want even one gangster to run away.’

b. ne, aby \(#[\text{StrictNPI} \text{ani jeden gangster}]/\text{ten gangster unikl!}\)
‘Not that #[even one gangster]/the gangster run away!’

**Keywords:** negation, subjunctive, alternatives, Slavic languages

**References**


Czech is a language in which Anglicisms form a numerous and productive lexical subclass. Their phonological adaptation (Calabrese & Wetzels, 2009) is mostly based on the phonological approximation principle, where each phoneme is replaced by its nearest counterpart within the limits of native phonotactics, e.g. $\text{jazz} /\text{ʤæz}/ \rightarrow /\text{ʤɛs}/$. Other principles, such as the spelling pronunciation, original pronunciation or various kinds of analogies, are observed as well, though less frequently (Duběda et al. 2014). Czech is resistant to foreign phonological elements except those which have already been integrated into the phonological system, e.g. /f/, /ɡ/, /oː/ (Romportl 1978).

However, a number of aspects, both static (the phonological makeup of Anglicisms) and dynamic (high phonological variability, irregular mapping between spelling and pronunciation) make it possible for a language user to identify these words as Anglicisms and to treat them, at least partly, within a specific subsystem of Czech phonology.

The phonology of Czech Anglicisms can be thus regarded as a result of two antagonistic tendencies (integrative and modular).

The purpose of the present paper is to develop and substantiate this conception by means of the following claims:

1. Phonological approximation, which corresponds to a projection of the English phonological system onto the Czech system, is rather regular (Kučera & Zeman, 1998). In a sample of 225 frequent Anglicisms used with original spelling, the phonological form recommended by dictionaries and pronunciation handbooks (Romportl, 1978; Hůrková, 1995) can be predicted for almost 80% of the entries.

2. In real usage, however, the different adaptation principles are frequently in competition, making Anglicisms phonologically more variable than native words, as can be demonstrated and quantified by the recently published Database of Pronunciation Variance of Foreign Words (Duběda, in print). In this survey, 300 phonetically problematic entries (including 148 Anglicisms), selected according to their frequency in a large corpus of Czech and integrated into carrier sentences, were pronounced by 300 native speakers. For Anglicisms, the average number of different phonological variants per entry is 12.1 (14.5 for proper names).

3. The spelling form is latently present in the mental lexicon of the language user and may interfere if the phonological form is not fully activated.

4. The pronunciation of Anglicisms in Czech can be best modelled by assuming a specific phonological subsystem, though deeply embedded in native phonology, which is activated each time a word is identified as an Anglicism. This system can be considered a “shadow phonology” of English.

5. Within this subsystem, Anglicisms are “negotiated” against well-established loans, but also against the knowledge of English, which is the most commonly spoken foreign language in the Czech Republic.

6. There is evidence for an active phonological treatment of loanwords, e.g. phonological overshoots (Robert pronounced as [ˈroubr̩t]), or the phonological contamination of other borrowings by English pronunciation rules (the French wine brand Sauternes pronounced as [ˈsɔtɾn]).
Specific phonological subsystems may be hypothesised for other sufficiently represented lexical strata in Czech, such as Latinisms, Hellenisms or Gallicisms.

**Keywords:** Czech; Anglicisms; loanword phonology; loanword adaptation

**References**
Nocturnal and Vizard as Late Modern English Metonymy-Based Synonyms of Prostitute

Bożena Duda
University of Rzeszów
bozenaduda@yahoo.co.uk

For well over a decade it has been acknowledged that, as Panther and Radden (1999:9) put it, “metonymy is not, as has often been taken for granted, merely a matter of the substitution of linguistic expressions but a cognitive process that evokes a conceptual frame”. A fair number of contemporary cognitive linguists “share the assumption that metonymy is a cognitive phenomenon underlying much of our ordinary thinking and that the use of metonymy in language is a reflection of its use” (Panther and Radden 1999:2). One of innumerable plausible linguistic realisations of metonymic mappings is the subject matter of the ongoing analysis, namely the formation of synonyms of prostitute. From the diachronic perspective, metonymy has been part and parcel of the conceptualisation patterns observable in the ontological realm of prostitution since Anglo-Saxon times. Already the body of Old English (O.E.) lexical items comprising beþæcestre, forligerwif and portcwene betray their metonymic nature, as well as the Middle English (M.E.) synonym strumpet. In turn, among the Early Modern English (E.Mod.E.) synonyms of prostitute metonymic conceptualization patterns may be observed in the semantics of such lexical items as stew, streetwalker, walk-street, night-shade, night-trader, waistcoateer, fling-dust/stink and baggage. The body of lexical data to be studied in the foregoing includes such metonymically conditioned Late Modern English (L.Mod.E.) synonyms of prostitute as nocturnal, night-walker, town-woman, town-miss, market dame, kennel-nymph, crack, vizard, mobbed-head and barber’s chair, extracted from the Historical Thesaurus of English (and testified in a number of historical lexicographic sources).

In what follows, we shall analyse the metonymic mappings within the ontological realm of prostitution with the use of the idealized cognitive model (ICM) theory as understood by Lakoff (1987). Following Kövecses (2000:34), it needs to be stressed at this point that, like conceptual metaphors, conceptual metonymies “may have stability over time”, which may not necessarily be seen at the lexical level (or through linguistic manifestations, to use Kövecses’s (2000:34) terms). This is to say that while metonymic patterns may be universal across time, their linguistic realisations may undergo alterations as a result of the changing reality, be it sociological, cultural or technological. In order to provide evidence for the stability of metonymic patterns over time the L.Mod.E. lexical items will be compared with the synonyms from the earlier historical periods.

Keywords: metonymy, onomasiology, synonyms of prostitute, euphemisation, idealised cognitive models (ICMs)
Selected References


Aims: Our aim is to explain in the generative framework the word order variation of the Udmurt sentence and the different interpretations associated with the word order variants.

Data to be explained: In Udmurt, the most neutral order of a transitive sentence answering the questions 'What's new?' and 'What did XY do?' is S(Adjunct)OV. E.g.:

context1: 'What's new?'
context2: 'What did Sasha do?'

(1) Saša kinoňeatryn T'ERMINATOREZ uchkiz.
    Alex cinema-INE Terminator-ACC watch-PST.3SG
    'Alex watched the Terminator in the cinema.'

In contexts eliciting subject or object focus also other variants can appear. According to Tánczos (2010), sentences with object focus are either SOV (2a) or SVO (2b), while sentences with subject focus show OSV (3a), SVO (3b) or OVS (3c) (i.e., foci appear either left-adjacent to the verb or sentence-finally). Moreover, our data suggest that OSV order is also possible with object foci (2c) and SOV order with subject foci (3d). Thus, we argue that a further, sentence-initial focus position exists, as well.

(2) (a) Saša kinoňeatryn T'ERMINATOREZ uchkiz.
(b) Saša kinoňeatryn uchkiz T'ERMINATOREZ.
(c) T'ERMINATOREZ Saša kinoňeatryn uchkiz.

(3) (a) T'ERMINATOREZ kinoňeatryn SAŠA uchkiz.
(b) SAŠA uchkiz kinoňeatryn T'ERMINATOREZ.
(c) T'ERMINATOREZ kinoňeatryn uchkiz SAŠA.
(d) SAŠA kinoňeatryn T'ERMINATOREZ uchkiz.

We have also found evidence of all-new sentences with a clause-final subject, e.g.:

Context: 'What's new?'

(4) Tunne mi dory vuoz kuno.
    today 3PL.NOM to arrive-FUT.3SG guest.NOM
    'A guest will come to us.'

We seek answers to the following questions: Are there any syntactic and semantic/pragmatic differences between the initial, preverbal and clause-final focus positions? Are these focus slots A or A-bar positions? Is any of them associated with the features [+contrastive] and/or [+exhaustive]? Do non-specific subjects conveying new information occupy the same position as specific, topical subjects? Or do specific subjects undergo topic movement?
**Methods of investigation:** Our research relies on data collected from Udmurt informants. They had to make grammaticality judgements about sentences that contained a focused item in different positions. The sentences were situated in focus-eliciting contexts, e.g., various types of wh-questions, sentences with a superlative predicate, etc. We also obtained information on their interpretative differences (the exhaustivity and the contrastivity of their foci). We used the exhaustivity tests of Szabolcsi (1981), i.e., in the case of (3c), we asked if the sentence *Səša kinořeatryn Terminatorz uchkız, Mašajen Saša öz* “It was Sasha who saw Terminator in the cinema, not Sasha and Masha”) was contradictory. A further exhaustivity test involved numerically modified noun phrases; we asked if their numeral was interpreted as ‘exactly n’ or ‘at least n’. The presence of the feature [+contrastive] was tested in alternative questions, and polar questions with corrective answers. We tested if specific, topical subjects and non-specific subjects occupy the same position.

We asked subjects to evaluate the different word order variants on a five-point Likert scale.

**Results:** Our data suggest that the left-peripheral focus position is only licensed in contrastive contexts. It is an operator position in the CP domain derived by A-bar movement. Interestingly, it appears to be available only to arguments. The preverbal focus is neither [+exhaustive] nor [+contrastive], although it can assume an exhaustive or contrastive meaning pragmatically. We concluded that it is an in situ information focus. Objects can be focussed in this position without any movement. Preverbal subject foci are derived either by the leftward movement of the object into Spec, TopP or by the rightward extraposition of the object. Sentence-final foci are not necessarily [+exhaustive] or [+contrastive], either. We tentatively assume that they are derived by rightward extraposition. It needs further testing if a constituent extraposed to the right can also function as a topic, and whether in the Russified Udmurt of the younger generations, a postverbal object can be an object in situ.

The non-specific subjects of unaccusative verbs appear either immediately preverbally, or in the clause-final focus position; they cannot occupy the left-peripheral position of specific subjects. This suggests that the left-peripheral position occupied by the subjects e.g., in (1) and (2a–c) is a topic position in Spec,TopP.

**Keywords:** Udmurt; information structure; word order; focus; topic

**References**


Overview: using explain as a case study, we argue that syntax can be eliminated in accounting for selection of DPs vs. CPs.

Ambiguities with explain: Pietroski (2000) claims that the interpretation of a sentence with explain depends on the category of its complement. In (1a), the CP complement specifies the content of the explanation (the explanans reading), whereas in (1b) the DP complement specifies what is explained (the explanandum reading). Pietroski captures this distinction by claiming that CPs and DPs are assigned the CONTENT and THEME thematic roles, respectively. This commits Pietroski to a lexical ambiguity view of explain, but this phenomenon can also be observed with verbs such as describe. In (2a), the CP specifies the content of the description. In (2b) the nominal complement specifies what is described.

Special quantifiers: we provide evidence from Moltmann's (2013) special quantifiers that Pietroski’s empirical generalization is incorrect: DPs may receive both explanans and explanandum readings, as shown by (3). Special quantifiers have wider, hitherto unnoticed consequences for selection. Examples like (4) standardly show that think strictly selects for a CP complement. Grimshaw (1979) accounts for this in terms of c-selection; Pesetsky (1982) in terms of case (believe assigns case, whereas think does not). Neither approach can be correct, since think allows a special quantifier as its complement (5). That special quantifiers are DPs is evidenced by (6) and (7): they do not survive under passivization, and they can appear as the complement to a preposition, unlike CPs.

Analysis: Following Moulton (2015), we propose that that-clauses denote properties of individuals with propositional content, of type <e,t>. We eliminate the basic type-distinction between individuals and events, instead implementing them as sortal distinctions in De. Consequently, explain denotes a property of events of type <e,t>, and the verb and that-clause combine intersectively, returning the set of explaining events with the CONTENT specified by the embedded clause. DPs headed by content nouns are of type e, and therefore must be integrated via the neo-Davidsonian thematic function THEME. Special quantifiers are free to range over events or facts, and therefore may combine with the verb directly, or via THEME. This elegantly derives Pietroski’s observations without invoking the syntactic category of the complement. I explore further ramifications for the relation between the verb and its complement in the talk, as well as spelling out the compositional semantics in detail.
(1a) Nora explained [CP that Fido barked].  (1b) Nora explained [DP the fact that Fido barked].
(2a) Nora described [CP that the CEO was wearing a tie]
(2b) Nora described [DP the CEO's attire].
(3) John explained something namely – [[DP the fact that Fido barked] | [CP that Fido barked]]
(4) John {*thinks | believes} [DP the rumour that Mary is moving to Germany].
(5) John {thinks | believes} [DP something].
(6) It was believed {*[DP the rumour] | *(DP something) | [CP that Fido barked]}
(7) John hoped for {[[DP Mary's return] | [DP something] | *[CP that Fido barked]}

**Keywords:** c-selection, s-selection, syntax-semantics interface, clausal complements, complementation, event semantics

**References**
Moulton, Keir. 2015. “CPs: copies and compositionality.” *Linguistic Inquiry*.
Overview: we examine Left Branch Extraction (LBE) (Ross 1967, Corver 1990) in Slavic. Based on evidence from non-intersective readings of adjectives, we argue that 'LBE' reflects two distinct strategies: (i) literal extraction from the edge of the Noun Phrase, and (ii) a base-generated long-distance dependency.

LBE: extraction of adjectival modifiers from the DP is impossible in, e.g., English (1), but possible in many Slavic languages, such as Serbo-Croatian (SC) (3) and Czech (5) (see Bošković 2005 for an overview).

Non-intersective readings: some adjectives are ambiguous between intersective and nonintersective readings. In (4a) the intersective reading entails the Olga is both a dancer, and beautiful. The non-intersective reading entails that Olga dances beautifully, but not that she is beautiful per se. Similarly, under their non-intersective readings, (2) entails that Meri barely works, and (5) entails that Marie has known her friend for a long time.

Analysis: we follow Larson (1998) in claiming that certain nouns introduce an event variable, which certain adjectives can modify. The event variable is closed off at a fixed position in the NP by a covert event quantifier GEN, partitioning NP into the eventive domain and noneventive domain. Non-intersective adjectives are merged below GEN, and intersective adjectives are merged above. This explains why the non-intersective reading is unavailable when (i) the modifier is a predicate (4b), and (ii) there is an intervening intersective adjective (4d). It is not bled by DegP inversion (4c), rather the moved DegP can reconstruct to its base position below GEN. (4c) suggests that if LBE is literal extraction from the NP (as in Bošković 2005) the non-intersective reading should be available, since the moved DegP should reconstruct to its base-position below GEN. This is true for Czech (5) but not for SC (3) (we provide a cross-linguistic survey in the talk). We argue that in Czech-type languages, LBE is literal extraction from the NP, but in SC-type languages, it is a scope-marking construction involve base-generation of a wh-operator in SpecCP. In the talk, we provide additional evidence for these two strategies from intervention effects (Beck 2006).
(1) *(Dep How beautiful) is Olga [DP a t dancer]?
(2) [NP Koliko teškog radnika) je Meri zaposlila t i ?
    how heavy non-worker.ACC be.3P.SG Meri.NOM hire.3P.SG.PAST
    *interective, non-interective
(3) [Dep Koliko teškog) je Meri zaposlila [NP t) radnika]?
    How heavy be.3P.SG Meri.NOM hire.3P.SG.PAST non-worker.ACC
    *interective, non-interective
(4a) Olga is [DP a [GEN [Dep beautiful] dancer]]. (4b) [DP That [GEN dancer]] is [Dep beautiful]
    *interective, non-interective
(4c) [DP [Dep How beautiful] a [GEN t) dancer]) is Olga t i ? *interective, non-interective
(4d) Olga is [DP a [Dep beautiful] [Dep talkative] [GEN dancer]]. *interective, non-interective
(5) [Dep Jak starého) Marie včera potkala [NP t) přitele]?
    How old.ACC Marie yesterday met friend.ACC
    *interective, non-interective

References

Keywords: Left branch extraction, Slavic, syntax-semantics interface, adjectival modification, locality
How Formal Lexical Entries Produce Paradigms: The French Clitic System

Joseph E. Emonds
Palacky University, Olomouc, Czech Republic
jeemonds@hotmail.com

Systems of pronominal clitics for arguments and adverbial adjuncts of verbs in Romance languages have many regular and formal properties that widely accepted grammatical models have not been able to satisfactorily account for. Most recent analyses center on some of these properties and leave several others to the side (Adger and Harbour 2007; Cardinaletti 2008; de Cat (2002); Emonds 2001; Perlmutter 1971). Empirically, French provides a typical system.

Their analysis in formal grammar needs to account for:

- Their internal word orders, which do not reflect the order of phrasal arguments. They often conform to the Mirror Principle (Baker 1985).
- Whether they are proclitic or enclitic, with proclitic status favoured. More generally, the devices needed to account for their word-internal positions (Lieber 1980).
- Their ‘structural distance’ allowed between a clitic and the argument it replaces (Embick and Noyer 2001; Emonds 1999).
- Two opposing tendencies, to have the same form as strong form pronouns and partial departures from this pattern (Kayne 1975).
- The fact that a language’s clitic system is ‘complete’ or doesn’t exist at all.
- The possibility that some clitics are actually phonologically empty (Rizzi 1986).
- A recurrent pattern whereby grammatical number, unlike case, doesn’t affect clitic ordering.

Taken together, these properties point to a type of analysis in which groups of clitics are analyzed as single lexical entries inserted in Phonological Form, with allomorphs specified by extensive use of the parentheses/brace notations. In French, the groups include adverbial clitics, clitics marked for person, and third person non-reflexive clitics. Lexical entries for each group will be formalized (cf/ Borer 1985) and it will be explained how their formal properties predict their empirical patterns (as in e.g. Veselovská and Vos 1999).

Keywords: adverbial clitics, Borer’s Conjecture; French clitics; lexical entries; paradigm pressure; person clitics

References


Reporting Verb *Said* and Its Czech Translations

*Lenka Fárová*

Charles University, Prague, Czech Republic
lenka.farova@ff.cuni.cz

The question of shifts in translating reporting verbs, especially English verb *said*, into Czech has recently occurred in several studies. Corness (2009, 171) points out that at least some Czech translators of English fiction tend to distort narrative structure and style of the original text by introducing semantically and stylistically loaded reporting verbs instead of neutral *said*. This neutral reporting verb is however in about 10% of cases accompanied by an adverb of manner.

As Czech translations influenced by the stylistic tradition introduced mainly by Levý (1998, 144) tend to avoid repetition, it can be expected that translators will render the modifying adverbs accompanying the verb *said* with reporting verbs in which the semantic information of the adverb is somehow included (e. g. *said accusingly* – *vyčetl/vyčetla*, *said apologetically* – *omlouval/omlouvala se*).

The investigation is based on a sub-corpus created from the parallel corpus InterCorp, version 8 (available at [http://www.korpus.cz/intercorp](http://www.korpus.cz/intercorp)) containing fiction originally written in English. The size of the sub-corpus is about 11 million words. The study focuses mainly on the most prototypical pattern, i.e. *said* followed by an adverb ending in *-ly*.

The most common adverbs modifying reporting verb *said* are *quietly, softly, slowly and quickly*, but more than 500 different adverbs occurred in this position in total, of which 200 were unique occurrences. An analysis of translation equivalents shows that most of the translations keep the structure of the original (verb + adverb) and the expected use of the more expressive verbs that Czech otherwise commonly uses is in fact quite rare in these cases. The closest equivalent of *said*, which is *řekl/řekla*, is used dominantly (35%). It is interesting to note, that in some cases the reporting verbs used as equivalents do not correspond to the semantic meaning of the original, sometimes amplifying (*said feebly* – *hlesl chabě*) or even shifting it (*said nervously* – *stěžovala si nervózně)*.

**Keywords:** Corpus/Computer linguistics, Contrastive Linguistics, Translation studies

**References**


New Roles and Architecture for Gender: Evidence from Semitic, Berber, and Romance

Abdelkader Fassi Fehri
Mohammed V University, Rabat, Morocco
abdelkaderfassifehri@gmail.com

Since at least Brugmann (1897) for Indo-European, or Brockelmann (1910) for Semitic, Gen(der) has been shown to express diverse multiple meanings other than sex (or animacy), including individuation, collectivity, abstractness, quantity, size. It is not solely ‘arbitrary’/formal, or ‘redundant’. Further meanings and structures (including discursive functions) have been identified, including qualitative evaluation, perspectivization, and speech act role modification or performativity, to name a few.

This polysemny (and polyfunctionality) and the differentiated structural multitude are not expected in the current reductive typology/theory of Gen, which sees it as essentially (a) a nominal class marking device, (b) semantically sex-based, and (c) syntactically reflected in gender agreement through sex/animate controllers (Corbett 1991; Kibort & Corbett 2008; Dahl 2000; Kihm 2005), despite its well-acknowledged complexity. We propose instead that only an integrative and constructional view of Gen is appropriate for taking into account its polyfunctional nature, and providing room for more ‘unorthodox’ syntax. Furthermore, the fact that gender agreement patterns and controllers are syntactically diverse and non-uniform is compatible with such a view.

Assuming a Minimalist Distributed Morphology model of Grammar a la Chomsky (1995) and Halle & Marantz (1993), we depart from the view that Gen is confined to a dedicated syntactic position (GenP in Picallo (2008), or nP in Kihm (2005) or Kramer (2014)), and interpreted basically as male/female (Percus 2010). Rather, Gen is distributed over various layers of the DP or CP architecture, in the spirit of Steriopolo & Wiltschko (2010), Déchaîne (2014), Pesetsky (2013), or Ritter (1993), or in SAP (making use of Speas & Tenny’s 2003, or Hill’s 2014 cartography).

Distinct layers of Gen are then postulated and motivated: (a) conceptual Gen; (b) n Gen; (c) Cl(assifier) Gen; (d) Num(ber) Gen; (e) DP/CP Gen, or (f) Speaker Gen. Most of these distinct positions give sources to unorthodox meanings. Syntactically relevant interpretable Gen features include ± fem, ± indiv, ± group, ± small/big, ± good/bad, ± endearing, etc.

Since Gen is shown not to be limited to ‘referential-tracking’ (Heath 1983), or perspectivization of referents, extending to expressives, perspective shifts, or illocutionary/speech acts functions, a unique semantics/pragmatics based on sex appears to be highly inappropriate. Our polysemic/polyfunctional hypothesis provides a more appropriate integrative treatment. Building on contributions by Jurafsky (1996), Dressler & Barbaresi (1994), Körtvélyessy (2014), and Grandi & Körtvélyessy (2015) with regard to the semantic treatment of evaluatives, we propose that Gen can be (distributively) hyperonymic, embracing the diverse and structurally organized and related meanings or functions found cross-linguistically, the sex (or animate) meaning being only a hyponym.

The following cross-linguistic data illustrate partially the wide range of meaning/function diversity a Gen typology/theory should be able account for (‘fem’ marked by –at, -t or –a):

Arabic
(1) (a) nahl ‘bee’ → nahl-at ‘bee-unit’, “a bee” (singulative)
    (b) xayyaat “tailor” → xayyaat-a ‘tailor-fem’, “group of tailors” (plural)
(c) raḥil “traveller” → raḥhaal-at ‘traveller + augment+ fem’
i. “an extremely big traveller” (intensive);
ii. “a famous big traveller” (evaluative)

(2) yaaʔab-at-i!
Oh father-fem-mine
Oh my beloved father! (performative)

Hebrew
(3) (a) neqam-a “vengeance” (abstract);
(b) oni → oniy-y-a “a ship” (singulative);
(c) daag “a fish” → dagg-a “a collection of fish” (plurative)

Berber
(4) afus “hand” → t-afus-t “small hand” (diminutive)

(5) tamšmašt-a “a single apricot” (individuative)

(6) argaz “man” → t-argaz-t «mannish female» (endearment or contempt)

Romance
(7) (a) kortell-o → kortell-a ‘large kitchen knife’ (augmentative)
(b) barc-o → barc-a ‘small ship’ (diminutive)
(c) aceitun-o ‘olive tree’ → aceitun-a ‘olive-unit’ (singulative)

Keywords: polysemous gender; distributed; singulative; plurative; performative

References
Spatial reference has always been debated in linguistics. Many authors have developed models which aim to account how spatial expressions are syntactically encoded in natural languages (Koopman 2000, Svenonius 2010, Den Dikken 2010, Pantcheva 2011, Lestrade et al. 2011, among many others). Prima facie natural languages seem to display a lot of variation in this respect, employing specialized case-endings (Finnish, Hungarian, Lezgian, etc.), spatial adpositions (Italian), or both (Pashto). This fact is not surprising if we take case-infections to be the morphological counterpart of adpositions (Fillmore 1968). We focus on inflected locative adpositions in Uralic languages. Building on primary data from Finnish (1) and on a fine-grained cross-linguistic survey taking into account 16 Uralic varieties we show that a two-tiered structure as that of the tree in Figure (1) can capture the cross-linguistic facts.

(1) Auto kadun keskellä.  
Car-NOM street-GEN centre-ADE  
‘The car in the middle of the street’  

Finnish

What are labeled adpositions in Finnish (and elsewhere in Uralic) are best characterized as Axial-Parts (Svenonius 2006). In a Figure/Ground configuration (Talmy 1991, 2000) involving spatial Ps the Ground-complement can be assumed as the possessor of the axis (the AxPart P) taken to evaluate the location of the Figure, introduced by means of the pP node in (1b). Axial Parts are nominal in nature in Uralic and their projection is the ‘external argument’ of the genitive –n in (1b). Many languages beyond the Uralic family treat regional constructions in the same manner as dependent possessed nominals (e.g. Inuktitut and Uzbeki, Johns & Thurgood 2011). Thus, the relation between AxPart and the complement of the genitive predicate in (1b) is a kind of ‘Part-Whole’ relationship (Fábregas 2007), notated here as (⊆), following Manzini & Franco (2016) who have treated (genitive/dative) case morphemes as (elementary) predicates introducing a relation between the argument they select and another argument (contra the characterization of Asbury 2010, who assumes that such items are D-like elements). We aim at addressing the proximity of dative/genitive and locative specifications, corresponding to frequent syncretic lexicalizations of dative and locative. Possession/Part-Whole is often identified with a location in the literature, cf. in particular Freeze (1992), Lyons (1967). Recently, Boneh & Sichel (2010) take the part-whole relation to be the conceptual core of partitives and of inalienable possession - however they factor out alienable possession, treated as a locative relation, since in the language they consider (Palestinian Arabic), material possession is lexicalized by a locative preposition. Here we suggest that the –l-/s locative morphemes of Finnish (Ylikoski 2011) could be analysed as locative specializations of the fundamental oblique (⊆). We construe locatives as a specialization of the part-whole relation, roughly ‘x included by y, y a location’, where different locatives introduce different restrictions on inclusion. What the previous discussion implies is that there is a parallelism between the syntactic configuration/relationship involving
the Axial Part and the Ground and the Part-Whole relation. This idea is confirmed by the fact that natural languages consistently employ the same strategy to lexically encode a Part-Whole and an Axial Part-Ground relation (Franco et al. in preparation). Uralic data support this view.

![Diagram of syntactic structure](image)

**Figure (1)**

**Keywords:** Locatives, Part-Whole, Axial Parts, Oblique Case, Uralic.

**References:**
Raijmund G. Piotrowski (1974, cf. Altmann 1983) suggested that language change is described by a formula which was used in other branches of science, chiefly in biology, as a model of growth processes. The aim of this paper is to confront Piotrowski law with empirical data gathered from a diachronic corpus of Polish. The corpus contains ca 9 million running words, spanning from 1380 to 1774. In a commonly accepted periodisation, around 1546 commences the Middle Polish period, which lasts until mid 18th century. Thus the corpus covers more than this epoch. It is rather opportunistic, consisting chiefly of belles-lettres, with a handful of sermons, registers of rural community courts as well as learned texts.

Of the changes to be examined three are examples of epenthesis: a change of na- > naj, a prefix which is a marker of superlative, and two isolated epentheses: [s] in wszystko > wszystko ‘all’, [d] in barzo > bardzo ‘very’, and a change większe > większy ‘bigger’.

The corpus was divided into time brackets. We experimented with different spans of these brackets – between 20 and 50 years, yet they were all shifted by 10 years, e.g. 1550-1570, 1560-1580 ... 1740-1760, 1750-1770 or 1560-1610 etc. This procedure can be seen as a kind of smoothing, however its motivation is twofold. On the one hand, the time-span should be wide enough so as to guarantee a balanced (or rather varied) sample which is cumulated as a data point. On the other hand, in order to assure feasible analysis the number of data points should be as high as possible. Yet, if each data point represents a longer time-span, than outermost texts within it are much more distant than two texts in neighbouring stretches. The proposed procedure assures both a higher number of data-points and rather varied group of texts which represent a data point. However, what’s most important is that we diminish an effect of a Procrustean bed that is cutting the data into arbitrary slices.

Of these four examples two are described with a Piotrowskian curve with a reasonable fit. The evolution of wszystko doesn’t even slightly resemble the curve in question, although we observe a shift from a rather limited amount of the new forms up to 0,7% at the end of the examined period. These four cases should be confronted with the elision of [l] in albo which was an unsuccessful innovation. The share of the innovative form raised rapidly to 50%, oscillated for some 60 years between 50% and 27%, and again fell to almost 0% in our corpus. The archaic form is the only accepted one in modern Polish.

Our paper supports the observation that the law named after its discoverer is a strong tendency with a considerable number of exceptions, rather than a linguistic law in a strict sense.

References

**Keywords:** Diachronic change, Piotrowski’s Law, Quantitative linguistics
Resumptive Pronouns in Polish co Relative Clauses

Wojciech Guz

John Paul II Catholic University of Lublin, Poland
wguz@o2.pl

This paper discusses the problem of resumption in Polish object relative clauses introduced by the relative marker co, as in (1):

(1) gdzie ta mucha co ją zabiłem?
where this flyNOM co sheACC I killed
‘Where is that fly (that) I killed?’

In the literature dealing with resumptive pronouns, different accounts vary significantly as to the basic question of when the resumptive pronoun is to be expected. Some say it is always obligatory in object clauses (Giejgo 1981; Kardela 1986; Bondaruk 1995; Mykowiecka 2001), others seem to make the same assumption without stating it explicitly (Gołąb and Friedman 1972; Topolińska, Grochowski, and Karolak 1984, 345), still others allow for reservations to be made to the presence of the resumptive (Fisiak, Lipińska-Grzegorek, and Zabrocki 1978; Broihier 1995; Szczegielniak 2004; Lavine 2003). Besides the contrasting claims put forward by authors, virtually all previous analyses have been based on introspection only.

The present study addresses the use and non-use of resumptive pronouns (accusative and non-accusative) by means of qualitative and quantitative analysis of authentic conversational spoken Polish – a language variety in which co relatives typically occur. Additionally, introspection is utilized to reinforce and complement corpus data when we consider the relative acceptability or felicitousness of the presence or absence of resumptives. The focus of the analysis is on discovering and describing patterns of usage.

The corpus examples were retrieved from a collection of 1.6 milion words of conversational spoken Polish by an exhaustive search of all occurrences of the word co. Each occurrence was manually inspected and relevant tokens of the relative marker use of co were collected. The sample consists of 204 object relative clauses, which were analysed for factors relevant in resumption.

As is shown, the relatives are used in two broad configurations – unmarked (most frequent (81%), with null resumptives and inanimate referents) and marked (with overt resumptives and human referents). Both scenarios are linked to distinct strategies of case-recovery. The presence of the pronoun itself is one such strategy. In contrast, the omission of the pronoun is often (74%) accompanied by case-matching effects between the head and the trace which facilitate the omission. It appears that co relative clauses display a case-matching effect, whereby utterances in which the case form of the head matches that of the trace (te and tę in (2)) sound better – better integrated – than utterances with non-matching forms (ta in (2)). In relatives where the cases do not match, a resumptive may be used to resolve the awkwardness, as in (3).
Another typical property of co relatives is their preference for encoding definiteness of referents, as opposed to the readiness of który clauses to signal indefiniteness. This is evidenced by the frequent co-occurrence of co clauses with head-internal demonstratives (85%), as in (1). Interestingly, these head-internal demonstratives can also render resumptive pronouns unnecessary, thus constituting another factor relevant in resumption.

Keywords: spoken Polish, relative clauses, resumptive pronouns, case-matching, case-recovery

References
Comparing Naive Discriminative Learning, Sublexicon Phonotactics, and Analogical Learning

Matías Guzmán Naranjo and Elena Pyatigorskaya

Leipzig University, Leipzig, Germany
matias.guzman_naranjo@uni-leipzig.de; elenaln@gmx.de

Introduction. To explain how speakers choose between different diminutive forms in Russian masculine nouns (-ik, -chik, -ok), Gouskova, Newlin-Łukowicz, and Kasyanenko (2015) (from now on as GNK) argue for a model of lexicon/morphology where speakers store sublexicons of nouns with a sort of diacritic that indicate the diminutive form they take, and then use some form of phonotactic generalization over those sublexicons that they then use to produce novel forms. At its heart, this is an analogical model of morphology (Bybee 2010), but the implementation is much more complex and theory dependent than simple analogical models.

Objective. We wanted to compare GNK’s approach with a more traditional analogical model (implemented with a perceptron), and for reference also compare these two with a Naive Discriminative Learning approach (Arppe et al. 2014).

Methodology. We used the data set provided by GNK in their study, which consists of around 1300 masculine diminutive nouns derived with one of the three forms mentioned above. We expanded this list by extracting their frequency from Google ngrams. We then fitted a traditional NDL model using 2- and 3-graphs (sequences of two and three letters) within the nouns, and a simple analogical model using hand picked predictors: last phoneme of the word, position of stress, stressed vowel, last vowel, number of vowels, and length. Finally, we compared both models with GNK’s model.

Results. First of all, we found a clear flaw in GNK because they fit and evaluate their multinomial regression model on the same data set, which means that they are not really making predictions on new data. However, if one directly tests the predictions made by the grammars they generated with the UCLA phonotactics learner, the result is that the grammar (a) is more or less at chance level (the accuracy is no different, or inferior than the no information rate, and kappa = 0.073), (b) makes absolute predictions and completely misses the fact that the alternation is probabilistic (as their own study on nonce words shows). Similarly, the NDL model trained with 2- and 3-graphs performs at chance level both for categorical predictions and probabilistic predictions. Compared to these, the hand picked analogical model performs above chance (kappa = 0.26), and for probabilistic predictions the analogical model outperforms the NDL (lower brier scores, lower mean square errors, and higher correlation coefficients).

Finally, including frequency information reduced the performance of both the NDL and analogical models, which is specially interesting, because it indicates that types are more important for this particular alternation than tokens are (that is, high frequency types are not more relevant for the generalizations than low frequency types).
**Conclusion.** We conclude that a simple analogical (perceptron-based) model can better capture speakers performance when forming Russian diminutives than the alternative approaches.

**Keywords:** morphology, analogy, Russian diminutives, alternations

**References**
A New Syntactic Analysis of Dutch Nominal Infinitives

Kateřina Havranová

Palacky University, Olomouc, Czech Republic
havranova@email.com

This paper deals with Dutch nominalizations, namely with two types of nominal infinitives, and tries to draw parallels between the Dutch and English constructions. Traditionally, four types of constructions are considered to be Dutch nominalizations (GE-nominals, ING-nominals, and bare and determined nominal infinitives) since they all fulfill two basic criteria: Firstly, they inherit the denotation (namely the state of affairs) of the verb they are derived from and, secondly, they inherit the argument structure of that verb if interpreted as process (complex event) nominals.¹

However, the main focus of this paper are the nominal infinitives (NIs) of two types. The first type are bare nominal infinitives (NI-Bs) (1), the second type are determined nominal infinitives with the definite article het (NI-Ds) (1b).

(1) a Boeken lezen is interessant.
   books read is interesting
   ‘Reading books is interesting.’

   b Het lezen van boeken is interessant.
   the read of books is interesting
   ‘The reading of books is interesting.’

I demonstrate that although their external syntax is basically the same – they both have the distribution of NP/DPs fulfilling the syntactic functions of subjects, objects, PP-objects or adverbials – their internal structures differ. Both types of nominal infinitives also inherit the argument structure of the verb as well as its denotation, but they denote the action of the verb only as a process (event) and never as a result. Furthermore, I compare their nominal and verbal properties and examine how they behave with respect to different criteria such as the form of an internal object, determiners, pluralization, quantification, modification, etc.

Dutch nominal infinitives have been studied previously in detail by both traditional and generative linguists (e.g. Hoekstra, 1985; Zubizarreta & Van Haaften, 1988; Looyenga, 1992; Hoekstra, 1999; Schoorlemmer, 2002; Reuland, 2011; Broekhuis, 2012, etc.) but they have always been “notoriously difficult to analyze” (Schoorlemmer, 2002). Although these authors have managed to describe their behaviour extensively, they have not been able to show formally that they are in fact very systematic constructions or how they fit into a broader theoretical linguistic context. I depart from previous approaches in my treatment of nominal infinitives and propose a new analysis using only one syntactic operation that

¹ The terms process/complex event and result nominals are used in line with Grimshaw (1990).
combines Categorial Switch, a language particular phenomenon previously described by Panagiotidis & Grohmann (2009), and Merge, which is part of Universal Grammar, into a single step. This elegant solution also helps to explain how both types of Dutch nominal infinitives can combine with their objects in two different ways.

**Keywords:** (Dutch nominalizations, the nominalization process, Dutch nominal infinitives, external distribution, internal syntax)

**References**
Phimsawat (2011) proposes that the null inclusive generic pronoun in Thai has no φ-features.

(1)  

\[
\text{hàak Ø wáicay khôn Ø cà cùncay ʔeeŋ} \quad \text{[Thai]}
\]

‘If one trusts acquaintances one will regret it later.’

This explains (a) why it has inclusive generic reference, and (b) why it’s null. No φ-features means unrestricted reference, and that is what inclusive generic reference is: it includes the speaker, the addressee, and other people. It’s null because there are no φ-features and no root to spell out. So what is it if it has no φ-features? Just N, a minimal noun? But can there be a categorial feature N (or n), when there is no root?

The Thai null inclusive pronoun can be compared with the Finnish null inclusive pronoun.

(2)  

\[
\text{Tässä voi Ø istua.} \quad \text{[Finnish]}
\]

‘One can sit here.’

This pronoun has the φ-features 3SG, shown by the agreement on the finite verb. Biberauer et al (2010) shows that this is proper agreement, not default agreement. This suggests that the null subject has φ-features just because they are needed to assign values to the unvalued φ-features of T, and it can be null because the φ-features are spelled out on T. The Finnish inclusive generic pronoun has another characteristic: it can refer to humans only: (3) can only be about humans. It cannot include for example plants (along with humans).

(3)  

\[
\text{Sitä kasvaa hyvin jos saa hyvää hoitoa ja paljon ravintoa.} \quad \text{[Finnish]}
\]

‘One grows well if one gets good care and much nutrition.’

Why the feature [human]? We will explore the idea that it’s a default value of N, a minimal root. We propose that there can be no N without either semantic features (a root) or phonetic features (the latter to allow for expletives). Furthermore, there can be no φ-features without N.

The minimal φ-feature specification is 3SG. It follows that the null inclusive generic pronoun in Finnish is 3SG and restricted to human reference. The connection between agreement and human reference is confirmed by the observation that the generic construction in Thai is not restricted to human reference. (4) can generalize over humans, animals, and plants.

(4)  

\[
\text{thâa dâayráab khwaamrák khwaam?awjaysay kôo cà too rew} \quad \text{[Thai]}
\]

‘If one gets love and care, one will grow faster.’ [plants, animals, and humans]

The same is true of Japanese, Korean, Mandarin, Sinhalese, Vietnamese, all agreementless languages, while it is restricted to humans in Hebrew, Brazilian Portuguese, Bengali, Assamese, all languages with agreement. This suggests that the inclusive pronoun in the agreementless languages isn’t syntactically represented at all. However, the reflexive in (1) indicates that there is a syntactic null generic subject in this Thai sentence. If this can be
further substantiated for all the languages concerned, our initial hypothesis will need to be reconsidered.

**Keywords:** generic, inclusive, human, agreement

**References**


Lexical and Orthographic Distances between Czech, Polish, Russian, and Bulgarian – a Comparative Analysis of the Most Frequent Nouns

Klára Jágrrová, Irina Stenger, Roland Marti, and Tania Avgustinova
Saarland University, Saarbrücken, Germany
{kjagrova, avgustinova}@coli.uni-saarland.de
{ira.stenger, rwmslav}@mx.uni-saarland.de

We examine the lexical distances between four selected Slavic languages with special attention to orthography. The underlying assumption is that not only the share of cognates but also the transparency of orthography may increase the mutual intelligibility of related languages. This study is based on methods of Heeringa et al. (2013) for determining lexical and orthographic distance between related languages in correlation to geographic distance. The latter is, however, beyond the scope of our analysis. We modify their method as follows.

<table>
<thead>
<tr>
<th>Empirical basis</th>
<th>Heeringa et al.</th>
<th>Our approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source of data set</td>
<td>British National Corpus</td>
<td>Frequency lists from the national corpora of Czech, Polish, Russian, Bulgarian</td>
</tr>
<tr>
<td>Material</td>
<td>translations of the 100 most frequent English nouns</td>
<td>selection of the 100 most frequent nouns of CZ, PL, RU, BG</td>
</tr>
<tr>
<td>Language family</td>
<td>Romance, Germanic, Slavic</td>
<td>Slavic</td>
</tr>
<tr>
<td>Slavic languages</td>
<td>Czech, Slovak, Polish, Slovene, Croatian, Bulgarian</td>
<td>Czech, Polish, Russian, Bulgarian</td>
</tr>
</tbody>
</table>

Instead of translating the 100 most frequent nouns from the British National Corpus, which turn out to include instances that are considerably rare in other languages (e.g. form, effect, police), we systematically use original resources only in order to establish a more representative picture of the individual languages.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Heeringa et al.</th>
<th>Our approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation</td>
<td>in two steps:</td>
<td>CZ → PL, RU, BG</td>
</tr>
<tr>
<td></td>
<td>1. English → each language</td>
<td>PL → RU, BG, CZ</td>
</tr>
<tr>
<td></td>
<td>2. each language → other languages of the same group</td>
<td>RU → BG, CZ, PL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BG → CZ, PL, RU</td>
</tr>
<tr>
<td>Cyrillic to Latin transliteration</td>
<td>web application <a href="http://bg.translit.cc/">http://bg.translit.cc/</a></td>
<td>according to DIN 1460</td>
</tr>
</tbody>
</table>

In both approaches the decision process in case of polisemy of potential cognates is handled similarly: (i) if there are several possible translations, cognates are preferred; (ii) in case of multiple cognates, the orthographically closest is taken; (iii) no false friends, i.e. there is at least one overlap in meaning. For example, for the Czech word oko (eye) belonging to the 100 most frequent Czech nouns and having two possible translations in Russian, глаз and око, we choose the cognate око (arch. eye), because глаз, which is the 14th most frequent noun in
our Russian list, does not have a Czech cognate. Similar cases are observed in all investigated language pairs.

<table>
<thead>
<tr>
<th>Distance measures</th>
<th>Heeringa et al.</th>
<th>Our approach¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical</td>
<td>number of non-cognates for each combination</td>
<td></td>
</tr>
<tr>
<td>Orthographic</td>
<td>average Levenshtein distance in the cognate pairs</td>
<td></td>
</tr>
<tr>
<td>Geographic</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

While Heeringa et al. find no asymmetries on the lexical level in their test set, our results are lexically asymmetric in all language combinations, e.g. CZ-RU 27% (Czech decoder of Russian stimulus) vs. RU-CZ 21% (Russian decoder of Czech stimulus). Eventually, we got a smaller orthographic distance between BG and RU 13.47% (Cyrillic script) than between CZ and PL 37.91% (Latin script) with further asymmetries depending on the decoding direction for both pairs.

**Keywords:** receptive multilingualism, Slavic, lexical distance, orthographic distance

**References**


**Sources of lexical material:**


¹ We wish to thank Varvara Obolonchykova for her support in the automatic calculation of the orthographic distances.
Subextraction in Dative Clitic and Differential Object Marking Constructions

Ángel L. Jiménez-Fernández
University of Seville
ajimfer@us.es

1. Introduction, the problem and our goal. In current research on the structure of DPs in Differential Object Marking and Dative Clitic Constructions, there has been an explosion of proposals suggesting that the preposition a present in both accusative and dative objects is not a true P in Spanish, but a morphological marker (Demonte 1995; Cuervo 2003; Ormazabal & Romero 2013; Pineda 2013, a.o.). Adopting the idea that this P does not project into a PP but rather occupies the K(ase) head above DP, in this talk I analyse subextraction out of both accusative and dative prepositional DPs in psych constructions in Spanish Wh-questions, as illustrated in (1).

(1) a. ¿De qué edificio dices que no conoces [a ningún vecino]? 
   ‘Of what building do you say that you don’t know any neighbours?’

b. ¿De qué partido crees que no les ha gustado [a muchos votantes] la nueva normativa? 
   ‘Of what party do you think many voters didn’t like the new regulations?’

However, the Spanish data are far from clear in that by assuming that a P is opaque for extraction (Abels 2003), the prediction follows that Experiencers introduced by P a should induce island effects, contrary to facts. To solve this problem, in line with Haegeman et al. (2014), I argue that island effects in objects introduced by a in Spanish are multifactorial. I extend their analysis by suggesting that one of the factors mitigating islandhood is the functional character of some prepositions (Riemsdijk 1978), and that the P a in dative and accusative DPs is a functional preposition which heads a Kase Phrase, motivated by the case properties of P. Being endowed with an Edge Feature, this KP is a weak phase which allows subextraction.

2. Analysis. I claim that the dative and the accusative are transparent for movement since the preposition a is just a case-assigning element, resurrecting the suggestion that this kind of DP projects into a Kase Phrase, a functional projection independently proposed by Loebel (1994), Lyons (1999), Abney (1987), and Lamontagne & Travis (1986), as a category separate from D. By adopting the distinction between weak and strong phases (Bianchi & Chesi 2014; Jiménez-Fernández 2009), I propose that, as prepositional, this KP is a phase, but of a weak nature, whose specifier can be used as an escape hatch for extraction, because K has an Edge Feature. Lexical PPs, by contrast, are strong phases and no Edge Feature is carried by P. This predicts that under minimalist premises subextraction will be available in KP but not in PP. This prediction is borne out, as observed when comparing (1a-b) and (2):

(2) *¿De qué congreso crees que esa propuesta depende [de los organizadores]? 
   ‘Of what conference do you think that proposal depends on the organisers?’

Keywords: subextraction, dative clitic constructions, Differential Object Marking, Kase Phrase, phases
Selected References
The Hungarian Focus: Presuppositional Content and Exhaustivity Revisited

Tamás Káldi¹, Anna Babarczy² and Ágnes Bende-Farkas³

¹³Research Institute for Linguistics (HAS); ²Department of Cognitive Science (BME)
  kaldi.tamas@nytud.mta.hu; babarczy@cogsci.bme.hu; agnesbf@gmail.com

The study looks at factors behind the exhaustiveness effects of focus in general and in the Hungarian pre-verbal focus (preVF) in particular. Some theorists argue that the exhaustiveness of preVF is truth-conditional and thus invariably present (e.g. Szabolcsi 1994, É. Kiss 1998, Kenesei 2006). Others, however, claim that the exhaustive interpretation arises through an implicature (e.g. Wedgwood 2005) and is context dependent. Experimental work (e.g. Geröcs et al. 2014, Onea & Beaver 2011, Kas & Lukács 2013) has supported the latter view showing that the exhaustive interpretation is optional in a variety of behavioural tasks.

One objection to these experiments is that lexically marked focus sentences (only-F) were used as a control condition even though only-F sentences are believed to have a different presuppositional content than their preVF counterparts. While in 'John ate only the apple' the presupposed content is that 'John ate the apple', and the asserted content is that 'it is exclusively the apple', in 'It is the apple that John ate' the presupposition is that 'John ate something', while the assertion is that what he ate is 'the apple'. These differences may thus make the two sentence types incomparable. A second objection is that since the above studies used behavioural methods, they could only measure the 'end-product' of the process of interpretation but provided no insight into the process itself. We addressed these issues in the present experiments.

First, in an online survey 50 respondents completed preVF and only-F sentences with an ending that directly reflects either (a) on the exclusiveness of the object or (b) on the identity of the object. If the above theorising is correct, (a) should be preferred in the case of only-F, and (b) in the case of preVF sentences. Our results do not support these predictions. While for only-F sentences respondents chose (a) in 97.5% of cases, for preVF sentences the choice of (b) was around chance level suggesting that subjects are not necessarily sensitive to the presuppositional content of the preVF construction. Thus we conclude that preVF may be compared to only-F as a baseline.

Second, in two visual-world eye-tracking experiments participants listened to a pre-recorded preVF or only-F sentence and saw four images: an exhaustive scenario, a non-exhaustive image and two distractors. The task was to choose one image in the first experiment (N = 16), and any number of images in the second experiment (N = 30) that best matched the sentence. The results reveal that in the forced choice setup, subjects invariably chose the exhaustive image, but with no limit on the number of choices, both the exhaustive and non-exhaustive images were chosen 34.17% of the time. In the forced choice task, eye-tracking data revealed a gradually increasing preference for the exhaustive picture in both sentence conditions but in the free choice task, this trend was observed only for only-F sentences while looking preferences remained at chance level in the preVF condition even if an exhaustive response was given (Figures 1 & 2). This pattern suggests that first, task context plays a crucial role; and second, the non-exhaustive interpretation is not automatically exclu-
ed as the sentence is processed. Thus our results give further support to the pragmatic view of preVf interpretation.

![Figure 1](image1.png)

**Figure 1**

Linearly increasing trend over IPs: $F(2, 15) = 4.45, p = 0.34$. No main effect of Sent. Type.

![Figure 2](image2.png)

**Figure 2**

Main effect of Sent. Type: $F(1, 29) = 15.20, p = .001$. No significant trend in IP.

**References**


Introduction. The categorial distinction between ‘given’ and ‘new’ is frequently made in the literature on information structure. While the notions of ‘givenness’ and/or ‘newness’ have been extensively debated (e.g., Schwarzschild 1999; Kučerová 2007; Šimík and Wierzba 2015), their relation to the more general notion of ‘information status’ has received less attention (e.g., Prince 1981; Nissim et al. 2004; Riester 2008). The present paper reports the results of two online experimental studies on the syntax-semantics interface, which together show that not all information can be classified as exclusively ‘given’ or ‘new’.

Methodology. Experiment I (n=45) assumed a one-way design with three levels of independent variable. The phenomenon studied was the sensitivity of the Czech particle že in embedded wh-questions to the level of givenness of the proposition expressed by the answer to the question. Each test item consisted of a short story followed by a lead-in question and two sentences. The stories had three versions. Version I (see (1a)) forced the informant to interpret the embedded clause as a request for repetition. Version II (see (1b)) and Version III (see (1c)) forced the informant to interpret the embedded clause as a request for new information. The labels given to the resulting three conditions were ‘Given’, ‘Implied’ and ‘New’, respectively. The sentences that followed took the form of audio recordings and presented a minimal pair consisting of a particular wh-question with že and a corresponding wh-question without že (see (2)). The elicitation method used in the experiment was Forced Choice. Experiment II (n=70) was a follow-up study testing the phenomenon in matrix wh-questions for Implied and New conditions.

(1) One day, Peter talked to Martha about the purchase of her new car. Peter knew that Martha bought a new car, but, because she had been in a difficult financial situation for some time, he did not know where she got the money to buy it.
   (a) During their chat, Martha mentioned that she had to borrow the money, and then she said who lent it to her. When she was saying who it was, an ambulance was passing by, and Peter heard only ‘… lent me the money for my new car’. Peter was interested in knowing who lent Martha the money, so he asked her about it.
   (b) During their chat, Martha mentioned that she had to borrow the money, but she did not say who lent it to her. Peter was interested in knowing who it was, so he asked her about it.
   (c) Peter was convinced that somebody lent money to Martha, and, since he was interested in knowing who it was, he asked her about it.

(2) Petr chetěl vědět, kdo (že) Martě půjčil peníze na její nové auto.  
   Peter wanted know who že MarthaDAT lent moneyACC for her new car  
   ‘Peter wanted to know who lent Martha money for her new car.’

Results. The statistical method used to analyse the data was a version of Generalised Linear Mixed-effects Model. In Experiment I, there was a highly significant difference (p<.001) between Given and Implied/New. Interestingly, there was no significant difference (p>.05) between Implied and New (see Figure 1). In Experiment II, there was a highly significant difference (p<.001) between Implied and New (see Figure 2). (Note: The higher the value on the vertical axis, the higher the preference for že in the relevant condition (e.g., 20 = ‘an average of 20% of all trials in the relevant condition’).

Conclusion. The results of Experiment I are compatible with the ‘standard’ bi-categorial classification system (i.e., [given – new]). The results of Experiment II are compatible with a
classification system that makes either sub-categorial distinctions (i.e., [given – [new [implied – brand new]]]) or further categorial distinctions (i.e., [given – implied – new]). The results suggest that the representation of discourse with respect to ‘speaker knowledge’ needs to be more granular than usually assumed in the realm of information structure.

**Keywords:** syntax-semantics interface; information status; givenness; common ground

Figure 1 – Estimated means, including confidence intervals, for Experiment I.

![Figure 1](image1.png)

Figure 2 – Estimated means, including confidence intervals, for Experiment II.

![Figure 2](image2.png)

**References**


Counterscopal Morphology

Paul Kiparsky
Stanford University, Stanford, CA, USA
kiparsky@stanford.edu

The Mirror Principle (Baker 1985) stipulates that morpheme order and the constituent structure of words reflect semantic scope. The Mirror Principle can be violated when morphotactic constraints impose counterscopal affix orders and constituency. In this talk I show that such mismatches correlate with particular types of exceptional phonological and allomorphic behavior, and propose an analysis of them that explains this correlation between semantics, morphology, and phonology, as well as the Mirror Principle itself. The supporting data comes from inflection and derivation in morphologically rich languages.

I assume a generative lexicalist morphology, although the arguments could also be adapted to realizational frameworks such as Distributed Morphology. The key assumption is that morphological operations are cyclically interleaved with the processes that assign them a semantic and phonological interpretation. Thus a morphological combination (such as affixation and compounding) undergoes lexical phonology and semantics before the next morphological operation. This has the conceptual advantage of unifying the Mirror Principle with the cyclicity of rule application on the P-side. It also derives the basic morphological locality and inward sensitivity constraints (Carstairs 1987), and automatically excludes lookahead allomorph selection. And it correctly predicts that phonological allomorph selection depends only on the cyclically derived phonological form of their bases—not on the underlying form, or on the output form.

The mutual scope of derivational affixes can be DISTINCTIVE (the causative of a desiderative is distinct from the desiderative of a causative), while the mutual scope of inflectional affixes is PREDICTABLE: case invariably scopes over possessors, and possessors over number, similarly mood > tense > aspect, etc. Mismatches with morphological order and constituent structure occur with both types of scope relations (see (1b), (2b)). However, the mismatches show very different syndromes of correlated phonological and morphological anomalies, which diagnose distinct morphological structures.

Mismatches involving predictable scope result from affix bundling (“co-affixation”, “simulflication”): the counter-scopally ordered affixes are added in one step, viz. [Stem [ Affix, Affix, ] ]. Evidence for this structure comes from anticyclicity in phonology (no cycle on Stem+Affix,, cycle on Affix,+Affix,), from bidirectional dependencies in morphology (Affix, can depend on Affix, and vice versa), and from level-ordering paradoxes and bracketing paradoxes. The bundled affixes remain distinct morphemes but form a single morphotactic unit with the distributional properties of a portmanteau morpheme. I provide examples from Hungarian (see (3)), Finnish, Tundra Nenets, Upper Sorbian, and English.

Mismatches involving reversal of distinctive scope, on the other hand, cannot have this structure, since the semantic interpretation is not recoverable from an affix bundle. In such cases the morphotactics and semantic interpretation jointly force a derivation in which the affixes are inserted and interpreted in the derivational order that reflects their scope, but
placed in the sequence dictated by morphotactic constraints. These force affixes to be inserted below the previously introduced affixes that they scope over (Kiparsky 1982, cf. the morphological head operations of Aronoff and Richards’s notion of syntactic “tucking in”). This is confirmed by the phonological and morphological data. Phonological evidence comes from cases where Root + Affix, undergoes phonological processes before a scopally higher Affix splits them (“endocyclicity”, Hyman and Orgun 2005), and from outwardly-sensitive allomorphy selection. Examples will be cited from Karimojong vowel harmony (Lesley-Neumann 2007), Sanskrit prefix-suffix interactions, Enets (Yenisei-Samoyed), and Modern Greek.

(1) a. ci-kâr -ṣ -ay -a -ti ≠ ci-kâr -ay -iṣ -a -ti
   make -Desid -Caus -Pres -3Sg make -Caus -Desid -Pres -3Sg
   ‘causes to want to make’ ‘wants to cause to make’ (Sanskrit)

   cry Caus Applic stir Caus Applic stir Applic Caus
   ‘[ cause to cry] with ]’ ‘cause [to stir with]’ ‘cause [to stir with]’ (Chichewa)

(2) a. köpek -ler -im -de ‘in my dogs’ (Turkish)
   dog -Pl -my -Locative

   b. koir -i -ssa- ni ‘in my dogs’ (Finnish)
   dog -Pl -Nessive -my

(3) Hungarian: affix bundling resolves level-ordering violations

   a. (a) rohan -t (b) rohan -hat- ott (Törkenczy 2004)
      run -Past run -Modal -Past
      ‘he ran’ ‘it is/was possible that he ran’ (not ‘it was possible that he would run’)

   b. Semantics: (((run)Past)Possible)

   c. A level-ordering paradox: Modal -hat is a level 2 (word level) ending, because it attaches without a linking vowel to any stem that ends in a single consonant. Past tense -(V)t(t) is a level 1 ending, because its initial linking vowel depends on the last consonant of the stem.

   d. Affix bundling analysis: /hat/ level 1, hat-ott level 2, rohan-hat-ott

References
The research aims to investigate several features of inaugural addresses of the presidents of the United States. Although many analyses have been done in this field, most of them are qualitative or concentrated to content analysis (e.g. Lim 2004; Liu 2012). Our approach is strictly quantitative based on the last methods of the contemporary quantitative linguistics (Popescu et al. 2009; Čech et al. 2015, Kubát & Milička 2013). The goal of the paper is to observe the presidential speeches from a viewpoint of stylometry indices and to discover whether political and historical circumstances (wars, financial crisis, ideology, etc.) influence the style of inaugural addresses; analogically to findings presented by Čech (2014).

Specifically, vocabulary richness, thematic concentration and text activity are computed and differences among texts are statistically tested (Čech et al. 2015; Kubát & Milička 2013; Zörnig et al. 2015). Moreover, the thematic concentration also allows extracting so called thematic words, i.e. words which represent main topic(s) of text. The thematic words (TW) can be viewed as an alternative to keywords (Čech et al. 2015). The advantage of TW lies in the fact that those words are based solely on the frequency structure distribution of the text; no reference corpus is needed for the analysis. These three indices were chosen especially due to (a) their high efficiency of automatic text classification (genre analysis, authorship attribution, etc.), (b) their independence on text length and (c) simple linguistic interpretation. The combination of the three methods allows both to investigate the style of the particular presidential speeches in powerful linguistically comprehensive view and to observe the development trends of the specific genre of inaugural addresses during the more than 200 years long history. The corpus comprises inaugural addresses of all US presidents from George Washington to Barack Obama (57 texts in total).

The obtained results are linguistically and historically interpreted and compared with previous works which dealt with presidential speeches (Savoy 2010; Lim 2004; Liu 2012; Tuzzi et al. 2010; Čech 2014).

Keywords: stylometry, presidential speeches, vocabulary richness, thematic concentration, activity of text

References


φ-Feature Resolution in a DP Coordination:  
A Case Study at the Syntax-Semantics Interface

Ivona Kučerová  
McMaster University, Hamilton, ON, Canada  
ivona@alum.mit.edu

Most of the current literature agrees that the label of a conjoined DP is determined both by semantic and morpho-syntactic features (Farkas and Zec 1995, King and Dalrymple 2004, Heycock and Zamparelli 2005, a.o.), even though strictly morpho-syntactic (e.g., Marušić et al. 2015) and semantic approaches have been proposed as well (e.g., Lasersohn 2013). Less is known about the formal connection between the morpho-syntactic and semantic features. This paper provides novel empirical evidence that the connection at the level of the CI interface is established via PERSON, modeled as [±PARTICIPANT] (Nevins 2007). The formal connection between syntactic PERSON and semantic PARTICIPANT allows for a direct CI association of a PERSON feature with an index. The data comes from DP coordinations with numeral constructions in Slavic, primarily Czech.

The puzzle: Farkas and Zec (1995) observe that the conjoined DP φ-features (as identified by predicate agreement) are identical to φ-features of a pronoun anaphoric to the coordination. The pattern shows that there is a connection between a pronominal φ-feature resolution and φ-feature resolution on coordinated DPs. Czech (but also Polish and Slovenian) numeral constructions with 5 and above (5&up) are a striking exception. Even though their coordination yields obligatory singular agreement (failed agree; Preminger 2009), the corresponding anaphor is in plural, (1).

(1) Pět chlapců a pět dívek se sešlo/*šli v klubu.  
five boys. GEN and five girls. GEN REFL met. N.SG/* M.PL in club  
*Ono/ oni. . .  
it. SG.N /they. M.PL  
‘Five boys and five girls got together in the club.’ DEFAULT AGREE

5&UPs are also the only DPs that in coordination optionally trigger closest conjunct agreement (CCA), (2). If, however, the other conjunct is 1st or 2nd person, the agreement must be plural, (3).

(2) Děvčata a pět chlapců šli/šlo/*šla do ZOO.  
girls. N.PL and five boys gone. M.PL/N.SG/N.PL to ZOO  
‘Girls and five boys went to the ZOO.’ CCA OR PLURAL

(3) Já/ty a pět chlapců jsme/jste šli/*šlo do ZOO.  
I/you and five. NOM boys. GEN.PL AUX.1/2.PL gone.M.PL/ gone.N.SG to ZOO  
‘I/you and five boys went to the ZOO.’ PLURAL
The proposal: I argue that the key to the pattern lies in the observation (i) that 5\&UPs are not labeled for φ-features (they trigger only a default agreement; their anaphoric pronouns must agree in semantic features, never in N.SG; they cannot licence secondary predicates; data omitted for reasons of space), and (ii) plural agreement (based on boolean conjunction) arises only if the other conjunct provides a PERSON feature (obligatory for [+PERSON], optional for [−PERSON]). I follow Farkas and Zec (1995) and others in that features of a coordinated DP are computed as a combination of semantic and morpho-syntactic features. The primary semantic feature is PERSON, modeled as [±PARTICIPANT] (Nevins 2007), which allows for a direct CI association with an index (obligatory for [+PERSON]/ [+PARTICIPANT]). Semantic coordination is based on indices associated with individual conjuncts (matching indices ⇒ SG; non-matching ⇒ PL). Semantic plurality is obligatory if at least one of the conjuncts is [+PERSON]. 5\&UPs lack a PERSON feature but if they combine with [+PERSON], agreement is still PL because the [+PERSON] conjunct provides an index, and there is no matching index on the 5\&UP conjunct, (3). If the coordination consists solely of 5\&UPs, there is no PERSON feature to compute semantic agreement. In addition, since 5\&UPs are numberless, the system cannot calculate the number from morpho-syntactic features either, and the only plausible agreement is post-syntactic agreement with the closest conjunct (Bhatt and Walkow 2013), (1). If one conjunct is [−PERSON], the system can either track the PERSON feature (⇒ PL; syntactic agree), or morpho-syntactic features (⇒ CCA; post-syntactic), (2).

Keywords: numerals; Agree; coordination; Czech

References
The (Not So) Inverse Linking Pattern in Plains Cree

Yuriy Kushnir
Leipzig University, Leipzig, Germany
yuriy@lingvistas.com

Background and claims. In this work, I argue that the so-called inverse agreement pattern observed in the Algonquian language Plains Cree is, in fact, a system displaying the more common typological feature called differential case marking, influenced by an intrinsic salience hierarchy built into the grammar. By re-analyzing the direction markers of Plains Cree as case markers, I assert that there is no (stipulative) morpho-syntactic feature of the type [±INV] as assumed in previous work.

Data. The most striking typological feature of Cree is the so-called inverse agreement pattern (Klaiman 1992, Zúñiga 2006) driven by the following language-specific salience hierarchy: 2 » 1 » 3 » 3’ » 3”. A basic example can be seen below:

(1) ki-waapam-aa-w-ak
   2-see-DIR-3.ANIM-3PL
   ‘You see them.’

(2) ki-waapam-iko-w-ak
   2-see-INV-3.ANIM-3PL
   ‘They see you.’

The direction markers marked in bold indicate whether the argument configuration is aligned properly with the salience hierarchy (i.e. whether the agentive argument is more discourse-salient than the patient). The nature of these markers is, however, questionable since some of them encode not only directionality, but also φ-features of the respective argument DPs (3-6). Also, it is not entirely clear how ‘inverseness’ is formally implemented: most previous work either states the descriptive generalization without any formal derivation or has to postulate a stipulative feature of the kind [±INV] (e.g. Stump 2001).

(3) ki-waapam-i-n
   2-see-DIR(2›1)-SG
   ‘You see me.’

(4) ki-waapam-iti-n
   2-see-INV(1›2)-SG
   ‘I see you.’

(5) ni-waapam-aa-w
   1-see-DIR(SAP3)-3.ANIM
   ‘I see him.’

(6) waapam-ee-w
    see-DIR(3)-3.ANIM
    ‘Heprox sees him/them.’

If we juxtapose these latter examples with (1) and (2), it becomes evident that the direction markers reflect various properties of the arguments. The directionality marking system is thus not at all ‘clean’. I therefore argue that what hides behind these affixes are actually fairly straightforward case markers whose distribution is driven by the inherent salience hierarchy outlined above. This analysis continues along the guidelines outlined in Aissen 2002 and Wunderlich 2005, and provides a deeper account for the agreement pattern in Plains Cree than that described in traditional reference grammars (Dahlstrom 1991, Wolfart 1973) or work dedicated to inverse languages.

Analysis. I assume that structural cases assigned in syntax are further decomposed into binary features, one of them being [±subj]. An argument marked for [±subj] is deemed to be in the Ergative case; [–subj] is the Accusative, and a marker in the case slot which is
underspecified for [±subj] is considered to be in the Absolutive (unmarked) form. This creates the following tripartite system: Abs ↔ Erg ↔ Acc.

In a language with differential case marking (DCM), various arguments will or will not bear overt case morphology depending on their internal properties. If we assume the inherent hierarchy mentioned in the paragraphs above, it may be postulated that SAP (speech act participant) arguments are prototypical agents, while third-person actants are expected to be patients. In this configuration, the interpretation is trivial and therefore no case marking is needed for either one of the arguments. However, if one (or both) of the arguments licensed by the verb is/are atypical, e.g. 3›SAP / SAP›SAP or 3›3, a case marker must be used to reflect this.

In order to implement my idea, I use the Distributed Morphology (DM) framework. Before a vocabulary item may be inserted into a terminal, the features of a functional head are manipulated by the morphological apparatus in order to enable the hierarchy effects mentioned above. Two crucial impoverishment rules deleting case features apply as follows:

1. +subj → Ø / { +part } 2. –subj → Ø / { –part }

The feature [±part] distinguishes between SAP and non-SAP arguments. These two rules insure that prototypical subjects and objects are not marked for case. As for the actual vocabulary items, they are represented by several straightforward case markers:

1. { +subj –part } ↔ /-iko/ 3.ERG 4. { +part } ↔ /-aa/ 1/2.ABS
2. { –subj +part –2 } ↔ /-i/ 1.ACC 5. { –part } ↔ /-ee/ 3.ABS
3. { –subj +part (+2) } ↔ /-iti/ 2.ACC

The fact that the ergative is not marked in 3.prox › 3.obv configurations can be captured by a further impoverishment rule which effectively removes the [±subj] feature in these contexts based on the fact that the object is even lower on the hierarchy scale than the subject.

Conclusion. In this paper, I have shown how a relatively simple case-marking system can provide an adequate account for the inverse pattern observed in Plains Cree. This re-addresses the challenge of analyzing such systems in terms of their formal behavior on deeper representation levels.

Keywords: morphology, inverse, algonquian, case, argument alignment.

References
Nasal Place Neutralization in Classical Nahuatl: A Derivational OT Account

Michał Kuźmicki
University of Warsaw, Poland
Mikaelk89@gmail.com

1. Introduction. The paper offers a phonological account of a previously unrecognized pattern in the distribution of nasal segments in Classical Nahuatl. I argue that their behaviour is driven by place restrictions on coda nasals. However, there is an asymmetry between morpheme-internal and morpheme final codas. I show that standard Optimality Theory is unable to handle these patterns but that the problem can be easily solved if the possibility of level distinction in the evaluation of output forms is admitted.

2. Descriptive generalizations. The two nasal stops in Classical Nahuatl, bilabial [m] and alveolar [n], contrast before vowels but elsewhere their distribution is restricted. On the basis of data gathered from descriptive sources (Lockhart 2001, Campbell & Karttunen 1989), three generalizations may be drawn: (a) Morpheme-internally, nasals agree in terms of place with the following consonant unless it is dorsal, when [n] surfaces, see (1); (b) Word-finally, only [n] is found. This static generalization is supported by alternations, whereby underlying /m/ delabializes when followed by a word boundary, as in (2); (c) Across morpheme boundaries, NC sequences diverge: [n] can appear before any segment, whereas [m] is only found before other bilabials. Thus, underlying ns are invariant (3) but underlying ms alternate with [n], (4).

(1) pampa [pampa] ‘on behalf of’; intlā [intla] ‘when’; cenkah [senka] ‘very, greatly’
(2) nemi [nemi] ‘He/she lives’ – nen [nen] ‘He/she lived’
(3) connaca [k+on+maka] ‘He gives it to him’ – conhuica [k+on+wika] ‘He takes it away’
(4) ampāquih [am+p:ki+ʔ] ‘You (PL) are happy’ – annemi [an+nemi+ʔ] ‘You (PL) live’

3. The problem. The lack of restrictions in the prevocalic position, coupled with the fact that syllables in Classical Nahuatl are maximally CVC, suggests that the observations above can be attributed to a version of the Coda Condition constraint (Itô 1986) that penalizes a nasal segment’s Place if it is not associated with the syllable onset (NCC). If place assimilation is viewed as place sharing, observation (a) can be explained by ranking NCC above IDENTITY constraints that require place faithfulness. Assimilation to dorsals is blocked due to an undominated segment inventory constraint banning the velar nasal, *[ŋ]. In this case, the least marked alveolar nasal emerges as a result of the universal ranking *PL(LAB) » *PL(COR). However, the ranking required to analyze place assimilation, shown in (5), yields incorrect results for observations (b) and (c). For word-final nasals, it predicts that the faithful candidate should always win, as both word-final [n] and [m] violate NCC. For sequences spanning two morphemes, it predicts that not only underlying ms but also ns should assimilate.

(5) HAVEPLACE, *[ŋ], NCC » IDENT(PLACE) » *PL(LAB, DOR) » *PL(COR)

To derive the correct results for observations (b) and (c), it is necessary to use a version of NCC relativized to labials, while the more general version of NCC should be ranked low:

(6) HAVEPLACE, *[ŋ], NCC(LAB) » IDENT(PLACE) » *PL(LAB, DOR) » *PL(COR), NCC

Although this ranking works for observations (b) and (c), it predicts that morpheme-internally, /n/ should not assimilate to the following labial sounds, contrary to the facts.

4. Proposal. The ranking paradox described above can be solved if the difference between the stem level and word level is recognized in OT (Kiparsky 2000, Rubach 2000, Bermúdez-Otero 2003). At stem level, high-ranked NCC results in placeless nasals in the coda position, both word-internally and word-medially (Table 1) if HAVEPLACE is low ranked and place sharing is penalized by *MULT-LINK (Rubach 2000, 288). At word level, NCC and *MULT-LINK are demoted, while NCC(LAB) and HAVEPLACE are ranked higher (Table 2). This not only yields the correct results for observations (a–c) but also avoids the directionality problem discussed, e.g. by Wilson (2001) and McCarthy (2008).
Table 1. DOT derivation: stem level (hypothetical inputs used for compactness)

<table>
<thead>
<tr>
<th></th>
<th>MULT -LINK</th>
<th>*[ŋ]</th>
<th>NCC</th>
<th>IDENT (PL)</th>
<th>*PL (LAB)</th>
<th>*PL (COR)</th>
<th>NCC (LAB)</th>
<th>HAVE PLACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>//panpa//</td>
<td>a. pan.pab</td>
<td>![nst]</td>
<td>![nst]</td>
<td>**</td>
<td>*</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
</tr>
<tr>
<td></td>
<td>e. panp.pab</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
</tr>
</tbody>
</table>

Table 2. DOT derivation: word level (hypothetical inputs used for compactness)

<table>
<thead>
<tr>
<th></th>
<th>MULT -LINK</th>
<th>*[ŋ]</th>
<th>NCC</th>
<th>IDENT (PL)</th>
<th>*PL (LAB)</th>
<th>*PL (COR)</th>
<th>NCC</th>
<th>HAVE PLACE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b. nen</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
</tr>
<tr>
<td></td>
<td>c. neN</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
</tr>
</tbody>
</table>

Table 2. DOT derivation: word level (hypothetical inputs used for compactness)

<table>
<thead>
<tr>
<th></th>
<th>MULT -LINK</th>
<th>*[ŋ]</th>
<th>NCC</th>
<th>IDENT (PL)</th>
<th>*PL (LAB)</th>
<th>*PL (COR)</th>
<th>NCC</th>
<th>HAVE PLACE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b. on.pan.pab</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
</tr>
<tr>
<td></td>
<td>c. on.pam.pab</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
</tr>
<tr>
<td></td>
<td>e. om.pam.pab</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
</tr>
</tbody>
</table>

Table 2. DOT derivation: word level (hypothetical inputs used for compactness)

<table>
<thead>
<tr>
<th></th>
<th>MULT -LINK</th>
<th>*[ŋ]</th>
<th>NCC</th>
<th>IDENT (PL)</th>
<th>*PL (LAB)</th>
<th>*PL (COR)</th>
<th>NCC</th>
<th>HAVE PLACE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b. am.nen</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
<td>![nst]</td>
</tr>
</tbody>
</table>

**Keywords:** Classical Nahuatl, Optimality Theory, nasals, delabialization, assimilation

**References**


Legate (2014) proposes a coarse-grained Initiator “macro-role” in her analysis of non-canonical voice constructions (see also Ramchand 2008). I argue below for a finer-grained treatment of causers based on Legate’s own discussion of accusative survival in Ukrainian and Icelandic passives. My principal claim is that non-volitional causers are not external arguments and, as such, are not licensed by features of a Voice head. Further evidence for distinguishing non-volitional causers from the more general Initiator role comes from the restricted distribution of causers in Russian and Lithuanian “pain verbs” which also occur with accusative in the absence of an external argument. In the case of pain verbs, I show that VP-internal, non-volitional causers do not always find an (appropriate) interpretation when merged in the specifier of VoiceP. Note first the Transitive Impersonal Passive in Ukrainian (1) and its apparent counterpart in Icelandic (2), referred to by Jónsson 2009 and Sigurðsson 2011 as the “New Passive”. In both constructions accusative appears on the thematic object and the participial morphology is default.

(1) Tabir bulo zajnato amerykans’kym vijs’kom.
   camp:ACC was occupied:PASS American troops-INST
   ‘The camp was occupied by American troops.’ [Lavine 2013:187–194]
(2) Það var skoðað bílinn af bifvélavirkjanum.
   there:EXPL was inspected:PASS car.the:ACC by car.mechanic.the
   ‘The car was inspected by the mechanic.’ [Jónsson 2009: 294]

Legate’s analysis falsely predicts that the passive-like construction in Ukrainian (3), involving a Natural Force Causer, should carry over to Icelandic (4) if the Natural Force Causer is just one instantiation of the all-inclusive Initiator (Legate 2014: 94). In (4) we see, instead, that the Icelandic New Passive fails with a Natural Force Causer.

(3) Derevo bulo vypaleno soncem.
   wood:ACC was:PST burned:IMP sun:INST
   ‘The wood was burned/warped by the sun.’ [Lavine 2013: 187–189]
(4) *Það var rekið báttinn á land af vindinum.
   there:EXPL was driven:PASS boat.the:ACC to land by wind.the
   [Intended: ‘The boat was driven ashore by the wind’] [H. Sigurðsson, p.c.]

Natural Force, it follows, is not licensed by (Legate’s) Voice. This suggests that accusative assignment in Ukrainian (3) might have to do, instead, with a lower v-CAUSE head and its associated feature of causation. Pylkkänen (2008) observes that transitive causative constructions occur in the absence of an external argument (and Voice), citing the Japanese Adversity Causative as an example (90–91). The patently non-passive Adversity Causative
can even occur with a "by-phrase" that names the causing event, just as Ukrainian 
soncem:sun-INST names the causer in (3). On the present analysis, in which causers are not all treated as syntactically equivalent, implicit Agents, as in the passives in (1–2), are introduced by Voice, whereas Voice plays no role in introducing non-volitional causers. I suggest below that the latter identify v-CAUSE as active and accusative-assigning. It follows that Ukrainian (3), in point of fact, is not passive, despite its passive-participial morphology and surface similarity with the Ukrainian passive in (1). Now consider the pain verbs in Russian and Lithuanian (5–6). In Russian (5b), if we merge the causer (‘hunger’) in Spec,VoiceP as a generic Initiator, where it receives nominative, the idiosyncratic reading is lost and the sentence is pragmatically odd. In Lithuanian (6b), higher merger of the causer (‘intense light’) is impossible, suggesting the absence of a VoiceP in this construction altogether. That is, some non-volitional causer constructions have no canonical transitive counterpart, which provides further evidence against the identical treatment of causers and other Initiators.

(5)a. Život krutit ot goloda. b.*#Golod krutit život.
   stomach:ACC turn:IMP from hunger hunger:NOM.3SG turn:PRS.3SG stomach:ACC
   ‘My stomach is churning from hunger’
(6)a. Nuo ryškios šviesos jam skaudėjo akis.
   from intense light he:DAT hurt:PST.3 eyes:ACC intense light:NOM
   ‘His eyes hurt from the intense light.’
   [Intended: ‘Intense light hurt his eyes’]

Keywords: Initiator, non-volitional causer, Voice, v-CAUSE, passive

References
The Influence of Sequential Iconicity on the Ordering of Czech Adverbial Clauses

Michal Láznička
Charles University, Prague, Czech Republic
michal.laznicka@ff.cuni.cz

In an analysis of temporal adverbial clauses in English, Diessel (2008) concludes that sequential iconicity is a strong predictor of clause order. This term refers to the tendency for the order of events in discourse to reflect the perceived order of events in real world (e.g. Haiman 1985). Such situation is illustrated in 1., where the emptying preceded the lighting both in reality and in discourse. However, the subordinate clause (SC) may also precede the main clause (MC), the order of events being marked by the connective než ‘before’.

1. Dopil sklenici, než si zapálil. ‘He emptied the glass, before he lit a cigarette.’
2. Než si zapálil, dopil sklenici. ‘Before he lit a cigarette, he emptied the glass.’

There are only a few studies addressing the principles governing clause order in Czech, most of the authors agreeing that information structure plays an important role in the positioning of SCs whose order relative to the MC is not fixed (e.g. Havránek and Jedlička 1970). However, there is no empirical research on the role of different potentially important factors influencing clause order.

In this paper, I follow Diessel’s analysis and elaborate on Chudobová’s (2011) pilot corpus study by addressing the influence of iconicity on clause order. In the analysis, I focus on the connectives než, dříve/předtím než ‘before’; až, potom/poté co ‘after’. These are particularly interesting because the connectives provide information about the order of events by marking the SC as either preceding or following the events depicted in MC, at the same time their ‘time reverse’ potential goes against the iconicity principle. These connectives are contrasted with zatímco ‘while’ and protože ‘because’. While the former marks simultaneity of events, the latter is used in causal adverbial clauses and has no inherent temporal meaning. The “cause event” may thus precede, follow or happen simultaneously with the MC event.

To provide a complex picture of the clauses’ behavior, I use samples of 500 instances for each connective extracted from the corpus SYN2015 (Křen et al. 2015), treating the ‘before’ and ‘after’ connectives both as a single construction. I code the individual sentences for following variables: SC position (ante-, inter-, or postposition), conceptual order (iconic, non-iconic), relative length of SC (number of words in SC divided by number of words total), SC complexity (single clause, other subordinate), sentence complexity (other MC or other clause subordinated to the MC), discourse status of event (given, new), subject identity (same, other), verb lemma, verb tense, aspect, and mood. This rich annotation allows to assess the role of iconicity and all the other potentially relevant factors using regression models. It also provides a general profile of the constructions especially with respect to TAM configurations of both MC and SC verbs. Based on the literature, I predict to find a combined effect of discourse status, complexity, and iconicity, or, in the case of the ‘before’ and ‘after’ connectives, “reversed” iconicity, as suggested by a preliminary corpus query.

Keywords: clause order – Czech adverbial clauses – sequential iconicity
References
Czech Habitual Morpheme -va- as an Expression for Coding Gaps in Time Reference

Eva Lehečková and Jan Křivan
Charles University in Prague, Czech Republic
Eva.Leheckova@ff.cuni.cz; Jan.Krivan@ff.cuni.cz

The paper focuses on the status of Czech imperfective verbs formed by a grammaticalized morpheme -va- that conveys the meaning traditionally called habitual. First part of the paper concentrates on the description of the morpheme coding the past time reference. Second part focuses on a generalized account of -va- regardless of the tense category.

Two accounts of Czech -va- markers are at hand. First, there is a prevalent claim in existing literature on habituality in Czech (e.g. Kopečný 1962, Danaher 2003, among others) that habitual -va-verbs in the past tense denote a special meaning of a distant past (cf. (1) from Danaher 2003). Second, Plungian and van der Auwera (2006) in their cross-linguistic investigation of a specific class of “idle” past markers introduced the temporal concept of a “discontinuous past” (markers coding discontinuity of the past and the present).

(1) V dětství jsme často chodí-li / chodi-li k babičce.
As children AUX:1PL often go-HAB-PST:PL(IPFV) / go-PST:PL(IPFV) to grandmother
“As children, we often used to visit / visited grandmother.”

We assume that the latter account is adequate for Czech, too. In order to verify it, we carried out a corpus survey based on SYN2015 (Křen et al. 2015) where a set of habitual verbs was searched for. We have extracted a random sample of 1500 concordances and coded the dataset for a number of relevant formal and semantic features (e.g. lemma, tense, structural and event type of the verb and its temporal and measure adjuncts). A cluster analysis (see e.g. Gries 2009, Levshina 2015) was then performed so as to show what distinct subtypes of -va- verbs emerge from the data.

With the support of our corpus analysis, we argue that in contrast to languages that consistently express the distant past meaning by a specific marker (such as Makaa and Choctaw), Czech -va- verbs do not encode any objectively perceivable distant past meaning opposed to a recent past meaning. We show that Danaher’s account is essentially wrong and it does not account for empirical evidence properly. The distant past meaning in Czech is only one of several contextually induced components of a complex meaning of a habitual marker -va-. We suppose that the language specific past -va- coding in Czech is essentially based on the discontinuity temporal semantics suggested by Plungian and van der Auwera.

Following a well-known Reichenbach’s schema (1966 [1947]) we describe the crucial temporal meaning of the -va-form in the past as a gap between the time of speech (S) and the time of secondary reference (R, denoting the moment when a habit ceased to exist). Usage of -va-forms in Czech depends on two properties that vary according to cognitive and extra-linguistic conditions: (i) variable of time-span of E, (ii) variable interval between R and S.

Crucially, from a general perspective the properties (i) and (ii) allow to account for the whole range of habitual usage of -va- marker in Czech, irrespective of the value of the tense attribute. We predict that the -va- marker shares the identical inherent feature in all its
occurrences – a gap in time reference (for the future the time between S and R, for the past the time before R, in both cases R denoting the moment when a habit was initiated). Finally, we discuss this prediction and its consequences in a typological perspective of languages with “discontinuous past” markers.

**Keywords:** habituals, Czech, distant past, time reference points

**References**


This pilot study investigates the cross-linguistic asymmetries in the construals of social interaction, focusing on the so-called T/V-pronouns (e.g. French *tu* and *vous*, German *du* and *Sie*, Russian *ty* and *vy*) in twelve European languages from the Indo-European and Finno-Ugric families (Bulgarian, Czech, Dutch, French, German, Greek, Romanian, Russian, Spanish and Swedish; Finnish and Hungarian) . These construals represent an elusive object of investigation because the communicative situations where one or the other pronoun is used are difficult to compare cross-culturally. Film subtitles in different languages offer a convenient solution of these problems because the situations of communication between film characters can serve as convenient comparative concepts (Haspelmath 2010) and represent diverse social relationships.

To compare the pronouns, I first identify approximately 200 contexts where the pronoun *you* is used in the English version of the online user-made subtitles of ten films which represent different genres. Next, I identify the personal pronouns used in the translations. I also code the relationships between the speaker and addressee (age, social status, level of familiarity, etc.) based on an in-depth analysis of multimodal evidence from the films. On the basis of these data I perform a series of multidimensional analyses (Multiple Correspondence Analysis with supplementary points and Multidimensional Scaling), creating token-based semantic maps (e.g. Wälchli & Cysouw 2012) and type-based constructional maps. These methods allow me to pinpoint the differences and similarities between the language-specific pronouns and compare the ways the languages ‘carve up’ the pragmatic space.

The results of this pilot study can be used for development of questionnaires that would allow for a fine-grained analysis of the pronouns, and for formulation of hypotheses that can be tested on authentic (non-translated) data. I will also discuss the implications of this study for translation theory and practice.

**Keywords:** personal pronouns, parallel corpus, probabilistic semantic map, multivariate statistics

**References**

---

**A Pilot Study of T/V-Pronouns in European Languages Based on a Parallel Corpus of Online Film Subtitles**

*Natalia Levshina*

Leipzig University, Leipzig, Germany

Natalia.Levshina@uni-leipzig.de
Use of Parallel Texts as a Methodology in Cultural-Cognitive Linguistics: A Case Study on English-Chinese and English-Czech Parallel Texts

Wei-lun Lu
Center for Chinese Studies, Masaryk University, Czech Republic
weilunlu@gmail.com

The present paper addresses the usefulness of parallel literary texts in studying metaphor variation across languages and cultures, which is expected to help define cognitive linguistics as a multi-disciplinary field by attracting scholars in contrastive linguistics, corpus linguistics, linguistic typology, cultural linguistics, stylistics, translation, and literary studies.

The role of cultural context has been extensively recognized in metaphor studies. Different languages and cultures provide their users with different socio-cultural environments, with which the speakers understand, conceptualize and interact with the world around them. Accordingly, the metaphor (or the source concept involved therein) that different languages and cultures may utilize in talking about the same event is bound to vary.

Past research on the variation of metaphor has been methodologically based on use of introspection, mono-lingual corpora and experimental methods. However, as language production is heavily influenced by all sorts of context, there has been no way of studying the interaction of culture and metaphor by controlling for the same linguistic, physical and social context, while keeping the language production contextualized. In view of this problem, I believe that use of parallel texts (translations) constitute an efficient methodological opportunity for studying metaphor across languages and cultures in a contextualized way—If we see the author and the translator(s) as sensible text producers, then by keeping the same most other contextual factors, including linguistic, physical, social context, production mode and genre, researchers are allowed to empirically study the role played by the cultural context in metaphor use in stretches of discourse where all text producers try to get across highly similar (if not identical) messages. However, use of parallel texts (or translation) in studying metaphor has not received much attention, with only few exceptions (Chapter 7 in Kövecses 2005; Rojo and Ibarretxe-Antuñano 2013; Schaffner 2004; Tabakowska 1993).

Accordingly, the paper will present various parallel samples from English with translations from and to Chinese and Czech, using world masterpieces such as Alice in Wonderland (by Lewis Carroll) and The Republic of Wine (by Mo Yan), in order to testify the usefulness of the methodology in cognitive linguistics.

Keywords: metaphor, cultural variation, parallel texts, literature

References
Polyfunctionality and German Modals

Dagmar Machová
Tomas Bata University in Zlín, Czech Republic
machova@fhs.utb.cz

The paper argues that in German (and also in English), there is a link between formal properties of modal elements and modal polyfunctionality, i.e. the ability of a morpheme to express simultaneously both epistemic and deontic modalities. More precisely, the paper proposes the hypothesis that polyfunctionality of German modals triggers the absence of agreement in the singular paradigm in 1st and 3rd persons – see the following set, comparing a lexical verb machen ‘make’ and a polyfunctional modal verb müssen ‘must’:

(1)  

<table>
<thead>
<tr>
<th>machen ‘make’</th>
<th>müssen ‘must’</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) ich mache</td>
<td>ich muss Ø</td>
</tr>
<tr>
<td>(b) du machst</td>
<td>du musst</td>
</tr>
<tr>
<td>(c) er macht</td>
<td>er muss Ø</td>
</tr>
</tbody>
</table>

To test the hypothesis, the paper analyzes semantic and formal properties of several German verbs. These include dürfen ‘can’, können ‘can’, mögen ‘will’, müssen ‘must’, sollen ‘should’, wollen ‘will’, brauchen ‘need’, werden ‘become’, wissen ‘know’, and lassen ‘let’.

The paper proves that there is, indeed, a consistent correlation between the two properties. First, modal verbs dürfen ‘can’, können ‘can’, mögen ‘will’, müssen ‘must’, sollen ‘should’, wollen ‘will’ are all polyfunctional and demonstrate gaps in the agreement. The paper stresses that the absence of a full agreement paradigm is not automatically related to the preterite-present origin of the modal verbs; see polyfunctional wollen ‘will’, which was originally a regular verb with a regular agreement paradigm. Second, the paper shows that werden ‘become’ is modally polyfunctional and at the same time it demonstrates the same agreement paradigm as other central modals. Furthermore, brauchen ‘need’ has recently acquired polyfunctionality, and as a result, in some dialects, it demonstrates predicted gaps in the agreement paradigm – see the examples in (2).

(2)  

<table>
<thead>
<tr>
<th>werden ‘become’</th>
<th>brauchen ‘need’</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) ich werde</td>
<td>ich brauch Ø</td>
</tr>
<tr>
<td>(b) du wirst</td>
<td>du brauchst</td>
</tr>
<tr>
<td>(c) er wird Ø</td>
<td>er brauch Ø</td>
</tr>
</tbody>
</table>

Third, the verb lassen ‘let’ has got a modal meaning similar to English deontic can, however, it is monofunctional, i.e. it does not express epistemic modality. As a result, it demonstrates a regular agreement paradigm. The paper also tries to provide an explanation of the semantic and formal behavior of the marginal modal wissen ‘know’, which is monofunctional, even though it demonstrates gaps in the agreement paradigm.
Keywords: modal polyfunctionality; agreement; German

References
N Morphology and Its Interpretation: Romance Feminine/Plural –a

M. Rita Manzini and Leonardo M. Savoia

Università di Firenze, Italy

rmanzini@unifi.it; lsavoia@unifi.it

Standard Italian has feminine plurals in –a alternating with masculine singulars. These are count plurals, cf. tre dit-a ‘three fingers’, with a distinctive semantics, denoting ‘a plurality of weakly differentiated parts’ (Acquaviva 2009). The –a plurals are few in Italian, but they are quite productive in Central/Southern Italian varieties, for instance S. Gregorio Matese in Fig. 1. Loporcaro & Paciaroni (2011), based on Central/Southern Italian, treat the alternations [masc sg] vs. [fem pl] as defining a third, ‘neuter’ gender. Acquaviva (2009), based on Italian, argues that –a plurals are a “lexeme deriving process”.

We contribute two sets of data to this discussion, from Central Calabrian (Fig. 2) and Romansh (Surselva) (Fig. 3). We assume that lexical roots are category-less (Marantz 1997) predicates; they have an argument that is bound by a D/Q operator (Higginbotham 1985). The functional layer closest to the root consists of a Class(ification) system: gender (and number, Borer 2005) in Romance/Germanic, nominal class in Bantu, classifiers in Chinese (Crisma et al. 2011). The Class system contributes to the nominal categorization of the expression, be it the Marantzian n (Kihm 2005), N Asp (Déchaine et al. 2014), or just N (Manzini & Savoia 2005). More controversially, we take all manifestations of N/Class morphology (gender/number, nominal class, classifiers) to have interpretive content. We work towards two conclusions: (A) Italo-Romance –a is an N/Class morpheme (inflectional); (B) the interpretive contribution of –a confirms the semantically contentful nature of N/Class in general.

Central Calabrian. Central Calabrian varieties like Iacurso (Fig. 2) allow us to pose the question of –a plurals independently of gender, since [fem] and [masc] are differentiated only in the singular. The plural has the gender-neutral realization –i on most nouns, on adjectives and on functional categories of the noun – alternating with the realization –a on some nouns. Acquaviva’s (2009) semantic characterization of –a plurals as consisting of ‘weakly differentiated parts’ appears to holds for Calabrian as well, witness l-i jiðit-a ‘the fingers’.

Let us assume that –i and –a are pure plurals. Manzini & Savoia (2011) impute to the plural (specifically to Italian –i) a content that they notate [\{\}]; [\{\}] says that subsets can be partitioned off the set denoted by the lexical base and is therefore essentially Borer’s (2005) Div. Let us say that in Calabrian, the –i morphology is associated with the [\{\}] content. The –a inflection will be associated with the same content, and with a further restriction to ‘weakly differentiated’ members, as in Fig. 4. Thus –a is an inflectional N/Class morpheme and at the same time it carries semantic content – which governs its agreement patterns.

Sursilvan. Sursilvan varieties of Romansh such as Vattiz (Fig.3) have an –s plural which combines both with bare [masc] bases and with –a [fem] bases. The same semantics that we have so far imputed to –a plurals appears to be associated with singular –a forms, alternating with bare bases, witness l-a det-a ‘the fingers set’. The singular status of –a is confirmed by agreement with the verb, e.g. la deta ai liunga ‘the fingers IS long’ (with distributive reading).

In present notation, –s is the [\{\}] morpheme, i.e. a plurality of atomic individuals. Following Chierchia (2010), a mass singular is also a plurality of sorts, namely a whole made up of parts. Thus a singular mass noun is like a plural count noun in that both include a multiplicity – namely a multiplicity of individuals, or a multiplicity of parts. In this perspective, we are not surprised that the Romance –a morphology can turn up denoting both a ‘plurality of poorly differentiated parts’ (Central/Southern Italian) and a ‘mass’/collective singular (Sursilvan). We stress this continuity by our notation of the ‘mass’/collective singular
as \([\subseteq: \text{part/whole}]\) as in Fig. 5. Assuming that \([\text{[fem]}]\) is the Elsewhere setting of \(-a\), the correct \([\text{[fem]}]\) agreement patterns are derived.

**Central Italian: consequences.** In Central Italian, the \(-a\) morpheme is \([\subseteq: \text{‘weakly differentiated’ members}]\), as in Calabrian, agreeing with respect to the \([\subseteq]\) property. At the same time, it is not gender neutral. Now, we have just postulated that, \(-a\) is \([\text{[fem]}]\) by Elsewhere. We assume that this is true independently of whether \(-a\) is singular, as in Sursilvan, or plural, as in Central Italian, yielding \([\text{[fem]}]\) agreement in both instances. As time allows, we will consider whether there are more principled treatments of the ‘syncretism’ of \([\subseteq]\) and \([\text{[fem]}]\) attested by \(-a\).

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>sg</td>
<td>r-u jatt-u</td>
<td>l-a καf-f-a</td>
<td>r-u περ-e</td>
<td>l-a nof-e</td>
<td>r-u riut-u</td>
</tr>
<tr>
<td>pl</td>
<td>r-i jatt-i</td>
<td>l-e καf-f-e</td>
<td>r-i περ-i</td>
<td>l-e nof-i</td>
<td>r-e reut-a</td>
</tr>
</tbody>
</table>

Fig. 1 Inflection classes of S. Gregorio Matese (South-Central Italy)

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>sg</td>
<td>l-u kurti-ŋ-u</td>
<td>l-a rot-a</td>
<td>l-u melun-ẹ</td>
<td>l-a cav-ẹ</td>
<td>l-u jiōt-ŋ-u</td>
</tr>
<tr>
<td>pl</td>
<td>l-i kurti-ŋ-i</td>
<td>l-i ruŋt-i</td>
<td>l-i mlun-i</td>
<td>l-i cav-i</td>
<td>l-i jiōt-a, l-i kurti-ŋ-a</td>
</tr>
</tbody>
</table>

Fig. 2 Inflection classes of Iacurso (Central Calabria, Italy)

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>sg</td>
<td>iñ maŋ</td>
<td>l-a rod-a</td>
<td>l-a mail-a ‘the apples (as a whole made of apple parts)’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pl</td>
<td>iñ-s mail-ts</td>
<td>l-as rod-a-s</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3 Inflection classes of Vattiz (Lumnezia Valley, Switzerland)

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>sg</td>
<td>l-a reut-a</td>
<td>l-a rod-a</td>
<td>l-a mail-a ‘the apples (as a whole made of apple parts)’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pl</td>
<td>l-as reut-a</td>
<td>l-as rod-a-s</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4 DP Iacurso (Central Calabrian)

Fig. 5 DP Vattiz (Romansh)

Fig. 6 DP S. Gregorio Matese (South-Central Italy)
In 1985 Talmy first posits a typological distinction according to the way languages express Path and Manner in the description of Motion events. In most Indo-European languages (with the exception of Romance languages), motion verbs encode the Manner of motion, while the Path is expressed in what he calls a satellite, hence satellite-framed languages. In the second typological pattern it is the verb that encodes the Path. “If a Co-event of Manner … is expressed in the same sentence, it must be as an independent, usually adverbial or gerundive type constituent”, and is often omitted (2000, 49). This was confirmed by psycholinguistic experiments, which revealed a difference in the “rhetorical style” in the description of motion events in the two types of languages (e.g. Berman and Slobin 1994, Slobin 1996). What was challenged, however, was the fact that all languages should fall neatly into one of the patterns (e.g. Slobin 2006). Differences were also observed between languages belonging to the same group (e.g. Ibarretxe-Antuñano 2004, Hijazo-Gascón and Ibarretxe-Antuñano 2013, Filipović 2007, Slobin et al. 2014). Translatological research then focused on the use of translation strategies when the source and target languages differ in their semantic typology, and investigated the amount of information lost in translation (Slobin 2005).

This pilot study relies on what Chesterman (2003) calls T-universals, i.e. “claims about the way translators use the target language” (318), namely quantitative deviations from the target language norm (Altenberg and Granger 2002, 40). Czech (satellite-framed, but satellite is a bound morpheme, which might lead to restrictions, cf. Filipović 2007 about Serbo-Croatian) is used to contrast English (satellite-framed, satellite is a free morpheme) and Spanish (verb-framed) in their expression of boundary crossing events. Tokens of verbs with the satellite v- [in] denoting a boundary crossing were downloaded from three subcorpora of Czech fiction (cca 9.5 million text positions each), created within the parallel translation corpus Intercorp: 1. Czech source texts, 2. Czech translations from English and 3. Czech translations from Spanish. Preliminary findings reveal indeed a higher number of different verb types and tokens in Czech translations from English than from Spanish, which reflects the typological difference. This happens in spite of the fact that “the information about the Manner of motion can be spontaneously added (Ibarretxe-Antuñano and Filipović [2013, 270] about interpreting witness interviews from Spanish to English). Using classification from Slobin et al. (2014), the paper then compares correspondences of selected Czech verbs in the Spanish and English source and target texts.

The aim of the study is not only to contribute to the on-going investigation of the typological difference, but also (given that Czech has not been yet systematically studied with respect to this typological difference) to initiate a discussion about the way Czech speakers encode motion events.
**Keywords:** parallel corpus, motion, lexicalization patterns, semantic typology

**References**


In this presentation we address the issue of the recognition of random strings of symbols, where ‘random’ is to be interpreted as ‘not being derived by means of a grammatical formalism’. We will present a quantitative procedure to identify random strings from strings that are the output of a formal grammar based on Type-Token Ratios (TTR), series of n-gram analysis in combination with Zipf’s law and measurement of specific features and their comparison (Zipf, 1949). The grammars we will consider all fall within the Chomsky Hierarchy of formal grammars and languages (Chomsky, 1959). This is relevant for the study of both formal and natural languages, because we can address questions about the nature of the computational system that underlies human language by analyzing the quantitative properties of its output. What we propose is a simple method that can distinguish random strings from non-random strings with relative success. It has been tested for various types of texts ranging from natural languages, encrypted texts, computer binary code, DNA strings, sound, and various types of randomly generated strings (e.g., background radiation). The results are stable and provide good measure for randomness. Thus, we have a heuristic procedure to establish whether a string is random or not.

Prospects for future research on making the heuristics more fine graded in order to distinguish between levels in the Chomsky Hierarchy will also be presented. The ultimate aim of this method is to find a way to infer the type of Chomsky-normal / Greibach-normal grammar which generated a particular set of strings based on quantitative properties of said strings.

**Keywords:** String randomness; n-gram analysis; Type-Token Ratio; formal grammar; Zipf’s law

**References**
In this talk I use Slovenian to explore the parallel behaviour of questions with multiple wh-movement, (1), and questions with short movement (i.e. wh-questions in which at least one wh-phrase moves to the clause initial position but one moves to a clause internal position, cf. Citko (2010)), (2). Based on the similarities, I argue that in both cases wh-phrases undergo wh-movement to a ‘Periphery’ – short wh-movement to the Low Periphery in the sense of Belletti (2004), and multiple wh-movement to the Left Periphery in the sense of Rizzi (1997).

In Slovenian, we can observe parallel behavior of non-initial wh-phrases in questions with multiple wh-movement, (1), and clause internal wh-phrases in questions with short movement, (2). In both (i) the order of the focused and the non-initial wh-phrases is free, (3a, b) vs. (3c, d), (ii) the order of the topic and the non-initial wh-phrases is free, (iii) in both there is no strict word order between the topic, focus, and wh-phrases, (4a, b) vs. (4c, d), and (iv) just as to the Left Periphery, more than one wh-phrase can be short moved, (1) vs. (2).

(1) Kdo je kaj kdaj Martinu dal?
   who AUX what when Martin.DAT give
   ‘Who have what to Martin?’

(2) Kdo je Martinu kaj kdaj dal?
   who AUX Martin.DAT what when give
   ‘Who have what to Martin?’

(3) (a) Kdaj je komu TONETA Miha predstavil?
    when AUX who.DAT TONE.ACC Miha.NOM introduce
    ‘When did Miha introduce TONE to whom?’

   (b) Kdaj je TONETA komu Miha predstavil?

   (c) Kdaj je Miha komu TONETA predstavil?

   (d) Kdaj je Miha TONETA komu predstavil?

   (e) #TONETA je Miha kdaj komu predstavil?
Maja je izpostavila Toneta virusu ebole.
‘Maja exposed Tone to the Ebola virus.’

(a) Zakaj je kdaj BOŽOTA temu virusu Janez izpostavil?
why AUX when Božo,ACC this virus,DAT Janez,NOM expose
‘Why and when did Janez expose Božo to this virus?’

(b) Zakaj je temu virusu BOŽOTA kdaj temu virusu BOŽOTA Janez izpostavil?

(c) Zakaj je Janez kdaj temu virusu BOŽOTA/BOŽOTA temu virusu izpostavil?

(d) Zakaj je Janez temu virusu kdaj BOŽOTA/BOŽOTA kdaj izpostavil?

Based on this, I argue that short wh-movement is parallel to the movement of non-initial wh-phrases in multiple wh-fronting which I assume, following Mišmaš (2015a, b), is movement to the Left Periphery. Given the data, I propose that all non-initial wh-phrases move to Wh-Projections (either in the Low or the Left Periphery), which means that short movement is wh-movement to the Low Periphery in Slovenian (but see Citko (2010) for Polish). Based on the data, I show the structure of the Low and the Left Periphery in Slovenian, both of which can host Wh-, Focus, and Topic Projections in which the wh-, focus, and topic phrases are located.

This proposal enables us to treat multiple wh-questions with short wh-movement in the same manner as questions with multiple wh-fronting.

**Keywords:** Low Periphery, Left Periphery, wh-movement, Slavic

**References**


A Corpus-Based Study of Gradience of a Complex Preposition Category: 
The Case of Polish na rzecz, do spraw and w celu

Urszula Modrzyk
Institute of the Polish Language, Polish Academy of Sciences, Cracow, Poland
urszula.modrzyk@ijp-pan.krakow.pl

Grammaticalization is considered a gradual process (Bybee 2011, Hoffmann 2005, Poplack 2011, Torres Cacoullos & Walker 2011). In effect, a synchronic ‘gradience’ of categories can be observed (Anderson 2001, Traugott & Trousdale 2013), which means that at any moment constructions belonging to one category can vary according to the degree of their grammaticalization. The aim of this study is to observe the degree of grammaticalization of three semantically similar Polish phrases, which show evident features of a complex preposition: na rzecz ‘for the benefit,’ do spraw ‘for’ and w celu ‘in order to’.

There is an established array of tests used in ‘armchair’ linguistics to decide whether a certain Prepositional Phrase is in fact a complex preposition (Milewska 2003, Nowak 2008). Another aim of this paper is to discuss the possibilities and limitations of a realisable adaptation of these methods for a corpus study. The data used here was obtained from the National Corpus of Polish (NKJP).

The tests which were adapted for the purpose of this paper examine the following:
– discontinuity of phrases,
– possibility of relativization of the nominal constituent of the complex preposition,
– the question if pronouns, which function as complements, occur in post-prepositional forms (which would suggest the prepositional character of a given phrase)
– possibility to form questions, for example, na czyją rzecz? ‘for whose benefit’, and in consequence, to interpret the relation between ‘rzecz’ and its subordinate as one of possession,
– the question if a superordinate phrase opens a place for the whole complex preposition or only for its prepositional constituent,
– the question if a subordinate phrase is a complement of the whole complex preposition or if it is just a subordinate of the nominal constituent,
– semantic substitution – determining if the phrases can be replaced by a single preposition.

In the case of the first four tests, the number of occurrences of particular types of phrases (eg. continuous vs discontinuous) in the NKJP is considered. In the next two tests, I have measured the strength of the collocation between the complex preposition and its subordinate and superordinate phrases compared to the collocational strength between each of the constituents of the complex preposition separately and the adjacent words. Concerning semantic substitution, I have checked if there are examples of the most frequent left and right collocates with single-word prepositions in-between.

In addition to the tests listed above, I have examined the semantics of the phrases by analyses of semantic classes of their strongest collocates. Also the use of the phrases within proper names was examined, via analysis of their spelling (according to the rules of Polish orthography, prepositions in proper names should be spelled in lower-case).
The results of the study prove that the phrases in question are placed at different stages in the process of grammaticalization. Therefore, they can be seen as an example of ‘gradience.’ The most visible symptom of this ‘gradience’ is a difference between the results of syntactic and semantic analyses of the phrases. Another interesting outcome of the tests is the fact that the results of grammaticalization can suggest unevenness of the process – some of the phrases examined are more grammaticalized in bureaucratic texts.

Keywords: complex preposition, gradience, grammaticalization, Polish

References
Gang Effects with Left-Branch Extraction: Arguments for Weighted Constraints

Andrew Murphy
Universität Leipzig, Germany
andrew.murphy@uni-leipzig.de

Problem: A long-standing puzzle in the literature on wh-movement in Slavic is why multiple left-branch extraction (MLBE) seems to be impossible in languages that have multiple wh-fronting (MWF) and left-branch extraction (LBE) independently (Fernandez-Salgueiro 2006, Grebenyova 2012). One would expect that if a language allows for both MWF (1) and LBE (2), it should allow these two processes to combine and permit MLBE. In languages with both processes (e.g. Polish, Russian, Serbo-Croatian), MLBE is entirely impossible (3).

(1) Ktoj kogo, tī priglasil tī na užin? (2) Čju which whom invited to dinner [NP tī mašinu] whose he bought [NP tī mašinu]?

(3) *Kakoj, čjuj [NP tī aktër] kupil [NP tī mašinu]?

which whose actot bought [NP tī mašinu] (Russian; Grebenyova 2012)

Following the standard approach in Bošković (2005), LBE is possible if languages lack a DP phase, there is no obvious reason why MLBE should be excluded. There have not been many attempts to explain this odd fact, and all existing proposals claim that LBE must involve a different kind of movement (Fernandez-Salgueiro 2006, Grebenyova 2012).

Proposal: Instead, I argue that the ban on MLBE is a case of cumulative constraint interaction in Harmonic Grammar. To model this, I adopt a constraint-based model of syntax with cyclic optimization (cf. Harmonic Serialism; McCarthy 2010, Heck & Müller 2013) and weighted constraints (cf. Harmonic Grammar; Legendre et al. 1990, Pater 2009). Crucially, operations are only permitted if they are harmonically improving, i.e. a movement operation is only possible if it does not violate a more important constraint than non-movement. In HG, constraints bear a numerical weight and violations are deducted from the base Harmony Score of 0. In a language allowing LBE, there is a constraint WH-CRIT stating that all wh-phrases must be in the specifier of a licensing head (C[wh]). There is another constraint, LBC (Left-Branch Condition), banning LBE. If LBE is possible, the weighting $w$(WH-CRIT) > $w$(LBC) must hold. Thus, in Tableau 1 the candidate which does not move (a) has a worse harmony score than (a) and loses. Furthermore, if this language also allows MWF, then the constraint WH-CRIT must have a higher weight than *MULTSPEC, which punishes the creation of multiple specifiers. The weighting $w$(WH-CRIT) > $w$(*MULTSPEC) allows for MWF (T2). With these rankings in place, we would expect MLBE to be possible, since $w$(WH-CRIT) > {LBC, *MULTSPEC}. HG is unique in that it allows for cumulative constraint interaction or ‘gang effects’. Since we have weighted constraints, two constraints can ‘gang up’ to overthrow a higher one. In Step 1, LBE is licensed since it is harmonically improving to incur a violation of LBE to avoid a violation of WH-CRIT. For the second movement step (that creates an additional specifier), the summed weights of *MULTSPEC and LBC are higher than a single violation of WH-CRIT. This movement step is ruled out since it is not harmonically improving. This, with weighted constraints the ban on MLBE comes for free. Further evidence for this system comes from the suprising emergence of superiority effects with LBE in languages.
ordinarily lacking superiority (Fernandez-Salgueiro 2006). It will be shown that the cyclic optimization model proposed here can also capture these facts unlike parallel optimization.

**Keywords:** syntax, left-branch extraction, superiority, optimality theory

Tableau 1

<table>
<thead>
<tr>
<th></th>
<th>WH-CRIT</th>
<th>LBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ([CP \ C_{wh} \ldots [VP \ [DP \ w h \ NP]]])</td>
<td>-1</td>
<td>-2</td>
</tr>
<tr>
<td>b. ([CP \ C_{wh} \ldots [VP \ [DP \ w h \ NP]]])</td>
<td>-1</td>
<td>-1.5</td>
</tr>
</tbody>
</table>

Tableau 2

**Step 1:**

<table>
<thead>
<tr>
<th></th>
<th>WH-CRIT</th>
<th>*MULTSPEC</th>
<th>LBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ([CP \ C_{wh} \ldots [VP \ [DP \ w h \ NP]] \ldots [DP \ w h \ NP]])</td>
<td>-2</td>
<td>-4</td>
<td></td>
</tr>
<tr>
<td>b. ([CP \ C_{wh} \ldots [VP \ [DP \ w h \ NP]] \ldots [DP \ w h \ NP]])</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

**Step 2:**

<table>
<thead>
<tr>
<th></th>
<th>WH-CRIT</th>
<th>*MULTSPEC</th>
<th>LBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ([CP \ w h_1 \ C_{wh} \ldots [VP \ [DP \ t_1 \ NP]] \ldots [DP \ w h \ NP]])</td>
<td>-1</td>
<td>-2</td>
<td></td>
</tr>
<tr>
<td>b. ([CP \ w h_1 \ C_{wh} \ldots [VP \ [DP \ t_1 \ NP]] \ldots [DP \ t_2 \ NP]])</td>
<td>-1</td>
<td>-3</td>
<td></td>
</tr>
</tbody>
</table>

Tableaux 3 & 4

**References**


Copredication is standardly defined as a syntactic construction in which two predicates simultaneously apply to the same argument (Asher 2011). The nominal within the DP consequently has no fixed interpretation, but takes multiple contradictory predicates, as when event and object readings apply to lunch in (1), creating an ‘impossible’ entity.

(1) Lunch was delicious but took forever.

This presentation will explore the acceptability properties of copredication and to what extent they can inform debates both about the cost of putative type-shifting operations which generate copredications and formal theories of its structure. Across two acceptability judgement Experiments, we test whether copredication in book-, lunch- and city-type nominals is difficult across-the-board or depends on sense order. Psycholinguistic studies claim type-shifting operations are costly in environments such as complement coercion. Experiment 1 tested whether copredication, and hence type-shifting, is costly in a 2 (Nominal Number; 1 or 2) x 2 (Sense Type; related or distinct) design. (2a-d) shows example sentences for each condition:

(2) (a) Brad said that the absorbing and readable journal was part of the office block.
(b) Brad said that the prosperous and readable journal was part of the office block.
(c) Brad said that the absorbing magazine and readable journal were part of the office block.
(d) Brad said that the prosperous magazine and readable journal were part of the office block.

A 2 x 2 by-subjects ANOVA revealed a significant effect of Sense Type ($F(1,47) = 32.410, p < .001, \eta^2_p = .408$) but no effect of Nominal Number nor an interaction effect. Our data indicates that sentences that coordinate different senses, whether attributed to one or two nominals, are significantly less acceptable than those coordinating semantically related senses. The results further indicate that this acceptability difference across sense type did not interact with the number of nominals. From this, we can conclude that copredication is not costly.

Frisson (2015) detected an effect of polysemous sense order on single nominals, but due to inadequate controls it is not clear whether the effect is specific to copredication. As a result, Experiment 2 tested whether copredication costs are dependent on sense order by manipulating Sense Order (Concrete-Abstract, Abstract-Concrete) in addition to Nominal Number (1 or 2) (see 3a-d):

(3) (a) Stewart thought that the creased and expensive bill was being talked about at the table.
(b) Stewart thought that the expensive and creased bill was being talked about at the table.
Stewart thought that the *creased book* and *expensive bill* were being talked about at the table.

(d) Stewart thought that the *expensive book* and *creased bill* were being talked about at the table.

The by-subjects ANOVA showed that acceptability is not only modulated by coordinating different senses ($F(1,51) = 5.429$, $p = .024$, $\eta_p^2 = .096$), but also by nominal number ($F(1,51) = 4.044$, $p = .050$, $\eta_p^2 = .073$). Concrete-Abstract items were significantly more acceptable than Abstract-Concrete items, and 2 nominals were more acceptable than 1. There was no significant interaction between sense order and nominal number. This sense order effect (already present in Frisson 2015) was not exclusive to copredication, although the average scores indicate that the effect was more severe for copredication relative to non-copredication. Comparing Abstract-Concrete ordering across nominal number reveals that this type of sense integration on a single nominal is marginally more costly than sense dispersal over multiple nominals ($p<.07$). The data suggests that a type-shifting operation interacting with sense order is likely responsible for the greater processing costs associated with copredication. Since acceptability was not modulated purely by sense frequency/dominance but rather sense type (with Abstract-Concrete sentences being universally the least acceptable across all nominal types), an ‘Extended Mereological’ (EM) model is consequently proposed under which multiple senses are stored on single representations of complex nominals and ease of interpretability is based on a semantic type hierarchy, with processing being easiest (and acceptability being highest) when semantically more complex senses follow less complex ones. This may explain the contrast in (4a-b), with the PROCESS sense in (4b) preceding the PHYSICAL sense:

(4) (a) The translation that lies on the table was difficult.
   (b) The translation that was difficult lies on the table.

Presenting data on *newspaper*-type nominals, the standard treatment of copredication as involving composite dot objects is also rejected. A new type, an ‘Encompassed Type’, is proposed under which nominals like *newspaper* and *city* are able to host more than two senses, in contrast to Asher (2011) and Gotham’s (2015) claim that copredication is restricted to ‘dual-aspect nouns’ (see 5a-b). These data also support an EM model since three unrelated senses can be attributed to a single nominal and accessed without degraded results.

(5) (a) The well-written newspaper that I held this morning has been sued by the government.
   (b) The low-lying and polluted city is known for its reactionary stance on equal marriage.

This presentation will consequently present novel hypotheses about the structure of copredication by discussing a number of theoretical and behavioural considerations.

**Keywords:** copredication; psycholinguistics; generative lexicon; type-shifting; concepts

**References**


Gotham, Matthew. 2015. “Copredication, Quantification and Individuation.” PhD diss., University College London.
The aim of the present paper is to explore a methodological issue related to the use of parallel corpora, i.e. corpora composed of original and translated texts. As it was repeatedly stated, the language of translation presents certain particularities that may be called “laws” (Toury 1995), “tendencies” or “translation universals” (Baker 1993), in particular the simplification, the normalization and the explicitation. We are aware of the fact that the concept of “translation universals” may be contestable (Malmkjær 2012 or Zanettin 2013), but in this paper we will use this term in order to explain some shifts in segmentation in phrases appearing in the translations in the Czech-English-French part of the parallel corpus InterCorp. At the same time, we will show that these shifts may be the result of a global translation strategy potentially effecting the style of the original text.

The manual analysis of 4 000 non1:1 segments extracted from the Czech-French (and French-Czech) part of the parallel corpus InterCorp (imaginative and informative texts, Nádvorníková – Šotolová, forthcoming) revealed that a large number of the changes of the segmentation in phrases may be rated either as simplification (a long phrase is split in two or more phrases, considered more “digestible”), or to normalization (long or short phrases are split or linked to form a phrase of a “normal” length). The important observation was that these shifts occur in both directions of the translation, i.e. independently from the source language:

(1) J’ai fait le chemin à pied. J’ai voulu voir maman tout de suite. [I went there on foot. I asked to be allowed to see Mother at once, …] (Albert Camus, L’Étranger/The Stranger, 1950) > Došel jsem tam pěšky a hned jsem chtěl vidět maminku. [I went there on foot and I asked to be allowed to see Mother at once.] (transl. by Miloslav Žilina 1969)


In this paper, we intend to verify this tendency by adding a third language to the analysis (English) and by observing the non1:1 segments in the English-Czech and English-French language pairs.

**Keywords:** sentence splitting – translation universals – methodology – parallel corpus
References


The talk discusses Hungarian relatives and interrogatives with multiple wh-elements. We aim to show that by assuming that vocabulary insertion takes place after syntactic operations one can achieve a more unified theory of relative pronouns, interrogatives and quantifiers.

The talk focuses on the analysis of constructions as in (1).

(1) (a) Aki amit talált, megette.
    a. who a. what found ate
    “Everybody ate whatever they could find.”
(b) Ki mit talált?
    who what found?
    “Who found what?” (= what did everybody find)
(c) *Aki talált amit, megette.
(d) Ki talált mit? (≠ what did everybody find)

In cases like this, the analysis of É. Kiss (1993) faces problems. For interrogatives, she proposes that the second wh-element is moved to a FocP position while the first one moves to a quantifier position to account for its universal interpretation. One problem here is why a fronted wh-element can only be interpreted as a universal quantifier if it is accompanied by another fronted wh-element. Another problem is why a universal quantifier itself cannot occupy the position of the first wh-element. Moreover, the difference between multiple interrogatives and multiple relatives demonstrated in (1c) and (1d) is difficult to account for.

A different approach by Lipták (2000) argues for an adjunction structure where one relative pronoun is adjoined to the other. While this addresses the first of É. Kiss's problems, it merely describes the context in which the relevant interpretation is possible but does not account for why or how this interpretation is achieved. Lipták’s proposal also does not address the difference between interrogatives and relatives.

Our goal is to demonstrate that there is no need to assume any kind of dependency between the two wh-elements. From our perspective, the question is not why the wh-element is interpreted as a universal quantifier, but why an underlying universal quantifier is spelled out by a wh-element in these cases. In essence, we claim that realizing a fronted quantifier as a quantifier is blocked in these contexts, and their realization as wh-elements is the next best alternative.

We argue that the basic difference between multiple relatives and multiple interrogatives is that the former necessarily involves multiple quantification while the latter may not. (1d) is possible as it contains no underlying quantifier and hence a non-distributive reading is achieved. Instead the example concerns two underlying interrogative pronouns, which are
both spelled out faithfully. (1c) is ungrammatical as it contains a non-fronted quantifier which is spelled out by a non-optimal wh-element.

As a result, there is no need to assume that there is any kind of dependency between the two wh-elements; that is, no movement is stipulated that adjoins one of them to the other. In addition, the adoption of the late insertion mechanism allows for a more general common analysis of these relative, interrogative and quantificational constructions.

**Keywords:** relative clauses, interrogatives, quantification, late vocabulary insertion

**References**
Long-Distance Wh-Extraction in Polish. Problems of Structure and Derivation

Marcin Orszulak
University of Wroclaw, Wroclaw, Poland
marcin.orszulak@uwr.edu.pl

The issue. The aim of the presentation is to discuss puzzling facts about Polish long-distance wh-extraction. Although in the literature there is a debate about the acceptability of such structures in Polish, which obstructs a unified theoretical explanation (Willim 1989, 116), researchers seem to agree that the extractions out of tensed indicative complements give degraded results (see [1]), whereas the extractions out of infinitival complements are acceptable (see [2]) (Giejgo 1981; Kardela 1986; Willim 1989; Witkoś 1995). Furthermore, extractions out of subjunctive complements constitute a problematic case; they seem to be allowed only in the case of verbs that also select for infinitival complements (see [3]) (Willim 1989).

[1] *[CP Coi [TP Maria [VP sądziła, [CP t₁ że [TP Piotr [VP wcześniej studiował t₁]]]]]? what Mary thought that Peter previously studied
“What did Mary think that Peter studied previously?”

[2] [CP Coi [TP pro [VP chcesz [CP t₁ [TP PRO [VP kupiś t₁]]]]]? what pro want₂SG PRO buy_INF
“What do you want to buy?”

[3] [CP Coi [TP Maria [VP chce, [CP t₁ że-by [TP Piotr [VP kupił t₁]]]]]? what Mary wants that-SBJV Peter buy
“What does Mary want Peter to buy?”

Proposal. In the present analysis we will describe the structural properties of long-distance wh-questions by means of Split CP (Rizzi 1997). Following Lubańska (2005), we claim that in Polish wh-questions are derived via focus movement and we will show that the specifier of Focus Phrase is a landing site for Polish wh-phrases, which is evidenced by e.g. sentences in which topics precede wh-phrases and long wh-scoping constructions (Willim 1989; Orszulak 2011). Next, we will map our structural observations onto the phase model of derivation (Chomsky 2001; 2008). Following Gallego (2009, 136), we will assume that Force° should be treated as a strong phase head and thus Polish wh-phrases, which land in spec-FocP, so in the complement of a strong phase head that is further spelled out and thus impenetrable, are no longer accessible for further derivations (see [4]).

“What did Mary notice that Peter chose t₁?”

In spite of the landing site, in Polish the extraction out of subjunctive and infinitival complements is still possible. Therefore, we will propose that phasehood is sensitive to the properties of tense (in line with Gallego (2007, 11) reasoning about subjunctives). In detail, we will suggest that a strong phase should exhibit a referentially independent tense with absolute interpretation (evaluated in connection with the utterance time, see Comrie 1985; Wurmbrand 2006). However, this is not found in the case of infinitives and extractable
subjunctives in Polish, whose tense is dependent and unrealised (Wurmbrand 2006; Tomaszewicz 2008). Consequently, infinitival and subjunctive CPs are weak phases, which accounts for the extraction facts in Polish (possible implications of the proposed analysis for other locality phenomena will also be discussed).

**Keywords:** Polish, wh-movement, Split CP, phase syntax

**References**


Pluractional markers are ‘frequently reduplicative, most often derivational rather than inflectional [morphemes]’ which ‘attach to the verb to indicate a multiplicity of actions, whether involving multiple participants, times or locations’ (Lasersohn 1995:238, 240). Greenberg (1991) cites examples from Arabic, where the pluractional verb form involves reduplication of the median consonant of the neutral form (e.g. kasar ‘break’ \(\rightarrow\) kassar break.PL’). There is a substantial body of literature on pluractionality/event plurality, most of which falls in the domain of formal semantics (e.g. Cusic 1981, Carlson 2001, Faller 2008, Filip & Carlson 2001, van Geenhoven 2004, Laca 2004, Landman 1996, 2000, Tovena & Kihm 2008, Wood 2007). This presentation explores the morphosyntactic properties of pluractionality involving structure and derivation based on data from Arabic.

The foundational idea is the parallelism Greenberg (1991) draws between ‘verbal [plural] number’ and ‘nominal [plural] number’, which translates structurally into a Num(node) for CP-VP that parallels the Num-node of DP-NP (Ritter 1991). Moreover, Num has a CARD(INALITY) component, the valuation of which relative to the object of the verb or a VP-modifier (the contextual factors) turns out to play a critical role in determining pluractional interpretation.

(1) \([DP D \ldots [\#P [CARD]] [\#'] [\#N] \ldots]\)
(2) \([CP C \ldots [\#P [CARD]] [\#'] [\#V [VP V DP] [XP]]\ldots]\)

Greenberg (1991) highlights three properties of pluractionality in Arabic, argued here to follow from (2) in combination with verb-merger and Agree(ement) incorporating feature valuation. First, it involves a specialised verb form, called the intensive in the Afroasiatic tradition (Newman 1991), which derives from the neutral form by addition of a plural number morpheme that is generally expressed by some form of reduplication such as median consonant germination in Arabic. This property is shown to implicate merger of the verb into \# which serves as input to non-concatinative derivation at PF.

Secondly, pluractionality typically involves plural objects (e.g. kassar kibbe\(\text{ye}:t^*/kibbe\text{ye}\) ‘break.PL glasses/a glass’), giving rise to a false similarity with object agreement (Newman 1991, Durie 1986, Corbett 2000). Depending on the nature of the event, pluractionality can also involve a singular object, giving rise to the so-called ‘repetition meaning’ (e.g. xabbat\(l\)-be\(b\) ‘repeatedly hit.PL the door’). These properties are shown to follow from whether \# in (2) enters derivation unvalued or already valued in combination with the mechanism of Agree applying to \# and the object when \# enters derivation unvalued.

The third property, traditionally described with the term ‘exaggeration’, is that plural objects are normally understood to have a high cardinality value (e.g. a high number of broken glasses). It turns out that this is a default interpretation that applies in the absence of a cardinal modifier, including one with a low value such as ‘five’, either associated with the
object (e.g. kassar xames kibbe:ye:t ‘break.PL five glasses’) or a VP-modifier in repetition contexts (e.g. xabbatˁ l-be:b xames marraat ‘[hit.PL the door] five times’). These facts are shown to implicate the valuation of the CARD component in (2) in terms of Agree relative to the object, a VP-modifier or nothing in the case of default valuation.

References


Gender and the ‘Y’ Question

Jeffrey Keith Parrott

Palacký University, Olomouc, Czech Republic
jkparrott@gmail.com

In recent years, generative morphosyntax has seen renewed interest in formulating theoretical models of grammatical gender, aka noun class (e.g., Acquaviva 2009; Puskar 2015, a.o.). In this presentation, I consider gender within the framework of Minimalist (Chomsky 2013) Distributed Morphology (DM, Halle & Marantz 1993; Embick & Noyer 2007), focusing on the idea that roots are acategorial, with no phonological, semantic, or grammatical features in the pre-syntactic Formative list, distinguished in syntax only by a unique (numerical) index that links root terminals to their exponent in the Vocabulary list at PF and their interpretation in the Encyclopedia list at LF (e.g., Harley 2014, a.o.). Among other benefits, accepting acategorial roots provides an appealing account of variation in idiomatic meanings at the level of the verb phrase (e.g., *beg the question* ‘assume the conclusion’ or ‘raise the question’) and complex word (e.g., *enormity* ‘great depravity’ or ‘enormousness’); these and similar cases can be modeled as individually differing Encyclopedia entries for root terminals in particular morphosyntactic contexts.

The non-lexicalist DM architecture advocated here (1) is a strong implementation of the classic ‘Y’ model (e.g., Chomsky 1986:68, a.o.). Therefore it immediately raises what may be called the ‘Y’ question for gender. Cross-linguistically, nouns are ‘assigned’ a gender on the basis of the LF semantics of the nominal root (usually but not only biological sex, hence ‘Y’ also alludes to the chromosome), the PF morphophonology of the nominal root, or a combination of such properties (Comrie 1999; Corbett 1991; 2013). Gender agreement and Impoverishment (Nevins 2011; Parrott 2015) provide clear evidence that gender features must be present in syntax and morphology; gender features are evidently present at PF, where they condition allomorphy. In addition to the complex facts of semantically based assignment, it seems reasonable to suppose that gender features are present at LF on the grounds that gender may trigger semantic presupposition (e.g., Sudo 2012). We thus are faced with the serious but overlooked ‘Y’ question of how to associate gender features with semantic and phonological features when LF and PF have no access to each other after spellout, and roots have no semantic or phonological features prior to syntax.

I partially follow Kramer (2014; 2015, a.o.) in taking gender features to be bundled with categorizing little n terminals in the Formative list. In syntax, Merge can freely combine gendered little n terminals with indexed roots or complex terminals. However, I depart from Kramer (a.o.) in suggesting that gender features on n are in fact always potentially interpretable at LF, just in case they are specified in a root’s Encyclopedia entry. Gender features are drawn from a set of universal grammatical features that plausibly includes [±animate], [±human], [±female], [±male], [±animal], [±plant], or [±edible], inter alia. The bundling of these features with categorizing n terminals in the Formative list (e.g., n[+human, +female]) is an acquisition task; the child must also learn which indexed roots are Merged with which gendered n terminals (e.g., [n [√AUNT] n[+human, +female]]) by attending to externalized forms at PF, as well any interpretations at LF. From the perspective of
Minimalist DM, the ‘Y’ question is answered the same way for gender as it is for roots, namely, with syntax and lists but without a lexicon.

**Keywords:** gender, Distributed Morphology, Minimalism, roots, theoretical morphosyntax

**Selected References**

- Puskar, Zorica 2015. 'Interactions of gender and number agreement: Evidence from Bosnian/Croatian/Serbian', University of Leipzig.

(1) A Minimalist, DM ‘Y’ model

```
<table>
<thead>
<tr>
<th>Formative list (roots, feature bundles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax (Merge, Agree) ➔</td>
</tr>
<tr>
<td>Spellout ➔</td>
</tr>
</tbody>
</table>

← Morphology (Impoverishment, i.a.)

Encyclopedia list, LF

Vocabulary list, PF
```
The Analysis of the Evenki Analytic Negation

Elena Rudnitskaya

Institute of Oriental Studies, Moscow, Russia
erudnitskaya@gmail.com

In most varieties of minimalism, adjunction, in particular, right-adjunction, is not defined or even banned, and substituted for other operations, e.g. for Remnant Movement (following (Kayne 1994) – for instance, in the Cartographic approach (Rizzi 2004, 5)), or for concatenation: see (Hornstein, Nunes 2008; Ko 2015), the Bare Phrase Structure framework. However, there are cases, such as extraposition in German (Büring, Hartmann 1995) and postverbal arguments in Japanese or Korean (Cecchetto 2004: 174-175, Ko, Choi 2009) that show connectivity effects which can be better accounted for by right adjunction to TP // CP following movement rightwards than by Remnant Movement or concatenation.

We show that the analysis of the analytic verb negation in Evenki can be couched in the generative framework not only in the (Kayne 1994) head-initial structure, as in (Kaizer 2006) for Finnish, but also using the head-final syntactic tree and movement rightwards followed by right-adjunction. Evenki is a head-final language: it has only postpositions, and the head in a neutral case goes after the complement in most verb-argument and verb analytic constructions. However, information structure factors cause head-initial word order, especially with sentential arguments. (Holmberg 2000, 137), following (Kayne 1994), accounts for the OV – VO variation in Finnish using the Remnant Movement.

This account cannot explain the grammaticalized negative verb construction NEG [FIN] + V [PTCP] in (1), in which NEG is the finite negative auxiliary that always precedes the participial V, except for the emphatic “inverted” constructions” (Nedjalkov 1994).

(1) a-gki-n hula-ra (Malye jazuki...)
NEG-PST.ITER-3SG miss-PTCP.NEG
“did not miss” ‘any wild deer’ – meaning ‘always made a well-aimed shot at wild deer’

Which problems do these data pose for the (Kayne 1994) approach? Suppose, all the head-final constructions in Evenki were derived via the Remnant Movement. The NEG [FIN] + V [PTCP] construction is head-initial, so no special operations are needed to derive it. Since NEG is finite, it raises to T. Example (2) shows another instance of the negative analytic construction, in which the entire background clause follows NEG. In (2), NEG further moves to the Left Periphery, into the leftmost projection: e.g. PolP, according to (Kaizer 2006), for similar data in Finnish.

(2) a-ča-w [bi ješo alagu-m-da-ra škola-du] (Malye jazyki…)
NEG-PST-1SG 1SG yet teach-PASS-IPFV-PTCP.NEG school-DAT
{when the war started,} “I WAS NOT studying at school yet”

However, the subject (bi) in an example similar to (2) can precede NEG – that cannot be accounted for if PolP is the highest projection in the clause. There are other data for which the Remnant Movement analysis is not the only possible one and not always explanatory: Evenki, similar to Korean (Ko, Choi 2009; Yun 2013; Yun 2014), has destressed postverbal arguments. In the head-initial construal, the VO order (even with a destressed O) is supposed to be basic, whereas the authors cited show (for Korean) that it is derived. Extending this
account to Evenki, and departing from these and other data, we develop the analysis of the NEG [FIN] – V [PTCP] construction couched in the head-final tree structure and involving right-adjunction after movement rightwards.

**Keywords**: syntax, right-adjunction, head-final, analytic negation, Evenki

**References**
http://homepage.univie.ac.at/daniel.buring/phpsite/content/allpapers.html#thing Accessed December 29, 2015.


The paper examines English translation counterparts of Czech sentences containing the particles copak, jestlipak and kdepak. These particles, which share the expressive and intensifying postfix -pak (Dokulil et al. 1986; Komárek et al. 1986), are elements of the third syntactical plan, i. e. “components which place the content of the sentence in relation to the individual and his special ability to perceive, judge and assess” (Poldauf 1964, 242). The third syntactical plan is fully developed in Czech but to a much lesser degree in English. This paper has a twofold aim: a) to identify and describe English means which perform the same communicative functions as Czech sentences containing the -pak particles; b) to further specify the functions of the Czech particles. The data examined come from the fiction core of the parallel corpus InterCorp.

We use the methodology of contrastive analysis, which allows a comparison of meanings that stem from the same notions and serve the same communicative functions but are conveyed by different means in the respective languages, since “linguistic structure is language-specific while the cognitive and functional-communicative substance which constrains it is potentially universal“ (Boye 2012, 7; cf. Haspelmath 2010).

Concerning the English correspondences of -pak particles, our study shows a general (cf. Boye 2012) tendency in English to rely on grammatical means of expression, mainly various types of interrogative sentences (Dušková et al. 2012; Biber et al. 1999), where Czech prefers lexical means (particles).

To pursue our second aim, the English counterparts are used as markers of communicative functions of the Czech sentences (Malá 2013). The postfix -pak is shown to be a polyfunctional indicator of communicative function (Grepl, Karlík 1998): the -pak particles have content/speaker-related functions (markers of emotional evaluation and deliberative meaning) as well as communication/addressee-oriented functions (cf. Kranich, Gast 2015). The latter comprise the functions of appeal (Štícha 2013), establishing contact (Karlík, Nekula, Rusínová 2000) and triggering certain pragmatic presuppositions (Hirschová 2013; Levinson 1983).

The data further suggest that where the Czech particles indicate epistemic modality (Komárek et al. 1986) or the communicative functions of objection, reproach, disagreement or expressing surprise (Grepl, Karlík 1998), in English the negative meaning tends to be expressed explicitly (e. g. by negative declarative clauses or introductory negative expressions such as not at all).

**Keywords:** the third syntactical plan, translation correspondences, Czech particles, epistemic modality, means of establishing contact
References


A V Raising Account of Verb-Initiality in Isbubukun Bunun

Chao-Kai Shih
National Tsing Hua University, Taiwan
caillet56@hotmail.com

This paper aims to investigate the derivation of verb-initial word order of Bunun with special reference to the Is(bu)bukun dialect, spoken in Namasia and Taoyuan in Kaohsiung City, Taiwan. Bunun is one of the most widespread Austronesian languages spoken in Taiwan. It is proposed that the verb-initial word order of this language is attributed to V-to-C movement and [EPP] on T.

There have been proposals for the derivation of verb-initiality from generative perspective. In particular, Wu (2013) proposes that i) the initial predicate is actually a VP/XP, which has been promoted to the specifier of InflP/IP, which is proposed to be merged between TP and vP; ii) all the arguments, even the promoted subjects in the NAV context, stay within vP; iii) the object undergoes extraction out of the VP and ends up in an intermediate functional projection, namely AgroP, to check its Case feature. However, Wu’s analysis fails to capture the following syntactic properties of Isbubukun Bunun: i) the illicit adjacency of the vP complement to its restructuring predicate, as in (1a-b); ii) the separation of the indicative negator ni and its lexical verb complement, as in (2); iii) the separation of the motion verb and its PP complement, as in (3). In other words, the extraction of the aforementioned complements that do not require Case licensing is left unexplained.

Moreover, under Wu’s analysis, NegP is sandwiched between TP and InflP and thus the word order is predicted to be Fut-Neg-V-S-O only, which is obviously contrary to the fact. A possible solution to this problem would be that VP remains in situ here and object does not raise to [Spec, AgroP]; however, the Case feature on the object would thus fail to be checked and it is unclear how the formal feature on Infl0 which triggers VP fronting is deleted.

Our arguments for V/X raising analysis are as follows. First, [Spec, TP] may be filled either by the external argument or non-subject DP, as evidenced by the possible positions of the postverbal temporal adjuncts, as in (4). Second, the temporal adjuncts without the topic marker hai are disallowed in the preverbal position, which gives evidence that the highest verb or predicate head ends up in a position higher than T. See (5). More specifically, the landing site may be in Fin0 given that the verb cannot precede the complementizer tu (hence not Force0) and that the verb cannot be followed by the topic phrase (hence not Top0 or Foc0). See (6).

This paper proposes the phrase structure for Bunun: [CP C0 [TP ___ T0[EPP] [vP ___ DP_EA v0[EPP] ([Appl DP_AO Appl0) [vP ___ V0 DP_O ___][])]])]. Under Minimalist framework, the derivation is as follows: i) EPP on v may attract the internal or applied argument to its edge; ii) [EPP] on T attracts either the external or the internal argument to its specifier; iii) the verb undergoes successive head movement all the way to an empty C0 (Fin0).

What is more, the proposed V-to-C movement can account for the verb-initiality in Squliq Atayal (R’uyan dialect) as well, which is also one of the most widespread Austronesian languages spoken in Taiwan. First, the fact that the indicative negator ini’ can
license the subject NPI \textit{ana ima’} indicates that this negator ends up in CP domain, as in (7a). When the fronting of the negator is blocked by a higher head, say \textit{wal} in Asp$^0$ or \textit{bali} in C$^0$, the negator remains in its base position Neg between T and v, and thus fails to command the subject NPI; hence the ungrammaticality of (7b) and (7c).

Lastly, a Case parameter is tentatively proposed that V-to-C movement is one of the overt indicators of strong morphological features which determine Case spell-out at PF.

(1) a. \textit{asa} a subali=a \textit{ma-baliv} mas iskaan
   AV.want NOM PN=DET.NOM AV-buy OBJ fish
   ‘Subali wants to buy fish.’
   
   b. \textit{asa-un}=ku a iskaan=an \textit{ma-baliv}
   want-PV=1SG.GEN NOM fish=DET.NOM AV-buy
   ‘Subali wants to buy that fish.’

(2) \textit{na}=\textit{ni} a saia \textit{sipungul-un} ma=zaku
   IRR=NEG NOM 3SG.NOM forget-PV GEN=1SG
   ‘He will not be forgotten by me.’

(3) kau-dadaza a utung=a sia lukis
   AV.go-up NOM monkey=DET.NOM P tree
   ‘That monkey went up to a tree.’

(4) kislaupa (\textit{takna}) a subali=a (\textit{takna}) mas babu
   AV.stab yesterday NOM PN=DET.NOM yesterday OBJ pig
   ‘Subali stabbed a pig yesterday.’

(5) \textit{takna} *(\textit{hai},) kislaupa a subali=a mas babu
   yesterday TOP AV.stab NOM PN=DET.NOM OBJ pig
   ‘Subali stabbed a pig yesterday.’

(6) m-intatulun a subali \textit{tu ma}<i>baliv \textit{mas} ahil=tan
   AV.regret NOM PN COMP AV<EXP>buy OBJ book=DET.OBJ
   ‘Subali regrets having bought that book.’

(7) a. \textit{ini’} rsani yapit \textit{ni’} yumin \textit{ana ima’}
   NEG bring.CV.NIND flying.squirrel GEN PN any who
   ‘Yumin doesn’t bring flying squirrels to anyone.’
   
   b.*\textit{wal} ini’ rsani yapit \textit{ni’} yumin \textit{ana ima’}
   PRF NEG bring.CV.NIND flying.squirrel GEN PN any who
   ‘Yumin didn’t bring flying squirrels to anyone.’
   
   c.*\textit{bali}=nya’ (\textit{wal}) ini’ rsani yapit \textit{ana ima’}
   certainly.not=3S.GEN PRF NEG bring.LV.NIND flying.squirrel any one
   ‘It’s certainly not the case that he didn’t bring flying squirrels to anyone.’

\textbf{Keywords:} word order; verb/predicate-initiality; VP/XP raising; V/X raising

\textbf{References}
Mirativity and Focus in DP

Masaharu Shimada and Akiko Nagano
University of Tsukuba, Tsukuba, Japan; Tohoku University, Sendai, Japan
shimada.masaharu.fu@u.tsukuba.ac.jp; tmg0002as@gmail.com

With the notion of Mirativity, we often refer to an utterance which expresses the speaker’s emotion of surprise, conveying new information unexpected to the speaker. Spanish verbless predicative clauses discussed by Gutiérrez-Rexach and González-Rivera (2014, 102-103), exemplified in (1), can be characterized as bearing a mirative interpretation.

(1)  
\[\text{Extremadamente caro}]_{\text{AP}, \ \text{aquel vino}}.\]  

‘Extremely expensive, that wine.’

(1) contrasts with a usual predicative sentence in (2) in that a be-verb is dropped and the adjective is preposed in (1), and that the predicate in (1) has an evaluative/affective interpretation followed by a topic DP and constitutes a new information focus:

(2)  
\[\text{Aquel vino ES extremadamente caro}\]  

‘That wine is extremely expensive.’

Within a framework of generative linguistics, Cruschina (2011, 105) structurally distinguishes contrastive focus, CFoc, and information focus, IFoc. CFoc and IFoc occupy different positions, as illustrated in (3):

(3)  
\[
\text{CFOcP}\{\text{CFoc}\} \ldots \text{TopP} \ldots \text{IFocP}\{\text{Wh/IFoc}\} \ldots \text{FinP}
\]

Cruschina argues that fronting to the IFoc position is a necessary condition to derive a Mirative interpretation in Sicilian. This fronting is called mirative fronting. Adopting Cruschina’s idea, we can analyze (1) in such a way that the verbless adjectival phrase *extremadamente caro* ‘extremely expensive’ is fronted to the spec position of IFocP.

In the tradition of generative linguistics, the parallelism between clausal structures and nominal structures has been studied. It has been argued, for example, that CP and DP share the same internal structures, and that the same syntactic operation as movement can occur within them. As summarized in Abor et al. (2010), their parallelism holds in terms of information structure as well. Both CP and DP have topic and focus projections within them.

If the mirative fronting analysis for clausal structures is on the right track, we predict that the mirative fronting can be observed in nominal structures as well. Based on the data on prenominal modification in Japanese, this paper will show that this is indeed the case, and that CFoc and IFoc positions can arise in DP.

We treat such a Japanese example of prenominal modification as (4):

(4)  
\[
[\text{DP} \left[ \text{Rel pro}_1 \text{taka-i} \right] \text{wain}_1)]\]  

\[\text{pro expensive-Cop wine} \quad \text{‘expensive wine’}\]

In (4), the noun *wain* ‘wine’ is modified by the relative clause containing the adjective *taka-i* ‘expensive.’ The unmarked interpretation in (4) is just a normal adjectival modification. Interestingly, as observed in traditional Japanese linguistics, (4) can have another interpretation with a different intonation. In the unmarked interpretation, *taka-i* ‘expensive’ and *wain* ‘wine’ are pronounced with their original accents kept. In the other interpretation, the expression is pronounced starting with the low tone and ending up with the rising
intonation. In the latter case, (4) can express the speaker’s surprise of the expensiveness of the bottle of wine in front of her. This is just the mirative interpretation.

This duality of the modificational construction in Japanese can be explained if DP has an IFoc position above the normal modifier position and the mirative interpretation is derived as a result of the (vacuous) fronting of the relative modifier [pro taka-i] ‘expensive’ to the IFoc position within DP.

(5) \[ \text{DP} \left[ \text{IFoc} \left[ \text{Rel pro}_1 \text{taka-i} \right] \text{t}_2 \text{wain}_1 \right] \]

Interestingly, when the modified noun conveys old information with a mirative reading of the whole phrase, it precedes the adjectival modifier, attached with the nominal affix -sa, as shown in (6):

(6) \[ \text{kono wain no taka-sa!} \]

This paper pursues the possibility that (6) is also derived from the structure in (3). The modifiee kono wain ‘this wine’ is fronted over the modifier into the TopP position, changing the inflectional ending of the adjective form -i to -sa.

Aboh et al. (2010) give several examples in which contrastive focus, that is, CFoc, is involved in DP. Also, the contrastive interpretation of the SALAD salad expression observed by Ghomeshi, Jackendoff, Rosen and Russell (2004) may be attributed to CFoc in DP. In contrast, according to Aboh et al. (2010)’s summary, there seem to be not so many phenomena suggesting the existence of information focus, that is, IFoc, in nominal structures. The nominal modification with the mirative interpretation in Japanese provides new evidence for IFoc in DP.

**Keywords:** adjectives; cartography; fronting; Japanese

**References**


Emotions Translated: Enhancing a Subjectivity Lexicon  
Using a Parallel Valency Lexicon  

Jana Šindlerová and Aleš Tamchyna  
Charles University, Prague, Czech Republic  
sindlerova@ufal.mff.cuni.cz; tamchyna@ufal.mff.cuni.cz

Subjectivity lexicons collect words determining evaluative character of a text. For a good analysis of evaluative features in texts, we need reliable and exhaustive resources. So far, the resources existing for Czech are usable, but far from perfect. In this paper, we present an attempt to enhance a Czech subjectivity lexicon using a parallel valency lexicon.

The Czech Sublex (Veselovská 2013) is a lexical resource for Sentiment Analysis tasks. Basically, it provides subjectivity clues, i.e., positive and negative evaluative words. The Czech Sublex was obtained by automatic translation of a freely available English subjectivity lexicon (Wilson et al. 2005); the automatic translation was carried out using a parallel treebank CzEng.1.0  (Bojar and Žabokrtský 2006) and the lexicon was cleaned manually then. The version of Czech Sublex we use in our study includes 4625 evaluative expressions, of which 1549 are verbs.

For enlarging the range of verbs included in the Czech Sublex, we explore (automatically) the bilingual valency lexicon CzEngVallex 1.0 (Urešová et al. 2015). Since verbs commonly constitute the heart of the evaluative state expressed in a sentence, the advantages of using the CzEngVallex for enhancing the Sublex data are the following:

a) CzEngVallex covers a reasonable amount of verbal vocabulary used in media discourse and provides reliable and structured information about valency participants of the verbs (the participants usually correspond to the roles of the Source and Target of the evaluation);
b) the individual verb senses are structured as separate units, allowing us thus to differentiate between emotional and non-emotional meanings of the same lexical unit;
c) the lexicon is interconnected with a sentence-aligned parallel treebank, the Prague Czech English Dependency Treebank, allowing us thus to check immediately the relevance of the outcome.

Exploration of the Czengvallex produced 1166 Czech verb translations unmet in the previous version of the Czech Sublex, which correspond to 578 unique lemmas. Such a number was unexpected considering the fact that the original Czech Sublex, containing 1549 verbs, was built on substantially larger parallel treebank data. The reasons probably lie in the following:

a) the new verbs represent a “raw” outcome and, as it appears, manual clean-up is needed;
b) many aspect counterparts and prepositional variants to other verbs in Sublex are included.
In the paper we present an analysis of the newly acquired verbs and comment on their validity considering evaluative meaning. Mainly, the question of the relation between evaluative meaning and inherently positive/negative semantics is discussed. Also, we make general statements about the distribution of Source and Target roles between the individual verb frame participants and about the propagation of negative or positive meaning to the individual evaluative roles.

**Keywords:** parallel treebank; bilingual valency lexicon; subjectivity; evaluative meaning

**References**
Imageability and Past Tense: Imageability Effects Are Not Limited to Irregular Verbs

Filip Smolík
Institute of Psychology, Czech Academy of Sciences, Prague
smolik@praha.psu.cas.cz

Imageability is the extent to which a word elicits internal sensory image of its referent. It is known to affect various aspects of word processing and acquisition. Recent research also suggested that imageability of words facilitates the processing and acquisition of inflected forms in these words. Highly imageable irregular verbs are processed faster than less imageable ones; this relation is weaker or absent in regular verbs (Prado & Ullman, 2009; Dye et al., 2013). Plural forms of highly imageable nouns appear in children’s language earlier than plurals of less imageable words, even after controlling for frequency and word age of acquisition (Smolík, 2014). The contribution presents two studies that examined the effects of imageability to further evaluate its relation to morphological processing and acquisition.

One study examined the data from the Past Tense Inflection Project (Cohen-Shikora et al., 2013). This includes data on production times for 2200 English verbs in the speeded past tense generation task, as well as ratings of imageability and other variables. The initial analysis revealed that imageability is related to faster past tense production in adults (Cohen-Shikora et al., 2013). The present analysis demonstrates that this relation is complex. There is a three-way interaction so that the effect is strongest in irregular verbs with low frequency. However, even in regular verbs, some effect of imageability can be found.

The second study examined acquisition data from the Manchester corpus of child English. Survival analysis of first usages of past tense, controlling for input frequency and age of acquisition of the uninflected form, revealed that past tense forms appear earlier in highly imageable verbs than in less imageable ones. However, this effect was only present in regular verbs and interacted with frequency. The effects of imageability were only present in words with higher frequency. This is the opposite effect of frequency than observed for past tense generation in adults. It points out that even though imageability affects both morphological acquisition and production, the exact mechanisms in each of these domains are different.

Overall, the findings indicate that semantic properties of words affect the acquisition of grammatical morphology in these words. Unlike the previous research findings on adult production (Prado & Ullman, 2009), the current results show that the effects are not limited to irregular verbs. This poses questions about the mechanism of the effect. It is suggested that the effect is due to facilitative effects of imageability in working memory, rather than long-term memory retrieval of stems.

Keywords. Imageability, past tense, grammatical development, psycholinguistics, morphology.
Figure 1: Model predictions for the relation between response time in the past tense generation task (Study I using the Past tense inflection project data). Predictors are verb regularity, imageability and frequency (TLFRQ).

References
Grammar or Social Understanding? Both Mental State Language and Grammatical Skills Predict Children’s Use of Personal Pronouns and Verb Inflections

Filip Smolík and Veronika Bláhová
Institute of Psychology, Czech Academy of Sciences, Prague
smolik@praha.psu.cas.cz; blahova@praha.psu.cas.cz

The emergence of personal pronouns has been considered as a sign of emerging social understanding and self-other differentiation (e.g., Courage, Edison, & Howe, 2004; Lewis & Ramsay, 2004). However, personal pronouns are also linguistically complex words with strong grammatical roles. It is thus a question whether the initial use of personal pronouns in children marks advances in social understanding, or whether it reflects linguistic, especially grammatical development. The same question may be asked for other devices of person reference, such as conjugated verb forms in languages that inflect verbs for person. The present study addressed these questions by examining the relation between early use of first and second person pronouns, verb inflections, mental state language (MSL), and general linguistic development in a group of children acquiring Czech.

Language samples of 58 children aged 29 to 31 months were analyzed. The samples were recorded when children played with their mothers during a lab visit. In addition, lexical and grammatical comprehension tests were administered, a short picture-vocabulary task, and an experimental task based on TROG (Bishop, 2003). Language samples were searched for all occurrences of words from three lists. 8 forms of the first person pronouns, 8 of the second person pronouns, 37 different words referring to mental states, which included references to sensory perception, emotions, desires, and cognitive states. In addition, number of occurrences of all first-and second-person present tense forms of verbs were counted.

Regression analyses used the number of occurrences of first and second person pronouns as the dependent variable. The predictors were the mean length of utterance (MLU) and the number of mental state language (MSL) tokens. The effect was significant for both MLU ($\beta=0.36, p=0.003$) and MSL ($\beta=0.36, p=0.002$), suggesting that each predictor has a unique contribution to the prediction of personal pronoun use. To ensure robustness against outliers and distributional assumption violations, the analysis was repeated using ordinal regression treating the dependent variable as an ordinal scale. This confirmed the concurrent significant effects of MSL ($p=0.022$) and MLU ($p<0.001$). The use of verb conjugation for person reference showed similar relations to MSL and MLU as pronouns, with both language complexity and mental state language having independent effects on person reference.

The results confirm that children’s references to persons using pronouns and verb forms are related both to language development and to the development of social understanding. This validates previous research relying parental report data (Markova & Smolik, 2014). Second person verb forms are especially strongly related to social understanding. At the same time, there is always a clear relation to measures of grammatical development. Studies using person reference to assess social understanding in children should always take into account the overall linguistic development.
Keywords. Pronouns, person conjugation, grammatical development, mental state language.

Table 1: Descriptive statistics for the key variables

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU</td>
<td>1.97</td>
<td>0.53</td>
<td>1</td>
<td>3.07</td>
</tr>
<tr>
<td>Pronoun tokens (1st and 2nd person)</td>
<td>2.71</td>
<td>3.64</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>Verb tokens (1st and 2nd person)</td>
<td>12.42</td>
<td>9.79</td>
<td>0</td>
<td>46</td>
</tr>
<tr>
<td>Grammar comprehension (TROG)</td>
<td>9.89</td>
<td>3.82</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>Lexical comprehension (picture vocabulary)</td>
<td>17.54</td>
<td>4.42</td>
<td>9</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 2: Correlations between the key variables

<table>
<thead>
<tr>
<th></th>
<th>MLU</th>
<th>Pron.</th>
<th>Verb.</th>
<th>TROG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pronoun tokens (1st and 2nd person)</td>
<td>0.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb tokens (1st and 2nd person)</td>
<td>0.58</td>
<td>0.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammar comprehension (TROG)</td>
<td>0.53</td>
<td>0.33</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Lexical comprehension</td>
<td>0.16</td>
<td>0.27</td>
<td>0.25</td>
<td>0.63</td>
</tr>
</tbody>
</table>

References


Metaphors, Metonymies, and Blends in English–Hungarian Equivalents of an Online ICT Glossary

Réka Sólyom

Károli Gáspár University of the Reformed Church in Hungary, Faculty of Humanities, Department of Hungarian Linguistics, Budapest, Hungary
solyomreka@hotmail.com, solyom.reka@kre.hu

The paper focuses on the semantic features of Hungarian and English ICT (Information and Communication Technology) words in a functional-cognitive framework.

The corpus of the analysis is given by the Hungarian and English words and phrases of an online ICT glossary compiled in Hungarobox Learning Partnership (a Grundtvig Lifelong Learning Programme) (www.hungarobox.eu), containing 365 items for learners at levels B1–B2 in six languages (English, Estonian, Finnish, German, Hungarian, Italian), and additional sample sentences and exercises. As a member of one of the groups who compiled the glossary, the author has experiences in elaborating the lists of English and Hungarian words and phrases as well as testing the translations and the related exercises with university students in different countries.

During compiling the glossary and during testing, it has become clear that various processes concerning meaning construal and understanding govern and can lead to the successful existence and usage of technical terms, among which many neologisms occur in the case of ICT language (e.g. beteggel/betaggel: ‘tag in’, felhő: ‘cloud’, szörföl: ‘surf the internet’). Therefore, with the help of words and sample sentences from the corpus, the analysis concentrates on the following questions: 1) What kind of metaphorical and metonymical processes can be witnessed in the field of Hungarian ICT words? 2) What kind of semantic similarities and differences can be proved between the Hungarian and English equivalents? 3) How can these metaphorical and metonymical processes of meaning construal help language users understand and use ICT words? 4) Are there any traces for other mental processes (e.g. blending) that can lead to more complex semantic systems within the analysed field? Referring to related researches on the ICT field (cf. Meyer–Zaluski–Mackintosh–Foz 1998, Izwaini 2003), the study gives an insight into the semantic characteristics, matchings and differences of the analysed English–Hungarian equivalents.

The analysis is based on a functional-cognitive approach, and it argues for the importance of metonymical (cf. Kövecses–Radden 1998) and metaphorical (cf. Lakoff–Johnson 1980, Kövecses 2015) features and relations (Dirven 2003) in the case of many items in the glossary. It claims, too, that being creative while using neologisms (cf. Benczes 2010, Sólyom 2013, 2014), these features can help build various mental spaces (Fauconnier 1994 [1985]) and may lead to conceptual integration (blending) (Fauconnier–Turner 1998) in order to make human understanding and language use more plausible and effective.

Keywords: ICT language, metaphor, metonymy, blending, neologism
References
Benczes, Réka. 2010. ”Setting limits on creativity in the production and use of metaphorical
and metonymical compounds.” In Cognitive Perspectives on Word Formation, edited by
Dirven, René. 2003. “Metonymy and metaphor: different mental strategies of
conceptualisation.” In Metaphor and Metonymy in Comparison and Contrast, edited by
University Press.
Science 22 (2): 133–87.
Izwaini, Sattar. 2003. “A corpus-based study of metaphor in information technology.” In
Proceedings of the Interdisciplinary Workshop on Corpus-Based Approaches to
Figurative Language held in conjunction with the Corpus Linguistics 2003 conference
UCREL technical paper number 18, edited by John Barnden, Sheila Glasbey, Mark Lee,
Oxford: Oxford University Press.
University of Chicago Press.
Internet Terms in English and French.” In Euralex’98 Proceedings, vol. II., edited by
Thierry Fontenelle, Philippe Hiligsmann, Archibald Michiels, André Moulin and
neologisms.” Studia Linguistica Hungarica (Formerly Annales Sectio Linguistica). The
Linguistic Journal of the Faculty of Humanities, Eötvös Loránd University, Department
Sólyom, Réka. 2014. “Semantic Features of Hungarian Neologisms With the Prefix Be:
Analysis of Some Frequently Used Verbs (e.g., bevállal).” AHEA – Hungarian Cultural
www.hungarobox.eu
A Sociolinguistic Study of Hybrid Language Practices, Attitudes towards Language and the Role of Language in Identity Formation of Latina Students at a Public University in Los Angeles

Eva Švejdarová
ŠKODA AUTO University, Mladá Boleslav, Czech Republic
svejdarova@is.savs.cz

This paper is based on data collected as part of a larger sociolinguistic study focused on hybrid language practices, attitudes towards language and the role of language in the formation of identity among Latina students in Los Angeles, California, United States. Mixed methods with a special focus on qualitative methodologies of sociolinguistic research were employed to examine attitudes and preferences towards language use and the impact on identity formation among 72 Latina students. In particular, five students were selected and monitored irregularly for nearly seven years to answer the following research question – “How did these five key participants use language to negotiate their identities?” Qualitative methods of data collection were used in order to investigate those informants’ language attitudes towards the use of English, Spanish and hybrid language practices such as code-switching. The longitude of the study provided fruitful data that added a very important dimension to the research, reflecting not only the myriad of identities assumed by one person, but also the extent to which human beings are able to shift opinions over time.

Issues surrounding ethnic and social identity as well as language choice were the bases for the questionnaires used. The negotiation of language choice and practice, while at the same time exploring the meanings and attitudes towards those linguistic choices, were discussed with the five key Latina participants using interviews, audiotaped conversations, as well as participant observations. Data for this particular study were selected on base of frequency of appearance that was based on coding.

As the language constitutes a central feature of human identity, when we hear someone speak, we immediately make assumptions about gender, education level, age, profession, and place of origin. Beyond this individual matter, a language is a powerful symbol of national and ethnic identity (Spolsky 1999, 181). Appel and Muysken (1987) point out the fact that languages are not only objective, socially neutral instruments for conveying meaning, but are linked with the identities of social or ethnic groups has consequences for the social evaluation of, and the attitudes towards languages (Appel and Muysken 1987, 16).

One of the salient themes of the investigation proved to be the interaction between English and Spanish, of which the result is a linguistic phenomenon called Spanglish1. The study proves that most of the Latina university students that took part in this investigation had a positive attitude toward the use of Spanglish, of which such finding might prove very important for Latino/a students. If Latina bilinguals maintain the same level of consciousness

---

1 In the study, as well as in the questionnaires, the term Spanglish is used to embrace the complex continuum of language situations and is used as a common denominator for the interaction between English and Spanish language.
relative to language practice and choice, they can then influence others’ behavior and attitude toward them by deliberately changing their language choice.

To conclude, using the voices of the participants, this case study uncovered important factors contributing to the relationship between language and identity as they relate to Latina students, as well as exposed the language ideology (Kroskrity in Duranti 2004, 497) behind such processes.

**Keywords:** hybrid language practices, language attitudes, identity, Latina, qualitative research

**References**


Stance and “Saying” Routines in Expert-Lay Communication

Magdalena Szczyrbak
Institute of English Studies, Jagiellonian University
magdalena.szczyrbak@uj.edu.pl

The verb *say*, it can be argued, is one of the stancetaking resources which interactants have at their disposal during high-stake confrontational encounters. Even more so, given that apart from its core meaning of pronouncing words or sounds and expressing a thought or opinion, *say*, whose stancetaking status might not be immediately evident, perspectivises the information provided by speakers or, put differently, indicates speakers’ claim to epistemic priority (cf. Clift 2006). As such, it also plays a role in formulating legal narratives and, consequently, in constructing evidence in legal-lay communication.

Drawing on stance research (du Bois 2007) as well as analyses exploring the pragmatics of *say* (Craig and Sanusi 2000; Brinton 2008; Holt and Johnson 2010; Johnson 2014), the current study examines the patterns of use involving *say* in 30 FBI/police interview transcripts (totalling 271,544 words) from a homicide investigation. The aim of the analysis was threefold: 1/ to determine the frequencies of selected *say* forms (e.g. like *I say/said, you know what I’m saying, when you say, and you say*); 2/ to explore the speaker-form correlation, i.e. to establish how legal professionals (police officers/FBI detectives) and laypersons (witnesses) deploy *say* in interaction and, finally, 3/ to examine the stancetaking potential of selected *say* forms, including their role in shifting standpoints, challenging standpoints, reconstructing reality and marking standpoint continuity.

In what was conceived as a corpus-assisted discourse study (Partington et al. 2013), the Concord function of WordSmith Tools was applied to determine the frequencies of selected *say* realisations. During the qualitative stage of the analysis, in turn, individual instances of *say* were analysed manually to determine their distribution across turns of talk. The investigation showed that individual *say* realisations are used to mark shifts in discourse perspective as well as for intersubjective positioning. It was also discovered that pragmatic *say* can be linked to the enactment of power and discourse control (e.g. *and you say, when you say*) or to the speaker’s low power and uncertainty (e.g. *you know what I’m saying*). Overall, the study revealed that rather than use *say* as a “neutral” communication verb, the interview participants use *I*- and *you*-oriented *say* forms to point to the source of information, enabling “hearers to interpret the status of information with respect to some person’s epistemological assessment of it” (Mushin 2001, 34), i.e. to express epistemological stance.

Keywords: police interviews, *say*-markers, stance

References


As observed among others by É. Kiss (1989, 2002), Hungarian verbs show what is called definiteness agreement with the object: if the object is definite, it is the definite conjugation of the verb that appears, and we have indefinite conjugation when the object is indefinite.

When an infinitive takes a definite or indefinite object there is no difference in the verb forms. In certain well definable cases the finite verb taking the infinitival clause as its argument shows definiteness agreement with the object of the infinitive. These data are usually accounted for by assuming clause union (Den Dikken 2004) or Long Distance Agreement (LDA) taking place between the finite verb and the object of the infinitive (É. Kiss 2002). Both approaches assume that agreement is between the finite verb and the object DP. When the infinitive has no object the finite verb selecting the infinitive shows the indefinite agreement pattern.

Our research question concerns the exact nature of agreement: its trigger and what exactly agrees with what. We cannot say that agreement depends on the argument structure of verbs: it is not only verbs also taking DP objects that can agree. Agreement does not exclusively depend on the presence of a definite object either: there are verbs that fail to agree with it. The data in (1) can shed some light on the agreement patterns observed. In (1ab) containing akar ‘want’, a verb that also shows definiteness agreement when finite, the verb fog ‘will’ also shows agreement for definiteness. In (1cd) the verb fél ‘be afraid’ is one not showing definiteness agreement. As a result, fog ‘will’ cannot show definiteness agreement with the object of the infinitive embedded into the non-agreeing infinitival clause either.

(1)  

a. Péter fog/*fogja akarni nézni egy filmet.  
Peter will.INDEF/DEF to.want to.watch a film.ACC  
‘Peter will want to watch a film.’

b. Péter *fog/fogja akarni nézni a filmet.  
Peter will.INDEF/DEF to.want to.watch the film.ACC  
‘Peter will want to watch the film.’

c. Péter fog/*fogja félni nézni egy filmet.  
Peter will.INDEF/DEF to.be.afraid to.watch a film.ACC  
‘Peter will be afraid to watch a film.’

d. Péter fog/*fogja félni nézni a filmet.  
Peter will.INDEF/DEF to.be.afraid to.watch the film.ACC  
‘Peter will be afraid to watch the film.’

These data indicate that agreement is not the result of LDA between the finite verb and the object of the infinitive, and clause union does not necessarily have to be assumed either (at least in order to account for agreement). Definiteness agreement seems to have a cyclic nature: the type of the infinitive also has an effect on the availability of definiteness agreement in the main clause. Our conclusion is that the infinitive also agrees with the object
covertly, and that the verb selecting the infinitive agrees not with the object itself, but with the definiteness feature of the infinitive selecting it, if the infinitive has one. When the infinitive lacks such a definiteness feature (like in the case of fél ‘be afraid’) the result is default indefinite agreement. This proposal also accounts for those patterns where the infinitive has no object: it can be argued to have a default indefinite feature, this is what the agreeing finite verb agrees with. The implementation of this difference can be a feature-based analysis operating with default features, or, following Bartos (1999), a structural difference between definite and indefinite constituents can be assumed.

Keywords: Hungarian, infinitive, object agreement, definiteness

References
Contact-induced structural change has long been an area of heated debate in contact linguistics. It is traditionally assumed (e.g., Lass 1980) that language change is caused by internal evolution and thus rule-governed and regular. Under this assumption, systematic changes induced by language contact are unexpected because externally-motivated changes would be unpredictable (Poplack & Levey 2010). Even if one adopts the view that contact-induced change may affect the structural (e.g., morphological and syntactic) domains, whether such change comes about as an indirect consequence of lexical borrowing (King 2000, Sankoff 2002) or results from application of morphosyntactic rules (Thomason 2006, forthcoming) or mechanisms (e.g. Harris & Campbell 1995) remains unsettled. In this talk, we use the new expression bucuo V ‘good to V’ in Taiwan Mandarin discussed in Tseng (2003), Kao (2008), Lien (2008), and Cheng (2014), illustrated in (1), as a case study to show that it serves as an instance of rule borrowing, thus in support of direct involvement of morphosyntactic rules in contact-induced change.

(1) a. Zhege xingren bing bucuo chi.
   this.CL almond cake good eat
   ‘This almond cake tastes good.’ (Tseng 2003:105)

b. Zhengbu dianying zhende shi hen bucuo kan.
   whole.CL movie really SHI very good see
   ‘The whole movie is indeed really very good to see.’ (Tseng 2003:105)

c. Zheben shu kanqilai bucuo du. (Kao 2008:224)
   this.CL book read.QILAI good read
   ‘This book appears to be fun to read.’

We first show that the bucuo V ‘good to V’ sequence is a word and is not generated in the syntax proper. Then we demonstrate that the morphological structure of bucuo-V ‘good to V’ is unique for the grammar of Taiwan Mandarin because Mandarin in general lacks adjectives with a modifier-head structure composed of a disyllabic plus a monosyllabic morpheme (cf. Pan 2010, Dong 2014). Rather, the morphological structure of bucuo-V ‘good to V’ patterns more with that of its equivalent(s) bebai/bephai-V ‘good to V’ in Taiwan Southern Min (or TSM) than with that of any other existing word in Taiwan Mandarin. Thus the emergence of this new expression is unlikely to be motivated by language-internal mechanisms but rather more likely by language contact with TSM. We then argue that a morphological rule that generates bebai/bephai-V ‘good to V’ in TSM is responsible for deriving bucuo-V ‘good to V’ in Taiwan Mandarin when the rule is borrowed or transferred from TSM to Taiwan Mandarin via language contact. The results of this study lend support to
Thomason’s (2006, forthcoming) theory of rule borrowing as well as provide evidence for the view that syntactic change may result from syntactic borrowing.

References


Preverbal Focus and Syntactically Unmarked Focus in Hungarian: A Comparison

Enikő Tóth and Péter Csatár
University of Debrecen, Debrecen, Hungary
toth.eniko@arts.unideb.hu; csatar.peter@arts.unideb.hu

Background. In Hungarian preverbal/identificational focus (PVF) is marked by stress, and the focussed constituent moves into a preverbal position, while information focus is syntactically unmarked (SUF), the focussed element only receives prosodic prominence and remains in situ. It is generally accepted that an exhaustive interpretation is associated with PVF, while SUF is not exhaustive, it only conveys new and non-presupposed information (É. Kiss 1998). Recently, the inherent exhaustivity of PVF has been questioned both from an experimental and pragmatic perspective. Among others, Kas & Lukács (2013) showed that empirical data do not support the syntactic-semantic model of focus interpretation, and Gerőcs et al. (2013) argued that the exhaustivity of PVF is only a pragmatic implicature. Zimmermann (2008) suggested that focus constructions in various languages are triggered by discourse-semantic factors, but Hungarian might be an exception to this generalization. Skopeteas & Fanselow’s (2011) experiments strengthened the exceptional status of Hungarian and showed that Hungarian PVF is not sensitive to contextual or discourse factors.

Aim. The aim of the present study is to collect further data regarding the exhaustive interpretation of PVF and the role of contextual factors. Moreover, since Surányi (2011) challenged É. Kiss’s (1998) view of SUF and argued that SUF might also be exhaustive, we also wanted to compare the acceptability of PVF and SUF in various contexts.

Experiment. We carried out two pilot studies, where participants rated the target sentences on a 1-6 Likert scale embedded in a context, which was presented in writing and also illustrated by a picture. We tested two factors with two levels: context (exhaustive vs. non-exhaustive) and discourse expectability of the focussed constituent (expected vs. unexpected). The target sentences always formed an answer to a wh-question, and the focussed constituent was the object. 8-8 participants filled in a questionnaire containing 20 target sentences.

Results. The results of the pilot studies imply that both PVF and SUF receive significantly higher ratings in exhaustive contexts, while the discourse expectability of the focussed constituent does not influence the ratings in either case (PVF: Friedman-test: $\chi^2(3) = 90.736$, $p < 0.01$, post-hoc tests also showed significant differences between critical conditions; SUF: Friedman-test: $\chi^2(3) = 60.75$, $p < 0.01$, post-hoc tests also showed significant differences between critical conditions).

Discussion. Our findings provide empirical evidence in favour of Surányi (2011), who suggested that SUF can also receive an exhaustive interpretation in Hungarian, at least in a context which strongly supports exhaustivity. The results, if reproduced by the coming experiments (where the target sentences will be presented auditorily in order to exploit the prosodic features of the constructions under investigation), will also question the inherent exhaustivity of PVF included in the standard view. Therefore, the findings support those
views which claim that exhaustivity of PVF arises only as a pragmatic implicature, and since both PVF and SUF received similar ratings, the same might be true for SUF, as outlined in Surányi (2011). The results also confirm Zimmermann’s (2008) and Skopeteas and Fanselow’s (2011) suggestion that the interpretation of PVF in Hungarian does not depend on the discourse expectability of the focussed constituent.

**Keywords:** focus, exhaustivity, experimental pragmatics

**References**


Information Structure and Inverse Quantifier Scope Interpretation

Gergő Turi\textsuperscript{1} and Balázs Surányi\textsuperscript{1,2}

\textsuperscript{1}Research Institute for Linguistics of Hungarian Academy of Sciences, Budapest, Hungary
\textsuperscript{2}Pázmány Péter Catholic University, Budapest, Hungary
turi.gergo@nytud.mta.hu; suranyi.balazs@nytud.mta.hu

Background and objective. The preferred choice of interpretation assigned to scopally ambiguous sentences containing quantified phrases (QPs) is known to be affected by the interaction of disparate properties, including surface structure, grammatical/thematic functions, the semantic type of the QPs involved, and contextual cues, among others. A significant share of the impact that the sentence-context relation bears on actual scope readings is mediated by Information Structure (IS) (Partee 1990, 1999; Herburger 1997), which includes the IS status of the scope-taking QP itself. Topic status has long been claimed to be associated with wide scope (Kempson and Cormack 1981, Reinhart 1983, Erteschik-Shir 1997, Portner and Yabushita 2001). The effect of Focus status, however, is less clear. It has been varyingly linked either to narrow scope (Kitagawa 1990, 1994; Diesing 1992, Kratzer 1995, Krifka 2001, Cohen and Erteschik-Shir 2002) or to wide scope (Williams 1988; Deguchi and Kitagawa 2002, Ishihara 2002, 2010; or either of the two, depending on contrastiveness, Erteschik-Shir 1997).

Experiment. To address the divergence among the latter views, we carried out a Likert-scale based sentence-picture matching experiment in Hungarian.

Participants judged whether a sentence containing a post-verbal, D-linked, subject universal QP headed by ‘each’, presented on a screen, can express a particular targeted scope interpretation that was simultaneously triggered by both its context and the picture (Fig. 1). On the targeted reading the QP had either inverse wide scope over a pre-verbal number phrase, or surface narrow scope (Wide/Narrow). The universal QP was either a new contrastive (corrective) focus, or it was given in the context (Focus/Given). We added controls (which were expected not to license the intended scope readings) and fillers to this basic 2x2 design. Counterbalancing and randomization was applied as usual.

Results and discussion. Judgment $z$-scores were analyzed using a LMEM, and yielded the following results: the IS status of the sentences induced a significant main effect, however neither Scopal, nor interactive effects were found. This means that (i) both when the QP is focused and when it is given, both scope readings are equally easily available to it (corresponding to cca. 4 on the 5-point scale), and (ii) sentences with a focused QP are less acceptable than those containing a given QP. The low rating of controls provides clear evidence of participants’ sensitivity to (un)availability of scope in the task. We conclude from these findings that the availability of an inverse scope interpretation to a universal QP exhibits a high degree of autonomy from IS.

A follow-up experiment showed that the independent reason why focus on QP led to relative degradation (on both scope readings) was that target sentences containing a focused post-verbal QP were semantically more complex than target sentences in which the post-verbal QP was given.

Keywords: Information Structure, Quantifier Scope, Hungarian
Figure 1: A sample picture of the dialogues (IS: Focus, Scope: Narrow)

Figure 2: Results in raw scores

Selected references
Subject-Complement Asymmetry in Copular Clauses: Investigating the Realization of the ‘General’ NP across Different Copular Types

Wout Van Praet
UNamur, Namur & KU Leuven, Leuven, Belgium
wout.vanpraet@unamur.be

Ordinary ‘non-equative’ copular clauses are characterized by a difference in lexicosemantic and/or referential generality between subject and complement (Lyons 1977; Heycock and Kroch 1999; Mikkelsen 2005). The aim of this paper is to investigate the nature of this hyp(er)onymical relation by looking at the form-function correlation of the lexicosemantically/referentially more ‘general’ NP (e.g. a musician in (1)) across different copular types.

(1) David is a musician. ['specific’ NP] be ['general’ NP]

A sample of 1,200 copular examples from WordbanksOnline was first analysed into two typological distinctions. The first compared specificational vs. ascriptive clauses (i.e. with the general NP being construed either as a variable for which a value is specified, or as an abstract category that is ascribed to the subject respectively); the second distinguished between definite vs. indefinite clauses (i.e. based on the (in)definiteness of the ‘general’ NP). This resulted in a cross-classification into four types (300 examples each):

(2) (a) In those days, a popular remedy was laudanum. (WB) [indef. specificational]
(b) I wondered if German efficiency was really a myth. (WB) [indef. ascriptive]
(c) The purpose of a school is to make people inquisitive. (WB) [def. specificational]
(d) Wyness was the founder [...] of The Olympic Club. (WB) [def. ascriptive]

All examples were then analyzed for the semantics of the general NP’s head noun and – if any – its modifiers. A first semantic distinction was made between ‘abstract’ vs. ‘concrete’ head nouns (i.e. denoting concepts relatively remote from perception vs. physical objects respectively (Cruse 2004)).

Secondly, if the general NP contained premodifiers, a quantitative analysis was used to determine the number of premodifiers in each example. A qualitative analysis sorted those premodifiers in different functional zones of premodification (i.e. ‘postdeterminers’, ‘attributes’ and ‘classifiers’), which reflect an abstract-to-concrete continuum (Feist 2012).

Thirdly, in case of postmodification, the length of the postmodifier zone was measured based on word count. The length of the postmodifier slot was taken as an index of the lexicosemantic/referential specificity of the ‘general’ NP: longer postmodifiers have a more restrictive effect on the reference set of the class denoted by the head noun more as well as on the lexicosemantic specificity of the head noun.

Finally, an analysis of the syntactic category of the ‘specific’ XPs allowed me to determine the typical categories associated with each copular type as well as the degree of syntactic variation between them. A distinction was drawn between ‘nominal’ categories
(nouns, pronouns) and more ‘relational’ (verb-like) categories (‘infinitives’, ‘wh-clauses’, ‘that-clauses’, etc.). This distinction provides an insight into the referential specificity of the ‘general’ NP: a larger degree of syntactic variation in the ‘specific’ XP slot is more likely to result in a higher frequency of ‘relational’ categories, which – unlike ‘nominal’ categories – typically do not denote concrete entities and, therefore, typically have more abstract meaning. Consequently, the more abstract the meaning of the ‘specific’ XP, the more likely it is that the ‘general’ NP has abstract meaning as well.

To conclude, this paper investigates the realization of the ‘general’ NP across different copular types, looking into (i) the meaning of the head noun, (ii) its potential premodifiers, (iii) its potential postmodifiers and (iv) the nature and degree of syntactic variation of the ‘specific’ XPs with which it is linked. The lexicosemantic/referential analysis of the ‘general’ NP will, in turn, allowed me to investigate how the ‘asymmetric’ relation between the subject and the complement is construed in different copular types.

**Keywords:** copular clauses; subject-complement asymmetry; (in)definite nominal reference; lexicosemantic/referential meaning, specification/ascription

**References**


Specificity and Past Participle Agreement in Catalan: A Diachronic Approach

Jorge Vega Vilanova
University of Hamburg, Germany
jorge.vega.vilanova@uni-hamburg.de

Past Participle agreement, PPA, (1) was common in all Old Romance languages, but is getting lost in the course of time: nowadays we find it e.g. in Italian and French, but not in Spanish, Portuguese and Romanian, and as for Catalan, PPA seems to be optional (e.g. Lois 1990).

(1) Creus que ja l’ haurà vista, la pel·lícula? (Catalan)
‘Do you think he/she has already seen this film?’

In this paper, I will provide a new analysis of PPA in Catalan. On the basis of newly collected data I will show that the loss of PPA can be related to other phenomena than auxiliary selection or the syntactic position of the object. I will demonstrate that PPA is connected to differential object marking (DOM) and clitic doubling (CLD), two phenomena which are interrelated to specificity. Assuming that specificity is encoded together with verbal aspect within the functional category AgrO (Krifka 1989, Leiss 2000 etc.), I will suggest that PPA and CLD/DOM are two complementary means of satisfying formal features (case and aspect/specificity) hosted in AgrO. Old Catalan data seem to confirm this claim.

The loss of PPA has usually been related to reanalysis of the participle and grammaticalisation of the auxiliary verb (e.g. Macpherson 1967). Building on this view, Lois (1990) proposes a positive correlation between auxiliary selection and PPA. However, Catalan (with PPA, but no auxiliary selection) seems to contradict her analysis. Other approaches emphasize on the agreement relation between the object and the participle by means of movement to the functional category AgrO (Kayne 1989). Looking at Old Romance, however, poses a challenge for these accounts: on the one hand, object shift does not explain the obligatoriness of PPA (objects in situ also trigger PPA); on the other hand, these accounts cannot identify an adequate trigger for the loss of PPA in Modern Romance languages.

My corpus provides evidence for the fact that the first cases of optional PPA involve non-specific objects (2), whereas agreement with definite specific objects remains stable at least until the 16th century (3). So, specificity seems to be the trigger of the language change.

(2) car oït he moltes coses (3) haveu cobrada la vostra raó
‘since I heart many things that…’ ‘you recovered your common sense.’

As has often been argued, specificity is also responsible for CLD and DOM, which links them to PPA/object-shift (Tsakali & Anagnostopoulou 2008). The emergence of CLD, indeed, coincides with the decline of PPA. Further differences between Old and Modern Romance, however, have to be connected to case-assignment: when word-order got fixed in Old Romance (Fischer 2010), morphological case marking first weakened and finally it disappears. Consequently, I will claim that more local mechanisms for object case checking
(e.g. V-incorporation; cf. López [2012]) emerge and movement to AgrO ceases to be an option, and no agree relation is established with the object.

**Keywords:** Agreement; Past participle; Language change; Specificity; Case-assignment

**References**


Using Universal Dependencies for Cross-Linguistic Comparison of Syntactic Features in Evaluative Structures

Kateřina Veselovská
Charles University, Prague, Czech Republic
veselovska@ufal.mff.cuni.cz

This paper is to present a unique project of parallel treebank annotation and a case study of its practical use. We describe the project of Universal Dependencies, explain the methodological approach, illustrate the basic principles of parallel annotation and introduce the universal annotation scheme. Also, we provide a contrastive syntactic analysis of evaluative sentences in typologically different languages.

Over a long time, linguistics has been witnessing a growing interest in dependency analysis related to availability of richly annotated treebanks for many languages. However, annotation styles of the treebanks often differ and it is thus difficult to provide reliable studies concerning cross-linguistic comparison of various linguistic features conducted on large syntactically annotated datasets.

The project called Universal Dependencies\(^1\) is to develop cross-linguistically consistent treebank annotation for many typologically different languages. Its annotation scheme is based on an evolution of Stanford dependencies (de Marneffe et al. 2006, 2008, 2014), Google universal part-of-speech tags (Petrov et al. 2012), and the Interset interlingua for morphosyntactic tagsets (Zeman 2008). The general concept of the project is to provide a universal inventory of categories and guidelines to facilitate consistent annotation of similar constructions across languages, while allowing language-specific extensions when necessary.

In the current study, we use Universal Dependencies to compare syntactic features in evaluative structures of selected languages. Based on the contrastive analysis, we create universal rules for evaluated entity detection in given structures (such as e.g. “If you find an evaluative adjective which is a part of a verbonominal predicate, the subject of its governing verb should be an evaluated entity.” – cf. The food was great). The universal rules can be further employed e.g. in the so-called opinion target identification, i.e. a computational application in which the evaluated entities need to be automatically identified in natural language texts. Moreover, the analysis of syntactic features in evaluative sentences is also applicable within other subtasks of sentiment analysis, i.e. automatic extraction of emotions from text.

**Keywords:** universal dependencies, syntactic analysis, evaluative structures

**References**

de Marneffe, Marie-Catherine, MacCartney, Bill and Christopher D. Manning. 2006. *Generating typed dependency parses from phrase structure parses.* In Proceedings of LREC.

\(^1\) See https://universaldependencies.github.io/docs/.


The study concerns the analysis of Czech quantifiers considering their morphosyntactic patterns (including some aspects of interpretation). Its theoretical framework is following Pesetsky (2013), who proposes that case is a categorial feature of a lexical head, e.g. N (N-Case, Genitive), copied to the relevant domain via derivation or a primeval case assigned by a respective categorial suffix. The framework also assumes one grammatical system including syntactic derivation as well as morphology - i.e. I will propose specific structures (and derivations) for the Czech data as they can be deduced from language specific morphological patterns using (i) the theory of Feature Assignment (i.e. a locality restricted feature spreading) modified for the attested language specific data and (ii) the universal blocking of the case assignment and feature spreading caused by head movement and/or by a phrasal border.

Concentrating on structures comprising two Czech quantifiers spousta ‘plenty’ and málo ‘little/few’ (and some others) I will provide a summary and statistics of data extracted from Syn 2010 (Czech National Corpus), the study by Veselovská (2001) and Jiranová (2008). The Czech structures with quantifiers appear with complex agreement patterns and they allow a certain level of variety as e.g. below in (a) and (b), some of which are apparently problematic even with native speakers (and the prescriptive rules).

I will demonstrate the acceptable (=attested) varieties and consider the existing structures from the point of view of economy of derivation. I will show that the proposed analysis is able to predict the occurrence of existing paradigms: the structures which are equivalent w.r.t. economy criterion do compete in the collected corpus (the speakers use both varieties), while the structures which are structurally unambiguous do not show statistically relevant variety.

(a)  
Přišlo /-la  spousta lidí.  
arrivedSN/SF plentyNOM peopleGEN
‘There arrived a lot of people.’

(b)  
Mluvil s  pár svědky / párem svědků  
spokeSM with coupleNOM witnessGEN [couple witness]INSTR
‘He spoke with a couple of witnesses.’

My analysis will be based on contrasting the Czech quantified structures (QP) with projections of prototypical nominals (DP). I will propose the derivation of the quantificational structures within the above theoretical framework, showing that two theoretically predicted kinds of genitive exist inside the Czech nominal structures - (i) one primeval genitive, which is the case preserved inside DP when D was deprived of its case assigning potential: this case is typical for the structures containing quantifiers (including
Czech group nouns, some quantifiers and numerals); and (ii) the genitive assigned during the derivation by a categorial head N (under extended sisterhood relation) which is typical for standard double DP patterns with adnominal genitives (including the Czech container nouns). I will demonstrate the distinction between the two patterns as it correlates other characteristics of the quantified projections.

References


In standard entailment-based typologies of adjectives, modifiers such as *fake* or *false* are commonly classified as privative adjectives (e.g. Kamp 1975). The compositional behavior of partitives is puzzling since they display the inference pattern in (1) \[ [AN] \cap [N] = \emptyset \], i.e. denotations of composite constructions involving such modifiers do not overlap with denotations of modified nouns. Unlike in cases of intersective or subsective adjectives, a sentence such as (2) *That is a fake gun* does not entail (3) *That is a gun*, but rather its negation, i.e. (4) *That is not a gun*. Surprisingly, despite the inference pattern in (1)–(4) sentences such as (5) *Is that gun real or fake?* are felicitous.

Due to their unique behavior partitive adjectives have been recognized as a challenge for compositionality. To account for the data such as (5), it has been recently proposed by Partee (2009, 2010) that in fact modifiers such as *fake* are subsective adjectives which coerce a modified noun. For instance, the noun *gun* in the expression (6) *fake gun* is coerced to mean both guns and non-guns. Nevertheless, as pointed out by Oliver (2013) this approach seems to predict that (2) should be felicitous also when talking about a fork instead of a toy gun; this is obviously not the case though. On the other hand, Del Pinal (2015) argues that compositionality can be saved assuming dual content semantics of common nouns which differentiates between E-structure, i.e. the extension of a noun, and C-structure, i.e. “core facts” associated with the noun’s stereotype. In this paper, I propose an alternative approach to privative adjectives based on the distinction between tokens and types introduced in standard literature on genericity (e.g. Carlson 1977).

Building on the analysis of relational adjectives by McNally & Boleda (2004), I propose that privative adjectives are in fact properties of type-level individuals. Following Krifka (1995), I assume a broad notion of ‘concept’ to cover both conventional kinds and entities which do not need to be well-established, but are not spatio-temporally bounded just as kinds proper. In other words, kinds constitute a subset of a more general class of concepts. The core of the analysis is the following one. Privatives do not target objects, but rather concepts. They establish a “peripheral” subclass within a concept. Such a “peripheral” subclass differs from the “central” (or “core”) one in that their members are not stereotypical and often lack some property typically associated with a concept in question. As a result, (6) refers to a type-level entity which is taxonomically associated with a concept ‘gun’ although it does not shoot.

The analysis seems to be supported by the fact that privatives fit well in generic contexts such as (7) *A fake gun is a kind of a gun which does not shoot* or (8) *All sorts of guns are strictly prohibited on board: regardless whether they are real or fake*. Furthermore, linguistic evidence in favor of the treatment of privative adjectives as properties of concepts comes from word order. Assuming that adjective ordering corresponds to the order of composition (cf. McNally & Boleda 2004), it is predicted that concept-modifying adjectives should occur closer to a modified noun than object-modifying adjectives. The contrast between such
phrases as (9) small fake gun and (10) #fake small gun seems to corroborate the proposed analysis.

**Keywords:** modification; privative adjectives; genericity; concepts

**References**
This paper shows, based on the distribution of bound (fake) indexicals in four Germanic languages, that binding is not sensitive, nor can it be assumed to be driven or mediated by functional heads as postulated in many current approaches to binding (Reuland 2001, 2005, 2011, Chomsky 2008, Kratzer 2009). Instead data are provided that argue for a return to the traditional view that binding requires a direct dependency between the antecedent and the bindee. I propose that this dependency is best formalized as Reverse Agree (Wurmbrand 2011, 2012, Zeijlstra 2012) constrained by a locality condition reminiscent of Rule H (Heim 1993, Fox 1998) and the concept of feature sharing proposed in Pesetsky and Torrego 2007.

Sentences such as (1) are possible in E(nglish) and D(utch), but require a third person possessive in G(erman) and I(celandic) to express a bound variable [BV] interpretation.

(1) I am the only one who did my best.

Accounts based on binding via functional heads cannot predict the language variation and also make incorrect predictions for constructions in which the order of antecedent and bindee is reversed but the functional heads involved are kept constant, as in (2) in G: to yield a BV interpretation, the antecedent must c-command the pronoun.

(2) weil {mein_BV Sohn} nur mich_{AC} {mein_BV Sohn} versorgt
since {my_{NOM} son} only me_{ACC} {my_{BV} Sohn} take.care.of.3.SG

I propose that the language variation in (1) originates in a difference in the morphological make-up of the head DP of the relative clause (in G also the relative pronoun): DP_REL shows gender distinctions in the singular in G/I but not in E/D (see Table 1). The semantic dependencies involved in (1) which yield [I, am the only x, such that x does x’s best] are: i) predication (I ➟ the only one), ii) relativization (movement of REL.OP or head NP, λ-abstraction), and iii) binding of the possessive. Following Pesetsky and Torrego 2007, multiple transitively linked Agree relations involve feature sharing. Thus in (1), I, the only one who, and my are linked, and all features common to these elements are shared. In G/I (3), the features shared are gender φ and number # (der/die einzige(n)), whereas in E/D (4), only # is shared (the one/ones). Assuming shared features cannot be overwritten, this correctly predicts that if I=female, the possessive pronoun in (3) must realize the shared φ, # features, yielding ‘her’ (3rd person [π] corresponds to the lack of person).

(3) I [1.F.SG] the only Ø who [# φ]_{G,I} my [# φ π] ➟ her
(4) I [1.F.SG] the only Ø who [# φ]_{E,D} my [# φ π] ➟ her or my

In E/D, only SG is shared, and the possessive pronoun therefore needs to acquire the remaining feature values in a different way. I suggest that this can happen via an additional feature sharing relation with an element linked (directly or indirectly) to the pronoun already through the chain of dependencies. Thus, in (4), the pronoun can draw φ or π from the features of the matrix subject—if the former, her arises, if the latter my arises. Lastly, to
exclude such ‘reaching up’ in G/I (3), I propose a locality constraint Rule $H_{PF}$: A variable $x$ cannot Agree with an antecedent $\alpha$, in cases where a more local antecedent $\beta$ (with the same binding index) could Agree with $x$ and share morphosyntactic features with $x$. I show that Rule $H_{PF}$ operates as a general constraint on feature transmission in binding and correctly predicts the distribution of (im)possible mismatches attested in gender agreement.

**Keywords:** binding, fake indexicals, Agree(ment), feature sharing, locality, mismatches

<table>
<thead>
<tr>
<th>Table 1</th>
<th>English</th>
<th>Dutch</th>
<th>German</th>
<th>Icelandic</th>
</tr>
</thead>
<tbody>
<tr>
<td>F.SG</td>
<td>the only one</td>
<td>de enige</td>
<td>die einzige</td>
<td>sú eina</td>
</tr>
<tr>
<td>M.SG</td>
<td>the only one</td>
<td>de enige</td>
<td>der einzige</td>
<td>sát eini</td>
</tr>
<tr>
<td>PL</td>
<td>the only ones</td>
<td>de enigen</td>
<td>die einzigen</td>
<td>þær einu</td>
</tr>
</tbody>
</table>

**References**
Wurmbrand, Susi. 2012. Agreement: Looking up or looking down? Lecture given in Agreement seminar, MIT, Cambridge, MA.
Semantic Gender Features of Nouns—Visible and Active in Syntax

Susi Wurmbrand
University of Connecticut
susi@alum.mit.edu

This paper offers new evidence for semantic gender features on nouns and featurally triggered semantic agreement.

German has formal and semantic gender, which can mismatch in certain cases. For instance, the noun Mädchen `girl’ is grammatically NEUT(er), but semantically FEM(inine). The noun Person `person’ is grammatically FEM, but can correspond to either a FEM or a MASC(uline) referent. I propose that nouns that semantically encode a specific gender (girl) involve a semantic gender feature as part of the lexical entry, whereas nouns that do not specify gender (person) do not involve a semantic gender feature.

First evidence for a different featural make-up comes from binding. When such nouns head NPs which bind pronouns, a difference arises: Mädchen, (1a), can trigger formal (NEUT) or semantic (FEM) agreement, whereas Person, (1b), can only trigger formal (FEM) agreement.

(1) a. Nur das Mädchen / Kein Mädchen hat sein / ihr Bestes gegeben formal/semantic
   Only the.NEUT girl / No.NEUT girl has /her best given
   b. Nur diese Person / Keine Person hat ihr / *sein Bestes gegeben only formal
   Only this.FEM person /No.FEM person has /his best given

Further evidence for the presence of a semantic gender feature on nouns like Mädchen can be drawn from nominal ellipsis. Nominal ellipsis comes in two types: N(P)-ellipsis, which involves PF deletion of a specific antecedent noun; and `deep’ ellipsis, which involves an abstract 0.N specified for [+ANIMATE] but no concrete noun. The two types of ellipsis can be distinguished by context. N(P)-ellipsis as in (2a) singles out one girl from a group of girls, whereas deep ellipsis as in (2b) does not restrict the comparison group to girls but the N is interpreted as an abstract PERSON/HUMAN. In (2a), agreement on the ellipsis remnant is determined by the features of the elided noun (via concord). In (2b), crucially, the ellipsis remnant cannot show formal (NEUT) agreement with the antecedent, but must show semantic FEM agreement. I argue that this is the result of obligatory semantic agreement with the antecedent: the semantic FEM feature of das Mädchen is transferred to 0[+ANIM], which then in turn triggers formal agreement/concord on the ellipsis remnant. Lastly (2c) shows that (2b) cannot involve a silent noun `person’, since a masculine subject, which would be compatible with a FEM DP when it contains the overt noun Person, cannot occur with a FEM remnant.

(2) a. Context: a group of girls
   Das blonde Mädchen ist das einzige Mädchen, das lacht
   the.NEUT blonde girl is the.NEUT only girl who.NEUT laughs
   b. Context: a group of men and only one girl
      Cannot mean: the girl is the only girl...
      Das Mädchen ist *das einzige / ✓ die einzige 0[+ANIM] die lacht.
      the.NEUT girl is *the.NEUT only / ✓ the.FEM only ONE who.FEM laughs
   c. Er ist {die einzige *(Person), die ... } / {der einzige (*Person), der }
      He is {the.FEM only *(person) who.FEM } / {the.MASC only *(person) who.MASC }

These data and analysis challenge the universal nature of Corbett’s (1979, 2006) Agreement Hierarchy since predicates are lower on the semantic agreement scale than relative pronouns, yet, the former do but the latter don’t allow semantic agreement. Furthermore, since semantic
agreement feeds into formal agreement in (2b), the semantic features involved need to be visible to the formal component. This can be achieved if nouns such as Mädchen (and as a result also the projecting NP/DP) have two sets of features—formal and semantic gender features.

**Keywords:** Agreement, features, mismatches, semantic gender
Testing Phonotactic Constraints: A Language Game Experiment

Joanna Zaleska
University of Leipzig
joanna.zaleska@uni-leipzig.de

Introduction. This paper reports on a pilot study involving an experimentally constructed language game, aimed as a contribution to the long-standing debate on the status of the Polish high central vowel, [ɨ]. The results suggest that the ban on word-initial [ɨ] is an accidental gap, and as such cannot be used as an argument for the allophonic status of the vowel. Additionally, they indicate that the [±back] feature spreads from a vowel onto the preceding consonant rather than vice versa, supporting the hypothesis that [ɨ] and [ɨ] are separate units.

Background. The surface vowel inventory of Polish contains 6 oral monophthongs, [a, ɛ, ɔ, u, i, ɨ]. The two high unrounded vowels, front [i] and central [ɨ], stand in complementary distribution: [i] may appear word-initially and after ‘soft’ (palatalized) consonants as well as prepalatals, (1a). [ɨ] may only appear after ‘hard’ consonants (1b).

(1) a. il [iw] ‘loam’ mila [miwa] ‘nice’ (fem) sila [ɕiwa] ‘strength’
    b. *[#i] myła [miwa] ‘she washed’ *[ɕi]

This has led some scholars (e.g. Czaykowska-Higgins 1988) to conclude that [ɨ] is a positional variant of /i/ that only appears after non-palatalized consonants. Others (e.g. Rubach 1984) recognize /ɨ/ as a separate underlying segment and treat the soft consonants as allophones of their plain equivalents, derived by spreading [–back] from the following high vocoids. The latter view reduces the total absence of [ɨ] in the word-initial position to an accidental gap.

Research questions. To determine whether the lack of word-initial [ɨ]s is a systematic gap, I used a Pig-Latin type game, which involved moving the word-initial onset to the end of the word and suffixing an /u/, thus, e.g. turning droga [dɾɔɡa] ‘road’ into [ɕɡadɾu]. If the ban is a systematic force in the language, we may expect speakers to realize [ɨ] differently (e.g. by changing it into [ɨ]) when the first vowel becomes word-initial in a word like ryba [rɨba] ‘fish’. The game also makes it possible to test whether the frontness of the vowel in a C′i sequence is a reflex of the soft quality of the preceding consonant. If so, the initial consonant in a word like piwo [pivɔ] ‘beer’ should retain its palatalization when moved. If, on the other hand, the [–back] feature originates on [ɨ], the consonant is expected to lose palatalization.

Method. The experiment was run with 20 native speakers of Polish, 10 male and 10 female, aged between 22 and 52 (mean age 31.2). It consisted of a learning phase, in which the participants trained to achieve fluency in the game by responding to words with non-high vowels provided verbally by the experimenter. In the production phase that followed, they had to produce a verbal response to items heard through headphones. The stimulus set consisted of 20 Polish disyllabic words of the shape (C)CV(C)CV, all of them frequently used nouns in nominative case (singular and plural). In 10 of the items, the vowel in the initial syllable was [i]; in the remaining 10, it was [ɨ]. The 20 experimental items were randomly interspersed with 60 fillers, which contained other vowels in the initial syllable but otherwise had the same characteristics as the experimental words (disyllabic, (C)CV(C)CV shape). The list of 80 items was presented to each participant in the same order. The responses were recorded and later submitted to perceptual inspection and transcription.
Results and discussion. In the experiment, the [i] vowel was realized differently in the transformed word and in the stimulus in 1% of the trials, whereas [ɨ] was modified in 16% of the trials (Fig. 1). It was not, however, fronted to [ɨ] but rather lowered to the mid position. This suggests that the gap is accidental for many speakers, and even those speakers who are aware of it do not associate [ɨ] with /i/. Secondly, whereas plain consonants remained non-palatalized in the transformed word in 100% of the trials, soft consonants were depalatalized in 91% (Fig 2). This provides strong support to the regressive palatalization hypothesis.

Figure 1. Proportion of vowels realized faithfully with respect to the stimuli, broken down by vowel quality in the first syllable of the stimulus. Each dot represents the proportion of ‘faithful’ responses for a single participant with dots jittered on the x axis to avoid overlap.

Figure 2. Proportion of consonants realized faithfully with respect to the stimuli, broken down by vowel quality in the first syllable of the stimulus. Each dot represents the proportion of ‘faithful’ responses for a single participant with dots jittered on the x axis to avoid overlap.

Keywords: artificial language game, surface palatalization, Polish, phonotactics

References
Contentives or the Alleged Universal Noun-Verb Distinction: An Asymmetric Perspective

Hedde Zeijlstra
Göttingen University, Göttingen, Germany
hzeijls@uni-goettingen.de

A traditional assumption is that languages minimally distinguish nouns and verbs, even if that distinction is morphologically invisible (cf. Halle & Marantz 1993, Baker 2003, 2008, Borer 2003, Croft 2005, 2009). For these scholars, this universal noun-verb distinction is given by UG. However, can it be proven that every language exhibits nouns and verbs? Take Samoan. In Samoan, all content words can systematically be used both verbally and nominally. Samoan alu in (1) may either mean 'to go' or 'the going', depending on the grammatical context: combined with a tense marker it obtains a verbal reading ‘go’; combined with an article, it yields a nominal reading ‘(the) going’ (cf. Mosel & Hovdhaugen 1992; Don & Van Lier 2013):

1. a. E alu le pasi I Apia. b. Le alu o le pasi I Apia
   PRES go the bus to Apia The go of the bus to Apia
   'The bus goes to Apia.' 'The going of the bus to Apia'

Are the two instances of alu in (1) two different, homophonous lexical items or not? Whereas Croft (2005) argues that they are, Mosel and Hovdhaugen (1992) have argued that such languages exhibit a single category, generally referred to as contentive, a supercategory above nouns and verbs. Samoan would then be a language that lacks nouns and verbs. Similar claims have been made for Mundari (Hengeveld and Rijkhoff 2005), Kharia (Peterson 2006) and Riau Indonesian (Gil 2013).

The central argument for postulating a universal noun-verb distinction is that the meanings of the verbal and nominal usages of alleged contentives do not follow compositionally. For instance, Samoan tutsi can mean 'to write', 'letter' and 'book'. It would be hard to come up with a compositional analysis that can derive these meanings from a universal base meaning purely on the basis of the linguistic environment. For this reason, Hengeveld and Rijkhoff (2005) have argued that in such flexible languages interpretation does not have to proceed in a strictly compositional fashion. For them, the underlying semantics is vague, so that both readings can come about, even though it is not predictable which meanings must be yielded. The existence of contentives thus seems to rely on the assumption that languages may allow non-compositional interpretations. However, this restricted view on compositionality is at complete odds with its original motivation in Frege (1892) and Janssen (1986); giving up compositionality immediately introduces the virtually unsolvable question as to why sentences in all other languages must be subject to compositional interpretation.

In this paper, I pursue a novel, asymmetric approach. So far it has always been assumed that if a language exhibits two major lexical categories, they must always be nouns and verbs. Thus, if Samoan has two different lexical items tutsi (to write and the writing/letter/book), one must be a noun tutsi (meaning 'writing/letter/book') and the other a verb tutsi (meaning 'to write'). No one, as of yet, has defended the view that some languages may have asymmetric categorical relations: i.e. languages that have contentives and nouns, but lack verbs, or
languages that have contentives and verbs, but lack nouns. But nothing principal excludes such languages. If, however, such asymmetries are not ruled out, it is predicted that in languages that exhibit this kind of flexibility, it should be possible that one of the two related meanings can be used both verbally and nominally (as it is a contentive), whereas the other meaning can only be used either verbally or nominally (as it is either a noun or a verb). This prediction is indeed borne out. As shown by Mosel & Hovdhaugen (1992) and Don & Van Lier (2013) Samoan tutsi, as well as a number of other examples, behave exactly like that. Tutsi has much richer nominal usage, varying from 'the writing' to 'the letter/book', than a verbal usage: tutsi in a verbal context can mean 'to write', but never 'to be a book'. This shows that there is a noun tutsi meaning ‘book’, and a contentive tutsi that can mean ‘to write’ (in verbal contexts) and ‘the writing’ (in nominal contexts). Hence, even though Samoan does exhibit two different lexical categories, it falsifies the claim that every language has nouns and verbs and confirms the claim that languages may exhibit contentives.

**Keywords:** contentives, nouns, verbs, UG

**References**
Croft, W. 2009. ‘Syntax is more diverse, and evolutionary linguistics is already here.’ *Behavioral and Brain Sciences* 32, 453-454.
What Can a Discourse Behavior of a Demonstrative Tell Us about Its Stage of Grammaticalization?

Magdalena Ziková
Charles University, Prague, Czech Republic
magdalena.zikova@ff.cuni.cz

It has been almost one hundred years since it was first explicitly stated that the Czech demonstrative *ten* exhibits some properties of a definite article (Mathesius 1926) and since then uncertainty about its grammatical status persists (Meyerstein 1972, Heine & Kuteva 2006). However, empirical evidence is still missing, partly due to the natural limitation of historical sources, partly due to the lack of adequate methodology.

The aim of this paper is to shed some light on the past development of the Czech demonstrative *ten* by analyzing its distribution in spoken discourse in contemporary Czech. The methodology is mostly inspired by the work done on discourse anaphora (Givón 1983) and similar studies carried out on other languages (Laury 1997). Anaphoric reference (reference to an entity previously mentioned in the discourse) is a common basis of referent identification and, as such, is standardly accompanied with a definite article. Moreover, anaphoric relations in contrast to other bases of identification, such as various kinds of common knowledge and conventional bonds, are relatively easy to delimit, as they are framed by the current discourse. In this way anaphoric reference can be used as a background against which the function of the demonstrative *ten* can be tested: if the demonstrative serves a definite-like function, it should generally mark all previously mentioned referents (apart from some specific usage types).

My analysis relies in systematic comparison of all anaphoric mentions which are marked with *ten* with those marked with a lexical NP without *ten* (leaving the pronoun reference aside). In this way I expect to detect factors which favor or disfavor the usage of the demonstrative, particularly discourse prominence and discourse activation of the referent – which are considered the key concepts in the research on anaphoric means (Givón 2001).

The data consist of three short narratives spoken by 15 subjects (8 females, 7 males). Their task was to retell a short non-verbal film story which they had seen immediately before. 630 anaphoric mentions appeared in the narratives in total and all of them were coded for factors, such as animacy, syntactic role, sentence position (*topic – focus* distinction), sentence type, anaphoric distance (distance from the previous mention of the referent which corresponds to the level of its discourse activation) and its persistence in the following discourse.

Several important findings arise from the quantitative analysis. First, the overall counts prove that the demonstrative is already a well-established means for marking anaphoric reference: out of 630 mentions, 400 of them are accompanied with *ten*. As the identity of the referent is recoverable from the context so that nothing really calls for the referent to be explicitly marked, this finding is far from trivial. Second, the distribution of the demonstrative in Czech is not equal. Its usage appears to consist of two distinct functions: (i) the *strongly* demonstrative one which serves to point to non-activated prominent referents in the *focus*
position, and (ii) the weakly demonstrative one which serves to identify weakly activated non-prominent referents in the topic position. The first usage tends to reintroduce main characters back to the story after some time they were not present in the narration (cf. Laury 1997), the second usage rather contextualizes objects involved in current actions. It seems in the end that out of the factors examined it is the overall prominence of the referent together with its informational value and local/prospective relevance which are predictive of the use of the demonstrative.

**Keywords:** grammaticalization; Czech; demonstrative; anaphora; discourse analysis

**References**