Programm der Arbeitsgruppen/Programme of the Working Groups (2021-02-23)

Benutzung des Programms/How to Use this Programme

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Mittwoch/Wednesday, 24.02.2021

13:45-15:45

AG	13:45-14:15	14:15-14:45	14:45-15:15	15:15-15:45
AG 1: Thomas Strobel, Helmut Weiß, *Grammatical gaps: Definition, typology and theory*	**Caroline Féry** (National and Kapodistrian University of Athens) The role of stress and metrical foot in ineffability in German (Keynote)		**Mariia Privizentseva** (Universität Leipzig) Restrictions on mixed agreement in Russian: Feature conflicts and ineffability in DM	**Andreas Blümel** (Göttingen) Filling the gap: In defense of periphrastic forms as cells in paradigms
AG 2: Ulrike Domahs, Angela Grimm, Mathias Scharinger, *Weak elements in prosodic acquisition and processing*	**Farhat Jabeen** (Bielefeld) Prosodic status of polar *kya* in Urdu/Hindi	**Mirjam Ernestus**, Keynote Speaker (Nijmegen) Reducing word pronunciation variants: Properties and processing		**Alina Lausecker, Angela Grimm & Petra Schulz** (Frankfurt) Truncation of weak syllables: early L2 learners behave like monolingual children

AG 3: Sven Kotowski, Ingo Plag, *The semantics of derivational morphology: Theory, methods, evidence*	**Shelly Lieber & Ingo Plag** The semantics of conversion nouns and *-ing* nominalizations: A quantitative and theoretical perspective	**Richard Huyghe, Alizée Lombard, Justine Salvadori & Sandra Schwab** Assessing the rivalry between French deverbal nouns in *-age, -ion* and *-ment* through the analysis of neologisms	**Verginica Mititelu, Svetlozara Leseva, Ivelina Stoyanova & Gianina Iordachioaia** The meanings of nominal vs. verbal zero affixes
AG 4: Kristin Kopf, Thilo Weber, *Free variation = unexplained variation? Empirical and theoretical approaches to optionality in grammar*	**Kristin Kopf & Thilo Weber** (IDS Mannheim) Introduction	**Freek Van de Velde,** invited speaker (KU Leuven) Didymophilia in language	**Yidong Yu** (Göttingen) Optionality and categorial properties: The case of optional plural marking in Yucatec Maya
AG 5: Hanna Fischer, Melitta Gillmann, Mirjam Schmuck, *Encoding aspectuality in Germanic languages — empirical and theoretical approaches*	**Frank Brisard** (invited speaker) The modal basis of progressive marking	**Maarten Bogaards** Beyond progressive aspectuality: Aspectual *aan-*constructions in Dutch	**Jianan Li** Diatopic and diachronic variations of the German *am*-progressive: A corpus-based investigation
AG 6: Katharina Schaebbicke, Heiko Seeliger, *Empirical approaches to canonical and non-canonical uses of negation*	**Heiko Seeliger & Katharina Schaebbicke** Introduction to the workshop	**Carolin Dudschig** (Tübingen) Processing accounts for negation in linguistic and non-linguistic domains _TOC_250004	**Beata Trawinski** Validating the Performativity Hypothesis to Neg-Raising using corpus data: Evidence from Polish
AG 7: Katrin Axel-Tober, Lutz Gunkel, Jutta M. Hartmann, Anke Holler, *On the nouniness of propositional arguments*	**Katalin É. Kiss**, Invited Speaker (Budapest) From parataxis via clausal adjunction to nouny subordination	**Paul Poirier** (Toronto) Japanese nominalizations and the copula	**Elizabeth Bogal-Allbritten, Keir Moulton & Junko Shimoyama** (Göteborg, Toronto, Montreal) Nouny propositions and their individual correlates: The view from Japanese

AG 8: Cherlon Ussery, Jóhannes Gísli Jónsson, Nicole Dehé, *Ditransitives across languages and frameworks*	**András Bárány** Object agreement and the structure of ditransitives across languages	**Milena Šereikaité** Dative case assignment and ditransitives in Lithuanian	**Vera Lee-Schoenfeld, Gabriele Diewald & Maud Kelly** Fragen kostet nichts: New corpus inquiries into German double-accusative verbs	**Jim Wood** Nominalizations of ditransitives in Icelandic
AG 9: Mingya Liu, Mathias Barthel, *The semantics and pragmatics of conditional connectives*	**Mingya Liu & Mathias Barthel** An experimental approach to the semantics and pragmatics of conditional connectives	**Invited Talk / Anastasia Giannakidou** Manipulation of nonveridical equilibrium produces negative bias in conditionals		**Juliane Schwab & Mingya Liu** *All that* in conditionals
AG 10 (10a) (Kurz-AG): Fabian Schubö, Sabine Zerbian, Sandra Hanne, Isabell Wartenburger, *Prosodic boundary phenomena*	**Isabelle Franz, Christine Knoop, Gerrit Kentner, Sascha Rothbart, Vanessa Kegel, Julia Vasilieva, Sanja Methner & Winfried Menninghaus** Prosodic phrasing and syllable prominence in spoken prose: Prediction from text and validation	**Ludger Paschen, Susanne Fuchs & Frank Seifart** Final and pre-final lengthening in 13 languages	**Ricardo Napoleão de Souza** Phonetic cues to IP-initial boundaries: Acoustic data from English, Spanish, and Portuguese	**Bistra Andreeva, Bernd Möbius, Omnia Ibrahim & Ivan Yuen** The effect of predictability on the duration of phrase-final syllables
AG 11 (10b) (Kurz-AG): Daniel Gleim, Marie-Luise Popp, *Edge-asymmetries in morphophonology*				
AG 12 (11a) (Kurz-AG): Martina Penke, Judith Schlenter, Elyesa Seidel, *Eye-tracking and language production*	**Martina Penke, Judith Schlenter & Elyesa Seidel** Introduction	**Gabriela Garrido Rodriguez, Sasha Wilmoth, Rachel Nordlinger & Evan Kidd** Sentence planning and production in two Australian free word order languages	**Xiaogang Wu & Johannes Gerwien** Linear vs. structural incrementality in the face of sentence production in context	**Yvonne Portele** Implicit perceptual priming in context: When the prominent patient meets the eye
AG 13 (11b) (Kurz-AG): Gerhard Jäger, Johann-Mattis List, *Model and evidence in quantitative comparative linguistics*				

AG 14 (Kurz-AG): Anja Müller, Katharina Turgay, *Grammatische Modellierung als Grundlage für sprachdidaktische Vermittlung*	**Sandra Döring** Grammatiktheoretische Überlegungen zum Schulunterricht	**Matthias Granzow-Emden** Sind Haupt- und Nebensatz noch zu retten? Ein Plädoyer für eine widerspruchsfreie Satzanalyse	**Elvira Topalovic & Benjamin Uhl** Das Stellungsfeldermodell im Sprachunterricht: Wie urteilen Deutschlehrer*innen über ein grammatisches Modell?	**Steffen Dyck** Satzgliedmodelle in Schulbüchern: Eine qualitative Untersuchung der Klassen 5 und 6
AG 15 (12b) (Kurz-AG): Martin Klotz, Anke Lüdeling, Anna Shadrova, *Contrastive corpus methodology for language modeling and analysis*				

16:30-18:00

AG	16:30-17:00	17:00-17:30	17:30-18:00
AG 1: Thomas Strobel, Helmut Weiß, *Grammatical gaps: Definition, typology and theory*	**Andrea D. Sims** (The Ohio State University) How do grammars leak? A close look at the idea of syncretism as repair for defectiveness (Keynote)		**Tabea Reiner** (Ludwig-Maximilians- Universität München) What counts as a gap? The case of typological hierarchies
AG 2: Ulrike Domahs, Angela Grimm, Mathias Scharinger, *Weak elements in prosodic acquisition and processing*			
AG 3: Sven Kotowski, Ingo Plag, *The semantics of derivational morphology: Theory, methods, evidence*	**Viktoria Schneider** Events in the semantics of non-deverbal nominalizations	**Sven Kotowski** Locative prefixes and nominal scalarity	**Olivier Bonami, Louise McNally & Denis Paperno** The meaning of derivation: Relations and scenarios
AG 4: Kristin Kopf, Thilo Weber, *Free variation = unexplained variation? Empirical and theoretical approaches to optionality in grammar*	**Claudia Felser & Anna Jessen** (Potsdam) Correlative coordination and variable subject-verb agreement in German: An experimental study	**Karolina Rudnicka** (Gdansk) **& Aleš Klégr** (Prag) Non-verbal number agreement between the distributive plural and singular: Exceptions or free variation?	**Merit Müller** (Tartu) Investigating morphosyntactic variation in a Uralic minority language: The Aanaar Saami conditional perfect
AG 5: Hanna Fischer, Melitta Gillmann, Mirjam Schmuck, *Encoding aspectuality in Germanic languages — empirical and theoretical approaches*	**Adam Tomas** (München) Grammaticalization in speech-islands: Possibilities and neglects	**Anna Saller** Periphrastic *tun* in Australian German: A past tense habitual marker?	**Nadine Proske** Pseudo-coordinated *sitzen* ('sit') and *stehen* ('stand') in spoken German: A case of emergent progressive aspect?
AG 6: Katharina Schaebbicke, Heiko Seeliger, *Empirical approaches to canonical and non-canonical uses of negation*	**Elisabeth Gibert-Sotelo** Affixal negation is not always negative: Evidence from Catalan and Spanish	**Boer Fu** Negation scoping and focus in Mandarin biased questions: A verum account	**Aurore Gonzalez & Justin Royer** Expletive negation and negative polarity: The view from Québec French
AG 7: Katrin Axel-Tober, Lutz Gunkel, Jutta M. Hartmann, Anke Holler, *On the nouniness of propositional arguments*	**Carlos de Cuba**, Invited Speaker (New York) Relatively nouny?		**Ellen Brandner** (Stuttgart) CP-complementation and selection

AG 8: Cherlon Ussery, Jóhannes Gísli Jónsson, Nicole Dehé, *Ditransitives across languages and frameworks*	**Xiaomeng Ma** Why no double objective construction in Shupamem	**Cherlon Ussery & Hjalmar Petersen** Scope in Faroese ditransitives	
AG 9: Mingya Liu, Mathias Barthel, *The semantics and pragmatics of conditional connectives*	**Magdalena Kaufmann & John Whitman** Conditional conjunctions informed by Japanese and Korean	**Muyi Yang** Iffy discourse: Japanese moshi in conditionals and nominal topics	**Paolo Santorio & Alexis Wellwood** Non-Boolean conditionals
AG 10 (10a) (Kurz-AG): Fabian Schubö, Sabine Zerbian, Sandra Hanne, Isabell Wartenburger, *Prosodic boundary phenomena*	**Gerrit Kentner, Isabelle Franz, Christine Knoop & Winfried Menninghaus** Pause duration and other prosodic boundary cues are not monotonically correlated	**Laurence White** Temporal prediction in speech segmentation is modulated by foregoing utterance length	**Xin Xie, Andres Buxó-Lugo & Chigusa Kurumada** An ideal-observer approach to structured talker variability in prosodic productions
AG 11 (10b) (Kurz-AG): Daniel Gleim, Marie-Luise Popp, *Edge-asymmetries in morphophonology*			
AG 12 (Kurz-AG): Martina Penke, Judith Schlenter, Elyesa Seidel, *Eye-tracking and language production*	**Sebastian Sauppe, Elisabeth Norcliffe, Kamal K. Choudhary, Agnieszka E. Konopka, Aitor M. Egurtzegi, Nathalie Giroud, Shikha Bhattamishra, Mahima Gulati, Gabriela Garrido, Damian E. Blasi, Ina Bornkessel-Schleswesky, Itziar Laka, Martin Meyer, Stephen C. Levinson & Balthasar Bickel** Case marking shapes the time-course of sentence planning: Crosslinguistic evidence from Hindi, Yélî Dnye, Japanese, Basque and Swiss German	**Mikhail Pokhoday, Yury Shtyrov & Andriy Myachykov** Attention and syntactic choice: Evidence from Russian and English	
AG 13 (11b) (Kurz-AG): Gerhard Jäger, Johann-Mattis List, *Model and evidence in quantitative comparative linguistics*	**Johann-Mattis List** Data in quantitative comparative linguistics	**Gerhard Jäger** Models in quantitative comparative linguistics	**Justin Power, Danny Law & David Quinto-Pozos** Methods and models in historical comparative research on signed languages
AG 14 (Kurz-AG): Anja Müller, Katharina Turgay, *Grammatische Modellierung als Grundlage für sprachdidaktische Vermittlung*	**Daniela Elsner** (invited talk) Empirische Befunde zum Einsatz grammatischer Modelle im Deutschunterricht		**Christina Noack, Anna Kurtz & Bastian Stöppler** Integrative Sprachbildung und sprachreflexive Vermittlung in der Grundschule am Beispiel des Projekts „wortreich"

AG 15 (12b) (Kurz-AG): Martin Klotz, Anke Lüdeling, Anna Shadrova, *Contrastive corpus methodology for language modeling and analysis*			

Donnerstag/Thursday, 25.02.2021

9:00-10:30

AG	9:00-9:30	9:30-10:00	10:00-10:30
AG 1: Thomas Strobel, Helmut Weiß, *Grammatical gaps: Definition, typology and theory*	Peter Gallmann (Friedrich-Schiller-Universität Jena) Flexivische Lücken bei Sprachbezeichnungen (Keynote)		André Meinunger (Leibniz-Zentrum Allgemeine Sprachwissenschaft ZAS) Auf Beugen und Brechen: Über Finitheit, wo sie eigentlich nichts zu suchen hat
AG 2: Ulrike Domahs, Angela Grimm, Mathias Scharinger, *Weak elements in prosodic acquisition and processing*	Katherine Demuth, Keynote Speaker (Macquarie University) The acquisition of weak elements: Lexical, morphological, and prosodic considerations		Christina Kauschke, Ulrike Domahs & Angela Grimm (Marburg, Freiburg) Schwa syllables in early language acquisition and speech and language disorders
AG 3: Sven Kotowski, Ingo Plag, *The semantics of derivational morphology: Theory, methods, evidence*	Matthias Irmer & Olav Mueller-Reichau The pragmatics of word formation: A case study on German *stoff*	Lea Kawaletz The polysemy of newly derived forms: An investigation of English -*ment* nominalizations	Natascha Elxnath On the interpretation of German A-V-*er*-constructions and the notion of concepts
AG 4: Kristin Kopf, Thilo Weber, *Free variation = unexplained variation? Empirical and theoretical approaches to optionality in grammar*	Chiara Fioravanti (Kiel) Der Abbau 'freier graphematischer Variation' in der Geschichte des Deutschen: Methodische Überlegungen zu einer Korpusuntersuchung	Vilma Symanczyk Joppe (Düsseldorf) Fakultative Verbvalenzen als freie Variation	Maud Westendorp & Björn Lundquist (Tromsö) The presence of light objects affects variable verb and subject placement in North Germanic
AG 5: Hanna Fischer, Melitta Gillmann, Mirjam Schmuck, *Encoding aspectuality in Germanic languages — empirical and theoretical approaches*	Torodd Kinn (invited speaker) Is pseudocoordination an aspectual construction?		Ermenegildo Bidese & Maria Rita Manzini Progressive and prospective in German dialects of Italy

AG 6: Katharina Schaebbicke, Heiko Seeliger, *Empirical approaches to canonical and non-canonical uses of negation*	**Hedde Zeijlstra** (Göttingen) Types of Negative Concord systems		**Cory Bill & Todor Koev** High negation questions are always polarity focused and sometimes contain VERUM
AG 7: Katrin Axel-Tober, Lutz Gunkel, Jutta M. Hartmann, Anke Holler, *On the nouniness of propositional arguments*	**Éva Dékány & Ekaterina Georgieva** (Budapest) Where propositional arguments and participial relative clauses meet		**Kalle Müller** (Tübingen) That relatives! and the relativization of *dass*-clauses in German
AG 8: Cherlon Ussery, Jóhannes Gísli Jónsson, Nicole Dehé, *Ditransitives across languages and frameworks*		**Alina Tigau & Klaus Von Heusinger** [Person] intervention effects with Romanian ditransitive constructions	**Klaus Von Heusinger, Diego Romero Heredero & Marco García García** Verb class and differential object marking in Spanish ditransitive constructions
AG 9: Mingya Liu, Mathias Barthel, *The semantics and pragmatics of conditional connectives*	**Markus Egg & Debopam Das** Signalling conditional relations	**Robert van Rooij & Katrin Schulz** A causal relevance analysis of (hidden) conditionals	**Niels Skovgaard Olsen & Peter Collins** Indicatives, subjunctives, and the falsity of the antecedent
AG 10 (10a) (Kurz-AG): Fabian Schubö, Sabine Zerbian, Sandra Hanne, Isabell Wartenburger, *Prosodic boundary phenomena*	**Naomi Peck, Kirsten Culhane & Maria Vollmer** Comparing cues: a mixed methods study of intonation unit boundaries in three typologically diverse languages	**Sandrien van Ommen, Natalie Boll-Avetisyan, Barbara Höhle & Thierry Nazzi** Prosodic boundaries in phrase processing, a click-detection study	**Nele Ots & Piia Taremaa** Effects of prosody and collocation frequency on language chunking
AG 11 (10b) (Kurz-AG): Daniel Gleim, Marie-Luise Popp, *Edge-asymmetries in morphophonology*			
AG 12 (Kurz-AG): Martina Penke, Judith Schlenter, Elyesa Seidel, *Eye-tracking and language production*	**Arrate Isasi-Isasmendi, Sebastian Sauppe, Caroline Andrews, Monique Flecken, Moritz Daum, Itziar Laka, Martin Meyer & Balthasar Bickel** Extracting event structure at a glance: The role of case during scene apprehension for speaking	**Emiel van den Hoven, F.-Xavier Alario & Audrey Bürki** Using eye-tracking to gauge the effect of phonological dependencies on planning	Discussion and closing remarks

AG 13 (11b) (Kurz-AG): Gerhard Jäger, Johann-Mattis List, *Model and evidence in quantitative comparative linguistics*	Harald Hammarström Language contact in the evolution of linguistic features	Abbie Hantgan-Sonko Partial cognate comparison and pre-settlement history of the Dogon ethnolinguistic group	Philipp Rönchen & Tilo Wiklund Why we need more study of methods, not data, in computational historical linguistics
AG 14 (Kurz-AG): Anja Müller, Katharina Turgay, *Grammatische Modellierung als Grundlage für sprachdidaktische Vermittlung*	Katharina Böhnert Modelle des Sprachwandels im Deutschunterricht: Sprachreflexive und fächerübergreifende Potenzial	Eva Breindl Grammatische Modelle im Unterricht Deutsch als Fremdsprache: Indirekte Evidenzen aus Lernersprachenforschung und linguistischer Lehrwerkanalyse	
AG 15 (12b) (Kurz-AG): Martin Klotz, Anke Lüdeling, Anna Shadrova, *Contrastive corpus methodology for language modeling and analysis*			

11:15-12:45

AG	11:15-11:45	11:45-12:15	12:15-12:45
AG 1: Thomas Strobel, Helmut Weiß, *Grammatical gaps: Definition, typology and theory*	Oliver Schallert (Ludwig-Maximilians-Universität München) Modals between defectiveness and overdifferentiation	Anja Hasse (Universität Zürich) & Patrick Mächler (Universität Zürich) Lücken in der Definitheit im Germanischen	Elisabeth Scherr (Universität Graz) Attraction of the void: The lack of aspect in German and its effect on language change
AG 2: Ulrike Domahs, Angela Grimm, Mathias Scharinger, *Weak elements in prosodic acquisition and processing*	Beat Siebenhaar, Keynote Speaker (Leipzig) Geolinguistic differences of reductions in standard intended German due to a rise of speech rate		Christoph Gabriel, Jonas Grünke & Nils Karsten (Mainz, Amsterdam) Getting rid of the German canonical trochee in L3 French intonation: Comparing monolingually raised German and bilingual Turkish-German learners
AG 3: Sven Kotowski, Ingo Plag, *The semantics of derivational morphology: Theory, methods, evidence*	Marco Marelli The distributional-semantics side of morphologically complex words: Modelling the processing of affixed words in vector spaces		Martin Schäfer Splitting -*ly*'s: Using word embeddings to distinguish derivation and inflection

AG 4: Kristin Kopf, Thilo Weber, *Free variation = unexplained variation? Empirical and theoretical approaches to optionality in grammar*	**Markus Bader** (Frankfurt a. M.) How free is the position of German object pronouns?	**Marek Konopka** (IDS Mannheim) Freie Variation und Fugenelemente: Theorie und korpuslinguistische Realität	**Nathalie Entringer** (Luxembourg) Freie Variation = Einbahnstraße? Konzeptionelle und methodische Überlegungen am Beispiel von morphologischer Variation im Luxemburgischen
AG 5: Hanna Fischer, Melitta Gillmann, Mirjam Schmuck, *Encoding aspectuality in Germanic languages — empirical and theoretical approaches*	**Jens Fleischhauer** The syntactic expression of prospective aspect in German	**Katharina Paul** Go for ingressivity	**Sarah Ihden** Aspectual meanings of the present participle in Middle Low German
AG 6: Katharina Schaebbicke, Heiko Seeliger, *Empirical approaches to canonical and non-canonical uses of negation*	**Beata Gyuris** Hungarian *nem-e* interrogatives: Marking the source of speaker bias	**Ljudmila Geist & Sophie Repp** *Yes* and *no* in responses to negative (biased) questions: Russian vs. German	**Elena Albu, Oksana Tsaregorodtseva & Barbara Kaup** Is negation more difficult than affirmation?
AG 7: Katrin Axel-Tober, Lutz Gunkel, Jutta M. Hartmann, Anke Holler, *On the nouniness of propositional arguments*	**Andreas Blümel & Nobu Goto** (Göttingen, Tokio) Reconsidering the syntax of correlates and propositional arguments	**Nikos Angelopoulos** (Leuven) Nouny clauses: The clausal prolepsis strategy	**Alassane Kiemtoré** (Stuttgart) A syntactic account of clausal complementation in Jula
AG 8: Cherlon Ussery, Jóhannes Gísli Jónsson, Nicole Dehé, *Ditransitives across languages and frameworks*	**Elena Callegari & Anton Karl Ingason** Topicalization: The IO/DO asymmetry in Icelandic	**Johannes Rothert** Investigating person-case effects in Standard German and Swabian	**Jóhannes Gísli Jónsson** Inversion in Icelandic ditransitives
AG 9: Mingya Liu, Mathias Barthel, *The semantics and pragmatics of conditional connectives*	**Mingya Liu & Yuting Wang** Jiu-conditionals in Mandarin Chinese	**Jiyeong Kim & Sung-Eun Lee** Past tense morphology and the choice of connectives in Korean counterfactual conditionals	**Natalia Zevakhina & Veronika Prigorkina** Conditional perfection in causal and conventional conditionals
AG 10 (10a) (Kurz-AG): Fabian Schubö, Sabine Zerbian, Sandra Hanne, Isabell Wartenburger, *Prosodic boundary phenomena*			
AG 11 (10b) (Kurz-AG): Daniel Gleim, Marie-Luise Popp, *Edge-asymmetries in morphophonology*	Marie-Luise Popp & Daniel Gleim (Leipzig) Introduction	**Yuni Kim**, invited speaker (Essex) Morphological symmetry, prosodic asymmetry: The case of Huave mobile affixes	

AG 12 (Kurz-AG): Martina Penke, Judith Schlenter, Elyesa Seidel, *Eye-tracking and language production*			
AG 13 (11b) (Kurz-AG): Gerhard Jäger, Johann-Mattis List, *Model and evidence in quantitative comparative linguistics*	**Erik Elgh** Theoretical (in)compatibilities of the comparative method and cladistics	**Gereon Kaiping & Natalia Chousou-Polydouri** Lexedata: Tying existing software to CLDF Wordlists	**Matías Guzmán Naranjo & Laura Becker** Controlling for geographical, areal, and family biases in typology
AG 14 (Kurz-AG): Anja Müller, Katharina Turgay, *Grammatische Modellierung als Grundlage für sprachdidaktische Vermittlung*			
AG 15 (12b) (Kurz-AG): Martin Klotz, Anke Lüdeling, Anna Shadrova, *Contrastive corpus methodology for language modeling and analysis*	**Anna Shadrova, Martin Klotz & Anke Lüdeling** Linguistic modelling and analysis	**Wander Lowie**, Keynote Speaker (Groningen) The group and the individual: Complementary dimensions of language development	

13:45-14:45

AG	13:45-14:15	14:15-14:45
AG 1: Thomas Strobel, Helmut Weiß, *Grammatical gaps: Definition, typology and theory*	**Ermenegildo Bidese** (University of Trento), **Andrea Padovan** (University of Verona) **& Alessandra Tomaselli** (University of Verona) Circumventing the 'that-trace' effect: Different strategies between Germanic and Romance	**Julia Bacskai-Atkari** (Universität Konstanz) Syntactic paradigms, markedness and similative markers in comparative and relative clauses
AG 2: Ulrike Domahs, Angela Grimm, Mathias Scharinger, *Weak elements in prosodic acquisition and processing*	**Isabelle Franz, Markus Bader** (MPI Frankfurt) & **Gerrit Kentner (**Frankfurt) The influence of rhythm on placing the German object pronoun	**Christina Domene Moreno & Barış Kabak** (Würzburg) What makes grammatical words "weak"? Disentangling semantic, morphosyntactic and prosodic factors via language-music mapping
AG 3: Sven Kotowski, Ingo Plag, *The semantics of derivational morphology: Theory, methods, evidence*	**Matías Guzmán Naranjo & Olivier Bonami** Distributional evidence for derivational paradigms	**Gianina Iordachioaia, Gabriella Lapesa, Sarina Meyer & Sebastian Pado** Difference of first attestation dates as evidence for directionality in zero derivation

AG 4: Kristin Kopf, Thilo Weber, *Free variation = unexplained variation? Empirical and theoretical approaches to optionality in grammar*	**Roser Giménez** (Barcelona) Verbal periphrases, deontic modality and teenagers: Free variation in non-standard spoken Catalan?	**Göz Kaufmann & Daniel Duran** (Freiburg) Von snoidel'n und vom hofdüütsch'en: Zur phonetischen Variation im Pomerano
AG 5: Hanna Fischer, Melitta Gillmann, Mirjam Schmuck, *Encoding aspectuality in Germanic languages — empirical and theoretical approaches*	**Sophie Ellsäßer** Temporal adverbs as aspectuality markers? On the grammaticalization of *als* and *viel* in German substandard varieties	**Lena Schmidtkunz** "Wi wir am leben in alle plantation": The aspect system in Unserdeutsch (Rabaul Creole German)
AG 6: Katharina Schaebbicke, Heiko Seeliger, *Empirical approaches to canonical and non-canonical uses of negation*	**Marta Tagliani** Slow and steady wins the race: Positive effects of the negated information on negative sentence comprehension in Italian dyslexic adults	**Sumrah Arshad** How negative concord licenses the acquisition of formal negation
AG 7: Katrin Axel-Tober, Lutz Gunkel, Jutta M. Hartmann, Anke Holler, *On the nouniness of propositional arguments*	**Imke Driemel & Maria Kouneli** (Leipzig) Verb-y and noun-y complementation in Kipsigis	**Vesela Simeonova** (Tübingen) Definitely factive
AG 8: Cherlon Ussery, Jóhannes Gísli Jónsson, Nicole Dehé, *Ditransitives across languages and frameworks*	**Ana Regina Calindro & Maria Aparecida Torres Morais** Preposition reanalyzes and ditransitive sentences in Brazilian Portuguese	**Gary Thoms** On the derivation of prepositional dative constructions in Irish and Gaelic
AG 9: Mingya Liu, Mathias Barthel, *The semantics and pragmatics of conditional connectives*	**Maria Cristina Lo Baido, Egle Mocciaro & Luisa Brucale** Conditional connection explored: The case of Sicilian cusà	**Laura Margarita Merino Hernández** Conditional constructions in Spanish: Overt connectives, ellipsis, and juxtaposition
AG 10 (10a) (Kurz-AG): Fabian Schubö, Sabine Zerbian, Sandra Hanne, Isabell Wartenburger, *Prosodic boundary phenomena*		
AG 11 (10b) (Kurz-AG): Daniel Gleim, Marie-Luise Popp, *Edge-asymmetries in morphophonology*	**Thomas Schwaiger** (Graz) The suffixing preference and the edge-asymmetry in reversal in reduplication	**Noah Elkins** (UCLA) Prefix independence as root-initial percept maximization

AG 12 (Kurz-AG): Martina Penke, Judith Schlenter, Elyesa Seidel, *Eye-tracking and language production*		
AG 13 (11b) (Kurz-AG): Gerhard Jäger, Johann-Mattis List, *Model and evidence in quantitative comparative linguistics*	**Annemarie Verkerk, Hannah Haynie, Russell Gray, Simon Greenhill, Olena Shcherbakova & Hedvig Skirgård** Revisiting typological universals with Grambank	**Johannes Dellert** Towards richer multi-source machine-readable etymologies
AG 14 (Kurz-AG): Anja Müller, Katharina Turgay, *Grammatische Modellierung als Grundlage für sprachdidaktische Vermittlung*		
AG 15 (12b) (Kurz-AG): Martin Klotz, Anke Lüdeling, Anna Shadrova, *Contrastive corpus methodology for language modeling and analysis*	**Natalia Levshina** A comparison of frequentist and Bayesian models of language variation: The problems of priors and sample size	

Freitag/Friday, 26.02.2021

11:45-14:15

AG	11:45-12:15	12:15-12:45	12:45-13:15	13:15-13:45	13:45-14:15
AG 1: Thomas Strobel, Helmut Weiß, *Grammatical gaps: Definition, typology and theory*	**Ralf Vogel**, Bielefeld Empirical determinants of grammatical gaps and grammatical inventions (Keynote)	**Fenna Bergsma,** Goethe-Universität Frankfurt A typology of case competition in headless relatives		**Ewa Trutkowski**, Bozen How sex and gender shape agreement in German relative clauses	**Kerstin Hoge**, Oxford Cross-Germanic variation in relative clauses with pronominal antecedents
AG 2: Ulrike Domahs, Angela Grimm, Mathias Scharinger, *Weak elements in prosodic acquisition and processing*					
AG 3: Sven Kotowski, Ingo Plag, *The semantics of derivational morphology: Theory, methods, evidence*					
AG 4: Kristin Kopf, Thilo Weber, *Free variation = unexplained variation? Empirical and theoretical approaches to optionality in grammar*	**Mathilde Hutin, Ioana Vasilescu, Lori Lamel, Yaru Wu** (Paris-Saclay), **Martine Adda-Decker** (Paris 3 Sorbonne-Nouvelle) **& Adèle Jatteau** (Lille) Modelling the realization of variable word-final schwa in Standard French	**Anja Hasse** (Zürich) Zur Stabilität flexionsmorphologischer Variation: Die Dativformen des unbestimmten Artikels im Zürichdeutschen	**Noah Diewald** (Ohio State University) Overabundance and the interface	**Tania Paciaroni** (Zürich) Paradigm splits across parts of speech	**Kristin Kopf & Thilo Weber** (IDS Mannheim) Abschlussdiskussion

AG 5: Hanna Fischer, Melitta Gillmann, Mirjam Schmuck, *Encoding aspectuality in Germanic languages — empirical and theoretical approaches*	Fabian Fleißner Non-encoding aspectuality in Old High German, or: Why are we failing?	Stephanie Hackert, Robert Mailhammer & Elena Smirnova Perfect constructions in English and German: Typologies and diachronic implications	Kathrin Weber Auxiliary variation in the aspect-tense system of Low German speakers	Katharina Zaychenko The influence of grammatical and non-linguistic factors on motion event descriptions: A cross-linguistic study	Hanna Fischer, Melitta Gillmann & Mirjam Schmuck Final discussion: Exploring new perspectives on aspectuality in Germanic languages
AG 6: Katharina Schaebbicke, Heiko Seeliger, *Empirical approaches to canonical and non-canonical uses of negation*		Henrik Torgersen Initial negation in Norwegian: A curious case of licensing	Giuseppe Magistro The integration of acceptability tests into diachronic syntax: The case of presuppositional negation	Chloé Tahar Expletive negation: From embedded speech-acts to embedded propositions	Katharina Schaebbicke & Heiko Seeliger Exploring the landscape of German polarity items and their licensing conditions
AG 7: Katrin Axel-Tober, Lutz Gunkel, Jutta M. Hartmann, Anke Holler, *On the nouniness of propositional arguments*	Richard Faure (Nizza) From D to N, CPs as nominals in Greek	Jürgen Pafel (Stuttgart) Argument clauses and definite descriptions	Frank Sode (Frankfurt/M.) On the nouniness of V2-clauses under preference predicates	Patrick D. Elliott, Invited Speaker (Cambridge, MA) Objects of attitude ascriptions	
AG 8: Cherlon Ussery, Jóhannes Gísli Jónsson, Nicole Dehé, *Ditransitives across languages and frameworks*		Matthew Tyler Thematic role and movement to subject position: Muskogean evidence for a 'deactivation'-based account	Einar Freyr Sigurðsson & Jim Wood High applicatives in Icelandic adjectival constructions	Kevin Kwong Null/deleted prepositions and the illusion of double object constructions in Cantonese	Breanna Pratley & Philip Monahan Can English idioms undergo the dative alternation? A priming investigation
AG 9: Mingya Liu, Mathias Barthel, *The semantics and pragmatics of conditional connectives*	Patrick Grosz, invited talk Emojis and conditionals: Exploring the super linguistic interplay of expressive modifiers and conditional meaning		Nina Haslinger & Viola Schmitt What embedded counterfactuals tell us about the semantics of attitudes	Magdalena Sztencel & Sarah E. Duffy Modals as a diagnostic for biconditional vs. material interpretations of conditionals	Final discussion

AG 10 (10a) (Kurz-AG): Fabian Schubö, Sabine Zerbian, Sandra Hanne, Isabell Wartenburger, *Prosodic boundary phenomena*					
AG 11 (10b) (Kurz-AG): Daniel Gleim, Marie-Luise Popp, *Edge-asymmetries in morphophonology*	**Alexander Martin & Jennifer Culbertson** (Edinburgh) A domain-general bias cannot explain the suffixing preference: Experimental evidence from English and Kîîtharaka	**Xinyi Wang & Itamar Kastner** (Edinburgh) The suffixation preference: Native language and information load in artificial language learning	**Jochen Trommer** (Leipzig) Tonal affixes and the status of autosegmental association conventions	**Daniel Gleim & Sören Tebay** (Leipzig) Edge-biases in mutation	**Ronald P. Schaefer** (Edwardsville, Illinois) **& Francis O. Egbokhare** (Ibadan) Tonal asymmetry for tense-aspect at verbal phrase edges
AG 12 (Kurz-AG): Martina Penke, Judith Schlenter, Elyesa Seidel, *Eye-tracking and language production*					
AG 13 (11b) (Kurz-AG): Gerhard Jäger, Johann-Mattis List, *Model and evidence in quantitative comparative linguistics*	**Verena Blaschke & Johannes Dellert** Correlating borrowing events across concepts to derive a data-driven source of evidence for loanword etymologies	**Miri Mertner** Modelling linguistic data in space using autologistic regression	**Viktor Martinovic** Loanpy: A framework for computer-aided borrowing detection	**Johannes Wahle** The effect of priors on tree topologies	**Final discussion**
AG 14 (Kurz-AG): Anja Müller, Katharina Turgay, *Grammatische Modellierung als Grundlage für sprachdidaktische Vermittlung*					

| AG 15 (12b) (Kurz-AG): Martin Klotz, Anke Lüdeling, Anna Shadrova, *Contrastive corpus methodology for language modeling and analysis* | **Felix Bildhauer, Elisabeth Pankratz & Roland Schäfer** Corpora, inference, and models of register distributions | **Christof Schöch, Julia Dudar, Cora Rok & Keli Du** Deviation of proportions as the basis for a keyness measure | **Giuseppe Samo** Machine learning and syntactic theory: Focus on German and German varieties | **Abschlussdiskussion & Abschied** |

Ersatzsprecher/Alternates:

AG 1: Thomas Strobel, Helmut Weiß, *Grammatical gaps: Definition, typology and theory*	**Sebastian Fedden** (Université Sorbonne Nouvelle – Paris 3) Morphological gaps and syntax: Agreement in Mian discourse	**Ekaterina Levina** (University of Texas at Austin) Doubled possessors: One gap filled twice	
AG 4: Kristin Kopf, Thilo Weber, *Free variation = unexplained variation? Empirical and theoretical approaches to optionality in grammar*	**Yidong Yu** (Göttingen) Optionality and Categorial properties: the case of optional plural marking in Yucatec Maya	**Jirayu Tharincharoen** (Erlangen-Nürnberg) Eigenschaften der syntaktischen Allostruktionen: Am Beispiel des deutschen *je-desto*-Gefüges	**Jakob Maché** (Universidade de Lisboa) Modelling free variation of linking elements after feminine noun stems in German
AG 7: Katrin Axel-Tober, Lutz Gunkel, Jutta M. Hartmann, Anke Holler, *On the nouniness of propositional arguments*	**Patrick Brandt** (Mannheim) The transfer of nominal (ordinary individual) to propositional (phenomenal individual) properties in German particle verb constructions	**Jan Wiślicki** (Warschau) S-selection and presupposition in quotational complementation	

AG 11 (10b) (Kurz-AG): Daniel Gleim, Marie-Luise Popp, *Edge-asymmetries in morphophonology*	**Marie-Luise Popp** (Uni Leipzig) Edge-asymmetries in affix order		
AG 14 (12a) (Kurz-AG): Anja Müller, Katharina Turgay, *Grammatische Modellierung als Grundlage für sprachdidaktische Vermittlung*	**Iris Rautenberg**: Zum Umgang mit syntaktischen Strukturen und Prozeduren bei orthographischen Entscheidungen: Ergebnisse einer explorativen Interviewanalyse		

Arbeitsgruppe 1

Grammatical gaps: Definition, typology and theory

Thomas Strobel & Helmut Weiß

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

The role of stress and metrical foot in ineffability in German

Caroline Féry
National and Kapodistrian University Athens
caroline.fery@gmail.com

In Fanselow & Féry (2002), we wished to identify a typology of ineffabilities that helps to understand in which domains of language ineffability arises. In this talk I will concentrate on the prosodic structure of some morphological processes and show that we also need a typology of ineffabilities in this tiny corner of grammar.

It is well-known that German has a 'perfect' prosodic word of the size of a syllabic (or optionally moraic) trochaic foot: a trochee has a strong-weak rhythmic structure. A moraic trochee consists of one syllable with at least two moras (preferably three). Various morphological processes aspire to achieve this trochaic form, and some processes even require it. So-called i- formations (1), suffixation of diminutive *-chen* or *-lein* (2) with umlaut, and reduplications (3), see Kentner (2017), are restricted to syllabic trochees.

(1) a. Well-formed i-formations: Fabian → Fabi, Andreas → Andi, Westdeutscher → Wessi
   b. Ill-formed i-formations: Wilhelm → *Wilhi, Gabriel → *Gabri, Ulrike → *Ulri
(2) a. Well-formed diminutives: Jahr → Jährchen 'year.DIM.' Bruder → Brüderchen 'brother'
   b. Ill-formed diminutives: Európa → *Europächen 'Europe, DIM.'
(3) a. Well-formed reduplications: Hinkepinke (<hink), Wirrwarr (<wirr), Mischmasch(<misch)
   b. Ill-formed i- reduplications: Ivonne [i'vɔn] → *Ivonnepivonne, *Gerhardperhard

By contrast, more mundane morphological operations, such as infinitives (4) and some derivational affixes, such as *-ig* in (5), have a strong preference for the trochaic pattern. The infinitive *bauen* is disyllabic even though the monosyllabic form *baun* would be well-formed, and in trochaic *segeln*, it is [l] that is syllabic and not the suffix [n], and also not both of them, as in Dutch. If trochaic form is impossible, as in *arbeiten*, the infinitive is formed anyway. Two highly frequent monosyllabic, and thus irregular, infinitives (*sein* 'to be', *tun* 'to do') are tolerated. The examples in (4a) show that German morphological operations restructure some words if this leads to a trochaic structure.

(4)   a. Trochaic infinitives: bauen (*baun) 'to build', segeln (*segelen, *seglen) 'to sail'
      b. Non-trochaic infinitives: arbeiten 'to work', sein 'to be'
(5)   a. Derivation with suffix -ig: Sonne 'sun' → sonnig 'sunny'
      b. No derivation with -ig: Kámera → ?kameraig
(6)   Nominal plural: Drama/Dramen, Méntor/Mentóren

Optimality Theory cannot account for all these cases in a unified way even though it can identify the trochee as the best prosodic word (see McCarthy & Prince 1993). It should always be possible to create a disyllabic trochee from a sequence of segments organized in syllables, for instance at the price of deleting or reorganizing segments. However, it is not what is observed. In the cases described below some formations are banned for a variety of reasons.

A distinction must be made between morphological operations such as infinitives that are needed whatever their prosodic form is and 'superfluous" i-formations or reduplications that can be dispensed if they cannot achieve prosodic well-formedness. There is thus a fundamental difference between prohibitions of ill-formed words in (1) to (3) and the tolerance of prosodically ill-formed words in (4) and (5). The difficulties for OT arise as a result of the different explanations needed to explain the gaps. What goes wrong in the cases (1) to (5) is dependent of the form itself. Finally, the kind of repairs leading to acceptable forms: deletions of consonants in (1), accent shift in (1) and (2), no repair at all in (3), change in syllabicity of sonorants in (4) and 'blocking' phenomena in (5) are impossible to account for in a uniform OT model.

References Fanselow, Gisbert & Caroline Féry. (2002) Ineffability in Grammar In: Resolving Conflicts in Grammar: Optimality Theory in Syntax, Morphology, and Phonology. S*pecial Issue 11 of Linguistische Berichte.* Kentner, Gerrit. 2017. On the emergence of reduplication in German morphophonology. *Zeitschrift für Sprachwissenschaft* 36(2). 233–277. McCarthy, John J. & Alan S. Prince (1993) Generalized Alignment. In Geert Booij & Jaap van Marle (eds.) *Yearbook of Morphology*. Dordrecht: Kluwer. 79-153.

Restrictions on mixed gender agreement in Russian: Feature conflicts and ineffability in DM

Mariia Privizentseva
Leipzig University
mprivizentseva@uni-leipzig.de

On the basis on two novel empirical arguments, I claim that restrictions on mixed gender agreement in Russian are due to the conflicting feature specifications on a noun that cannot be implemented by the morphological component. The phenomenon provides an instance of a grammatical gap arising from properties of the paradigm.

In Russian, some morphologically masculine nouns trigger feminine agreement if a referent is female. This is allowed in the singular nominative (1) and in the plural (2) but not in oblique singular forms (3); see Pesetsky (2013), Gerasimova (2019), *i.a.*

(1) xoroš-yj/aja vrač   (2) ob-o/e-im   vrač-am   (3) xoroš-emu/*ej vrač-u
     good-M/F   doctor   both-M/F-PL.DAT   doctor-PL.DAT   good-M.DAT/*F doctor-DAT

I provide two novel arguments showing that case number restrictions stem from inflection on the noun. First, nouns with mixed agreement belong to class I that has only masculine nouns, and feminine agreement is restricted to forms where nominal exponents are syncretic to class III that includes feminine nouns. Second, restrictions don't hold under ellipsis (4). Given that ellipsis is absence of Vocabulary Insertion (Merchant 2001), this shows that insertion of a nominal form causes ungrammaticality.

(4) Ja   pojdu tol'ko   k   xoroš-ej [_].
    I   will.go only   to   good-F.DAT
    {Discussing doctors...} I will go only to a good one (f.).'

The number case restrictions can be derived if declension is decomposed into gender ([±fem]) and an idiosyncratic feature of a lexical item ([±α]); see, e.g., Roca (1989), Harris (1991), Wiese (2004), and Caha (2019) for declension exponence targeting gender directly. Hybrid nouns have [−fem][+α] declension features and also [+fem] gender if they denote a female. Following Schütze (2003), Coon and Keine (2020), *i.a.*, contradictory features on one node are tolerated by syntax but problematic for Vocabulary Insertion. The conflict can be resolved only by a syncretic form underspecified for the contradicting features. Thus, semantic agreement is allowed only if a vocabulary item is underspecified for gender and compatible with [+α]. The structure is ineffable in other cases because the inserted exponent is incompatible either with the grammatical or with the semantic gender.

References: Caha, Pavel. 2019. *Case competition in Nanosyntax: A study of numerals in Ossetic and Russian.* Language Science Press: Berlin. Coon, Jessica, and Stefan Keine. 2020. "Feature Gluttony." *Linguistic Inquiry* 2(1): 1–82. Gerasimova, Anastasija A. 2019. *Var'irovanie soglasovatel'nyx xarakteristik v russkoj umennoj gruppe.* Master's thesis. Moscow State University. Harris, James. 1991. "The Exponence of Gender in Spanish." *Linguistic Inquiry* 22(1): 27–62. Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis.* Oxford: Oxford University Press. Roca, Iggy. 1989. "The organization of grammatical gender." *Transactions of the Philological Society* 87(1): 1–32. Pesetsky, David. 2013. *Russian Case Morphology and the Syntactic Categories.* The MIT Press: Cambridge, MA. Schütze, Carson. 2003. "Syncretism and double agreement with Icelandic nominative objects." In Lars-Olof Delsing, Cecilia Falk, Gunlög Josefsson, and Halldór Ármann Sigurðsson, eds. *Grammar in focus: Festschrift for Christer Platzack*. Department of Scandinavian Languages: Lund. 295–303. Wiese, Bernd. 2004. "Categories and Paradigms. On Underspecification in Russian Declension." In Gereon Müller, Lutz Gunkel, and Gisela Zifonun, eds. *Explorations in Nominal Inflection.* Berlin: Mouton. 321–372.

Filling the gap: In defense of periphrastic forms as cells in paradigms

Andreas Blümel
University of Göttingen
ablueme@gwdg.de

Within generative grammar the relationship between underlying syntactic structures and morphological expression is characterized for the most part by the following matches and mismatches: Morphological expression corresponds to the underlying syntax in that a word- like expression expresses a single syntactic head (i)+(iii), or diverges from its underlying syntax in that a word-like expression corresponds to two distinct syntactic heads (ii). (The analysis of finite synthetic verb forms (iii) follows Haider 2010: chapter 2 in denying that G(erman) has V- to-T.)

i) I) [$_{TP}$ *John* T$_{[3rd\ Sg]}$=*has* [$_{VP}$ *slept*]] → *Johnhas slept.*   ***English present perfect***

ii) [$_{TP}$ *John* T$_{[3rd\ Sg]}$=*-s* [$_{VP}$ *sleep*]] →*John sleeps.*   ***English affix hopping***

iii) *weil* [$_{VP}$ *Johannes* V$_{[3rd\ Sg]}$=*schläft*]   ***German synthetic finite verbs***
    since John   sleeps

The combinatorics between the two syntactic options and morphological realizational options highlights an important gap: The periphrastic expression of a synthetic syntax. To fill this gap, this paper follows the view that periphrastic verb forms in languages like Dutch and G can "occupy cells in morphological paradigms" (and Zwart 2017:29; cf. also Ackerman & Webelhuth 1998 *i.a.*). On the syntactic side, it recasts Bayer & Kornfilt's (1994) view of the G verb cluster as follows: T is a syntactic affix –

Morph \ Syntax	**Periphrasis**	**Synthesis**
Periphrasis	English auxiliary verbs	? → German verb cluster
Synthesis	English affix hopping	German simple verbs

not a free standing head like E T –, as is the verbal categorizer *v*. Together, they form an amalgam ⟨*v*, T⟩=INFL by External Pair Merge (EKS 2016). INFL Set Merges with the structure {(DP), R}, where R=Root, giving {{(DP), R}, INFL}, the DP being the internal argument IA. Since within INFL, T is affixed to *v*, ϑ-marking of the external argument EA can proceed in the standard fashion by *v*. Being affixal, INFL forces raising of R (cf. Chomsky 2015:9 on *v*), resulting in the structure {EA, {{(IA), R}, ⟨R, INFL⟩}}, where ⟨*v*, T⟩ is affixed to the host R. Thus G has a syntactically synthetic verbal complex, unlike E with its syntactically analytical verbal region [$_{TP}$ T [$_{vP}$ *v* [R ...]]]. This naturally captures (a) the elusive absence of VP-ellipsis in G in that T is not a free standing morpheme to license it and (b) *all finite verbs* raise to C in root contexts in G, whereas *only finite auxiliary verbs* raise to C in E. Problems dissolve of accounting for why extraposed CPs in G adjoin to VP, forming [$_{VP}$ [$_{vP}$ ...t_{CP}... V] CP], as evidenced by VP-fronting, but cannot surface between sentence-final V and the head of a (putative) TP-projection (cf. Haider 2010:61-63/67- 68; *pace* Wurmbrand & Bobaljik 2005). The labeling algorithm LA (Chomsky 2013) finds the amalgam ⟨R, INFL⟩ and determines it to be the label in {{(IA), R}, ⟨R, INFL⟩}, i.e. that set is a ⟨R, INFL⟩. A suggestive idea is that this categorical difference between the verbal phrase in G and E allows subjects to remain within it in G, while EPP-raising if forced in E (cf. Chomsky 2013). In this sense, the current approach contributes to, if not quite offers, an analysis of G which derives these properties from elementary principles. Other contrasts will be explored, like the possibility of scrambling, the absence of *that*-trace effects in G, and their impossibility in E, as well as suggestions on V-to- C in root contexts (V1/V2) in G.

References: EKS/Epstein, S., H. Kitahara & D. Seely (2016) Phase cancellation by External Pair Merge of heads. *The Linguistic Review* 33(1), 87-102. Haider, H. (2010) *The Syntax of German*, CUP. Zwart, J-W. (2017) An argument against the syntactic nature of verb movement. In *Order and structure in syntax 1: Word order and syntactic structure* (pp. 29-47). (Open Generative Syntax; Vol. 1). Berlin. Language Science Press.

How do grammars leak? A close look at the idea of syncretism as repair for defectiveness

Andrea D. Sims
The Ohio State University
sims.120@osu.edu

In the call for papers for this workshop, the organizers cite the famous quote by Sapir that 'all grammars leak,' by which Sapir meant that grammars are not neat and tidy and are instead prone to irregularities around the edges. In this talk I probe the relationship between inflectional defectiveness and syncretism. There is a tendency to view defectiveness as a way in which grammars leak and syncretism as a way to repair such a leak. In other words, defectiveness breaks grammatical functioning and syncretism restores it. I argue that this framing is too simplistic, at least for inflectional morphology. A close look reveals that while some cases can be described in this way, there is no *general* sense in which syncretism acts as a repair for paradigmatic gaps (instances of inflectional defectiveness). Drawing examples from a range of languages, I show that many examples of defectiveness-syncretism interactions do not reflect this dynamic. Instead, the range of interactions found is what is expected in general for inflectional formatives with intersecting distributions. I explore the implications of this fact, in particular the idea that defectiveness piggybacks on the regular functioning of a language's inflectional system. This contrasts with a general intuition that paradigmatic gaps are anomalous to the regular functioning of inflection.

It is intuitively appealing to view syncretism as a repair for defectiveness. After all, when word structure fails speakers still need to convey the meaning of the word *somehow*. Recruiting a wordform from another cell in the same paradigm is a natural strategy (Baerman 2004, Mithun 2010, Rice 2005). For example, Mithun (2010) observes that in Yup'ik, nouns that are defective in singular number morphology extend the dual form, using it in singular.

The result is singular-dual syncretism. Some formal models enshrine this relationship into analyses of syncretism, positing that when underlying representations are defective, syncretism emerges in the context of a requirement that all licensed sets of morphosyntactic values be expressed (Calabrese 2011, Müller 2011, Wunderlich 2001). This explains syncretism in terms of the need to repair defectiveness.

At the same time, largely overlooked is that interactions between syncretism and defectiveness are not limited to syncretism repairing defectiveness. Stump (2010) identifies three kinds of interaction: defectiveness following the distribution of syncretism, defectiveness overriding syncretism, and syncretism overriding defectiveness. These are shown schematically in (1)-(3), respectively.
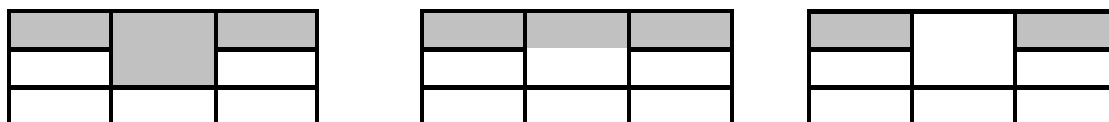

Figure: Schematic representations of paradigms showing possible distributions when syncretism and defectiveness have intersecting domains. Merged cells represent independently motivated syncretisms; shaded cells are defective.

Pattern (3) can be construed as syncretism repairing defectiveness. The others are initially surprising: Why should defectiveness spread beyond its distributional domain to encompass a syncretic form with an intersecting domain, as schematized in (1)? Why should *only one* of two syncretic cells be defective, as in (2)? If syncretism is a natural repair for defectiveness, we should expect the pattern in (3) in both cases. Yet examples of such interactions are scattered across the theoretical literature. In this talk I attempt to pull them together into a coherent picture, with examples from English, French, German, Greek, Icelandic, Italian, Itel'men, Mohawk, Romansh, Russian, and Spanish. I build on Stump's analysis to show that the interactions can be understood

as leaks at different points in the grammatical 'pipeline'. Ultimately, looking at the full range of ways in which syncretism and defectiveness interact leads to a view that while paradigmatic gaps represent places in which grammars are leaky, it is not clear that the leaks are inherently in need of repair, nor that syncretism often serves this purpose, despite conventional wisdom to this effect. At the same time, interactions between syncretism and defectiveness offer a window into *how* grammars leak, lending insight into those aspects of the grammar that linguists tend to find more pleasingly neat and tidy.

What counts as a gap? The case of typological hierarchies

Tabea Reiner
LMU Munich tabea.reiner@lmu.de

The talk addresses the question whether cut-off points on typological hierarchies represent the kind of gaps with which realistic meta theory is concerned. As an example, relativizability according to the Accessibility Hierarchy (AH, Keenan & Comrie 1977) will be discussed. The hierarchy is shown in (1).

(1) SU > DO > IO > OBL > GEN > OCOMP
read: if a language can relativize one of these positions it can also relativize all positions to the left

Like many other languages, Standard German does not cover the entire AH (Keenan & Comrie 1977:74–79). More precisely, it has a cut-off point before OCOMP, i.e. (2) is ungrammatical.

(2) *der   Mann, als   der   John   größer ist
      the   man   than   REL   John   taller is
     intended: 'the man who John is taller than'

If such cut-off points are viewed as gaps in the sense of realistic meta theory, examples like (2) are expected to evoke speaker uncertainty rather than plain ungrammaticality judgements. However, cut-off points on typological hierarchies represent definitional lacunae: a certain type of clause formation does not apply to a certain type of phrase. As such, they do make prime candidates for gaps in the sense of realistic meta theory.

Against this background, the more specific questions to be addressed in the talk include (but are not limited to):

- Are these gaps just "too" systematic? What does this mean?
- (How) Are they different from those gaps that Strobel & Weiß (2019) mention as systematic (p. 115) or from those they mention as arbitrary, yet explainable (p. 111–112)?
- Can we "promote" such gaps to the level of typological description, not granting them a place in individual mental grammars at all (Newmeyer 2004)? But, then, how can we capture the fact that speakers do have the relevant knowledge?

References: Keenan, Edward L., and Bernard Comrie. 1977. "Noun phrase accessibility and Universal Grammar". *Linguistic Inquiry* 8(1): 63–99. Newmeyer, Frederick J. 2004. "Typological evidence and Universal Grammar". *Studies in Language* 28(3): 527–548. Strobel, Thomas, and Helmut Weiß. 2019. "Von sprachlichen und nicht-sprachlichen Lücken". In Gerrit Kentner, Frank Kügler, Beata Moskal, and Stavros Skopeteas, eds. *Tweets 'n greets for Caroline Féry on the occasion of her farewell from Goethe- University Frankfurt*. Frankfurt. 111–117.

Abbreviations

AH	=	Accessibility Hierarchy
DO	=	direct object
GEN	=	genitive/possessor
IO	=	indirect object
OBL	=	major oblique case noun phrase
OCOMP	=	object of comparison
REL	=	relative pronoun
SU	=	subject

Flexivische Lücken bei Sprachbezeichnungen

Peter Gallmann
Friedrich-Schiller-Universität Jena
peter.gallmann@uni-jena.de

Zu den Besonderheiten der deutschen Nominalphrase (NP) gehört das Phänomen, dass die Flexion der einzelnen Bestandteile (Substantiv, Determinierer, Adjektive) von der NP- internen Struktur mitbeeinflusst wird. Dazu gehört die Wahl von starken und schwachen Endungen bei Adjektiven, aber auch Phänomene im substantivischen Bereich wie die Unterlassung der Kasusflexion in bestimmten Konfigurationen. Gewöhnlich wird davon ausgegangen, dass das Auftreten der „schwach" genannten Flexionsweise bei attributiven und substantivierten Adjektiven syntaktisch gesteuert ist, bei substantivischen Lexemen hingegen lexikalisch. Nun scheint aber auch bei Substantiven Syntaktisches zunehmend eine Rolle zu spielen. Und umgekehrt zeigen sich auch bei substantivierten Adjektiven lexikalische Besonderheiten.

Der Fokus des Vortrags wird auf Erscheinungen in einer Grauzone zwischen Adjektiv und Substantiv liegen, nämlich auf zwei Typen von Farb- und Sprachbezeichnungen: solchen mit adjektivischer und solchen mit substantivischer Flexion. Die beiden unterscheiden sich nicht nur leicht in der Semantik, sondern auch im syntaktischen Verhalten. Die Farbbezeichnungen sind syntaktisch unauffällig:

i.   In unserem Garten haben wir viel Grünes.
ii.   In unserem Garten haben wir viel Grün.
iii. In unserem Garten pflegen wir das Grüne.
iv.   In unserem Garten pflegen wir das Grün.

Bei den analogen Sprachbezeichnungen gibt es aber einige merkwürdige Lücken, wie die folgenden Beispiele zeigen.

v.   Das Deutsche (Englische, Spanische) ist eine dynamische Sprache.
vi.   Das *Deutsch (*Englisch, *Spanisch) ist eine dynamische Sprache.
vii. *Deutsches (*Englisches, *Spanisches) ist eine dynamische Sprache.
viii. Deutsch (Englisch, Spanisch) ist eine dynamische Sprache.

Dazu würde man gern mehr wissen. Es ist aber noch offen, ob sich der Vortrag auf die genaue Deskription der Phänomene beschränken muss oder ob sich auch schon Möglichkeiten der Explanation aufzeigen lassen.

References: Grauwe, Luc de. 2000. "Die Deklination des 'schwierigen' Deutsch(en). Zu einer defizienten Regel in den Grammatiken des Gegenwartsdeutschen." In: *Germanistische Mitteilungen* 52: 99–114. Wöllstein, Angelika, ed. 2016. *Duden. Die Grammatik.* Berlin: Dudenverlag.

Auf Beugen und Brechen: Über Finitheit, wo sie eigentlich nichts zu suchen hat

André Meinunger
Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS)
meinunger@leibniz-zas.de

Der Vortrag soll zwei „reale" Strukturen im Deutschen vorstellen und besprechen, die es so nicht geben sollte – möglicherweise „echte" Illusionen. Der Begriff „grammatische Illusion" wurde vor allem durch Hubert Haider (2011) populär gemacht. Angeregt durch Marga Reis' (1979) entdeckte „Stirnhorn-Illusion" wurden Verbal-Komplex-Konstruktionen wie (1) oder (2) zum Paradefall:

(1)   (Ich) hoffe, geholfen haben zu können.

(2)   Eine Pariserin namens Dimanche soll sich ein gewaltiges Stirnhorn operativ entfernt haben lassen.

Haider beschreibt ausführlich, was in diesen Verbalkomplexen seiner Meinung nach schief- läuft. Er schlägt vor, dass hier Notlösungen bzw. Reparatur-Konstrukte vorliegen, die zwar akzeptabel, aber nicht grammatikalisch wohlgeformt sind. Regeln, die diese Strukturen produzieren (würden), müss(t)en sich widersprechen. Vogel (2009) und Wurmbrand (2012) argumentieren dagegen und beschreiben diese Cluster als regelkonform gebildete Strukturen – unter Einbezug von Ideen wie Regel-Ordnung, im Sinne von Anwendungs- Hierarchie und Struktur-Sensitivität. In der Tat wäre in einer strengen Auslegung der Haider'schen Erklärung auch schon der „normale" Ersatzinfinitiv das Ergebnis konfligierender Regeln und somit ein Illusionskandidat. Aber auch unter einem erweiterten derivationellen Ansatz (à la Optimalität (Vogel) bzw. erlaubter minimaler Kontextsensitivität (Wurmbrand)) bleiben folgende zwei Strukturen (3) und (4) mysteriös.

(3)   ... habe keine kreative Phase im Moment! Bin froh, <u>wenigstens diesen Teil habe schreiben zu können</u>!

(3) und weitere erhobene und vorzustellende Belege sollten ungrammatisch sein, denn es handelt sich dabei um Infinitivkonstruktionen: Es gibt kein realisiertes Subjekt, es gibt keine subordinierende Konjunktion und es erscheint der Marker ‚zu'. Dennoch enthält die rechte Satzklammer einen finiten Bestandteil: ‚habe'. In Kombination sind diese grammatischen Merkmale in einem Teilsatz unvereinbar.

Die zweite Struktur ist ebenfalls „finit", wo Subjektkongruenz, also Finitheit, ausgeschlossen ist (Dang 2016, Gallmann 2018).

(4)   dass der von Alessandro <u>hätte ausgehen müssende</u> Wolfsburger Kombinationsfußball nicht zusande kam

Ein solches pränominales Partizipialattribut kann keine eindeutig finite Form ‚hätte' enthalten. Dennoch lassen sich solche Strukturen nachweisen und „verteidigen". Der Beitrag soll weitere Belege vorstellen, Analysevorschläge (Nicht-Indikativität als Semi-Finitheit; Potential der Oberfeld-Bildung) machen und zur gemeinsamen Diskussion zum Verhältnis von Grammatikalität, Akzeptabilität und Sprachwirklichkeit anregen.

References: Dang, T. T. H. 2016*. Das Partizipialattribut im Deutschen zwischen System und Norm.* Zur System konformität von PII + habend (Germanistische Linguistik 304). De Gruyter. Gallmann, P. 2018. Thi Thu Hien Dang: Das Partizipialattribut im Deutschen zwischen System und Norm. Zur Systemkonformität von 'PII + habend'. *ZS f. Rezensionen zur germ. Sprawiss*. 10, 1-2. 79-85. Haider, H. 2011. Grammatische Illusionen. *Zeitschrift für Sprachwissenschaft* 30: 223–257. Reis, M. 1979. Ansätze zu einer realistischen Grammatik. In: Grubmüller et al. (Hgg.): *Befund und Bedeutung.* Tübingen: Max Niemeyer Verlag, 1–21. Vogel, R. 2009. Skandal im Verbkomplex: Betrachtungen zur scheinbar inkorrekten Morphologie in infiniten Verbkomplexen des Deutschen*. Zeitschrift für Sprachwissenschaft* 28.2: 307–346. Wurmbrand, S. 2012. Skandal oder Illusion? Verbkomplexe in der Grammatiktheorie. Ms. bzw. H-O.

Modals between defectiveness and overdifferentiation

Oliver Schallert
Ludwig-Maximilians-Universität München
oliver.schallert@lmu.de

Modals verbs are an interesting empirical domain for studying morphological form-function mismatches and their interactions with other grammatical levels (in particular syntax and semantics). One facet of such deviations is the interplay between defectiveness and overdifferentiation, which bears on feature signatures and the way they are expressed (exponence) (cf. Corbett 2015: 153–154).

On the empirical base of different sources (grammatical descriptions, data from the Zwirner corpus[1], etc.), we investigate this relationship in the dialects of German (with some side-views on other Germanic languages/varieties). Dialects are an interesting testing ground for such an examination because they are in some sense more "natural" due to their status as primarily oral varieties (Weiß 2001). In particular, they allow us to tackle minimal system contrasts between sufficiently similar grammatical systems and use them as a basis for theoretical modeling (Moulton 1968; Rabanus 2010). We will be focussing on the following three aspects:

1. Due to their historical genesis (most members of this class derive from so-called "preterite-present" verbs), modals originally lacked preterite, participial, and infinitival forms (Birkmann 1987). This gap was subsequently closed by the emergence of forms modeled after the weak conjugation. Particularly in complex perfect forms, interactions with syntax lead to a range of morphological variants beside the well- known substitute infinitive (IPP, "infinitivus pro participio"). One example are specialized participles that combine different morphological strategies like suppletion or truncation, as evidenced by (1), and contrast with regular participles (Schallert 2014; see also Höhle 2006). In modern dialects in the southern parts of the German-speaking territories (with functional expansion of the perfect and concomitant "preterite loss"; cf. Fischer 2018), modals are more likely to preserve past tense forms, and it is an interesting question which grammatical factors allowed for this kind of overdifferentiation (with token frequency acting as an obvious catalyst).

   (1)   Schaumburg dialect [South Low German] (Bölsing 2011: 206, 208):
          hei hat kont-SUP loupen (regular participle: *ekont*)
          "He was able to run."

2. Besides expected gaps (e.g. modals don't have imperative forms, with some marked exceptions concerning *wissen* 'know'; cf. Reis 2001: 291, fn. 8), there seems to be a "finiteness gap" with respect to reportative/epistemic uses (Reis 2001: 293–299). While this question has been extensively discussed with regard to several Germanic standard languages, we know very little about the modality system of the dialects and its morphosyntactic expression.

3. Stem allomorphy triggered by different morphophonological processes (Umlaut, contractions) leads to overdifferentiation in parts of the paradigm, most notably in the present plural (cf. Nübling 2000, 2009). Interestingly, these patterns are linked by implicational relationships yielding different areal distributions (cf. Dammel and Schallert 2018). An open question, which will be addressed, is whether these forms are *morphomic* (Aronoff 1994) or figure as a coding device for highly relevant categories (in the sense of Bybee 1985).

References: Aronoff, Mark. 1994. *Morphology by itself. Stems and inflectional classes*. Cambridge, MA: MIT

---

[1]   Accessible via the DGD platform ("Datenbank für Gesprochenes Deutsch"): http://dgd.ids-mannheim.de/dgd/pragdb.dgd_extern.welcome [accessed November 16, 2020].

Press. Birkmann, Thomas. 1987. *Präteritopräsentia: Morphologische Entwicklung einer Sonderklasse in den altgermanischen Sprachen*. Tübingen: Niemeyer. Bölsing, Friedrich. 2011. *Niederdeutsche Sprachlehre: Plattdeutsch im Kirchspiel Lindhorst, Schaumburg-Lippe*. Hildesheim, Zürich, New York: Olms. Bybee, Joan L. 1985. *Morphology: a study of the relation between meaning and form*. Amsterdam: Benjamins. Corbett, Greville G. 2015. "Morphosyntactic complexity: A typology of lexical splits." *Language* 91(1): 145–193. Dammel, Antje and Oliver Schallert. 2018. "Modalverben in deutschen Dialekten: ein Testfall für die Modellierung morphologischer Variation." In Helen Cristen et al., eds. *Regiolekt – Der neue Dialekt? Akten des 6. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen* (IGDD). Stuttgart: Steiner. Fischer, Hanna. 2018. *Präteritumschwund im Deutschen. Dokumentation und Erklärung eines Verdrängungsprozesses*. Berlin, Boston: De Gruyter. Höhle, Tilman N. 2006. "Observing Non-Finite Verbs: Some 3V Phenomena in German- Dutch." In Patrick Brandt and Eric Fuß, eds. *Form, structure, and grammar: A Festschrift presented to Günther Grewendorf on occasion of his 60th birthday*. Berlin: Akademie Verlag. 55–77. Moulton, William G. 1968. "Structural dialectology." *Language* 44(3): 451–466. Nübling, Damaris. 2000. *Prinzipien der Irregularisierung: eine kontrastive Analyse von zehn Verben in zehn germanischen Sprachen*. Tübingen: Niemeyer. — 2009. "*Müssen, dürfen, können, mögen:* Der Umlaut in den Präteritopräsentia als transkategorialer Marker." *Beiträge zur Geschichte der deutschen Sprache und Literatur* 131: 207–228. Rabanus, Stefan. 2010. "Areal variation in morphology." In Peter Auer and Jürgen Erich Schmidt, eds. *Language and Space. An International Handbook of Linguistic Variation*. Berlin, New York: De Gruyter. 804–821. Reis, Marga. 2001. "Bilden Modalverben im Deutschen eine syntaktische Klasse?" In Reimar Müller and Marga Reis, eds. *Modalität und Modalverben im Deutschen*. Hamburg: Buske. 287–318. Schallert, Oliver. 2014. "IPP-constructions in Bavarian and Alemannic in comparison." In Günther Grewendorf and Helmut Weiß, eds. *Bavarian Syntax. Contributions to the Theory of Syntax*. Amsterdam: Benjamins. 249–304. Weiß, Helmut. 2001. "On Two Types of Natural Languages. Some Consequences for Linguistics." *Theoretical Linguistics* 27(1): 87–103.

Lücken in der Definitheit im Germanischen

Anja Hasse, Patrick Mächler
Universität Zürich, Universität Zürich
anja.hasse@ds.uzh.ch, patrick.maechler@uzh.ch

In den modernen germanischen Sprachen variieren die Ausdrucksverfahren zur Markierung der Definitheitsopposition deutlich: Während die westgermanischen Sprachen freistehende Artikelformen verwenden, bspw. das Deutsche oder Niederländische, machen die nordgermanischen Sprachen, bspw. das Isländische, von einem suffigierten Artikel oder einer Kombination der beiden Verfahren Gebrauch, bspw. das Schwedische.

Allerdings bestehen in einer Reihe von germanischen Varietäten Fälle, in denen die Opposition zwischen definiten und indefiniten Formen neutralisiert ist. Dadurch ergeben sich paradigmatische Lücken im System der Markierung dieser morphosyntaktischen Kategorie (vgl. Sims 2015). Welche Faktoren die Lückenkontexte definieren, ist einzelsprachlich unterschiedlich. Allerdings lassen sie sich systematisieren.

Neben phonologischen können morphosyntaktische und morphologische Faktoren greifen. Im urbanen Südostnorwegischen können Unterschiede zwischen definiten und indefiniten Formen nur im NEUTR.SG neutralisiert werden, *belte* [ˈb̥ɛltə] 'ein Gürtel' vs. *beltet* [ˈb̥ɛltə] 'der Gürtel', in anderen norwegischen Varietäten, wie etwa nordnorwegischen, zusätzlich auch im FEM.SG, *kåpa* 'ein/der Mantel' (vgl. Hanssen 1990: 148), im Färöischen im NOM/AKK.SG.NEUTR, *belti* [ˈb̥ɛ̧ti] 'ein Gürtel' vs. *beltið* [ˈb̥ɛ̧ti] 'der Gürtel' (vgl. Thráinsson et al. 2004: 92f.), in oberdeutschen Dialekten, beispielsweise im Schweizerdeutschen, im NOM./AKK.FEM.SG, *(d) Tapiokastèrchi* [ˈtɒpiokχɒˌʃtɛrχi] '(die) Tapiokastärke', sowie im gesamten PL, *(d) Täsche* [ˈtæʃːə] '(die) Taschen', und im Älvdalischen im DAT.PL mehrsilbiger Substantiva, im NOM/AKK.PL von Substantiva in Abhängigkeit von Genus und Flexionsklassen sowie im NOM/AKK.SG weniger Neutra (vgl. Åkerberg 2012).

Damit gibt es einzelsprachlich unterschiedlich viele Kontexte, in denen definite und indefinite Formen homophon sind, was unter anderem auch von der Komplexität des nominalen Flexionssystems sowie der Art der Definitheitsmarkierung abhängt. Entscheidend ist, in welchen syntaktischen Kontexten die homophonen definiten und indefiniten Formen auftreten können. So führt etwa die Linkserweiterung der Nominalphrasen, beispielsweise um Adjektive, dazu, dass die Unterscheidung nicht aufgehoben wird. Die Lücke fehlt aber nicht nur in bestimmten syntaktischen Kontexten. In jenen Kontexten, in denen sie fortbesteht, wird sie von Sprechern einiger Varietäten mit unterschiedlichen Mechanismen behoben, in anderen Varietäten ist sie dagegen diachron stabil.

In unserem Vortrag stellen wir dar, wodurch die Neutralisierung von Definitheit einzelsprachlich konditioniert ist. Dabei gehen wir auf die oben genannten Faktoren ein und skizzieren die phonologischen Prozesse, die dazu geführt haben, dass Definitheit nicht mehr in allen morphologischen Zellen des Paradigmas und syntaktischen Kontexten ausgedrückt werden kann. Damit sind die synchronen Lücken diachron bedingt (vgl. Reis 2017: 257). Schliesslich gehen wir der Frage nach, welche Auswirkungen diese Lücke der Definitheit in verschiedenen germanischen Sprachen auf das nominale Flexionssystem und die Einbettung von definiten NPs in die Syntax hat.

References: Hanssen, Eskil. 1990. "Nordland". In Jahr, Ernst Håkon, Hrsg. *Den store dialektboka*. Oslo: Novus Forlag. 141-155. Reis, Marga. 2017. "Grammatische Variation und realistische Grammatik". In Konopka, Marek & Wöllstein, Angelika, Hrsg. *Grammatische Variation. Empirische Zugänge und theoretische Modellierung*. Berlin & Boston: de Gruyter. 255-282. Sims, Andrea D. 2015. *Inflectional Defectiveness*. Cambridge: Cambridge University Press. Thráinsson, Höskuldur et al. 2004. *Faroese. An Overview and Reference Grammar.* Tórshavn: Føroya Fróðskaparfelag. Åkerberg, Bengt. 2012. *Älvdalsk Grammatik. Under medverkan av Gunnar Nyström*. Mora: Centrumtryck.

Attraction of the void: The lack of aspect in German and its effect on language change

Elisabeth Scherr
University of Graz
elisabeth.scherr@uni-graz.at

Contrary to its preceding language stages, (New High) German lacks one central component regarding its inherent verbal categories: Beside tense and mode it does not comprise aspect as an obligatory, inflectionally expressed category. This gap of morphological aspect marking, however, sets off compensatory strategies in German to encode the functional means of aspectuality. Inter alia it led to the strengthening of a strategy up to now less noticed in the functional sphere of aspectuality: the article system. For instance, in German nominal definiteness can lead to the perfectivization of the verbal action: *Ich esse.* (imp.) vs. *Ich esse den Apfel.* (perf.) (cf. Leiss 2000: 82). In the latter example it is evident that the definite article renders the referent limited or localizable, which as a consequence leads to the perfectivization of the verbal action. Likewise, the German article system also allows for a de-limitation of the nominal concept comparable to imperfective aspect. In this case the definiteness marker *d-*, generally setting off the deictic localization and the limitation of the nominal referent, gets omitted. This has been shown for clitization processes with prepositional phrases (*Ich muss zum Arzt.*) (cf. Nübling 2005). In these cases, a concrete, inclusive referent is absent, what is expressed instead is sheer activization (cf. Ágel 1996: 46) of a vague, delimited, imperfective concept.

With this in mind, my presentation will focus on a related yet understudied phenomenon up to now ascribed to dialectal communication settings in the Bavarian dialect area: the deletion of *d-* with the definite article in nominal phrases (cf. Weiß 1998: 47ff.):

1747 4022:	im summer	host	EH	**Øas**	**radl**	ghob
	in summer	have	anyway	[inflectional suffix]	bike	had

Whereas this drop of the definiteness-marker is indeed very frequent in dialectal speech, it occurs also in formal settings that rather elicit intended standard language use. In these latter contexts though, it is noticeable that the deletion of *d-* seems only acceptable when the nominal phrase is to express a delimited concept without concrete localization (abstract or common nouns, nominalized adjectives etc.) or in phrases with overall imperfective aspectuality. Thus, it clearly shows a functional motivation compatible with the theoretical outline explained above as it is regularly used for the activization of a delimited reference only. That is why instead of proclaiming loss or revocation of the German article system (cf. Leiss 2010; Nübling 2005), it should rather be argued for a functional enrichment that maybe has the lack of formal aspect marking in German as its driving force.

I will support these assumptions with quantitative and qualitative analyses of data from two corpora of spoken German that allow for valid evidence regarding the distribution of the phenomenon in question alongside the dialect-standard-axis but also regarding its implication for language change with the void of a lacking grammatical category at its center.

References: Ágel, Vilmos. 1996. "Finites Substantiv." *Zeitschrift für Germanistische Linguistik* 24: 16-57. Leiss, Elisabeth. 2010. "Koverter Abbau des Artikels im Gegenwartsdeutschen." In Torsten Leuschner, Tanja Mortelmans, and Sarah De Groodt, eds. *Grammatikalisierung im Deutschen*. Berlin/New York: de Gruyter. 137-158. Leiss, Elisabeth. 2000. *Artikel und Aspekt. Die grammatischen Muster von Definitheit*. Berlin/New York: de Gruyter. Nübling, Damaris. 2005. "Von in die über in'n und ins bis im. Die Klitisierung von Präposition und Artikel als ‚Grammatikalisierungsbaustelle'." In Torsten Leuschner, Tanja Mortelmans, and Sarah De Groodt, eds. *Grammatikalisierung im Deutschen*. Berlin/New York: de Gruyter. 105-132. Weiß, Helmut. 1998. *Syntax des Bairischen. Studien zur Grammatik einer natürlichen Sprache*. Tübingen: Niemeyer.

Circumventing the 'that-trace' effect: Different strategies between Germanic and Romance

Ermenegildo Bidese, Andrea Padovan, Alessandra Tomaselli
University of Trento, University of Verona, University of Verona
ermenegildo.bidese@unitn.it, andrea.padovan@univr.it, alessandra.tomaselli@univr.it

Cimbrian is a German(ic) VO variety that has lost the core property of the V2 phenomenon i.e. the well-known linear restriction found in Germanic V2 languages. However, it has retained mandatory V-to-Fin movement in root clauses.

As regards the null subject phenomenon, Cimbrian is clearly a non-pro-drop language (cf.1a) (see Bidese 2008, Bidese & Tomaselli 2018). It displays, nevertheless, a special kind of subject inversion which looks more Italian- than Germanic-like (see 1b): the subject follows the whole verbal complex.

(1) a. Gestarn   in balt   *di diarn/si*   hatt gesek   in   has
        yesterday   in.the wood   the girl/she has seen   the   hare
        'Yesterday, the girl saw the hare in the wood '

    b. Gestarn   in balt   hat-*(ta) gesek   {*di diarn*} in has   {*di diarn*}
        yesterday   in the wood   has-da.CL seen   {the girl} the hare   {the girl}
        'It was the girl that saw the hare yesterday in the wood '

The prediction (along the lines of Rizzi 1982) that the post-verbal (Italian-like) position of the subject is connected with the possibility of circumventing the 'that-trace 'effect is fully borne out in Cimbrian.

(2) a. Ber$_i$ gloabest-(t)o   az-*(**ta**)   khemm t$_i$?
        who   believe-you.CL   that-da.CL   come.SUBJ
        'Who do you think is coming? '

    b. Ber$_i$ gloabest-(t)o   az-*(**ta**)   habe   gesek t$_i$ in   has   in balt?
        who   believe-you.CL   that-da.CL have.SUBJ seen   the   hare   in.the wood
        'Who do you believe saw the hare in the wood? '

Interestingly, the overt complementizer (*az*) requires the obligatory presence of an expletive element -*da/ta*.

The aim of our paper is to explore the role of *da* in circumventing the 'that-trace' effect in Cimbrian and to compare it with other strategies to bypass the overt complementizer in both Germanic and Romance varieties (e.g. expletive *pro* in Bavarian and the well-known *que/qui* divide in French). Other phenomena like complementizer agreement (which turns out to be a complementary strategy to circumvent the 'that-trace' effect) will also be addressed. In particular, we will deal with lesser-studied South-Bavarian varieties such as the Tyrolean dialects which show interesting grammaticalization paths of the elements that encliticize onto C (the Upper Vinschgau Valley and the German Nonsberg).

References: Bidese, Ermenegildo. 2008. *Die diachronische Syntax des Zimbrischen*. Tübingen: Günter Narr. Bidese, Ermenegildo, Andrea Padovan, and Alessandra Tomaselli. 2020. "Rethinking V2 and Nominative case assignment: new insights from a Germanic variety in Northern Italy." In Rebecca Woods, and Sam Wolfe, eds. *Rethinking Verb Second.* Oxford: Oxford University Press. 575-593. Bidese, Ermenegildo, and Alessandra Tomaselli. 2018. "Developing pro-drop: the case of Cimbrian." In Federica Cognola, and Jan Casalicchio, eds. *Null Subjects in Generative Grammar: a Synchronic and Diachronic Perspective*. Oxford: Oxford University Press. 52-69. Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Berlin: De Gruyter Mouton.

Syntactic paradigms, markedness and similative markers in comparative and relative clauses

Julia Bacskai-Atkari
University of Konstanz
julia.bacskai-atkari@uni-konstanz.de

My talk investigates a peculiar part of the complementiser system in Germanic and beyond, showing that the relation between the individual clause types can be best modelled by stipulating the existence of syntactic paradigms, the members of which are ordered according to markedness. Gaps in the paradigm appear to be systematic in that they occur in the more marked (potential) slots. In Germanic languages, it is common for complementisers like *so* (*swa*) and its reinforced version *as* (*all + so* 'just like') to introduce not only similative clauses, (1a), but also relative clauses, (1b), degree equatives, (1c), and comparatives, (1d):

(1) a. *Se sæ heo onhefð… swa **swa** weall* 'The sea rises like a wall.'
    b. *and yrfan hi **swa** hi wyrðe witan* 'And let those inherit whom they know worthy.'
    c. *Mary is as tall **as** Peter is*.
    d. *Also this erbe haviþ mo vertues **as** endyue haþe*. 'This herb also has more virtue than endive has.'

I argue that the patterns in (1) constitute a syntactic paradigm, due to their syntactic similarities. The members of this paradigm are subject to diachronic changes in the complementiser along two major lines: (i) the morphological distinction among the individual members (comparable to phonological distinctions in inflectional paradigms) and (ii) analogical changes affecting the morphological properties of the complementiser (comparable to analogical changes in pronominal systems, e.g. the change from *h*-pronouns to *th*-pronouns in 3Pl in Middle English). For (i), the distinctions in present-day English provide support: similatives are introduced by *like*; relative clauses are introduced in some dialects by *as* but other dialects, including Standard English, use other complementisers or pronouns; equatives retain *as*; comparatives are introduced by *than*. For (ii), German shows cyclic changes from the demonstrative-based series (*so*, *als*) to the *w*-based series (*wo*, *wie*) throughout its history. The effects (i) and (ii) are truly paradigmatic in nature and differ from simple analogical extension. Regarding markedness, the unmarked pattern is (1a), which contains a complementiser expressing similarity (lexical meaning), while the lexical content is bleached in all the other constructions. In one direction, (headed) relative clauses are more marked due to the presence of a nominal head, overt or covert, [+rel] for shorthand. In the other direction, degree equatives and comparatives are more marked due to degree, [+deg], and comparatives are more marked than equatives due to inequality, [+ineq]. Reanalysis processes along the lines of (ii) spread from the unmarked case to the marked cases (see Jäger 2018 for comparatives; contrary to Haspelmath & Buchholz 1998 for relative clauses). Differentiations along the lines of (i) may be systematic (starting from the unmarked case) or system-external (starting as innovations in the marked cases, e.g. the use of non-similative- based complementisers such as English *that* or *than*). Regarding syntactic gaps, this model makes two predictions: (A) gaps arise in the most marked slots, i.e. relative and comparative clauses (e.g. the relative pronoun strategy and relative complementisers reanalysed from such pronouns; negative-based comparative complementisers and phrasal comparatives instead of comparative clauses), and (B) languages that lack constructions like (1a) also lack those in (1b)–(1d), but not the other way round (confirmed typological results), leading to similar results as systematic gaps in inflectional paradigms (e.g. the lack of morphological future tense but not of present tense). As I will show, both of these predictions are borne out.

References: Haspelmath, Martin & Oda Buchholz. 1998. "Equative and similative constructions in the languages of Europe." In Johan van der Auwera, ed. *Adverbial constructions in the languages of Europe*. Berlin: Mouton de Gruyter. Jäger, Agnes. 2018. *Vergleichskonstruktionen im Deutschen: Diachroner Wandel und synchrone Variation*. Berlin: De Gruyter.

Empirical determinants of grammatical gaps and grammatical inventions

Ralf Vogel
University of Bielefeld
ralf.vogel@uni-bielefeld.de

The identification of a *gap* in the grammatical system or inventories of a language (henceforth *language system*) presupposes the *observation* of such a gap in language use. My presentation is mainly concerned with three aims: i) attempt to understand the kinds of observations that lead to the conclusion that there must be such a gap; ii) make a proposal for the observational heuristics of identifying linguistic gaps, and iii) present a theory of the strategies that speakers employ when confronted with gaps in terms of linguistic creativity, in particular, *grammatical invention*.

The observation of a gap, by necessity, is theory-laden, as it is guided by expectations about the conventionalisation and use of forms. It is the non-fulfillment of such expectations (on what to observe) that leads to the identification of a gap by linguists. The best indicator of conventionalisation is corpus frequency. The thesis that some (expected) unit may not exist in a language predicts its absence from corpora and must be tested with corpus research.

But here we need to be careful and distinguish, following the tradition since Saussure, absence from speech and absence from the language system. Rare configurations may still have frequent solutions. This is typical in syntax. Consider (1): While *wh*-extraction out of an embedded clause is rather frequent, with the addition of further subordinate clauses the configuration gets rarer and rarer.

(1)   What do you think [that Paul thinks [that Anne thinks [that John thinks [that Mary suggested]]]] ?

This has never been seen as problematic, because the acceptability of cases like (1) can still be elicited. The assumption of a single rule of cyclic *wh*-movement is corroborated in the eyes of researchers. So despite its rarity, (1) exemplifies no gap, because it receives a standard solution. Assume now a hypothetical variant of English, English', where in cases with 4+ embeddings like (1) a new construction with *quid* (from Latin, 'what') would be used:

(2)   **Quid** do you think [that Paul thinks [that Anne thinks [that John thinks [that Mary suggested **it**]]]] ?

The source of this solution would have to be located in speakers' general linguistic competence. The *quid*-construction would then be analysed as a kind of ad hoc device outside of grammar (Reis 2017). **But:** there is no principal reason to exclude such an analysis for cyclic *wh*-movement in (1). Speakers might make up such a recursive rule at the very same moment when confronted with a 4+ embedded question. The reason why (1) appears less ad hoc to us is that (1) is constructed *by analogy* to simpler cases. A *failure of analogy* seems to be a further criterion for the identification of a gap. In turn, construction by analogy seems to be sufficient for us linguists to exclude the assumption of a gap. This might be a problematic epistemic aspect of our practice insofar as we obviously privilege construction by analogy over other modes of construction, where in fact these might just be strategies of grammatical creativity which are different but equal in status as solution strategies when confronted with rare configurations. What is necessary, in particular, are empirical criteria that allow us to distinguish cases where construction by analogy is a sign of conventionalisation from those cases where analogy is used as a mechanism of grammatical invention and I will discuss several such cases from the grammar of German in my presentation.

References: Reis, Marga. 2017. Grammatische Variation und realistische Grammatik. In Marek Konopka & Angelika Wöllstein (eds.), Grammatische Variation. Empirische Zugänge und theoretische Modellierung, 255–282. Berlin & Boston: de Gruyter.

A typology of case competition in headless relatives

Fenna Bergsma
Goethe-Universität Frankfurt
bergsma@em.uni-frankfurt.de

In case competition in headless relatives two aspects play a role. The first one is which case wins the case competition. It is a crosslinguistically stable fact that this is determined by the case scale: NOM < ACC < DAT (cf. Grosu 2003). A case more to the right on the scale wins over a case more to the left on the scale.

This generates the pattern shown in Table 1. The left column shows the internal case (the case assigned in the relative clause) between square brackets. The top row shows the external case (the case assigned in the main clause) between square brackets. The other cells indicate the case of the relative pronoun. When the dative wins over the accusative, the relative pronoun appears in the dative case. When the dative wins over the nominative, the relative pronoun appears in the nominative case. When the accusative wins over the nominative, the relative pronoun appears in the accusative case.

Table 1. Case competition

INT \ EXT	[NOM]	[ACC]	[DAT]
[NOM]	NOM	ACC	DAT
[ACC]	ACC	ACC	DAT
[DAT]	DAT	DAT	DAT

The second aspect that plays a role in headless relatives is whether the internal and the external case are allowed to surface when either of them wins the case competition. This differs across languages. There are four logically possible language types: (1) the unrestricted type, in which the internal and the external case are allowed to surface when either of them wins the case competition; (2) the internal only type, in which only the internal case is allowed to surface when it wins the case competition; (3) the external only type, in which only the external case is allowed to surface when it wins the case competition; and (4) the matching type, in which neither the internal case nor in the external case is allowed to surface when either of them wins the case competition.

As far as I am aware, only three of these possible patterns are attested in natural languages. In the description, I refer to the differ gray-marking in the table. The cells marked in light gray are the ones in which the internal case wins the case competition, the cells marked in dark gray are the ones in which the external case wins the case competition, and the unmarked cells are the ones in which the internal and external case match.

Old High German is an example of the non-matching type, in which relative pronouns in the unmarked, light gray and dark gray cells are attested. Modern German is an example of the internal-only type, in which relative pronouns in the unmarked and light gray cells are grammatical (Vogel 2001). To my knowledge, the external-only type is not attested. This would be a language in which relative pronouns in the unmarked and the dark gray cells are grammatical. Polish is an example of a language of the matching type, in which relative pronoun in only in the unmarked cells are grammatical (Citko 2013).

References: Citko, Barbara (2013). "Size matters: Multidominance and DP structure in Polish". Talk at the 44th Poznan Linguistic Meeting. Grosu, Alexander (2003). "A Unified Theory of 'standard' and 'transparent' Free Relatives". In: Natural Language and Linguistic Theory 21.2, pp. 247–331. Vogel, Ralf (2001). "Case Conflict in German Free Relative Constructions: An Optimality Theoretic Treatment". In: Competition in Syntax. Ed. by Gereon Müller and Wolfgang Sternefeld. Berlin: Mouton de Gruyter, pp. 341–375.

How sex and gender shape agreement in German relative clauses

Ewa Trutkowski
Free University Bozen-Bolzano
ewa.trutkowski@unibz.it

This talk aims at shedding some light on agreement and the resolution of feature mismatches in relative clauses (RCs) with 1st or 2nd person head nouns (HNs). That kind of RCs is characterized by a number of person mismatches, the one between HN and relative pronoun (RP) being of particular interest. As to this, Ito & Mester (2000) and Vogel (2007) claim that the presence of a resumptive pronoun (ResP) is necessary to resolve this mismatch.

However, Trutkowski & Weiß (2016) experimentally confirmed the view of Heck & Cuartero (2008) according to which such RCs come in two versions – with resumptive pronoun (ResP), cf. *Du, der du schläfst* (You-2sg, RP-sg.masc ResP-2sg sleeps-2sg) and without. In this talk, mainly the latter will be the focus of interest, cf. (1):

(1) a. *Du, der $^?$schläfst / schläft …*
      You-sg, RP-sg.masc sleep-2sg / sleep-3sg
   b. *Ihr, die schlaft / *schlafen …*
      You-pl, RP-pl sleep-2pl / sleep-3pl

In cases with ResP the finite verb within the RC (V.fin-RC) displays the same person and number features as the HN and the ResP. However, when the RC does not contain a ResP, the feature specification at V.fin-RC depends on the number of the HN: In the singular the V.fin-RC (preferably) displays 3rd person features; in the plural we find 1st or 2nd person agreement at V.fin-RC (depending on the particular person of the HN). From the above data the following questions emerge (inter alia): If we consider the V.fin-RC to be the 'target of agreement' (in the sense of Corbett 2003) – what is the 'source of agreement'? Can we have multiple sources of agreement? And if so: which 'sources' and features (from the main clause and the relative clause) are involved? In order to find an explanation for these questions, I presume that V.fin-RC can only spell out features borne by those elements that are part of the agreement chain (which consists of HN, RP and "co-congruent" elements which stand in an agreement relation with the latter). As person features of HN and RP are mismatched, and number alone is not sufficient enough, I assume that agreement with the V.fin-RC is enabled by sex and gender features. For the singular the outline of this idea reads as follows: Not only is gender overtly marked at the RP but indexicals are also equipped with covert sex features that are part of their reference assignment in the sense of Kaplan (1989a,b). The probability of this assumption is evinced by (2) where an RP relating to a singular indexical cannot be marked for neuter but must adjust to the sex of the indexical's (human) referent.

(2) *Du, der/die/*das schläft, kannst nicht kommen.*
   You-sg, RP-sg.masc/fem/neut sleep-3sg, can not come

In my talk I will substantiate the claim that gender and sex are crucial features in establishing agreement in this special kind of RCs with experimental evidence and explain the singular- plural difference which I link to the hypothesis that the German plural lacks gender.

References: Heck, Fabian & Juan Cuartero. 2012. Long Distance Agreement in Relative Clauses. In: Alexiadou, A., T. Kiss & G. Müller (eds.) *Local Modelling of Non-Local Dependencies in Syntax*, 49-83. Berlin. De Gruyter. Trutkowski, Ewa & Helmut Weiß. 2016. When Personal Pronouns Compete with Relative Pronouns. In: Grosz, P. & P. Patel-Grosz. *The Impact of Pronominal Form on Interpretation*. Berlin. De Gruyter, 135–166. Vogel, Ralf. 2007. „Ich, der ich . . .“ Seminar *Syntax und Morphologie*, winter term 2007/2008. Handout, Bielefeld University.

Cross-Germanic variation in relative clauses with pronominal antecedents

Kerstin Hoge
University of Oxford kerstin.hoge@mod-langs.ox.ac.uk

The present paper provides some support for two sources of ineffability, located in the lexicon and at the C-I interface, respectively. Adopting Minimalist assumptions, it considers the distribution of pronouns functioning as head nouns of relative clauses (RCs) in English, German and Yiddish. The three languages differ with respect to two dimensions: (i) the nature of the relative pronoun (a wh-word or complementiser in English and Yiddish; and a d- word in German[2]), and (ii) the acceptability of pronominal antecedents in restrictive and appositive RCs, which in (British Standard) English are allowed only in generic sentences or "Voldemort phrases" (Elbourne 2013; Zobel 2015; see Conrod et al. 2016 concerning contemporary American English), in German only in appositive RCs, and in Yiddish in both generic restrictive and appositive RCs, as summarised and exemplified below.

	Relative Pronoun	Restrictive RCs	Appositive RCs
English	wh-word; C	(✓)	*
German	d-word	*	✓
Yiddish	wh-word; C	(✓)	✓

(1) a. He who/that calls his brother a fool shall be damned.
   b. *I opened the door to him, who/that by the way came an hour late.

(2) a. ??Er, der seinen Bruder einen Narren schimpft, soll verdammt sein.
   b. Ich machte ihm, der übrigens eine Stunde später kam, die Tür auf.

(3) a. Zey vos/velkhe geyen oyf ayz zoln tretn laykht. [Yiddish]
        they COMP/who.PL walk on ice should tread lightly
        'Those who walk on ice should tread lightly.'
   b. Ikh hob im, vos er/velkher   iz agev   a sho   shpeter gekumen,
        I   have him COMP he/who.NOM.SG is by-the-way an hour later   come.PPL
                                                          oyfgeefnt di tir.
                                                          opened   the door.
        lit. 'I opened the door to him, who by the way came an hour late.'

The questions that arise in light of these data are: what rules out appositive RCs in English, which licenses restrictive postmodification of a pronoun; and why are restrictive RCs incompatible with pronominal antecedents in German when it is possible for a pronoun to occur with an appositive RC? In answer to these questions, this paper argues that (i) there is cross-linguistic variation in pronominal structure, with English and Yiddish (third-person) pronouns, unlike their German counterparts, possessing the extended functional structure of anaphoric definites (cf. Schwarz 2009; Patel-Grosz & Grosz 2017); and (ii) appositive relative clauses are parentheticals which are not syntactically integrated into their host and which, in the absence of a lexical head noun, can be interpreted only if they contain relative d-words (or, alternatively for Yiddish, resumptive pronouns) with a full phi-feature set. Once again, it is lexical gaps (in the case of German personal pronouns) and uninterpretability at the C-I interface (in the case of English appositive RCs) that appear to be involved as sources of ineffability.

---

[2] Relative clauses introduced by *welch-* will be ignored as these are typically confined to written language

References: Baker, M. 2008. The macroparameter in a microparametric world. In T. Biberauer (ed.), *The limits of variation*, 351-373. Amsterdam: John Benjamins. Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1-52. Cambridge, MA: MIT Press. Chomsky, N. 2008. On phases. In R. Freidin et al. (eds.), *Foundational issues in linguistic theory*: *Essays in honor of Jean-Roger Vergnaud*, 133-166. Cambridge, MA: MIT Press. Conrod, K., R. Tatman & R. Koncel-Kedziorski. 2016. We who tweet: Pronominal relative clauses on Twitter. *Corpus Linguistics Fest 2016 Proceedings*, Indiana University. https://linguistics.washington.edu/research/publications/we-who-tweet-pronominal-relative-clauses-twitter (Accessed on 20 September 2020). Elbourne, P. 2013. *Definite descriptions*. Oxford: Oxford University Press. Fanselow, G. & C. Féry. 2002. Ineffability in grammar. In G. Fanselow & C. Féry (eds.). *Resolving conflicts in grammars: Optimality Theory in syntax, morphology, and phonology*, 265-307. Hamburg: Helmut Buske Verlag. Patel-Grosz, P. & P. Grosz. 2017. Revisiting pronominal typology. *Linguistic Inquiry*, 258-297. Schwarz, F. 2009. *Two types of definites in natural language*. PhD dissertation, University of Massachusetts, Amherst. Zobel, S. Voldemort phrases in generic sentences. *Grazer Linguistische Studien* 83, 107-123.

Morphological gaps and syntax: Agreement in Mian discourse

Sebastian Fedden

Université Sorbonne Nouvelle sebastian.fedden@sorbonne-nouvelle.fr

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

(1)   máam=e   a-nâ'-n-o=be
                mosquito=SG.M   3SG.M1.OBJ-hit-REAL-3SG.F.SBJ=DECL
                'She hit the mosquito.'

(2)   máam=e   dob-ò-n-o=a
                mosquito=SG.M   3SG.M2.OBJ-take-SEQ-3SG.F.SBJ=MED
                'She picked the mosquito up and then ...'

(3)   máam=e   bou-n-o=be mosquito=SG.M   swat-REAL-3SG.F.SBJ=DECL
                'She swatted the mosquito.'

Due to the high type and token frequency of non-agreeing verbs the syntax is constantly encountering gaps. We should ask how such a system works in discourse; in particular, how the presence or absence of agreement relates to the overt vs. null realization of arguments. A hypothesis that has been put forward is the *Complementarity Principle* (CP) (see Kibrik 2011, Haig & Schnell 2016). This is a principle of economy which claims that null arguments are favoured by overt agreements and vice versa.

To check this hypothesis, I conducted a discourse study of Mian, using the procedures set out in Bickel (2003) and Nichols (2018). The table below summarizes the comparison of null vs. overt objects to agreement vs. non-agreement for Mian. I include the expected percentages under full complementarity (i.e. 0% of objects overt with agreeing verbs, 100% of objects overt with non-agreeing verbs) to help contextualize the actual figures (in boldface).

	Verb agrees with object			Verb does not agree with object
	system 1	system 2	combined	
null object	50	134	184	223
overt object	21	97	118	159
% overt object	**0.30**	**0.42**	**0.39**	**0.42**
% prediction from CP	0.00	0.00	0.00	1.00

Chi-square tests show that the percentages for system 1, system 2 and both systems combined, each in comparison with verbs that do not agree, do not differ significantly (at $p < .05$) from the percentage of non-agreeing verbs. **Thus the results give no support to the CP.** The gaps in the morphology are not compensated for by the syntax. This is further evidence for the independence of the morphological component.

References: Bickel, Balthasar. 2003. "Referential density in discourse and syntactic typology." *Language* 79: 708-736. Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press. Corbett, Greville G., Sebastian Fedden, and Raphael Finkel. 2017. "Single versus concurrent feature systems: nominal classification in Mian." *Linguistic Typology* 21: 209-260. Haig, Geoffrey, and Stefan Schnell. 2016. "The discourse basis of ergativity revisited." *Language* 92: 591-618. Kibrik, Andrej A. 2011. *Reference in discourse*. Oxford: Oxford University Press. Nichols, Johanna. 2018. "Agreement with overt and null arguments in Ingush." *Linguistics* 56: 845-863.

Doubled possessors: One gap filled twice

Ekaterina Levina

University of Texas at Austin ekaterina.levina@utexas.edu

In this talk, I will consider the asymmetric paradigm of Russian pronominal possessives (PrP). I will show that the strategies used to fill the gaps, which once existed, provide us important evidence for the crucial semantic and discursive characteristics of referential expressions which seek to be expressed in the language.

As in many other Indo-European languages (cf. Dobrovie-Sorin 2011, Kayne 2018), Russian shows a morphological split in the paradigm of PrP between 1st/2nd person possessives, on the one hand, and 3rd person, on the other. The former are results of adjectival derivation from pronominal bases (cf. Buslaev 1959, Lunt 2010). The latter are genitive forms of the correspondent personal pronouns. These two groups of PrP behave morphologically different: 1st/2nd possessives, as also denominal adjectival possessors, agree with possessees in gender, number and case (*nash-a* (*Katin-a) koshk-a* / our-**F.SG.NOM** (Katja.ADJZ-**F.SG.NOM**) cat-**F.SG.NOM**); 3rd person possessives, in the same way as genitive possessors, do not show any kind of agreement (*ih koshk-a (Kati)* /they.**GEN** cat-**F.SG.NOM** (Katja.**GEN**)).

This discrepancy in the paradigm is due to the late development of the 3rd person personal pronouns in Slavic (cf. Buslaev 1959, Volk 2014, Lunt 2010). The formation of 1st/2nd person PrP was already completed, and the underlying derivational process was no longer productive. Suppletive genitive forms of 3rd person personal pronouns were used instead. Yet, a while after, the second compensatory process has started: new adjectival possessives were derived from already existing 3rd person genitive forms. That is to say, adjectivization (different from 1st/2nd person possessives) was applied to already existing 3rd person PrP (second layer of possessivity marking was added); in this way, the lacking agreement between the possessor and the possessee was finally made to work (*ih koshk-a* /they.**GEN** cat-**F.SG.NOM** vs. *ih-nyaya koshk-a* / they-ADJZ.**F.SG.NOM** cat-**F.SG.NOM**).

Besides 3rd person PrP, all other kinds of possessors occurring in the prenominal position have unique, animate referents (cf. Dobrovie-Sorin 2011). Due to their pronominal nature, 3rd PrP usually also have unique referents, which, however, are not necessarily animates. The appearance of the second set of 3rd person PrP clearly shows a compensation strategy towards the paradigmatic harmonization. This process is interesting in two ways. Firstly, the morphological agreement features might be associated with the prenominal position where PrP appear. This can be explained with the information-structural requirement for the discourse-old elements (referents of PrP) to precede other elements facilitating on this way the cognitive processing of the information.

Secondly, the suffixes used for the derivation of adjectival possessors only occur with unique animate referents. The suffixes appearing in the second layer of possession marking on 3rd person PrP are different: they are used for the adjectival derivations from kinship terms that are not necessarily unique, but always animate (cf. Dobrovie-Sorin 2011). Interestingly, the second set of 3rd person PrP refers exclusively to animates. This semantic consistency of the derivational patterns clearly shows the high importance of the exact referential information about the possessor to be conveyed in a possessive phrase.

References: Buslaev, F.I. 1959. *Istoricheskaya grammatika russkogo* [Historical grammar of Russian language]. Moskva: Uchpedgyz. Dobrovie-Sorin, C. & I. Giurgea. 2011. Pronominal possessors and feature uniqueness. *Language* 87. 126–157. Kayne, R.S. 2018. Toward an understanding of apparent suppletion in Romance pronominal possessives. *Probus* 30(2). 251–276. Lunt, H.G. 2010. *Old church Slavonic grammar*. Berlin: Walter de Gruyter. Volk, V.S. 2014. Syntax of possessive pronouns and adjective derivation. *Acta Linguistica Petropolitana: Trudy Instituta lingvisticheskikh issledovanii*. X(2). 510–533.

Arbeitsgruppe 2

Weak elements in prosodic acquisition and processing
Ulrike Domahs, Angela Grimm & Mathias Scharinger

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Reduced word pronunciation variants: Properties and processing

Mirjam Ernestus
Centre for Language Studies, Radboud University
m.ernestus@let.ru.nl

In prosodically weak positions, words are often weakly articulated or produced with fewer segments or even with fewer syllables. English examples are *plice* for *police* and *yeshay* for *yesterday*. In this talk, I will discuss several acoustic and psycholinguistic studies documenting the properties of reduced word pronunciation variants and how we process them.

I will first discuss a single Dutch word. In Ernestus & Smith (2018), we studied in detail the Dutch discourse marker *eigenlijk* 'actually', which occurs in many pronunciation variants, ranging from full /ɛɪxələk/ to /ɛɪk/ or /ɛɪx/. This study is the first to show an effect of the rhythm of the sentence on the number of syllables of a word, which provides cues about when in the production process the number of syllables of *eigenlijk* is determined. Another important finding is that several variants of *eigenlijk* contain the phonotactically illegal consonant cluster /xk/, and that this cluster is typically relatively long. Both properties of the cluster may form cues for listeners to restore the word's full form.

However, listeners need experience to use these as well as other cues to overcome reductions. This appears most obviously from second language learners, who typically have little experience with reduced word pronunciation variants. In Ernestus et al. (2017), we found that advanced second language learners of Dutch not only make more errors with reduced word variants in a dictation task than native listeners, but they also seldom rely on semantic and syntactic information in the context or on subsegmental cues to overcome the reductions.

In two reaction time studies, we further documented the role of experience in the recognition of reduced word pronunciation variants. In Ernestus & Brand (2019), we focused on reduction patterns in French word final obstruent-liquid-schwa clusters (e.g. /trə/ as in *ministre* /ministrə/ 'minster'). We found that listeners more quickly recognize word variants displaying the highly frequent, highly reduced pronunciation of the cluster (e.g. /minis/) than the less frequent, less severely reduced pronunciation (e.g. /minisr/). In Ernestus & Brand (2018), we tested both native listeners and Dutch learners of French in a lexical decision task in which words were presented in full (e.g. *pelouse* /pəluz/ 'lawn') or without the schwa in the initial syllable (e.g. /pluz/). Native listeners' indication of how often they think a given word occurs in the two variants (reflecting their experience with the word) is a good predictor of how quickly they recognize each of the word's two variants. The same holds for learners, but, importantly, one group's estimations is a bad predictor for the other group's performance. This finding shows the role of the individual listener's experience in recognizing reduced words.

If listeners' ability to efficiently process reduced word pronunciation variants especially comes with experience, the question arises when native listeners can start acquiring these variants. In Lahey & Ernestus (2014), we compared the pronunciations of two Dutch adverbs (*helemaal* /heləmaal/ 'completely' and *allemaal* /ɑləmal/ 'all') between speech directed to 11/ 12 months old infants and to adults. The two words show the same variation in child and adult directed spontaneous speech, both in quantity and in quality (frequent variants being /heməl/ and /ɑmɑl/). From very early onwards, children hear reduced word pronunciation variants and the full range of reduction may help them to efficiently interpret the highly reduced ones.

In conclusion, there is much variation in how words sound in prosodically weak positions. Although many questions about how reduced word variants are recognized are still open, it is clear that listeners need experience to interpret the variety of cues that may help them overcome reduction. Native listeners start acquiring this experience in very early childhood.

References: Brand, Sophie, and Mirjam Ernestus. 2018. "Listeners' processing of a given reduced word pronunciation variant directly reflects their exposure to this variant: Evidence from native listeners and learners of French." *Quarterly Journal of Experimental Psychology* 71: 1240-1259. Brand, Sophie, and Mirjam Ernestus. 2019. "Understanding reduced words: the relevance of reduction degree and frequency

of occurrence." In *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*. 97-101. Ernestus, Mirjam, Myrthe Dikmans, and Ghislaine Giezenaar. 2017. "Advanced second language learners experience difficulties processing reduced word pronunciation variants." *Dutch Journal of Applied Linguistics* 6: 1-20. Ernestus, Mirjam, and Rachel Smith. 2018. "Qualitative and quantitative aspects of phonetic variation in Dutch eigenlijk." In: Francesco Cangemi, Meghan Clayards, Oliver Niebuhr, Barbara Schuppler, and Margaret Zellers, eds. *Rethinking reduction*. Berlin/Boston: De Gruyter Mouton. 129-163. Lahey, Mybeth, and Mirjam Ernestus. 2014. "Pronunciation variation in infant-directed speech: Phonetic reduction of two highly frequent words." *Language Learning and Development* 10: 308-327.

Truncation of weak syllables: Early L2 learners behave like monolingual children

Alina Lausecker, Angela Grimm, Petra Schulz
GU Frankfurt, GU Frankfurt, GU Frankfurt
Lausecker@em.uni-frankfurt.de, a.grimm@em.uni-frankfurt.de, p.schulz@em.uni-frankfurt.de

Table 1: Non-word construction based on Féry (1995)

Stress Patterns	Prosodic Structure	Non-Word Example	Real-Word Example
SWs	[[ˈ**dom**ga]$_F$[ˌfup]$_F$]$_{PW}$	**dom**gafup	**Cal**vados
WSW	[do [ˈ**gam**fu]$_F$]$_{PW}$	do**gam**fu	Ve**ran**da
sWS	[[ˌdogam]$_F$[ˈ**fulp**]$_F$]$_{PW}$	dogam**fulp**	Redun**danz**

Note: SWs = main stress on the antepenultimate syllable, WSW = main stress on the penultimate syllable, sWS = main stress on the ultimate syllable, F = Foot; PW = Prosodic Word.

Table 2: Numbers of truncation patterns in eL2-children

eL2-children	Age (years; months)	Exposure to L2 (months)	L1	Truncations	
				pretonic	*posttonic*
eL2_1	3;8	4	Russian	2	1
eL2_2	3;11	14	Russian	4	0
eL2_3	4;7	14	Russian	4	0
eL2_4	3;4	3	Russian	8	7
eL2_5	4;5	7	Mandarin	2	0
eL2_6	3;7	3	Punjabi	2	1
eL2_7	4;5	5	Portuguese	6	4
eL2_8	3;3	5	Arabic	19	4
eL2_9	3;4	5	Arabic	3	1
eL2_10	4;6	12	Igbo	4	0
total				**54**	**18**

Note: eL2 = early second language, L1 = first language.

Table 3: Foot structure of truncated items

Bisyllabic trochees			Iambs		Monosyllables
	Sw	Sə	wS	əS	S
SWs	10	8	-	-	0
WSW	23	2	0	0	0
sWS	-	-	20	3	6
total	33	10	20	3	6

Note: S = stressed syllable, w = weak syllable with full vowel, ə = weak syllable with reduced vowel, - = not possible.

Prosodic status of polar *kya* in Urdu/Hindi

Farhat Jabeen
University of Bielefeld farhat.jabeen@uni-bielefeld.de

The wh-word *kya* 'what' in Urdu/Hindi has a homophonous variant used optionally in polar questions. Butt et al. (2020) have shown that the wh-constituent *kya* has longer duration and is produced with a rising (LH) F0 contour as compared with polar *kya* which is always deaccented. However, further investigation has shown that polar *kya* may also be produced with a rising F0 contour. This raises questions against the existing understanding of polar *kya* as prosodically weak and suggests that polar *kya* may be produced as an Accentual Phrase (AP) with a rising contour. Moreover, as the overall intonation of polar questions varies on the basis of the position of the questioned constituent (Jabeen, 2020), it is unclear if the intonation and prosodic status of polar *kya* (as weak deaccented vs. as AP) is also affected by the position of the questioned constituent. This study aims to fill this gap and investigates the intonation contour associated with polar *kya* in Urdu/Hindi. It also reports if the intonation and prosodic status of polar *kya* varies due to its position and the position of the questioned constituent in the sentence.

 In order to answer these questions, I recorded eleven Urdu speakers from Pakistan. They were presented with ditransitive polar questions containing polar *kya* at sentence initial, medial, immediately preverbal, or final positions (1). For each position of polar *kya*, either a noun phrase or the verb was questioned. I measured relative duration (syllable duration/ sentence duration) and F0 contour of polar *kya* placed at different positions. Statistical analysis was carried out using Linear Mixed Effects Regression (Baayen et al., 2008).

(1) NP1 NP2 NP3 Verb
 (kjɑ) mo.nɑ=ne (kjɑ) no.mi=ko (kjɑ) t̪ə.mi:z sikʰ.vɑi (kjɑ)
 (what) Mona=Erg (what) Nomi=Acc (what) manner.F.Sg. teach.Caus.F.Sg (what)
 'Did Mona have Nomi taught manners?'

Results show that the position of the questioned constituent influences the intonation of polar *kya* regardless of the position of the particle itself. Polar *kya*, at any position, can be produced with a rising F0 contour or be deaccented. Moreover, the rising F0 contour is produced most frequently when polar *kya* is placed to the right of the questioned constituent. This tendency is stronger when polar *kya* occurs immediately before the questioned phrase. Following Jabeen and Delais-Roussarie (2019)'s claim that Accentual Phrases in Urdu/Hindi are produced with rising F0 contour, I claim that polar *kya*, when produced with a rise, forms an Accentual Phrase. This analysis is further supported by the application of three phonological processes on polar *kya* produced with an LH contour i.e. AP induced lengthening, elongation before an Intonational Phrase (IP), and deaccentuation after the questioned phrase.

 LMER analysis shows that, when produced with a rising F0 contour, polar *kya* has significantly longer relative duration as compared with its deaccented variants. This indicates AP based lengthening of polar *kya* produced with a rising contour. When produced immediately before an IP boundary, polar *kya* also exhibits pre-boundary lengthening.

 Moreover, polar *kya*, at the medial and immediately preverbal positions, is most frequently deaccented when placed after the questioned constituents (50% for medial; 64% for preverbal). Thus, the deaccentuation of polar *kya* results from its position after the questioned constituent and not due to its inherent accentlessness.

 This study shows that polar *kya* is not inherently weak and behaves prosodically like a regular AP and is subject to phonological processes. Its prosodic realization as an AP vs. its accentlessness results from its position with reference to the questioned constituent. These findings have implications for our overall understanding of weak elements and their prosodic status as resulting from factors other than their inherently weak status.

What makes grammatical words "weak"? Disentangling semantic, morphosyntactic and prosodic factors via language-music mapping

Christina Domene Moreno, Barış Kabak
University of Würzburg
christina.domene-moreno@uni-wuerzburg.de, baris.kabak@uni-wuerzburg.de

Based on their lexico-semantic, syntactic and information-structural properties, "low prominence" and "dependency" of grammatical words in comparison to lexical words have led to hypotheses that predict disparate behavior of these two in language production. These differences are evidenced by different omission rates or the order in which such words are accessed in pre-articulatory planning (Michel Lange et al. 2017). In phonology, grammatical words are assumed to be prosodically "weak" in some languages, suggested by their shorter duration and stresslessness, leading to their reduction in speech production and their cliticization. It is difficult to tease apart whether the source of their subordinate behavior in processing is due to their low *prosodic* or their low *lexico-semantic/information-structural* prominence. In this paper, we hypothesize that, if the differences between grammatical and lexical words are reflected in speech planning, we would also see differences in the way these words are mapped onto musical structures. This is because meter and melody in lyrical music have been shown to mirror prosodic prominence such as stress even in languages where the status of stress is dubious (e.g., Domene Moreno and Kabak, to appear). In this vein, Temperley & Temperley (2013) show that monosyllabic content (lexical) words are more likely to occur at strong positions than monosyllabic function (grammatical) words in French songs. This finding is, however, confounded by the stresslessness of function words in French. Indeed, Temperley & Temperley found the same pattern for stressed syllables as opposed to unstressed syllables in content words in French. Therefore, it remains to be seen whether the manifestation of prosodic weakness in music is modulated by the lexical vs. grammatical word dichotomy irrespective of accentual prominence. We test this by comparing *unstressed* syllables in lexical words to grammatical words with regard to their metrical and tonal targets in music. Using English children's songs (n=20), we investigate the relative degree of alignment in meter and melody, operationalized as metrical prominence (MM) and melodic peak (MP) respectively. We hypothesize that, since grammatical words possess no significant content, they may receive less prominence in music. On the other hand, as grammatical words are morphosyntactic "words", they are more likely to receive higher MM and/or MP scores than unstressed syllables in lexical words, our second hypothesis. Preliminary results based on a small sample of songs (n=5) reveal a trend towards robust differences in mean melodic peak between unstressed syllables in lexical words (3.94) and grammatical words (3.62), but no such trend for metrical prominence. In our study, we will corroborate our findings based on another type of prominence mapping: Music manifests higher pragmatic/emotional prominence in language via *duration*. Important words can be highlighted by longer notes in sacred harp music (Kelley 2017). Thus, unstressed syllables from lexical words may be more likely to be mapped onto long notes than those from grammatical words since grammatical words are ancillary to lexical words due to their lower semantic content.

References: Domeno Moreno, C. & Kabak, B. (to appear). Meter, melody, and stress in song: Theorizing prosody-music alignment through cross-linguistic evidence. For: M. Scharinger and R. Wiese (eds.), *Prosody from a Cross-Domain Perspective: How Language Speaks to Music (and Vice Versa)*. Kelley, R. T. (2017, December 31). Bound together: What makes an effective pairing of text and tune [Blog post]. Retrieved from http://originalsacredharp.com/2017/12/31/bound-together-makes-effective-pairing-text-tune/#footnote_5_8131. Michel Lange, V., Messerschmidt, M., Harder, P., Siebner, H. R., & Boye, K. (2017). Planning and production of grammatical and lexical verbs in multi-word messages. *PloS one*, *12*(11), e0186685. https://doi.org/10.1371/journal.pone.0186685 Temperley, N., and Temperley, D. (2013). Stress-meter alignment in French vocal music. *Journal of the Acoustical Society of America*, 134:1, 520-527.

The acquisition of weak elements: Lexical, morphological, and prosodic considerations

Katherine Demuth
Macquarie University katherine.demuth@mq.edu.au

Researchers have long realized that children omit certain 'weak' syllables in their everyday speech. This talk explores some of the patterns of early weak syllable omission found various lexical and morphological domains, showing how this shows systematic change over time. It then shows how the Prosodic Licensing Hypothesis can help make predictions regarding how this process will evolve, given the specific prosodic characteristics of a language at both the lexical and morphological level. Finally, it concludes by discussing experimental methods that may be useful in more systematically probing the development of speech planning processes. The findings have implications for both developmental theories of linguistic competence and for clinicians working with children exhibiting various types of language delay.

Schwa syllables in early language acquisition and speech and language disorders

Christina Kauschke, Ulrike Domah, Angela Grimm
University of Marburg, University of Marburg, University of Frankfurt
kauschke@uni-marburg.de, domahsu@uni-marburg.de, A.Grimm@em.uni-frankfurt.de

Word-final weak syllables form a part of the canonical trochee, which is a basic and dominant pattern of phonological words in languages such as English and German. In German, the production of weak syllables with schwa-vowels is fundamental for the realization of the canonical trochee. It is an understudied question whether or not the production of schwa in early phonological development is a demanding task, given that much work on the acquisition of the German phonological system has focused on consonants, rather than on vowels. Nevertheless, the acquisition of schwa syllables is a crucial step in the acquisition of the German phonological system and a prerequisite for lexical development.

For typically developing children, Kehoe & Lleó (2003) observed that children initially omit schwa or replace it by full vowels before they produce trochees with final schwa- syllables around age 2;6. This indicates a shift from initial augmentation to qualitative reduction. For children with atypical speech and language development, a few single case studies report severe problems in producing trochees with final schwa (Dümig & Frank, 2008, Kauschke, 2018).

In order to find further evidence for the importance of schwa syllables in typical and atypical speech and language development, we aimed at investigating the production of words with final schwa syllables in children with and without speech sound disorders in more detail by means of two sample studies. First, the corpus of Grimm (2008) - containing dense longitudinal production data of four typically developing children aged 1;2 to 2;1 - was analyzed with respect to the realization of schwa-vowels in disyllabic words. Results show that during their second year of life, children replaced schwa vowels by full vowels such as [a], [ɛ] or [i], before they produced schwa-vowels correctly at more than 70% by age 2, in line with Kehoe & Lleó (2003). The second study focused on three boys with speech sound disorders (aged 3;5, 3;7, and 4;6). A detailed analysis of their word productions, obtained by elicited speech production or imitation, revealed severe problems with schwa: the proportion of correct schwa vowels in weak final syllables ranged between 13% and 18% only. Schwa vowels were substituted predominantly by [a] in one child (['tɪsa] for Spritze ['ʃpʁɪtsə] 'injection') or [ɛ] in the other two children (['kɪkɛ] for Spritze).

The combined results of the two production studies suggest that augmentation of schwa syllables seems to be an intermediate stage of early speech development in German that may persist in children with severe phonological impairment. These children stagnate on very early developmental stages, impeding further progress in their speech and language abilities.

References: Dümig, S., & Frank, A. (2008). The syllable and schwa in first language acquisition: Normal and impaired development. *Frankfurter Linguistische Forschung, Sondernummer 11*, 65- 90. Grimm, A. (2008). The development of word-prosodic structure in child German: Simplex words and compounds (Dissertation). Universität Potsdam, Potsdam. Kauschke, C. (2018). Wortbetonung als Einstieg in die Therapie von Sprachentwicklungs- störungen. *Logopädieschweiz,* 3, 4-11. Kehoe, M., & Lleó, C. (2003). A phonological analysis of schwa in German first language acquisition. *Canadian Journal of Linguistics, 48*, 289-3.

Geolinguistic differences of reductions in standard intended German due to a rise of speech rate

Beat Siebenhaar
Universität Leipzig
siebenhaar@uni-leipzig.de

From 2006-2009 the Institut für deutsche Sprache Mannheim made recordings with 644 high school graduates in a total of 161 locations throughout the whole area where German is official and teaching language. One part of these recordings was the lecture of the "north wind and the sun". After the first reading, the participants were asked to read the text again, at a faster tempo. From these texts we wanted to obtain linguistic-geographical distributions of speech rate and reductions due to the increases in speech rate. Our first submission to the DFG was rejected, partly because the DFG review board considered it difficult to filter out diatopic variation from the reading data. Our second application was then accepted and I would like to use this presentation to show how much diatopic variation can be read out of the reading language.

First of all, it can be shown that the time needed to read the texts already shows a geolinguistic distribution. Moreover, the differences in the patterns between slow and fast reading aloud also indicate regional distributions of the reductions, which can be seen already in the pure number of sounds reproduced (Hahn & Siebenhaar 2019).

The measuring zero point for our study is a canonical pronunciation as given in the pronun- ciation dictionary (Duden-Aussprachewörtebuch, 6th edition of 2005 – the reference at the time the project started). In comparison to this canonical pronunciation, the recordings show regionally different areas of reduction, whereby the differences are even more obvious in many – but by far not all – cases due to the increase in speaking rate. It is particularly exciting to see that the strategies for increasing speech rate differ regionally. The three basic possibilities – 1) omitting sounds, 2) shortening sounds, 3) making sounds less precise – are used differently, which on the one hand depends on the respective different starting posi- tions, but on the other hand also shows sound-specific variations. This results in a patchwork of reduction phenomena across the entire German-speaking area, which, though, can be tied into an overall picture of various trends (Hahn in preparation). It is astonishing, however, that in addition to the three possibilities of reduction mentioned above and expected, a fourth pos- sibility is also used, namely a compensation, a more precise realisation of sounds. This is particularly evident in the realisation of vowels (Siebenhaar & Hahn 2019).

All in all, the linguistic geography of the read standard German shows that almost all sounds can be weak elements, either by omitting them completely, or by reducing them quantitatively or qualitatively. It is well known that some sounds react more sensitively to such reductions than others, however, the classification often shows geolinguistic distribution patterns.

References: DUDEN 6. *Aussprachewörterbuch*. 2005. 6. edt. Mannheim, Leipzig, Wien, Zürich: Dudenverlag. Hahn, Matthias. In preparation. Hahn, Matthias, and Beat Siebenhaar. 2019. "Spatial Variation of Articulation Rate and Phonetic Reduction in Standard-Intended German". In Calhoun, Sasha, Paola Escudero, Marija Tabain and Paul Warren, eds. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia*. Melbourne. 2695–2699. Siebenhaar, Beat, and Matthias Hahn. 2019. "Vowel space, speech rate and language space". In Calhoun, Sasha, Paola Escudero, Marija Tabain and Paul Warren, eds. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia*. Melbourne. 879–883.

Getting rid of the German canonical trochee in L3 French intonation: Comparing monolingually raised German and bilingual Turkish-German learners

Christoph Gabriel, Jonas Grünke, Nils Karsten
JGU Mainz, JGU Mainz, Universiteit van Amsterdam
christoph.gabriel@uni-mainz.de, jgruenke@uni-mainz.de, nils.karsten@student.uva.nl

Our contribution addresses the question of how monolingually raised German learners and bilingual Turkish-German learners of French as a foreign language acquire the intonational system of the target language. The three languages involved differ considerably at the prosodic level: German intonation is word-based and determined by pitch accents realized on the metrically strong syllables of lexically stressed words (Féry 1993). Regarding weak elements, it furthermore presents a contrast between unstressed and reduced syllables (e.g. *Tuba* [ˈtu.**ba**] 'tuba' vs. *Tube* [ˈtu.**bə**] 'tube'); the basic unit at the foot level is the so-called canonical trochee. French, by contrast, completely lacks lexical stress and presents a phrase-based intonation system, which is characterized by the obligatory marking of phrase- final syllables by means of a pitch accent; in addition, the left edge of the French Accentual Phrase (AP) can be signaled through a facultative initial F0 rise (Delais-Roussarie et al. 2015). Turkish, finally, occupies an intermediate position between German and French as it presents at least some words that display lexical stress (Göksel & Kerslake 2015: 26–39), but in the unmarked case assigns stress to the last syllable of prosodic words, which are marked by an initial L tone plus a final rise (Kamalı 2011; İpek & Jun 2013). This forms a striking parallel with the initial L and the final H of the French AP. As follows from these characteristics, global rhythmic patterns emerging from alternations between metrically strong and weak syllables are similar in French and Turkish, whereas German is characterized by a higher density of strong syllables. As a consequence, (monolingually raised) German learners of French must learn to suppress the assignment of pitch accents to content words in positions that are weak (i.e. not phrase-final) in the target language. Turkish-German learners, by contrast, should outperform German monolinguals when acquiring the weak elements of French intonation due to potential positive transfer from their heritage language, Turkish.

To test this assumption, we analyze read data (passage from a French textbook) recorded by 6 bilingual Turkish-German learners (ages: 15–17) who speak Turkish as a HL along with German. Monolingually raised German learners (n=8, ages: 15–17) as well as monolingual L1 speakers of Northern Standard French (n=3, ages: 21–23) serve as control groups. The analysis of syllable strength in our corpus shows that both the bilingual and monolingual learners differ from the French L1 speakers according to the prominence values assigned to each syllable by the software ANALOR (Avanzi et al. 2008; calculation based on acoustic parameters that are considered relevant for French). The scores expressing the deviation of the learners' productions from the French target values showed no significant difference between the two learner groups (p=.803), although the distribution of prominences in the bilingual data is closer to the L1 model regarding the phrasing of non-complex sentences.

As opposed to recent work on VOT production in Turkish-German learners of FFL (Gabriel et al. 2018), our expectations of positive transfer were thus only partly met for intonation. This suggests that suprasegmentals are less accessible to foreign language learners than segmental properties. In other words: Once a person has acquired the canonical trochee of L1 German, he or she is confronted with severe problems in getting rid of this property when learning a foreign language such as French, which displays a completely different intonation system. Positive transfer from the prosodic system of Turkish seems to require support by fostering prosodic awareness in the multilingually raised learners.

References: Avanzi, Mathieu, Anne Lacheret-Dujour, and Bernard Victorri. 2008. "ANALOR. A tool for semi-automatic annotation of French prosodic structure." In Plínio A. Barbosa, Sandra Madureira, and César Reis, eds. *Speech Prosody 2008.* Campinas: ISCA, 119-122. Delais-Roussarie, Elisabeth, Brechtje Post, Mathieu

Avanzi, Caroline Buthke, Albert di Cristo, Ingo Feldhausen, Sun-Ah Jun, Philippe Martin, Trudel Meisenburg, Annie Rialland, Rafèu Sichel-Bazin, and Hi-Yon Yoo. 2015. "Intonational Phonology of French. Developing a ToBI system for French." In Sónia Frota and Pilar Prieto, eds. *Intonation in Romance.* Oxford: Oxford University Press, 63-100. Féry, Caroline 1993. *German intonational patterns.* Tübingen: Niemeyer. Gabriel, Christoph, Marion Krause, and Tetyana Dittmers. 2018. "VOT production in multilingual learners of French as a foreign language." *Revue française de linguistique appliquée* 23, 59-72. Göksel, Asli, and Celia Kerslake. 2005. *Turkish. A comprehensive grammar*. London Routledge. İpek, Canan, and Sun-Ah Jun. 2013. "Towards a model of intonational phonology of Turkish. Neutral intonation." *Proceedings of Meetings on Acoustics 19* (ICA 2013 Montréal). Kamalı, Beste 2011. *Topics at the PF interface of Turkish*. PhD Diss., Harvard University.

The influence of rhythm on placing the German object pronoun

Isabelle Franz[1,2], Markus Bader[2] & Gerrit Kentner[1,2]
[1]Max-Planck-Institute for Empirical Aesthetics, [2]Goethe University Frankfurt
isabelle.franz@ae.mpg.de

What determines the variable placement of the object pronoun in the German middlefield? With three experiments, we assess the effect of linguistic rhythm on word order.

Rhythmic influences on word order have been documented (see, e.g. Anttila 2016 for a review), but they appear to be rather limited in scope (Kentner & Franz 2019). Here, we study word order preferences in embedded complementizer clauses (1). We systematically varied the stress pattern of the embedded subject (iambic, trochaic name) and of the embedded verb (initial stress, no initial stress). The examples in (1) show that, with the iambic subject (Marcél), the fronted object pronoun 'ihn', him, (underlined) yields a rhythmically less balanced structure than with the trochaic subject in (Márkus). Thus, we predict that trochaic embedded subjects (and verbs with no initial stress) promote sentences with a fronted pronoun.

(1a)   Der Júnge ságt, dass ihn Márkus/Marcél belúgt/áuslacht. (OS)
(1b)   Der Júnge ságt, dass Márkus/Marcél ihn belúgt/áuslacht. (SO)
        The boy says that Markus/Marcel is lying to / laughing at him.

The questionnaire with written stimuli like (1) confirmed that the non-canonical order (1a, with the object pronoun preceding the subject) becomes more acceptable when rhythmically well-formed. We are currently analysing data of two sentence production experiments (spoken and written, 64 stimuli as in Figure 1, 50 participants). Results in line with the questionnaire study would strengthen a bidirectional account of syntactic and phonological processing in sentence formation (Breiss & Hayes 2020).

Data will be analysed using general linear mixed effects models. The dichotomous response variable is the word order produced (SO vs OS). Predictor variables are the stress patterns of embedded subject and embedded verb. We will also analyse syllable duration, pitch and intensity (see Vogel et al., 2015, who showed phonetic effects of the rhythmic context for the unstressed pronoun 'es', it, in German, arguably provoked by the Rhythm Rule).
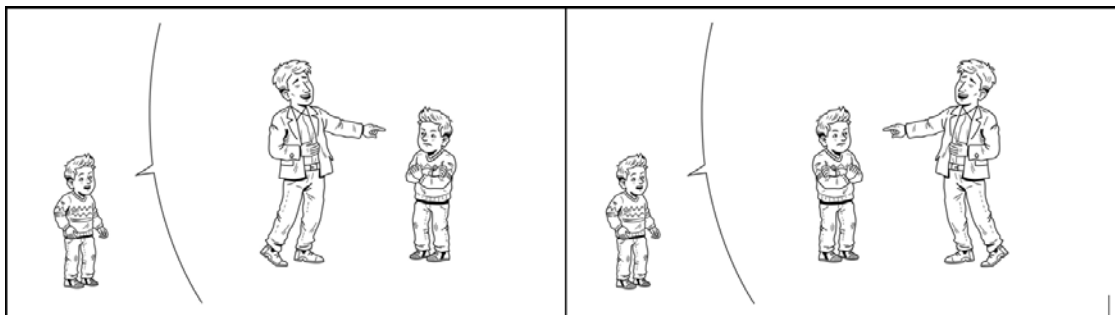

Figure 1. Stimulus example for a target sentence: 'Der Junge sagt, dass (ihn) Markus (ihn) auslacht.',*The boy says that Markus is laughing at him*. The left and right panels show the stimuli in their mirrored versions.

References: Anttila, A. (2016). Phonological effects on syntactic variation. Ann. Review of Linguistics 2. 115–137. Breiss, C., & Hayes, B. (2020). Phonological markedness effects in sentence formation. Language, 96(2), 338-370. Kentner, G., & Franz, I. (2019). No evidence for prosodic effects on the syntactic encoding of complement clauses in German. Glossa: a journal of general linguistics, 4(1), 1-29. Vogel, R. et al. (2015). Function words in rhythmic optimisation. In Ralf Vogel & Ruben van de Vijver (eds.), Rhythm in cognition and grammar: A Germanic perspective, 255–276. Berlin: De Gruyter.

Arbeitsgruppe 3

The semantics of derivational morphology: Theory, methods, evidence
Sven Kotowski & Ingo Plag

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

The semantics of conversion nouns and *-ing* nominalizations: A quantitative and theoretical perspective

Rochelle Lieber, Ingo Plag
University of New Hampshire, Heinrich Heine Universität Düsseldorf

This paper addresses a fundamental problem of derivational morphology: which meanings are possible for the words of a given morphological categories, which forms can be chosen to express a given meaning, and what is the role of the base in these mappings of form and meaning? In a broad empirical study we examine the extent to which two types of nominalizations in English – conversion nouns and *-ing* nominalizations – can express either eventive or referential readings, can be quantified as either count or mass, and can be based on verbs of particular aspectual classes (state, activity, accomplishment, achievement, semelfactive). Past literature (for example, Brinton 1995, 1998, Grimshaw 1990, Borer 2013) has suggested an association between conversion nominalization, count quantification, and referential reading on the one hand, and between *-ing* nominalization, mass quantification and eventive reading on the other. Using a subset of the data reported in Andreou & Lieber (2020), we give statistical evidence that the relationship between morphological form, type of quantification, and aspectual class of base verb is neither categorical, as the literature suggests, nor completely free, but rather is probabilistic. We provide both a univariate analysis and a multivariate analysis (regression and conditional inference trees) that show that the relationship among the variables of morphological form, eventivity, quantification and aspectual class of base is complex. Tendencies sometimes go in the direction suggested by past literature (e.g. *-ing* forms tend to be eventive), but sometimes contradict past predictions (conversion also tends to be eventive). We also document that an important role is played by the specific verb underlying the nominalization rather than the aspectual class of verb. Finally, we consider what the pattern of polysemy that we uncover suggests with respect to theoretical modeling, looking at syntactic models (Distributed Morphology), lexical semantic models (the Lexical Semantic Framework), Analogical Models, and Distributional Semantics.

The semantics of conversion nouns and *-ing* nominalizations: A quantitative and theoretical perspective

Assessing the rivalry between French deverbal neologisms in *-age, -ion* and *-ment*

Richard Huyghe, Alizée Lombard, Justine Salvadori, Sandra Schwab
University of Fribourg (Switzerland)
surname.name@unifr.ch

Rivalry between the three French nominalizing suffixes *-ion*, *-age*, and *-ment* has received a fair amount of attention in the last decade and is still a source of debate (see Martin 2010, Uth 2010, Wauquier et al. 2018, Fradin 2019, among others). While some authors consider that *-age*, *-ion* and *-ment* can be distinguished according to various factors, others argue that they are possibly overabundant suffixes related to the same underspecified semantic construction.

Most existing research examines morphological doublets (i.e. alternate nouns derived from the same base) and/or lexically well-established nouns. The present study focuses on a sample of neologisms ending in *-age*, *ion* or *-ment* that do not necessarily occur in doublets, insofar as lexicalized words may undergo semantic bleaching or idiosyncratic specialization that is independent of derivational semantics. The goal is to determine whether there are significant semantic differences between the three suffixes with regard to verbal input and nominal output properties.

The study is based on a sample of 300 deverbal candidates (100 per suffix) extracted from the frCOW16A corpus (Schäfer & Bildhauer 2012). Each pair of derived nouns and base verbs is annotated according to properties that could influence the rivalry between the three suffixes: transitivity, lexical aspect, semantic role assignment for verbal inputs; and semantic type, lexical aspect, semantic role assignment for nominal outputs. Assuming that derivational processes apply to semantically specified items, verbal and nominal lexemes are paired based on closest semantic proximity. The annotation of verb-noun pairs is conducted double blindly, and cases of disagreement are adjudicated with the help of a third annotator.

The influence of the annotated properties on suffix selection is statistically evaluated through conditional inference trees and random forests. The analysis of possible relationships between suffix selection and specific factors (e.g., verbal transitivity or agentivity, nominal semantic types) is refined by means of multinomial logistic regressions. A special attention is paid to the polysemous capacities of each suffix, as well as to their ability to form nouns that preserve or diverge from the semantic properties of the base verbs.

References: Fradin, Bernard. 2019. "Competition in Derivation: What Can We Learn from French Doublets in -Age and -Ment?" In Franz Rainer, Francesco Gardani, Wolfgang U. Dressler, and Hans Christian Luschützky, eds. *Competition in Inflection and Word- Formation.* Cham: Springer. 67–93. Martin, Fabienne. 2010. "The Semantics of Eventive Suffixes in French." In Monika Rathert and Artemis Alexiadou, eds. *The Semantics of Nominalizations across Languages and Frameworks*. Berlin, New York: De Gruyter Mouton. 109–40. Schäfer, Roland, and Felix Bildhauer. 2012. "Building Large Corpora from the Web Using a New Efficient Tool Chain." *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*: 486–493. Uth, Melanie. 2010. "The Rivalry of French -Ment and -Age from a Diachronic Perspective." In Monika Rathert and Artemis Alexiadou, eds. *The Semantics of Nominalizations across Languages and Frameworks*. Berlin, New York: De Gruyter Mouton. 215–44. Wauquier, Marine, Cécile Fabre, and Nabil Hathout. 2018. "Différenciation sémantique de dérivés morphologiques à l'aide de critères distributionnels." *SHS Web of Conferences* 46: 08006.

The meanings of nominal vs. verbal zero affixes

Verginica Barbu Mititelu, Svetlozara Leseva, Ivelina Stoyanova, Gianina Iordăchioaia
RACAI, IBL, IBL, University of Stuttgart
vergi@racai.ro, zarka@dcl.bas.bg, iva@dcl.bas.bg, gianina@ifla.uni-stuttgart.de

We study verb-noun zero-derivation in English from three perspectives: (i) the semantic relations between the members of zero-related and suffix-marked N-V pairs; (ii) the semantic classes that the Ns and Vs belong to, and (iii) the derivational direction inside a pair.

Resources and preliminary work. The study relies on two large lexicographic resources: the Princeton WordNet (PWN) and the Oxford English Dictionary (OED). PWN provides information about the semantics of 16,995 V-N pairs: they are labeled with one of a set of 14 morphosemantic relations (Fellbaum et al. 2009). Moreover, PWN can be harnessed so as to identify semantic clusters that favour the existence of derivationally (zero or affixal) related pairs reflecting a certain semantic relation. We use the OED to complement these data with information about the direction of the zero-derivation. The OED data comprise 5,921 N-to-V and 2,830 V-to-N conversion pairs. While PWN is known for dealing with word senses rather than words, such that all derivational pairs are established between words considered in one of their (clearly identifiable) meanings, OED also treats etymology at the sense level, where appropriate: a different derivational direction is indicated in those cases when a different sense of the word had a different evolution from the others. A semi- automatic analysis of the data showed that the zero affix prevails with some semantic relations (Result, Property, Undergoer, etc) and is the main competitor of the prevailing affix in the case of other relations (Agent, Material, etc). For different relations, semantic clusters (in the form of subtrees from the PWN hierarchy) can be identified for zero-derivation: e.g., a cluster of verbs of the semantic type *creation* in the Result relation with nouns of type *artifact*.

Theoretical hypotheses. We test two hypotheses on zero-derivation: (H1) Some N- V pairs show directionality ($cage_N$ > $cage_V$ vs. $chase_V$ > $chase_N$), while others do not (spy_V – spy_N; see Darby & Lahiri 2016); (H2) V-to-N conversion is semantically more systematic than N-to-V conversion (Kisselew et al. 2017). H1: Directionality in zero-derivation is often unclear (see Plag 2003 for an overview), which would support non-derivational analyses (Lieber 2004, Borer 2013). However, recent psycholinguistic studies argue that zero-derived words are morphologically complex like suffixal derivations, although some N-V pairs are categorially underspecified (Darby & Lahiri 2016). H1 predicts some PWN semantic relations to be specific to nominalization (V-to-N) or verbalization (N-to-V) – in comparison with suffix-based nominalizations and verbalizations – while others will be underspecified. For H2, Kisselew et al. (2017) found zero-derived nouns to show more regular semantics in relation to their bases than zero-derived verbs do. This resonates with theoretical proposals on zero- derived verbs as independent of the noun base in interpretation but not with similar proposals on zero-derived nouns (Borer 2013). Comparing PWN semantic relations for zero and overt nominalizing/verbalizing suffixes will allow us to test if the zero nominalizer resembles overt nominalizers more than zero verbalizers resemble overt verbalizers.

References: Borer, Hagit. 2013. *Taking Form*. Oxford: Oxford University Press. Darby, Jeannique & Aditi Lahiri. 2016. Covert morphological structure and the processing of zero-derived words. *The Mental Lexicon* 11.2: 186-215. Fellbaum, Christiane, Anne Osherson, and Peter E. Clark. 2009. "Putting Semantics into WordNet's "Morphosemantic" Links". In *Proceedings of the 3rd Language and Technology Conference*, Springer LNCS, vol 5603, 350-358. Kisselew, Max, Laura Rimell, Alexis Palmer, and Sebastian Padó. 2016. Predicting the direction of derivation in English conversion. In *Proceedings of the ACL SIGMORPHON workshop*, 93-98. Berlin. Lieber, Rochelle. 2004. *Morphology and Lexical Semantics*. Cambridge University Press: Cambridge. Plag, Ingo. 2003. *Word-Formation in English*. Cambridge: Cambridge University Press.

Events in the semantics of non-deverbal nominalizations

Viktoria Schneider
Heinrich-Heine-Universität Düsseldorf
Viktoria.Schneider@hhu.de

Research on nominalizations in which events are involved focuses on verbal bases, for example, *employ* as the base verb for *employee* or *reassure* as the base verb for *reassurement* (Barker 1998; Kawaletz & Plag 2015). However, nominalizing affixes often also attach to non-verbal bases (Bauer et al. 2013; Plag 1999, 2004). In the case of deverbal derivatives, the event is provided by the semantic representation of the base verb (see, for example, Plag et al. 2018; Löbner 2013). In contrast, with many non-verbal base words the derived noun cannot straightforwardly inherit an event from its base.

The problem is illustrated in (1). In (1a), for example, it is not immediately clear what kind of an event, if any, *debt* bequeaths. Similar problems arise with the bases in the example words in (1b), where it is unclear where the nature of the event and its representation would be inherent in the semantic representation of the respective base words.

(1)   a.   *The '**debtee**' is an old word for the creditor or payee.* (BNC)
        b.   *biographee, bankment, personage, astronomer* (BNC)

My goal is to investigate how non-verbal bases can be used in nominalizations to generate derivates which involve an event. This aim is reached by a detailed analysis of the semantic structure of the base words. It is possible to identify eventive elements in non-verbal bases and to show how these events and their participants are used by nominalizing suffixes to create the reading of the derivative. For example, my analysis of the noun *debt* as the base for the derivative *debtee* shows that actually two interrelated events, an *obligation-state* and a *paying-action*, are inherent in the base word. The reading of the derivative *debtee* in (1a) indicates a meaning shift from the base *debt* to the RECIPIENT of the *paying-action*.

I use frame semantics (see, for example, Löbner 2013: ch. 12) to model the inherent semantic representation of non-verbal base words and the meaning creation by the suffix (Plag et al. 2018; Andreou submitted). In this talk, I focus on non-deverbal nominalizations with nominal bases and the suffix *-ee*.

References: Andreou, Marios. submitted. Comparison-based modification in derivational morphology: diminutives, augmentatives and stereotype negation. https://frames.phil.uni- duesseldorf.de/wp-content/uploads/2020/01/comparison_based_modification.pdf. Barker, Chris. 1998. Episodic -ee in English: A Thematic Role Constraint on New Word Formation. *Language* 74(4). 695. Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford reference guide to English morphology.* Oxford: Oxford Univ. Press. Davies, Mark. 2004-. British National Corpus (from Oxford University Press). https://www.english-corpora.org/bnc/. Kawaletz, Lea & Ingo Plag. 2015. Predicting the Semantics of English Nominalizations: A Frame-Based Analysis of –ment Suffixation. In Laurie Bauer, Lívia Körtvélyessy & Pavol Štekauer (eds.), *semantics of complex words*, 289–319. Dordrecht: Springer. Löbner, Sebastian. 2013. *Understanding semantics,* 2nd edn. (Understanding language series). London: Routledge. Plag, Ingo. 1999. *Morphological Productivity: Structural Constraints in English Derivation* (Topics in English Linguistics /TiEL] 28). Berlin: De Gruyter. Plag, Ingo. 2004. Syntactic category Information and the semantics of derivational morphological rules. *Folia Linguistica* 38(3-4). Plag, Ingo, Marios Andreou & Lea Kawaletz. 2018. A frame-semantic approach to polysemy in affixation. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer (eds.), *The lexeme in descriptive and theoretical morphology*, 467– 486. Berlin: Language Science Press.

Locative prefixes and nominal scalarity

Sven Kotowski
Heinrich-Heine-Universität Düsseldorf
sven.kotowski@uni-duesseldorf.de

English locative prefixes such as *out-*, *over-*, or *under-* give rise to scalar-quantificational interpretations. As in (1)-(3), different such readings are attested cross-categorically (cf. Bauer et al. 2013: ch.16):

(1) **excess nouns** (verbs, adjectives)
overrespect; over-self-esteem (overcharge; overfond)

(2) **insufficiency nouns** (verbs, adjectives)
undernutrition; undergovernment (undereducate; underripe)

(3) **comparative verbs**
out-dollar s.o.; out-technology s.o. (overbid s.o.; underbake sth.; outrun s.o.)

Most studies on such forms focus on verbal structures, while little is known about both nominal bases and derivatives. Given the multitude of clearly nominal structures, this lack of studies on the interplay between (non-)scalar nouns (e.g. Morzicky 2009) and affix semantics is unfortunate. Looking at scalar information encoded in denominal derivatives, this paper aims at bridging this gap.

The study is based on roughly 2,000 corpus attestations of (de)nominal *out-*, *over-*, and *under-* derivatives (mainly from COCA; see Davies 2008). It focuses on the following: (i) what kind of scalar information can we derive from nouns, (ii) which kinds of noun (do not) occur as bases, and (iii) how does the constructional semantics of the word-formation processes interact with (i) and (ii) as well as contextual and/or world knowledge? Answers to these questions are not straightforward. Consider (4) and (5):

(4) Between 30 and 40 drug bosses have carved up the city and easily outgun the police with their arsenals. (OED)

(5) If he has a rifle and you try to outgun him with a pocket pistol rather than run for cover you are stupid. (iWeb)

An object noun such as *gun* serves as base for comparative *out-* and, as retrieved via contextual information, gives rise to two diverging scale-based interpretations: construal of a cardinality-scale as in (4), i.e. 'have more guns', and of a property-scale (with the dimension QUALITY) as in (5), i.e. 'use the gun better'.

I will present clear differences between the derivational processes. Comparative *out-* is highly promiscuous, allows for abstract and concrete as well as mass and count nouns as base, and gives rise to both cardinality and property scales. In contrast, the by far better part of (de)nominal *over-* and *under-*derivatives are based on abstract, scalar nouns or feature bases that are either deverbal or deadjectival themselves. I will model one process, *out-* prefixation, as a constraint-based lexeme-formation rules in frame semantics (cf. e.g. Plag et al. 2018) and account for its category-changing behaviour.

References: Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. The Oxford reference guide to English morphology. Oxford University Press, Oxford. Davies, Mark 2008. The Corpus of Contemporary American English: 400+ million words, 1990-present. Morzicky, Marcin. 2009. "Degree Modification of Gradable Nouns: Size Adjectives and Adnominal Degree Morphemes." *Natural Language Semantics* 17: 175–203. Plag, Ingo, Marios Andreou & Lea Kawaletz. 2018. A frame-semantic approach to polysemy in affixation. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer, eds. *The lexeme in descriptive and theoretical morphology*. Berlin: Language Science Press, 546–568.

The meaning of derivation: Relations and scenarios

Olivier Bonami, Louise McNally, Denis Paperno
U. Paris, U. Pompeu Fabra, U. Utrecht
olivier.bonami@gmail.com, louise.mcnally@upf.edu, denis.paperno@gmail.com

Two major trends in the recent history of theoretical morphology are the rise of interest in abstractive Word and Paradigm approaches (WP; see a.m.o. Blevins 2016) and paradigmatic approaches to word formation (see a.m.o. Bonami & Strnadová 2019).

Abstractive WP approaches put concrete words at the center of morphological theory; abstractions such as morphemes (and possibly lexemes) are secondary, if not superfluous. Paradigmatic approaches highlight the prevalence of morphological relatedness that does not reduce to the application of a productive process to a determinate base, and question the usefulness of such constructs both for theory and insightful description.

While this literature has successfully delineated a general architecture for, and modeled formal aspects of, morphological systems, it has so far failed to address semantic issues in detail. Work on derivational semantics still consistently makes assumptions at odds with WP-based approaches. Our goal is to show how recent efforts to connect conceptual and referential semantics (e.g. McNally & Boleda 2017), combined with a probabilistic approach to language use and interpretation, can be used to reconceptualize derivational semantics and more generally the semantic side of morphological relatedness.

Paradigms in derivational morphology are intuitively appealing but challenging to motivate semantically. We propose to ground morphological relatedness in *scenarios*, inspired in Fillmorean frames and defined by Erk & Herbelot (2020) as "larger settings" encompassing entities and events under a specific conceptualization. Pairs of related words are interpreted with respect to some scenario, with each word picking out one or more concepts, defined as representations used to categorize referents. Scenarios can then be combined with contextual factors to modulate concepts. We hypothesize that scenarios play a role in constraining affix-related polysemy.

Morphological families can be exploited semantically via two further assumptions. First, successful communication only requires that interlocutors coordinate on expressions at an intermediate level of granularity on which individual referents are paired with concepts, without precisely associating scenario-, concept- or referent-identifying functions with atomic parts of these expressions. Interpretation can be based on probabilistic inference from utterances, including parts that correspond to traditional morphemes. Our proposal is compatible with state of the art computational models built upon the neural attention mechanism. For example, an image captioning system of Cornia et al. (2020) attends to information from parts of the image as well as to parts of the previously generated text.Different parts of the material attended to can correspond to the three types of semantic objects: referents, categories, and scenarios; the parts might include word pieces but may also overlap. Second, we assume that complex word interpretation involves probabilistic reasoning based on relevant alternatives within morphological families.

We close by pointing to prospects for recasting existing accounts of derivational semantics within the assumptions of WP approaches.

References: Blevins, James P. 2016. *Word and paradigm morphology*. Oxford: Oxford University Press. Bonami, Olivier and Jana Strnadová. 2019. "Paradigm structure and predictability in derivational morphology." *Morphology* 29(2): 167-197. Cornia, Marcella, Matteo Stefanini, Lorenzo Baraldi and Rita Cucchiara. 2020. Meshed- memory transformer for image captioning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10578-10587. Erk, Katrin and Aurelie Herbelot. 2020. "How to marry a star: probabilistic constraints for meaning in context." arXiv:2009.07936. McNally, Louise and Gemma Boleda. 2017. "Conceptual vs. referential affordance in concept composition." In Yoad Winter and James Hampton, eds., *Compositionality and concepts in linguistics and philosophy*. Berlin: Springer. 245-267.

The pragmatics of word formation: A case study on German *stoff

Olav Mueller-Reichau, Matthias Irmer
University of Leipzig, OntoChem GmbH, Halle/Saale
reichau@uni-leipzig.de, matthias.irmer@ontochem.com

"The semantic interpretation of compounds is determined by pragmatic factors valid at the point of creation of the compound" ([1]:169). Hearer-friendliness requires a new category name to provide a (partial) description of the bearer of the name. Speaker-friendliness calls for a phonological string as short as possible. Successful reference is best maintained if the new name unequivocally describes a unique category. To meet these pragmatic requirements, the head of an (endocentric) compound is made up of a hypernym of the target category, while its modifier narrows down some of its aspects.

To pin down the pragmatic scaffold that guides speakers' behaviour, we work out the conceptual script knowledge [6, 3, 2] involved in German compounds headed by *Stoff* ("substance", "cloth"), and show how *Stoff*-prefixed compounds can be clustered into groups related to different aspects of that script. We claim that it is the script which mediates between the orthogonal needs of interlocutors, and which enables the use of a semantically under-specified expression for successful reference [5]. Regarding the descriptive meaning provided by *Stoff*, we propose that the use of this noun evokes a three-phased script involving three event frames [3]. The final state of the discovery phase one (realization of frame DISCOVER), optionally fol- lowed by a synthesis phase two (a finite cyclic realization of ADD frames), provides an occasion for the application phase three (frame APPLY) to start [4].

We collected the 100 most frequent "*stoff"-compounds from the DWDS corpus [8]. The data items fall into three groups corresponding to the phases of the script. In phase one compounds like *Sauerstoff*, the modifier names a characteristic property of the state holder. These are attributive compounds in the sense of Lieber ([5]:48), which "can receive just about any pragmatically conceivable interpretation". Phase two compounds either refer to the initial state holder of synthesis (Grundstoff) or to any entity newly introduced by ADD, albeit from different perspectives (Rohstoff, Inhaltsstoff). A special case is *Zusatzstoff*, which names a substance introduced after the final state of phase two, thus prolonging this phase of the script. The third phase, the application of substances resulting from previous phases, is exemplified by *Impfstoff*. For the class of compounds in which *Stoff* does not denote a material substance, we assume factual knowledge as a second, metaphorically derived meaning of *Stoff*. The as- sociated isomorphic script is a blend of two frames: the final state of an acquire-knowledge frame is at the same time the initial state of a convey-knowledge frame. In our data set we find examples like *Lesestoff, Lernstoff* or also *Konfliktstoff*.

Our approach sheds new light on the semantic mechanism behind analogy formation: the pragmatic scaffold provided by the conceptual script enables the hearer to track the in- tended meaning of new word formations [7]. As a case in point, consider, for instance, the compound *Hörstoff*, which mimics the pattern of *Lesestoff*.

References: Bauer, Laurie. 2017. *Compounds and Compounding*. CUP. Fanselow, Gisbert. 1981. Neues von der Kompositafront oder Zu drei Paradigmata in der Kompositionsgrammatik. *Studium Linguistik* 11:43–57. Fillmore, Charles. 1976. Frame semantics and the nature of language. In *Annals of the NY Academy of Sciences*, 20–32. Hobbs, Jerry. 1985. *On the coherence and structure of discourse*. Technical Report. CSLI, Stanford, 85-37. Lieber, Rochelle. 2016. *Compounding in the lexical semantic framework*, CUP, 38–53. Schank, Roger and Robert Abelson. 1977. *Scripts, Plans, Goals, and Understanding*. Erlbaum, Hillsdale. Schlücker, Barbara and Ingo Plag. 2011. Compound or phrase? Analogy in naming. *Lingua* 121:1539–1551. DWDS – Digitales Wörterbuch der deutschen Sprache. "Referenz- und Zeitungskorpus", www.dwds.de, accessed 2020-09-25.

The polysemy of newly derived forms: An investigation of English *-ment nominalizations*

Lea Kawaletz
Heinrich-Heine-Universität Düsseldorf
lea.kawaletz@hhu.de

It is well-known that many derivational affixes in English are able to produce more than one meaning (see e.g. Bauer et al. 2013, Lieber 2016). For example, nominalizations with the suffix –*er* can denote AGENT (*shooter*), INSTRUMENT (*opener*), or INHABITANT (*Londoner*), among others. Despite its pervasiveness, affix polysemy is rarely investigated, especially for ATK nominal–izations ('-*ation* and kin,' Borer 2013).

In this talk, I use corpus attestations with the ATK suffix -*ment* to examine affix poly- semy as a productive phenomenon. Specifically, I will tackle two research questions: Which readings can be produced by -*ment* in contemporary English? Which readings are available for each individual -*ment* derivative? What are the semantic contributions of the base, of the affix, and of the context?

My data set consists of 40 deverbal -*ment* neologisms extracted from the Oxford Eng- lish Dictionary (OED) and the Corpus of Contemporary American English (COCA). A total of 502 attestations were obtained from various corpora, and annotated using common semantic labels such as EVENT or PRODUCT. An example is given in (1), where *confoundment* is attest- ed in a RESULT STATE reading:

(1) I know a lot of our compatriots also feel the same angst, consternation and **confound–ment**. (GloWbE NEWS leadership.ng 2012)

My data shows that -*ment* productively produces a range of readings that is semantically diverse, but predictable given the semantics of the base verbs and the preferences of the suffix (see Plag et al. 2018, Kawaletz in prep). For example, if the base verb has a complex event structure, its -*ment* derivative can refer to either of the two subevents. Individually, each neologism is endued with an array of several possible readings, out of which one or more are selected in context. For example, *confoundment* can denote EVENT, CAUSE, and RESULT STATE, just like all -*ment* neologisms with semantically similar base verbs in my data set. In the context in (1), this polysemy is disambiguated to a RESULT STATE reading.

Affix polysemy is pervasive and productive, and it can be explained by way of a com-positional account that takes into consideration the contributions of the base, the affix, and the context in a principled way.

A major challenge in the research of affix polysemy has been and will be to spell out a (formal) account that is underspecified yet restrictive enough to allow for the kind of sys- tematic semantic diversity that can be found attested (see Lieber 2016, Kawaletz in prep).

References: Bauer, Laurie, Rochelle Lieber, and Ingo Plag. 2013. *Oxford reference guide to English mor- phology*. Oxford: Oxford University Press. Borer, Hagit. 2013. *Taking form: Structuring Sense vol 3.* Oxford: Oxford University Press. Kawaletz, Lea. In prep. *The semantics of English -ment nominalizations: A frame-based approach.* Heinrich-Heine-Universität Düsseldorf, PhD dissertation. Lieber, Rochelle. 2016. *English Nouns: the ecology of nominalization*. Cambridge: Cam- bridge University Press. Plag, Ingo, Marios Andreou and Lea Kawaletz. 2018. A frame-semantic approach to poly- semy in affixation. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo, Fiammetta Namer, eds. *The lexeme in descriptive and theoretical morphology*. *Empirically oriented theoretical morphology and syntax*. Berlin: Language Science Press. 467-486.

On the interpretation of German A-V-*er*-constructions and the notion of concepts

Natascha Elxnath
Universität Tübingen
natascha.elxnath@uni-tuebingen.de

My talk is concerned with the following fairly productive German word formation pattern:

(1) *Schönspieler*; *Schnelldenker*; *Langschläfer*
   'beautiful-player'; 'fast-thinker'; 'long-sleeper'

The construction is formed by means of an adjective (e.g. *schön*), a verb (e.g. *spiel-*) and the affix *-er*; and apparently, it is always interpreted in the same way: the adjective cannot refer directly to the noun, but needs to be understood in an adverbial way (e.g. *Schönspieler* cannot mean 'player who is beautiful' but only 'player who plays beautifully'). To account for this observation, the construction has traditionally been analysed as being built within two steps, namely compounding (e.g. *schön+spiel-*) followed by derivation (e.g. *schönspiel+er*) (e.g. Leser 1990; Fleischer & Barz 1995). Yet, there is one obvious morphological weakness in this approach: compounds with verbal heads such as **schönspiel-* are not attested in German (Motsch 1999). Moreover, as examples from the web show, the construction comes with a much wider meaning variation than previously assumed:

(2)   *Kaltläufer* ('cold-runner') = runner who prefers cold weather for training
(3)   *Schnelldampfer* ('fast-steamer') = steamer (ship) that goes fast
(4)   *Freitrinker* ('free-drinker') = person who drinks a lot of alcohol in public space

Thus, the aim of this talk is to provide an alternative formal-semantic analysis that allows for more flexibility. My hypothesis is that the constructions under discussion are usual A-N- compounds, and that the observed meaning restrictions in (1) are not due to structural reasons, but due to rules of concept formation.

Bücking's (2009) semantic form (SF) for A-N-compounds serves as starting point:

(5)   Template: $\lambda P \, \lambda Q \, \lambda x \, [P(x) \, \& \, R_{integral}(x, v) \, \& \, Q(v)]$
(6)   The SF involves a free variable v that leaves open to which entity the adjective refers to, and a relational variable $R_{integral}$ that governs the relation between this entity and the head noun x. In a particular context, the variables can then be specified. As in the following shown for *Kaltläufer*, A-V-er-constructions fit well within this template:
(7)   [[kalt]]: $\lambda u \, [\text{COLD}(u)]$; [[Läufer]]: $\lambda m \, \text{GENe} \, [\text{RUN}(e) \, \& \, \text{AGENT}(e, m)]$
(8)   [[Kaltläufer]]: $\lambda x \, \text{GENe} \, [\text{RUN}(e) \, \& \, \text{AGENT}(e, x) \, \& \, R_{integral}(x, v) \, \& \, \text{COLD}(v)]$ (instantiation of variables for reading (2): v = WEATHER and $R_{integral}$ = PREFER)

Most importantly, the SF is flexible enough to allow for further possible interpretations apart from (2) (e.g. *Kaltläufer* = someone who starts running although his body is still cold, i.e. v = BODY) and can also capture the meanings of the constructions in (1) (e.g. for *Schönspieler*: v = PLAYING STYLE). However, the template in its current form actually allows too many readings: the reading 'player who is beautiful' is also permitted.

In my talk, I will propose an additional pragmatic condition that restricts the possible candidates for v, based upon the idea that complex words must denote proper concepts. Further attested constructions (e.g. *Schön -maler* ('drawer'), *-fahrer* ('driver'), *-bauer* ('builder')) from DECOW (Schäfer & Bildhauer 2012) will serve as empirical basis for discussion.

References: Bücking, Sebastian. 2009. "How do phrasal and lexical modification differ? Contrasting adjective-noun combinations in German." *Word Structure* 2 (2): 184-204. Fleischer, Wolfgang, and Irmhild Barz. 1995. *Wortbildung der deutschen Gegenwartssprache*. 2nd edition. Tübingen: Niemeyer. Leser, Martin. 1990. *Das Problem der 'Zusammenbildungen'. Eine lexikalistische Studie*. Trier: WVT. Motsch, Wolfgang. 1999. *Deutsche Wortbildung in Grundzügen*. Berlin: De Gruyter. Schäfer, Roland, and Felix Bildhauer. 2012. "Building large corpora from the web using a new efficient tool chain." *Proceedings of LREC 8* (Istanbul): 486-493.

The distributional-semantics side of morphologically complex words: Modelling the processing of affixed words in vector spaces

Marco Marelli
University of Milano Bicocca marco.marelli@unimib.it

Since the seminal LSA proposal (Landauer & Dumais, 1997), distributional semantics has provided efficient data-driven models of the human conceptual system, representing word meaning through vectors recording lexical co-occurrences in large text. However, vanilla distributional models generate static descriptions of the semantic system, falling short of capturing the highly dynamical interactions occurring at the meaning level during language processing. In this presentation I will discuss possible ways to adapt distributional semantics in order to account for the internal structure of derived words.

In a functional perspective, affixes can be represented as matrices mapping stems into derived forms, and estimated from corpus data by means of machine learning techniques (Marelli & Baroni, 2015). As a consequence, derived-form meanings can be thought of as the result of a procedure which transforms the stem vector on the basis of the affix matrix (e.g., the meaning of "nameless" can be obtained by multiplying the vector of "name" with the matrix of "-less").

This architecture accounts for the remarkable human capacity of generating new words that denote novel meanings, correctly predicting semantic intuitions about nonce derived forms (e.g., "quick+ify"). Moreover, the proposed compositional approach, once paired with a whole-word route, provides a new interpretative framework for semantic transparency effects, which are here explained in terms of ease of the combinatorial procedure and strength of the transformation brought about by the affix (Marelli & Baroni, 2015; Günther, Smolka & Marelli, 2019).

However, an important shortcoming of such an approach is that it is particularly data-hungry: a specific function needs to be separately estimated for each single affix. I will discuss how this limitation could be addressed by building on more recent approaches focused on compounding (Marelli, Gagné & Spalding, 2017; Günther & Marelli, in press).
Such models induce a general compositional process that would not require characterizing each single affix in functional terms.

References: Günther, Fritz, Eva Smolka and Marco Marelli. 2019. "'Understanding'differs between English and German: Capturing systematic language differences of complex words." *Cortex* 116: 168-175. Günther, Fritz, and Marco Marelli. In press. "Trying to make it work: Compositional effects in the processing of compound 'nonwords'". *Quarterly Journal of Experimental Psychology.* Landauer, Thomas K., and Susan T. Dumais. 1997. "A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge." *Psychological review* 104(2): 211. Marelli, Marco, and Marco Baroni. 2015. "Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics." *Psychological review* 122(3): 485. Marelli, Marco, Christina L. Gagné and Thomas L. Spalding. 2017. "Compounding as Abstract Operation in Semantic Space: Investigating relational effects through a large-scale, data-driven computational model." *Cognition* 166: 207-224.

Splitting -*ly*'s: Using word embeddings to distinguish derivation and inflection

Martin Schäfer
Universität Tübingen post@martinschaefer.info

Whether the formation of adjective-*ly* from an adjective in English is best captured as a process of derivation or inflection has been widely debated, with Bauer et al. (2013, 536) "concluding that the evidence is inconclusive". As far as the semantics is concerned, the relationship between adjective and adjective-ly is seen as completely regular, with -ly carrying no lexical meaning (Giegerich 2012). In my paper, I explore whether a more finegrained look at the semantics involved does allow one to differentiate the regular cases into instances of derivation or inflection. Starting from the idea that adverbs prototypically modify events, I hypothesize that adjectives that are intrinsically event predicates behave different across the adjective-adjective-*ly* divide than adjectives that are intrinsically non- eventive: whereas the former require no semantic adaption and are thus fully inflectional in nature, the latter require a semantic adaption and are thus expected to show characteristics of derivation.

The hypothesis is tested using distributional semantics, building on the finding that inflectional contrasts are overall more stable across a class than derivational ones (Bonami and Paperno 2018), and exploiting differences between types of word embeddings associated with topical and functional similarity (using the BOW5 and DEPS embeddings from Levy and Goldberg 2014). The target difference, event vs. non-event predicate, is investigated by contrasting the adjective classes SPEED and HUMAN PROPENSITY.

This leads to the following expectations:

a.) On average, the relationship between adjective and adjective-*ly* should be more stable for pairs from the SPEED class, regardless of embedding.

b.) Distributional measures focused on functional similarity should make SPEED pairs more similar, but HUMAN PROPENSITY pairs less similar.

c.) Group internal similarities should be kept intact across members of the SPEED class, but not across the members of the HUMAN PROPENSITY class.

The results show a clear difference in the behavior of the word embeddings of the two semantic classes. All significant differences are in line with the expectations. The decreasing similarity across pairs for the HUMAN PROPENSITY class when switching from the BOW5 to the DEPS embeddings shows characteristics expected of derivations. The high correlation between the class internal similarities across SPEED pairs, and the low correlations for the HUMAN PROPENSITY class points to the inflectional character of the former and the derivational of the latter.

These results are significant because they show how capitalizing on known properties of distributional semantic measures of derivation and inflection allows us to understand the classificatory difficulty presented by -*ly*: It arises due to the conflation of different adjective classes that, individually, show clear characteristics of either derivation or inflection.

References: Bauer, Laurie, Rochelle Lieber, and Ingo Plag 2013. *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press. Bonami, Olivier and Denis Paperno 2018. "Inflection vs. derivation in a distributional vector space." *Lingue e Linguaggio* 17(2), 173–195. Giegerich, Heinz J. 2012. "The morphology of -*ly* and the categorial status of 'adverbs' in English." *English language and linguistics* 16(3), 341–359. Levy, Omer and Yoav Goldberg 2014. "Dependency-based word embeddings." In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore: Association for Computational Linguistics 302–308. Association for Computational Linguistics.

Distributional evidence for derivational paradigms

Matías Guzmán Naranjo, Olivier Bonami

LLF, CNRS, Université de Paris, LLF, CNRS

mguzmann89@gmail.com, olivier.bonami@gmail.com

Background Paradigmatic approaches to word formation (see a.m.o. Becker 1993; Bochner 1993; Booij 2010; Marle 1984; Štekauer 2014; Bonami & Strnadová 2019) build on the observation that morphological relatedness between lexemes is not always straightfowardly described in terms of a relation between an input and an output. Relevant phenomena include backformation, conversion pairs of undecidable directionality, parallel derivations with no synchronically available base (e.g. optimism~optimist), and mismatch between formal and semantic base. While such examples have been well-known for decades, traditional approaches to morphology do not attempt to account for them directly, on the assumption that they are too rare to play a role in the architecture of the system.

Hypothesis This abstract reports on an ongoing study that attempts to leverage distributional semantic methods applied to large lexical datasets to assess whether semantic evidence for morphological relatedness going beyond input-output relationships can be documented at the level of the system. The basic intuition is that, under a traditional rule- based view of derivation, the formal base of a derived lexeme should be the best predictor of its semantics within the derivational family. For instance, French nouns versement 'payment, instalment' and verseur 'container used for pouring, worker tasked with pouring' both have an easily recoverable semantic relationship with their base verser 'pour', but are not so obviously related with one another. We submit that any situation where we have strong evidence that something else than the formal base is a better predictor counts as evidence for paradigm structure.

Materials and methods Using lexical data from previous studies of derivation in French (Hathout and Namer, 2014; Tribout, 2010) as well as newly compiled data, we built a dataset of 8662 (base, derivative) pairs illustrating the use of 9 noun-forming processes, with at least 300 observations per process: -ment, -eur, -age, deverbal conversion, -ant, -ion, -iste and -isme. In parallel, we used the Gensim (Řehůřek, 2010) implementation of word2vec (Mikolov et al., 2013) to build a lemma-based vector space on the FrCow corpus (Schäfer, 2015; Schäfer and Bildhauer, 2012). The lexical dataset gives us access to sets of triples of lexemes (b, d1, d2), where (b, d1) illustrates a first process and (b, d2) a second process. Such triples can be seen as partial derivational paradigms. We generalize to all pairs of cells in these paradigms the proposition in Marelli and Baroni (2015) that the semantics of a derivational process be represented by a function from vector to vector. Specifically, we fit linear regressions to map the vector in one cell to the corresponding vector in another cell (i.e. we have 6 different models for a given set of triples). The resulting models represent the average semantic shift entailed by moving from one cell to another. We use the average cosine distance between predicted and actual vector as a way of evaluating how well the semantics of words in one cell is predicted predicts by those in another cell.

Selected results When we compare verbs to their derived -ment event nouns and - eur agent/instrument nouns, we find that the semantic relations are more predictable between base and derivative than between the two derived nouns. On the other hand, when we compare nouns in -isme and -iste to their bases, the semantics of the derivatives are more interpredictable than their relationship to their base is. Hence our method captures a well-known case of systematic paradigmatic relation (see e.g. Roché, 2011). In the talk we will show that other cases in our dataset fall between these two extremes, and discuss how other types of evidence for derivational paradigms can be assessed using similar methods.

predictor	predicted	cosine	predictor	predicted	cosine
base	-eur	0.633	base	-isme	0.746
base	-ment	0.675	base	-iste	0.704
-eur	base	0.637	-isme	base	0.573
-ment	base	0.688	-iste	base	0.561
-eur	-ment	0.6	-iste	-isme	0.813
-ment	-eur	0.614	-isme	-iste	0.79

Difference of attestation dates as evidence for directionality in zero derivation

Gianina Iordăchioaia, Gabriella Lapesa, Sarina Meyer, Sebastian Padó
University of Stuttgart, University of Stuttgart, University of Stuttgart, University of Stuttgart
gianina@ifla.uni-stuttgart.de, gabriella.lapesa@ims.uni-stuttgart.de, st166791@stud.uni-stuttgart.de, sarina-karen.meyer@ims.uni-stuttgart.de, pado@ims.uni-stuttgart.de

We investigate whether zero-derived nouns (ZeroNs: *to break > the break*) are morphologi- cally and semantically derived from verbs, like suffixed nominals (*break*$_V$ > *break*$_V$-**ing**$_N$/**∅**$_N$; Kiparsky 1982), or are built from categorially underspecified roots in parallel with verbs (*break*$_{N/V}$; Borer 2013). We build on recent psycholinguistic work which finds a derivational direction for some V > N pairs (*bite*$_V$ > *bite*$_N$), while others appear to be categorially underspecified (*guide*$_{N/V}$; Darby & Lahiri 2016). Iordăchioaia (2020) argues that ZeroNs built on change of state verbs behave like the former, while those built on psych verbs pattern with the latter. We study whether we can adduce corpus-based evidence for this dichotomy.

We rely on differences in first attestation dates. Directionality tests predict a derived word to be less frequent and later attested than the base (Plag 2003). That is, if a ZeroN is built on an underspecified root, its first attestation should be close to that of the V: once the root is lexicalized, both its categories will show comparable availability (=> a short "lag"). If a ZeroN is derived from a categorized verb, it will be attested later (=> a long "lag").

We analyze a dataset of 392 ZeroNs corresponding to better studied verb classes from the VerbNet (Change of state, Psych, Motion, Emission, and Communication). We use a standard linear regression model to predict the difference in attestation dates between ZeroNs and Vs, using as independent variables (a) relative frequency (ZeroN/V), (b) verb class, (c) root etymology (Germanic/Romance), (d) two distributional semantic measures of the relation between base and derivative, namely, cosine similarity (symmetrical) and difference in information content (base minus derived form, asymmetrical).

Applying backward elimination and multicollinearity checks for model selection, we obtain a model with significant effects for (a), (b), and (c). Two of our significant findings are: (1) Higher relative frequency corresponds to a smaller lag; (2) Change of state verbs show longer lag than Emission and Psych verbs. Finding (1) supports the intuition that derived forms generally establish themselves as independent lexical entries over time. Finding (2) confirms Iordăchioaia's (2020) theoretical contrast between Change of state and Psych ZeroNs. Roots of Emission Vs are known to express both events (~Vs) and objects as emitted substance (~Ns), which concurs with our finding that they are categorially underspecified just like psych roots. These results bolster our case for using difference in attestation dates as a measure of directionality in zero derivation.

The main negative result is that the distributional measures we selected to quantify directionality were not a significant predictor of attestation "lag". In the talk, we will report on a simplified prediction setup (predicting just the V attestation date) where we do find an effect of the distributional measures: higher information content correlates with earlier attestation.

References: Borer, Hagit. 2013. *Structuring sense: Volume III: Taking form*. Oxford: Oxford University Press. Darby, Jeannique & Aditi Lahiri. 2016. Covert morphological structure and the processing of zero-derived words. *The Mental Lexicon* 11.2: 186-215. Iordăchioaia, Gianina. 2020. Categorization and nominalization in zero nominals. Artemis Alexiadou & Hagit Borer, eds., *Nominalization: 50 Years on from Chomsky's* Remarks. Oxford: Oxford University Press. 231-253. Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In Harry van der Hulst & Norval Smith, eds., *The structure of phonological representations*. Dordrecht: Foris. 131-175. Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.

Arbeitsgruppe 4

Free variation = unexplained variation? Empirical and theoretical approaches to optionality in grammar

Kristin Kopf & Thilo Weber

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Didymophilia in language

Freek Van de Velde
KU Leuven freek.vandevelde@kuleuven.be

Proponents of the no-synonymy hypothesis have argued that unconditioned 'free' variation between alternating variants in language is exceedingly rare or non-existent. Indeed, synchronic variationist linguists have built lofty multivariate models that are able to explain large swaths of the variance in near-synonymous constructions (see Pijpops 2019, Ch.2 for a discussion), especially when taking into account social indexation of the forms: even if the distribution of alternating variants cannot 'residuelessly' be explained by language-internal (e.g. pronominality, lexical biases) and cognitive-semantic factors (e.g. animacy, topicality, priming) alone, the added explanatory power of social variables (SES, gender, region), often suggests the variants are prone to second-order indexing (in the sense of Silverstein 2003).

Diachronic variationists, on the other hand, have for the most part focused on cases in which an incoming mutant gradually takes over an older form (Blythe & Croft 2012), ousting the latter form from the language. This scenario indicates that a neat distinction on functional or social grounds of the two forms is not very stable. Moreover, while it is not unheard of that near-synonymous forms develop new meanings in a process of exaptation (Van de Velde & Norde 2016; De Smet & Van de Velde 2020), competing constructions can also converge over time, rather than divide the functional space among them (De Smet et al. 2018). Doubt with regard to a neat division of labour between competing constructions also comes from studies pointing out that synonymy and non-isomorphic tendencies can have advantages for the system (Van de Velde 2014; Fonteyn & Maekelberghe 2018), and that competing constructions can co-exist over a remarkably long time – a phenomenon that could be called 'didymophilia': a predilection or fascination for twins.

As I will argue, retaining a (small) residu of unexplained variance is advantageous, and this may be the reason behind the pervasiveness of variation.

References: Blythe, Richard & William Croft. 2012. "S-curves and the mechanisms of propagation in language change." *Language* 88(2): 269-304. De Smet, Hendrik, Frauke D'hoedt, Lauren Fonteyn & Kristel Van Goethem. 2018. "The changing functions of competing forms: attraction and differentiation." *Cognitive Linguistics* 29(2): 197-234. De Smet, Isabeau & Freek Van de Velde. 2020. "Semantic differences between strong and weak verb forms in Dutch." *Cognitive Linguistics*, (epub ahead of print). Fonteyn, Lauren, & Charlotte Maekelberghe. 2018. "Competing motivations in the diachronic nominalisation of English gerunds." *Diachronica* 35(4): 487-524. Pijpops, Dirk. 2019. *How, why and where does argument structure vary? A usage-based investigation into the Dutch transitive-prepositional alternation*. PhD, KU Leuven. Silverstein, Michael. 2003. "Indexical order and the dialectics of sociolinguistic life." *Language and Communication* 23: 193-229. Van de Velde, Freek. 2014. "Degeneracy: the maintenance of constructional networks." In Ronny Boogaart, Timothy Colleman & Gijsbert Rutten, eds. *The extending scope of construction grammar*. Berlin: De Gruyter Mouton. 141-179.Van de Velde, Freek & Muriel Norde. 2016. "Exaptation. Taking stock of a controversial notion in linguistics." In Muriel Norde & Freek Van de Velde, eds. *Exaptation and language change*. Amsterdam: John Benjamins. 1-35.

Optionality and categorial properties: The case of optional plural marking in Yucatec Maya

Yidong Yu
University of Göttingen
yidong.yu@uni-goettingen.de

In this paper, I propose a typological correlation between optionality in number marking and the properties of nominal category, based on a close examination of the empirical data from Yucatec Maya, a Mayan language spoken primarily in Yucatán Peninsula in Mexico.

The marking of plurality with the suffix *-o'ob* is essentially optional in Yucatec Maya, and this optionality is homogeneous throughout the entire system (1-2).

(1)   a. *le   x-ch'úupal-o'*   (2) *ka'a túul   nukuch tso'*
          DET FEM-girl-CL   two   CLF   big   turkey
          'the girl'/'the girls'   'two big turkeys'

       b. *le   x-ch'úupal-o'ob-o'*   (Monforte et al. 2010: 139)
          DET FEM-girl-PL-CL
          'the girls'   (Butler 2012: 34)

The motivation of the typological variation of obligatory and optional number marking in the world languages has been previously accounted for from three different angles: (i) variation in the morphosyntactic structure of the specific language (Witschko 2008), (ii) variation in the interpretation of the number morpheme (Borer 2005), and (iii) variation in the interpretation of the nouns (Chierchia 1998).

In Yucatec Maya, the morphosyntactic structure of plural marking does not determine the explicit grammatical expression of plurality. Butler (2011, 2012), following Witschko (2008), propose that the plural suffix in Yucatec Maya is merged as an adjunct modifier of the DP. Although this analysis is effective in accounting for the phenomenon of optional number marking in the language, it leaves the condition of such optionality unexplained.

The possibility that the Yucatec plural suffix *-o'ob* has various interpretations is also ruled out. In this respect, three parameters are examined: (i) degree of animacy (Smith-Stark 1974; Comrie 1981) of the host nouns, (ii) argument structure (Goldberg 1995; Jackendoff 2002) of the pluralized constructions, and (iii) constraints of numerical quantification (Xrakovskij 1997; Yu 2003) of the pluralized constructions. I show that the Yucatec plural suffix is fully grammaticalized, having only the grammatical meaning of simple plurality, and hence the motivation of the optionality in number marking in Yucatec Maya should not be located on the interpretation of the number morpheme.

Finally, I argue that the interpretation of Yucatec nouns motivates the explicit expression of grammatical plurality. I present evidence that Yucatec nouns behave like a combination of the English and Hindi types with respect to their generic interpretation (Carlson 1977; Dayal 1992, 2004; Deal and Nee 2018). Moreover, I propose that whether Yucatec nouns receive mass or count/apportioned interpretation is crucial to the overt marking of plurality (Quine 1960; Landman 1989; Grimm 2012).

References: Borer, Hagit. 2005. *Structuring Sense. In Name Only*, volume I. Oxford University Press, Oxford, UK. Carlson, Gregory N. 1977. *Reference to Kinds in English*. PhD thesis, University of Massachusetts, Amherst. Chierchia, Gennaro. 1998. "Reference to kinds across language." *Natural Language Semantics* 6(4): 339-405. Dayal, Veneeta. 2004. "Number marking and (in)definiteness in kind terms." *Linguistics and Philosophy* 27(4): 393-450. Grimm, Scott. 2012. "Degrees of countability: A mereotopological approach to the mass/count distinction." *Semantics and Linguistic Theory* 22: 584-603. Wiltschko, Martina. 2008. "The syntax of non-inflectional plural marking." *Natural Language & Linguistic Theory* 26(3): 639-694.

Correlative coordination and variable subject-verb agreement in German: An experimental study

Claudia Felser, Anna Jessen
Universität Potsdam, Universität Potsdam
felser@uni-potsdam.de, ajessen@uni-potsdam.de

Complex subjects joined by correlative conjunctions such as German *sowohl … als auch* ('both … and') or *weder … noch* ('neither … nor') often allow for variable number agreement with the finite verb. It is unclear whether the two agreement options can be said to be in free variation, however, and how factors such as the subject phrase's notional plurality, a conjunct's relative proximity to the verb, or the presence of negation in affect speakers' agreement preferences. For conjoined phrases involving plural noun phrases, Eisenberg (1989) and the *Duden* reference grammar observe that a plural verb is usually chosen if one of the conjuncts appears in the plural. It has also been suggested that verbal agreement tends to be determined by the conjunct closest to the verb (e.g. Klein, 2004).

We will report the results from two experiments that systematically assess German speakers' number agreement preferences with subjects joined by the correlative conjunctions *sowohl … als auch* or *weder … noch*, both of which have been argued to be semantically additive (e.g. Klein, 2004; Wurmbrand, 2008). Our study addresses the following questions: (i) How does the type of conjunction affect the acceptability of singular and plural verbs? (ii) If the two conjuncts differ in their grammatical number, how does their linear proximity to the verb affect speakers' preferences?

Experiment 1 was an untimed scalar acceptability rating task (n=65). The results showed that for double singular conditions, plural verbs were generally preferred, but this preference was significantly weaker for *weder … noch* compared to *sowohl ... als auch*. For conjuncts carrying conflicting number marking, we found a proximity effect for singular verbs: A singular second conjunct rendered singular verbs more acceptable compared to a singular first conjunct.

Experiment 2 was a speeded binary-choice sentence completion task (n=47) asking participants to choose between a singular or plural verb. Here we found no significant difference between the two connectors, with plural verbs being preferred for both. Conjunct proximity affected participants' verb choices such that a singular second conjunct elicited significantly more singular responses compared to a singular first conjunct.

Taken together, our results show that variation in number agreement with correlative coordination can only partly be accounted for by semantic differences between connectors or by conjunct proximity. While participants' show an overall preference for plural agreement, singular agreement is also acceptable if the second or both conjuncts are singular noun phrases, and more so for the negative conjunction *weder … noch*. We suggest that grammatical variation of this kind might best be captured by constraint-based models of grammar (e.g. Smolensky et al., 2014).

References: Duden - Das Wörterbuch der sprachlichen Zweifelsfälle: Richtiges und gutes Deutsch, edited by Dudenredaktion, Bibliographisches Institut GmbH, 2016. ProQuest Ebook Central, https://ebookcentral.proquest.com/lib/potsdamuni/detail.action?docID=4776694. Eisenberg, P. (1989). Grundriß der deutschen Grammatik. Stuttgart: Metzler. Klein, W.P. (2004). Koordination als Komplikation. Über eine strukturelle Ursache für die Entstehung syntaktischer Zweifelsfälle. Deutsche Sprache 32, 357–375. Smolensky, P., Goldrick, M., & Mathis, D. (2014). Optimization and quantization in gradient symbol systems: a framework for integrating the continuous and the discrete in cognition. Cognitive Science 38, 1102-1138. Wurmbrand, S. (2008). *Nor*: Neither disjunction nor paradox. Linguistic Inquiry 39, 511-522.

Non-verbal number agreement between the distributive plural and singular: Exceptions or free variation?

Karolina Rudnicka, Aleš Klégr
University of Gdańsk, Charles University in Prague
karolina.rudnicka@ug.edu.pl, ales.klegr@ff.cuni.cz

We investigate the topic of non-verbal number agreement, i.e. the agreement in number between the (formally or notionally) plural subject of a clause and a nominal clause element in the predicate part of this clause where the agreement may be viewed as an additional signal (and reinforcement) of the connection between them; compare sentences (1) and (2).

(1) 113 of the <u>soldiers</u> lost their <u>lives</u>, more than 100 were injured. (BNC, 1985-1994)

(2) Look at all of the new <u>people</u> that lost their <u>job</u> and (…) (COCA: 2012)

The paper focuses on contemporary English which seems to overwhelmingly prefer the so- called distributive plural occurring in a situation where "a set of entities [is] matched individually with individual entities in another set" (Quirk et al. 1985: 768). Sentence (1) is an example in which there is a direct correspondence between the number of the subjects and objects. This general tendency is, however, not without exceptions (cf. Sørensen 1985, Dušková et al. 2006), as example (2) shows.

The first aim of the paper is to provide a detailed account of literature-reported scenarios in which the general preference for the distributive plural is overruled (cf. Sørensen 1985 and Rappaport 2017), for example in the case of i) invariable idioms; ii) the indication of joint possession; iii) the intention to convey ideas of universal, abstract or figurative kind. The second aim is to offer fresh insights based on a corpus study (both quantitative and qualitative) of two structurally similar constructions, namely *lose one's life* and *lose one's job*, and the statistical data it provides on the distribution of the distributive plural and the distributive singular in these constructions. The third goal is to discuss the possible presence of free variation in some of the distributive singular and distributive plural uses, such as exemplified in (3) and (4):

(3) Those two <u>men</u> lost their <u>lives</u> and according to the Iraqi government so did two others from the Muslim family living nearby. (COCA: 2009)

(4) More than 65 <u>people</u> lost their <u>life</u> after a cruise ship sunk outside of the islands of Paros. (COCA: 2000)

On that account, the study suggests and discusses conditions for a case to be seen as an instantiation of free variation as understood by Cappelle (2009). Methodologically, the study draws upon corpus linguistics, computational linguistics and usage-based approaches. The data analysed in the study is extracted from the BNC and COCA.

References: Cappelle, Bert. 2009. Can we factor out free choice? In Dufter et al., eds. *Describing and Modeling Variation in Grammar*. Berlin and New York: Mouton de Gruyter: 183-201. Dušková, Libuše, et al. 2006 (3rd ed.). *Mluvnice současné angličtiny na pozadí češtiny* [A Grammar of Contemporary English against the Background of Czech]. Praha: Academia. Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Edinburgh: Longman. Rappaport, Jennifer. 2017. "Singular or Plural? Sometimes, It Depends". *The MLA Style Center: Writing Resources from the Modern Language Association*, 31 May 2017, https://style.mla.org/singular-or-plural/, accessed on 28 October 2020. Sørensen, Knud. 1985. The distributive plural and its limits. *English Studies* 66(4): 338-350.

Investigating morphosyntactic variation in a Uralic minority language: The Aanaar Saami conditional perfect

Merit Müller
University of Tartu
merit.muller@ut.ee

The Saami languages constitute the westernmost branch of the Uralic languages, forming a dialect continuum from central Sweden to the Kola peninsula in Russia. This paper focuses on Aanaar (Inari) Saami which is traditionally spoken in the northernmost part of Finland and, with a number of 350–450 speakers, is classified as a severely endangered language.

A structural property formerly shared by all Saami languages is the expression of the conditional perfect by two periphrastic constructions (see Bartens 1980). In Aanaar Saami, both constructions have been preserved (see examples 1a–b). The first construction consists of the auxiliary 'to be' in the conditional and the past participle of the main verb (hereinafter participle construction), the second of the auxiliary 'to be' in the indicative past tense and the infinitive of the main verb (hereinafter infinitive construction). In Aanaar Saami descriptions (e.g. Olthuis 2000: 90), the two periphrastic constructions are depicted as interchangeable.

(1)   a.   *Mun*   *liččim*   *moonnâm.*
           1SG   be.COND.1SG   go.PST.PTCP

      b.   *Mun*   *lijjim*   *moonnâđ.*
           1SG   be.PST.1SG   go.INF
                                       'I would have gone.'

The present study investigates the variation of the above-mentioned variables. The central research question is whether and to which extent the selection of one construction over the other can be explained by intra- and extralinguistic determinants. Furthermore, the variation will also be discussed from a diachronic point of view. The analysed data consists of three parts: (1) language samples collected by Erkki Itkonen in 1952; (2) the corpus of written Aanaar Saami texts (SIKOR) and (3) a survey which was conducted in spring 2020.

My presentation will outline how various factors (the lexical verb, the polarity and the type of clause as well as the dialect) play a role in the variation of the Aanaar Saami conditional perfect. Yet, the examined factors only function as constraints. In some settings, the two constructions appear to be interchangeable, i.e. they are subject to "free" variation.

The analysis also revealed ongoing changes in the variation and its determinants. While in the data collected in 1952, for example, the infinitive construction was favoured in the protasis and apodosis of conditional sentences, the type of clause does not seem to function as a determinant in contemporary Aanaar Saami. Moreover, the diachronic comparison showed a change in frequency: previously, the two competing constructions occurred at almost equal frequencies, but the newer data shows a clear shift towards the participle construction.

The Aanaar Saami conditional perfect functions as an example on how (partly "free") variation can change over time: in present days, the presumably older infinitive construction (see, e.g., Lehtiranta 1992: 92) is falling behind the participle construction in both, frequency and gradience. In my talk, I will illustrate how the underlying reasons for this development can be found in language-internal analogy as well as linguistic interference.

References: Bartens, Hans-Hermann. 1980. *Die Verwendung von Konditional und Potential im Lappischen.* Suomalais-Ugrilaisen Seuran Toimituksia = Mémoires de la Société Finno-Ougrienne 177. Helsinki: Suomalais-Ugrilainen Seura. Lehtiranta, Juhani. 1992. *Arjeploginsaamen äänne- ja taivutusopin pääpiirteet.* Suomalais- Ugrilaisen Seuran Toimituksia = Mémoires de la Société Finno-Ougrienne 212. Helsinki: Suomalais-Ugrilainen Seura. Olthuis, Marja-Liisa. 2000. *Kielâoppâ.* Inari: Sämitigge.

Der Abbau ,freier graphematischer Variation' in der Geschichte des Deutschen: Methodische Überlegungen zu einer Korpusuntersuchung

Chiara Fioravanti
Christian-Albrechts-Universität zu Kiel
fioravanti@germsem.uni-kiel.de

In meinem Dissertationsprojekt möchte ich anhand eines zeitlich-räumlich gestaffelten Textkorpus (14.–18. Jh.) die Entwicklung von einer vormodernen, variationsbasierten Schreibung hin zur modernen Orthographie nachverfolgen. Um die Geschichte der Variantenreduktion zu beschreiben, ist es sinnvoll, den Fokus auf die wortbezogene Variation zu legen, die nicht-normierte Schreibsysteme älterer Sprachstufen grundlegend charakterisiert und zugleich „unseren Vorstellungen eines geregelten Graphiengebrauchs am stärksten widerspricht" (Elmentaler 2012, S. 157). Denn während wir heute im phonetischen, lexikalischen und grammatischen Bereich sprachlicher Variation durchaus begegnen, basiert unsere Orthographie auf dem Prinzip der festen Wortschreibung. So wählen wir in der gesprochenen Sprache zwischen Varianten wie [køːnɪçʁaɪ̯ç] oder [køːnɪkraɪ̯ç], verwenden beim Schreiben aber die einheitliche Wortschreibung *Königreich*. Dagegen konnten noch bis ins 17. Jahrhundert im selben Text Schreibweisen wie *chunigreich*, *chůnich reich* und *Chůnig-Reich* nebeneinanderstehen. Hier wird ein Variationstyp sichtbar, die aus heutiger Sicht schwer interpretierbar ist. Hinter dieser ,freien graphematischen Variation' (vgl. Mihm 2007a, S. 201) werden in der Forschung meistens Reflexe phonetischer Varianten oder graphostilistische Motive vermutet (vgl. Mihm 2007b, S. 226f.; Voeste 2008, S. 27f.).

In meinem Vortrag möchte ich besondere methodische Herausforderungen diskutieren, die sich bei dem Versuch ergeben, ,freie' Variation zu identifizieren. Wichtig ist hier vor allem die Unterscheidung zwischen Varianten, die an strukturelle Bedingungen wie Folgekonsonanz oder Silbenstruktur ,gebunden' sind, und tatsächlichen ,freien' Varianten. Nur direkte Wortvarianten des Typs *pedewt/bedewt/bedewtt/bedeutt* ('bedeutet', 3. Sg. Präs. Ind.) können als ,freie' Varianten gelten.

Das Untersuchungskorpus, das meinem Projekt zugrunde liegen soll, setzt sich aus Textstichproben aus unterschiedlichen Regionen und Zeitstufen zusammen. Meine Analysen gründen also auf Daten, die auf unterschiedliche sprachliche Bezugssysteme zurückgehen. Aus diesem Grund stellt die Definition der innersprachlichen Kontexte, innerhalb deren ,freie' Variation angenommen werden kann, eine besonders schwierige Aufgabe dar, die sich auf die Untersuchungsanlage auswirkt.

References: Elmentaler, Michael. 2012. „Phonologie und Graphematik." In: Albrecht Greule, Jörg Meier und Arne Ziegler, Hrsg*. Kanzleisprachenforschung. Ein internationales Handbuch*. Berlin/Boston: De Gruyter.151-170. Mihm, Arend. 2007a. „Zur Neubestimmung des Verhältnisses zwischen Schreibsprachen und historischer Mündlichkeit." In Arend Mihm, *Sprachwandel im Spiegel der Schriftlichkeit. Studien zum Zeugniswert der historischen Schreibsprachen des 11. bis 17. Jahrhunderts*. Frankfurt am Main u.a.: Lang. 193-216. Mihm, Arend. 2007b. „Zur Deutung der graphematischen Variation in historischen Texten." In Arend Mihm, *Sprachwandel im Spiegel der Schriftlichkeit. Studien zum Zeugniswert der historischen Schreibsprachen des 11. bis 17. Jahrhunderts*. Frankfurt am Main u.a.: Lang. 217-230. Voeste, Anja. 2008. *Orthographie und Innovation. Die Segmentierung des Wortes im 16. Jahrhundert*. Hildesheim u.a.: Olms.

Fakultative Valenzen als freie Variation

Vilma Symanczyk Joppe
Heinrich-Heine-Universität Düsseldorf
symanczyk@hhu.de

Traditionell werden in der Valenztheorie solche Valenzen als fakultativ bezeichnet, die realisiert werden, aber auch weggelassen werden können. Jacobs (1994) modelliert diese Varianten mit und ohne die jeweilige Ergänzung als Alternativvalenzen, die in freier Variation stehen können. Viele dieser Alternativvalenzen sind dabei auf bestimmte Kontexte und Zusatzfaktoren beschränkt: Dazu gehören definite bzw. indefinite Interpretation der weggelassenen Ergänzung (Fillmore 1986), generische Kontexte (u.a. Härtl 2013) und der Satztyp (u.a. Jacobs 2016, Külpmann & Symanczyk Joppe 2015).

Der Vortrag basiert wesentlich auf einer umfangreichen Sprecherurteilserhebung zur Weglassbarkeit von Akkusativobjekten, die ich zwischen 2017 und Anfang 2020 an den Universitäten Wuppertal und Bonn durchgeführt habe. Die Testsätze in Kontexten involvieren 49 Verben in 22 Konstruktionen, darunter solche, die definite und indefinite Weglassungen, verschiedene Satztypen, Generizität, Kontrast, Durativierungen, Negation und Topikalisierungen der Verbalphrase testeten. Alle Kontexte wurden als Minimalpaare mit und ohne Akkusativobjekt getestet.

In den Daten lassen sich gleichermaßen konstruktionelle wie lexikalische Einflüsse erkennen – die Ergebnisse bleiben also weder über einzelne Verben hinweg noch über einzelne Konstruktionen hinweg stabil. Es findet sich jedoch eine Vielzahl einzelner Minimalpaare, in denen für ein bestimmtes Verb in einer bestimmten Konstruktion unabhängig von der Realisierung des direkten Objektes nahezu identische Testwerte erzielt werden. Diese Minimalpaare können entsprechend als freie Varianten voneinander aufgefasst werden. In (1) wird ein entsprechendes Beispiel aufgeführt:

(1) KONTEXT: Habitueller Satz; Verb: *bestellen*
     Sophie und ihre Freunde unterhalten sich darüber, wo sie ihre Bücher kaufen. „Früher hatte ich um die Ecke einen tollen Buchladen", berichtet Sophie.
     „Aber seit der zugemacht hat, bestelle ich meistens im Internet."   95,3 %
     „Aber seit der zugemacht hat, bestelle ich Bücher meistens im Internet."   95,8 %

Im Vortrag werden die Minimalpaare mit annähernd gleichen Ergebnissen (≤ 5% Differenz) unter die Lupe genommen. Dabei werden u.a. folgende Gesichtspunkte erörtert: 1) Treten die „freien Varianten" verb- bzw. konstruktionsübergreifend auf? 2) Betreffen sie nur hochfrequente Verben/Konstruktionen oder auch niedrigfrequente? 3) Existieren nur „freie Varianten" mit gleich hohen oder auch solche mit gleich niedrigen Akzeptabilitätswerten? 4) Korrelieren gleich hohe Akzeptabilitätswerte mit gleich hohen Korpusfrequenzen?

References: Fillmore, Charles. 1986. "Pragmatically controlled zero anaphora." In *Proceedings of the twelfth annual meeting of the Berkeley Linguistic Society*: 95–107. Härtl, Holden. 2013. "Generic rescue: argument alternations and the monotonicity condition." In Patrick Brandt und Eric Fuß, Hrsg. *Repairs – the added value of being wrong* (= Interface Explorations 27). Berlin/New York: de Gruyter: 95–130. Külpmann, Robert und Vilma Symanczyk Joppe. 2015. "Argument omission between valency and construction. Evidence for sentence type effects from acceptability rating studies." In Gerhard Jäger, Hrsg. *Proceedings of the 6th Conference on Quantitative Investigations in Theoretical Linguistics.* Universität Tübingen. Jacobs, Joachim. 1994. „Das lexikalische Fundament der Unterscheidung von fakultativen und obligatorischen Ergänzungen." In *Zeitschrift für Germanistische Linguistik* 22.3: 284–319. Jacobs, Joachim. 2016. „Satztypkonstruktionen und Satztypsensitivität." In Rita Finkbeiner und Jörg Meibauer, Hrsg. *Satztypen und Konstruktionen im Deutschen*. Berlin/Boston: de Gruyter: 23–71.

The presence of light objects affects variable verb and subject placement in North Germanic

Maud Westendorp, Björn Lundquist
UiT The Arctic University of Tromsø, UiT The Arctic University of Tromsø
maud.westendorp@uit.no, bjorn.lundquist@uit.no

The North Germanic Languages (NGL) share many core syntactic patterns, e.g. V2, VO and prepositions. There is however considerable variation within and between the NGLs with respect to the relative order of phrases in the middle field. In this talk, we discuss two cases of free word order variation in North Germanic: (1) placement of postverbal DP subjects w.r.t. negation in Norwegian (1a-b, contra strict Sub>Neg order in Faroese and Danish); and (2) placement of finite verbs w.r.t. sentence adverbials in assertive embedded clauses in Faroese (2a-b, contra e.g. Swedish, where verb placement has clear pragmatic effects).

(1) a. I går   barberte **bakeren**   {seg}   ikke   {*seg} med   barberhøvel
       yesterday shaved   baker.def {self} not   {self} with razor

   b. I går   barberte {*seg} ikke   **bakeren**   {seg}   med   barberhøvel
      yesterday shaved   {self} not   baker.def {self} with razor

   'Yesterday, the baker didn't shave (himself) with a razor.'

(2) a. Páll   segði at   hann **raki**   {sær} altíð {*sær} um morgunin
       Paul   said   that he   shaved {self} always {self} in morning.def

   b. Páll   segði at   hann {*sær} altíð **raki**   {sær} um morgunin
      Paul   said   that he   {self} always shaved {self} in morning.def

   'Paul said that he always shaves (himself) in the morning.'

We uncover a hitherto unknown factor that influences the choice of word order: the presence of a light pronominal object. In Norwegian and Faroese, such objects also appear in the middle field, and obligatorily precede sentence adverbs (1a-2a, Object Shift, Holmberg 1986), but crucially only if the finite verb and the subject also precede the adverbial (1b-2b).

Based on the results from a large-scale elicited (spoken) production study across Scandinavia (Lundquist et al. 2019), we show that the presence of a light object in the clause increases the likelihood of a DP shifting past a sentence adverb in Norwegian, and a finite verb in an embedded clause shifting past a sentence adverb in Faroese. That is, it looks like Object Shift probabilistically affects the verb- and subject placement. In the study, the target sentences were either intransitive or contained a light pronominal object. We find that in Norwegian (N=62, 10 target items), the subject is significantly more likely to shift across negation if accompanied by a light object (14% vs. 40%, $p < 0.01$, logistic glmer). In Faroese (N=33, 12 target items), we found that on average 40% of the verbs were shifted past a sentence adverbial. Here, the presence of a light pronominal objects increased the likelihood of the verb shifting across the adverb (35%–46%, $p<0.05$).

We propose that the Scandinavian TP/IP is basically flat. The word order is partly determined by strict linear ordering statements between pairs of constituents (e.g., order preservations such as Verb>Object, Subject>Object). In the absence of such ordering statements, the linearization algorithm chunks together frequent/recently activated bigrams first. Bigrams consisting of frequent function words will generally have a higher frequency that any pairs including lexical words. In the two cases discussed here, the pair Reflexive- Negation is most likely first chunked (in their most common order), forcing verb and subjects to precede the negation.

How free is the position of German object pronouns?

Markus Bader
Goethe Universität Frankfurt
bader@em.uni-frankfurt.de

Object pronouns in the German midfield present a case of free variation insofar as they can appear after or before the subject without any semantic or pragmatic effects, as illustrated in (1) (see Lenerz, 1992).

(1) Peter sagte, dass (ihn)   der   Lehrer   (ihn)   lobte.
     Peter said   that him.ACC the.NOM teacher   him.ACC praised
     'Peter said that the teacher praised him.'

In a corpus study (Bader, 2020), several factors known to determine word order were found to modulate the probability of object pronouns to appear in front or behind the subject. For certain factor combinations, one order was used almost obligatorily; for example, psych verbs together with inanimate subjects (e. g., *dass ihn das Buch langweilte* 'that the book annoyed him') resulted in over 90% OS order. In other cases, choice between SO and OS was approximately balanced (e. g., agentive verbs with definite subject, see (1)). A statistical model predicted the correct word order in about 76.7% of all cases, which is above the baseline of about 67.0% OS order, but far from perfect. The question then is how much of the unaccounted variation can be accounted for by factors not taken into account in the corpus study.

This presentation reports results from an ongoing experimental series addressing this question. The first two experiments investigated the same material in a production and an acceptability experiment. The production experiment required from participants to recall memorized main clauses as embedded clauses (see (2)).

(2) a. TARGET: Den   Regisseur hat der   (sehr faule) Mann gelangweilt
                 the.ACC   director   has the.NOM very lazy   man   bored
     b. PROMPT: Der Regisseur   hat gesagt, dass
                 the director   has said   that

Participants read out the target sentence and memorized it. After a key press, they read out the prompt and completed the complement clause by converting the target clause into an embedded clause. The experiment varied word order in the main clause (factor Prime) and the subject's length and animacy. The same sentences, with the factor Prime replaced by Order of subject and pronominal object, were rated in a magnitude estimation experiment. The results match only partially. The acceptability results mirror a strong production preference for OS order with animate subjects, but not an effect of subject length visible in the production data.

A further production experiment using the same recall method explored the possibility that some of the unaccounted variation is related to individual properties of speakers/writers. Two individual properties were investigated: working memory capacity and processing speed. Working memory capacity was measured using a reading span test (Unsworth, Heitz, Schrock & Engle, 2005). The results show a correlation between recall errors and reading span but not between word order choice and reading span. Processing speed, however, correlated with word order choice.

I will discuss the results with regard to the relationship between grammar and usage, in particular the online processes of sentence formulation.

References: Bader, Markus. 2020. The position of object pronouns in the German middlefield. *Linguistics 58*. 1059-1115. Lenerz, Jürgen. 1992. Zur Syntax der Pronomina im Deutschen. *Sprache und Pragmatik* 29. Unsworth, Nash, Richard P. Heitz, Josef C. Schrock & Randall W. Engle. 2005. An automated version of the operation span task. *Behavior Research Methods* 37. 498– 505.

Freie Variation und Fugenelemente: Theorie und korpuslinguistische Realität

Marek Konopka
Leibniz-Institut für Deutsche Sprache (Mannheim)
konopka@ids-mannheim.de

Variation kann im Umkreis gedanklicher Schemata gesehen werden, auf die auch Konzepte wie ‚Tertium Comparationis', ‚Genus proximum' und ‚Differentia specifica' oder auch ‚Kohyponymie' zutreffen. Im Idealfall ist bei freier Variation die Variantenwahl weder durch innersprachliche noch durch außersprachliche Faktoren bedingt. Diesem Idealfall wird oft die kombinatorische Variation entgegengestellt, bei der Varianten komplementär verteilt sind.

Der Unterschied kann als Opposition zwischen gemeinsamen und verschiedenen Kontexten modelliert werden, in die die Varianten einzusetzen sind. Strenggenommen liegt bei komplementärer Verteilung aber keine echte Variation vor, da eine Austauschbarkeit der vermeintlichen Varianten aufgrund der Kontextverschiedenheit nicht gegeben ist. Jede echte Variation ist wiederum automatisch frei: Sie ist nie endgültig determiniert, auch wenn manchmal Faktoren vorliegen, die eine Variante favorisieren.

Es wird dafür plädiert, anzuerkennen, dass es generell Vereinbarungssache ist, wie weit bei einer Untersuchung die Gemeinsamkeit der Kontexte gehen soll. So kann je nach Interessenlage etwa (1) eine grammatisch-strukturelle Äquivalenz (vgl. *Arbeit|s|weg* vs. *Heimat|ort*), (2) die Identität des Kotextes (vgl. *Umbruch|s|jahr* vs. *Umbruch|jahr*) oder auch zusätzlich (3) die Konstanz außersprachlicher Parameter wie Zeit, Raum oder Situation erwartet werden. Die Variation kann dann als frei im jeweils angesetzten Bereich erscheinen. Auf dem Weg von (1) nach (3) steigen im Übrigen die Chancen, dass man einer Variation innerhalb einer Einzelvarietät (Mikrovariation) und nicht Unterschieden zwischen Varietäten (Makrovariation) begegnet und dass eine Intra-Sprecher-Variation gefunden wird.

Die Einsetzbarkeit des Begriffs ‚freie Variation' wird anhand einer Korpusuntersuchung zum Auftreten von Fugenelementen in Komposita aus zwei Nomen (z. B. *Arbeit|s|weg, Ei|er|kopf, Arbeit|nehmer, Ei|weiß*) diskutiert. Der Untersuchung liegen 60.00 Kompositum-tokens aus dem Deutschen Referenzkorpus (DeReKo, vgl. Kupietz et al. 2018) zugrunde. Die Erstglieder der Komposita werden in drei Gruppen eingeteilt, je nachdem, ob sie ein prinzipiell kombinatorisches, weitgehend kombinatorisches oder ein prinzipiell variables Verfügungs-verhalten auszulösen scheinen. Die Gruppe mit variablem Verfügungsverhalten besteht aus Erstgliedern, die auf einen Konsonanten enden, und entspricht dem oben angeführten Variationstyp (1). Eine logistische Regressionsanalyse legt nahe, dass einige der potenziellen Einflussgrößen wie Art des Auslautkonsonanten, Silbenanzahl oder Häufigkeit des Erstglieds für das Verfügungsverhalten von Komposita von Bedeutung sind. Sie kann aber nur einen eher geringen Teil der Varianz in den Daten erklären. Somit scheinen insbesondere idiosynkratische Eigenschaften des Erstglieds die Variation zu beeinflussen. Man könnte hier von freier Variation sprechen, aber solche Eigenschaften erscheinen vielfach (wie bei *Ei|er|kopf* vs. *Ei|weiß*) zumindest teilweise historisch erklärbar. Und wie lässt sich eine historische Determiniertheit in die bisherigen Variationsüberlegungen integrieren?

Ein Ergebnis bahnt sich an: Freie Variation ist ein Konstrukt, das man pragmatisch einsetzen kann, vorzugsweise für Bereiche, die man bisher nicht erklären konnte. Man kann den Begriff ‚freie Variation' aber auch fallen lassen. Zumindest als Heuristik würde es der Forschung guttun.

References: Kupietz, Marc, et al. 2018. The German Reference Corpus DeReKo: New Developments – New Opportunities. In Nicoletta Calzolari et al. eds. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Miyazaki: ELRA. 4353-4360. Hansen, Sandra, Felix Bildhauer & Marek Konopka. Fugenelemente im Korpus: Regelhaftigkeit und Variation. Im Erscheinen.

Freie Variation = Einbahnstraße? Konzeptionelle und methodische Überlegungen am Beispiel von morphologischer Variation im Luxemburgischen

Nathalie Entringer
Universität Luxemburg
nathalie.entringer@uni.lu

Innerhalb der Variationslinguistik stellen die Kategorien „Noise" oder auch „freie Variation" aus konzeptioneller Sicht immer nur vorläufige Einordnungen dar, da man im Allgemeinen annimmt, dass eine definitive Kategorisierung nur noch nicht möglich ist, weil zum Beispiel die für die Variation entscheidende Variable noch nicht berücksichtigt wurde (cf. Cappelle 2009). Infolgedessen gilt die Kategorie „freie Variation" für die variationslinguistische Forschung als große Herausforderung. Gerade deswegen ist es unabdinglich, sich aus unterschiedlichsten Perspektiven mit diesem Konzept und auch mit dem als solches kategorisierte Sprachmaterial weiter auseinanderzusetzen. Dabei stellen sich vielerlei Fragen: Welche Methoden bzw. Methodentriangulation bietet sich zur Erforschung von freier Variation an? Wie lässt sich freie Variation konzeptionalisieren: Als (allgemeine) Wahlfreiheit (cf. Cappelle 2009), als Ausdruck individueller linguistischer Präferenzen (cf. Raumolin- Brunberg & Murmi 2011) oder als nicht-systematische Variation (cf. Ellis 1992; Song 2012). Unter welchen Umständen (u. a. Frequenz und Verteilung der Varianten) kann man überhaupt von freier Variation sprechen?

Dieser Beitrag versucht anhand eines umfangreichen Korpus aus luxemburgischen Sprachdaten, die mithilfe der mobilen Applikation *Schnëssen* seit April 2018 gesammelt werden (cf. Entringer et al. im Druck), diese Fragen zu beantworten. Der Fokus der variationslinguistischen Analyse liegt dabei auf der Morphologie, wobei beispielhaft drei Variationsphänomene im Mittelpunkt der Betrachtung stehen: die flexivische Variation des Superlativs *gréisst / gréisst-en-t / gréisst-en-t Haus* ‚größte Haus' bzw. des departizipialen Adjektivs *déi gefëllt / gefëllt-e Këscht* ‚die gefüllte Kiste' und die morphologische/prosodische Erweiterung von Präpositionaladverbien *dorun / dorun-ner* ‚daran'. Dabei zeigen sich zum einen unterschiedliche Erscheinungsformen von freier Variation, die theoretische Schlüsse zulassen, zum anderen methodische Herangehensweisen, die im Rahmen einer Beschäftigung mit freier Variation unabdinglich sind. Es wird deutlich, dass die unterschiedlichen Erscheinungsformen aus konzeptioneller Sicht eine Differenzierung zwischen unterschiedlichen Typen nahelegen. Methodisch zeigt sich, dass eine multiperspektivische Analyse unverzichtbar ist. Zum einen führt beispielsweise eine Erweiterung der Analyse der interindividuellen Variation durch die Analyse der intraindividuellen Variation (cf. Bülow, Scheutz, und Wallner 2019) zu weiteren Erkenntnissen. Zum anderen spielt die Akzeptanz der Varianten durch die Sprecher.innen bei der Kategorisierung der Variation als frei oder durch bestimmte Faktoren motiviert eine wichtige Rolle.

References: Cappelle, Bert. 2009. Can we factor out free choice? In Andreas Dufter, Jürg Fleischer, and Guido Seiler, eds. *Describing and Modeling Variation in Grammar.* Berlin. 183–201. Bülow, Lars, Hannes Scheutz, and Dominik Wallner. 2019. Variation and change of plural verbs in Salzburg's base dialects. In Antje Dammel, and Oliver Schallert, eds. *Morphological Variation. Theoretical and Empirical Perspectives*. Amsterdam. 95–134. Ellis, Rod. 1992. Learning to Communicate in the Classroom. A Study of two Learner Re- quests. *Studies in Second Language Acquisition* 14: 1–23. Entringer, Nathalie, Peter Gilles, Sara Martin, and Christoph Purschke. in press. Schnëssen. Surveying language dynamics in Luxembourgish with a mobile research app. *Linguistic Vanguard. A Multimodal Journal for the Language Sciences*. Raumolin-Brunberg, Helena, and Arja Nurmi. 2011. Grammaticalization and language change in the individual. In Heiko Narrog, and Bernd Heine, eds. *The Oxford Handbook of Grammaticalization*. Oxford. 251–262. Song, Lichao. 2012. On the variability of interlanguage. *Theory and Practice in Language Studies* 2 (4): 778–783.

Verbal periphrases, deontic modality and teenagers: Free variation in non-standard spoken Catalan?

Roser Giménez
Universitat de Barcelona and Laboratorio SQ-Lingüistas Forenses
rosergimenez@ub.edu

According to prescriptive Catalan grammar, the verbal periphrasis *tenir que* is not an acceptable alternative to the periphrasis *haver de* followed by an infinitive and the genuine periphrasis *tenir de*, also followed by an infinitive, "has lost its vitality and is not used in formal registers" ([GIEC] 2016: 951). Yet, these are all used by Catalan speakers in informal settings to express deontic modality (Sinner 2008: 534; Mier 1986; Stokes 2015). The first of these verbal periphrases with *tenir* rather than with *haver* originated from the contact with the Spanish *tener que*. Although the use of these forms in informal settings has been mentioned in the literature repeatedly (e.g., Hualde 1992: 325; Rodríguez-Vida 1997; Martínez Díaz 2002: 87), since Mier (1986) the use of these different verbal periphrases has seldom been observed from a variationist perspective. This study analyzes the frequency and distribution of these variants in relation to independent linguistic and social variables in a corpus of 138 recordings of a role-play activity gathered as part of a larger longitudinal research project.

The independent variables in this study include grammatical person, reflexivity, hetero-repetition and speakers' age, sex and native language (L1). Quantitative and qualitative methods are applied to test the hypothesis that most instances of these constructions can be explained by some combination of the independent variables, rather than as the result of free variation. If this is the case, the factors studied here (or some of them) may be included in our models of (Catalan) grammar as related to teenagers' choice of deontic verbal constructions. Thus far, variation in spoken Catalan, including this phenomenon, has been attributed to confronting standard (i.e. prestigious) varieties (e.g. Bibiloni 1998). However, preliminary results for some of the independent variables (speakers' age, sex and L1) indicate that the variaton observed may be (at least partially) explained by these factors, since (i) the speakers in this study produce a higher percentage of *tenir que* and *tenir de* at T1 (ca. 12 years old) than at T2 (ca. 16 years old), 44.82% and 20.44% respectively; (ii) female speakers in this study produce a higher percentage of *tenir* periphrases than male speakers, 36.07% and 28.02% respectively; and (iii) the L1 bilingual Catalan and Spanish speakers in this study produce a higher percentage of *tenir* periphrases than L1 monolingual speakers (L1Bi: 45.13%, L1Sp: 36.08%, L1Ca: 24.95%). Qualitative and quantitative analyses will be performed to assess any potential relation between the independent variables and the deontic verbal periphrases in the corpus.

References : Bibiloni, Gabriel. 1998. "Variació i estàndard." *Caplletra* 25: 163-172. Cabanes Fitor, Vicent. 1996. "Les perífrasis modals de necessitat-obligació i probabilitat en català. Seguiment diacrònic: segles XIII al XIX." *Caplletra* 20: 129-164. Hualde, José Ignacio. 1992. *Catalan*. London and New York: Routledge. [GIEC]. 2016. *Gramàtica de la llengua catalana*. Barcelona: Institut d'Estudis Catalans. Martínez Díaz, Eva. 2002. *Las perífrasis modales de obligación 'tener que + infinitivo' y 'haber de + infinitivo': Variación e interferencia en el español de Barcelona.* [PhD thesis]. Universitat de Barcelona. Mier, Jeanne Zang. 1986. "Estudi sociolingüístic de certs aspectes de la llengua catalana." *Treballs de Sociolingüística Catalana* 6: 33-112. Rodríguez-Vida, Susana. 1997. *Catalán-castellano Frente a Frente*. Barcelona: Inforbooks. Sinner, Carsten. 2008. "Castellano y catalán en contacto: oralidad y contextos informales (Spanish and Catalan in contact: orality and informal contexts)." *Oihenart* 23: 521- 543. Stokes, Craig R. 2015. "The use of Catalan verbal periphrases *haver de* and *tenir que* on Twitter." *Sociolinguistic Studies* 9.4: 445-466.

Von *snoidel'n* und vom *hofdüütsch'en*: Zur phonetischen Variation im Pomerano

Göz Kaufmann, Daniel Duran
Universität Freiburg, Universität Freiburg
goez.kaufmann@germanistik.uni-freiburg.de, daniel.duran@germanistik.uni-freiburg.de

Wir präsentieren Daten zur phonetischen Variation im Pomerano, der Sprache der Nachfahren der im 19. Jahrhundert nach Brasilien ausgewanderten Hinterpommern. Für die Variationsforschung ist dabei der Umstand, dass aufgrund der restriktiven Sprachenpolitik des *Estado Novo* das Standarddeutsche als Schul- und Kirchensprache früh wegfiel, besonders wichtig. Der Wegfall dieser überdachenden Varietät hat nämlich zu einem beeindruckenden Grad an (freier) Variation geführt (vgl. für die Syntax Kaufmann in Überarbeitung; für die Lexik Kaufmann 2017). Dass diese Variation in den weit voneinander lebenden Sprachgemeinschaften in Rio Grande do Sul, Espírito Santo und Rondônia vergleichbar ist, legt entweder einen europäischen Ursprung nahe oder einen zur Zeit der Auswanderung vorhandenen sprachlichen *drift*. Ein interessantes Beispiel ist dabei der Nasalverlust (vor Frikativen; vgl. Postma 2019: 52). Aus *Mensch* war schon in Europa *meisch* geworden, aber diese Tendenz setzte sich in Brasilien fort. So tauchen neben dem kanonischen Konditionalsubjunktor *wen* ('wenn') nicht nur flektierte Formen wie *wens*/*went* (2/3SG) auf, sondern auch denasalisierte Formen wie *wes*/*wet*. Auch silbenfinal erodieren Konsonanten. *Kan* ('kann.1/3SG') wird normalerweise zu *ka* und *häw* ('habe.1SG') zu *hä* (vgl. Postma 2019: 107 und 109). Oft kommt es daneben zum Austausch von Sonoranten. Statt *sloidel* ('Schlüssel') hört man oft *snoidel*/*sloiden*, statt *küüna* ('konnten.3PL') *küüla* und statt *traurig traulig*. Daneben führt die phonetische Ähnlichkeit von /f/ und /ç/ (vgl. *laughter*/*Gelächter*) zu Dubletten wie *kuuchen*/*kuufen* ('Kuchen'; *kauke* existiert nicht!) oder *hochdüütsch*/*hofdüütsch* ('Hochdeutsch'). Eine wichtige Frage ist nun, ob es sich hier um freie Variationen handelt oder ob es Faktoren gibt, die diese Variationen erklären. Anhand von binär-logistischen Regressionsanalysen konnten wir für einige Phänomene eine überraschende Entdeckung machen. Neben dem Alter, dem Geschlecht, der Schulbildung und der Position im Satz erweisen sich oft Standarddeutschkenntnisse als entscheidend – und dies, obwohl es seit 80 Jahren praktisch keine Rolle mehr spielt und obwohl unsere Gewährspersonen über keine oder nur geringe Kenntnisse dieser Sprache verfügen.

Nichtsdestotrotz scheinen diese Kenntnisse silbenfinale Konsonanten vor der Erosion zu bewahren. Ob sie dabei alleine kausal wirken oder ob dies – wie eine längere Schulbildung und ein höheres Alter – der Normorientierung eines allgemein konservativen Habitus entspringt, ist eine zentrale Frage. Allerdings kann dieser Habitus nicht alle Phänomene erklären. Diese Fälle könnten entweder als ein Typ von freier phonetischer Variation erklärt werden oder zeugen davon, dass wir die entscheidenden Faktoren noch nicht isoliert haben – es also keine freie Variation gibt. Interessant ist daneben, dass viele Gewährspersonen eine interne Variation aufweisen, was für parallele mentale Repräsentationen spricht. Unsere Analysen zeigen, wie universelle Lautwandelprozesse und soziale/linguistische Faktoren interagieren und zu komplexen Variationsmustern führen. Sie fußen auf der wohl umfangreichsten Datensammlung zum Pomerano, die aus mündlichen Übersetzungen von 61 portugiesischen Stimulussätzen durch 250 Gewährspersonen (etwa 15.000 Einzelsätze) und aus zwölf Stunden freier Gespräche besteht.

References: Kaufmann, Göz (2017). '"Sorvete und Tema is nich Dütsch": Zur lexikalischen Integration von Entlehnungen in drei deutschen Varietäten Südbrasiliens' in: Eller-Wildfeuer, Nicole, Péter Maitz und Alfred Wildfeuer (Hrsg.): Sprachkontaktforschung – explanativ (ZDL- Themenheft). Stuttgart: Steiner. 260–307. Kaufmann, Göz (in Überarbeitung). 'In the thick of it: Scope rivalry in past counterfactuals of Pomerano'. Postma, Gertjan (2019). A Contrastive Grammar of Brazilian Pomeranian. Amsterdam/Philadelphia: John Benjamins.

Modelling the realization of variable word-final schwa in Standard French

Mathilde Hutin, Yaru Wu, Adèle Jatteau, Ioana Vasilescu, Lori Lamel, Martine Adda-Decker

Université Paris-Saclay, CNRS-LIMSI, Université Paris-Saclay, CNRS-LIMSI, Université de Lille, Université Paris-Saclay, CNRS-LIMSI, Université Paris-Saclay, CNRS-LIMSI, Laboratoire de Phonétique et de Phonologie, Université Paris 3 Sorbonne-Nouvelle

mathilde.hutin@limsi.fr, yaru.wu@limsi.fr, adele.jatteau@univ-lille.fr, ioana.vasilescu@limsi.fr, lori.lamel@limsi.fr, madda@limsi.fr

French schwa is traditionally referred to as a weak or reduced vowel noted [ə] restricted to unstressed syllables and variably alternating with zero. It can surface word-internally as in [səmɛn], *semaine*, 'week', or word-finally as in [katχə], *quatre*, 'four'. In Standard French, it is considered a deletable lexical vowel when word-internal, but an epenthetic segment when word-final (Tranel 1981, Eychenne 2019, Hutin et al. 2020).

Since Mende (1880), the patterns for its realization have been extensively studied, but less has been said about its distribution exclusively in word-final position in Standard French. The reason behind this imbalance lies in the fact that studies of variation phenomena are necessarily limited by the data. The initial ones mostly relied on grammaticality judgments by one or few informants: Consequently, they explored only lexical schwas, for which the judgment is straight-forward. For example, native speakers of French know that they can pronounce *pelouse*, 'lawn' either as [pəluz] or [pluz] but not *blouse*, 'blouse' as *[bəluz]. Later studies, based on small corpora, did not allow extended detailed research. To provide a statistically reliable picture, such fine-grained variable phonetic phenomena are best investigated with a sufficient amount of tokens from natural data (Coleman et al. 2016).

We thus used three large corpora: 1) ESTER (Galliano et al. 2005) contains 80h of (semi)prepared speech (radio broadcast news) that we filtered to keep only ca. 40h of Standard French data; 2) ETAPE (Gravier et al. 2012) contains 13.5h of radio data and 29h of TV data, including debates and interviews; 3) NCCFr (Torreira et al. 2010) is comprised of 31h of face-to-face interactions between friends. Following the method described in Hallé and Adda-Decker (2011), an automated speech recognition (ASR) system for French (Gauvain et al. 2002, 2005) was used in forced alignment mode systematically allowing variants both with and without schwa. For example, the word *mode*, 'fashion' could be aligned with the transcriptions [mɔd] or [mɔdə] depending on whether the system judged that the coda was followed by a schwa or not.

A generalized linear model was applied to the data to measure the part of extra- linguistic factors such as speech style (Wu et al. 2016, 2017), gender (Wu et al. 2017, Purse 2019) and orthography (Durand and Eychenne 2004, Eychenne 2019, Purse 2019) as well as linguistic factors such as phonotactic constraints on the length of the consonantal sequence around schwa-site (Grammont 1894, Delattre 1966, Bürki et al. 2011, Wu et al. 2017), the quality of the word-final consonant (Hansen and Mosegaard-Hansen 2002) and the quality of the first segment of the following word (Dell 1970, Côté 2000).

We thus propose the first extensive description of word-final schwa after all obstruents of Standard French, i.e. /ptkfsʃbdgvzʒ/, based on more than 110h of speech, i.e. ca. 125.000 tokens, validated with a statistical model. This study is interesting for the knowledge it provides regarding word-final schwa in French, but also as an example of what large corpora and automated methodologies can bring to linguistic inquiry of fine-grained free variation.

References: Bürki, Audrey, Myriam Ernestus, Cédric Gendrot, Cécile Fougeron and Ulrich Frauenfelder. 2011. "What affects the presence versus absence of schwa and its duration: a corpus analysis of French connected speech". *Journal of the Acoustical Society of America*, 130, 3980-3991. Coleman, John, Margaret E.L. Renwick and Rosalind A.M. Temple. 2016. "Probabilistic underspecification in nasal place assimilation". *Phonology*, 33(3), 425-458. Côté, Marie-Hélène. 2000. *Consonant Cluster Phonotactics: A Perceptual Approach*. PhD- thesis, MIT. Delattre, Pierre. 1966. "Studies in French and comparative phonetics". La Haye: Mouton. Dell. François. 1970. *Les règles phonologiques tardives et la phonologie dérivationnelle du français*. PhD thesis, MIT. Durand, Jacques and Julien Eychenne. 2004. "Le schwa en français : pourquoi des corpus?"

Corpus n°3 « Usage des corpus en phonologie », 311-356. Eychenne, Julien. 2019. "On the deletion of word-final schwa in Southern French". *Phonology*, 36.3, 1-35. Galliano, Sylvain, Edouard Geoffrois, Djamel Mostefa, Khalid Choukri, Jean-François Bonastre and Guillaume Gravier. 2005. "ESTER Phase II Evaluation Campaign for the Rich Transcription of French Broadcast News." *Proceedings of Interspeech 2005*, 2453–2456. Gauvain, Jean-Luc, Gilles Adda, Martine Adda-Decker, Alexandre Allauzen, Véronique Gendner, Lori Lamel and Holger Schwenk. 2005. "Where are we in transcribing French broadcast news?" *Proceedings of Interspeech 2005*, 1665-1668. Gauvain, Jean-Luc, Lori Lamel and Gilles Adda. 2002. "The LIMSI broadcast news transcription system". *Speech communication 37* (1-2), 89–108. Grammont, Maurice. 1894. "Le patois de la Franche-Montagne et en particulier de Damprichard (Franche-Comté). IV : La loi des trois consonnes". *Mémoires de la Société de linguistique de Paris* 8: 53-90. Gravier, Guillaume, Gilles Adda, Niklas Paulson, Matthieu Carré, Aude Giraudel and Olivier Galibert. 2012. "The ETAPE corpus for the evaluation of speech-based TV content processing in the French language". *LREC Eighth international conference on Language Resources and Evaluation*. Hallé, Pierre and Martine Adda-Decker. 2011. "Voice assimilation in French obstruents: A gradient or a categorical process?". *Tones and features: A festschrift for Nick Clements*, De Gruyter, 149–175. Hansen, Anita B. and Maj Brit Mosegaard-Hansen. 2002. "Le e prépausal et l'interaction". In Anita B. Hansen and Maj Brit Mosegaard-Hansen (eds.), *Structures linguistiques et interactionnelles dans le français parlé*. Actes du Colloque international, Université de Copenhague du 22 au 23 juin 2001 (Études romanes 54). Copenhague: Museum Tusculanum Press, 89-109. Hutin, Mathilde, Adèle Jatteau, Ioana Vasilescu, Lori Lamel, Martine Adda-Decker. 2020. "Le schwa final en français standard est-il un lubrifiant phonétique?" *Actes de CMLF* 2020. Mende, Ad. 1880. *Étude sur la prononciation de l'e muet*. Londres : Trubner. Purse, Ruaridh. 2019. "Variable Word-Final Schwa in French: An OT Analysis". *University of Pennsylvania Working Papers in Linguistics*, Vol. 25, Issue 1, Proceedings of the 42nd Annual Penn Linguistics Conference, Article 22. R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria. http://www. R-project.org. Torreira Francisco, Martine Adda-Decker and Mirjam Ernestus. 2010. "The Nijmegen Corpus of Casual French". *Speech Communication*, Elsevier: North-Holland, 2010, 52 (3) Tranel, Bernard. 1981. *Concreteness in Generative Phonology*. Evidence from French. Berkeley: University of California Press. Wu, Yaru, Martine Adda-Decker and Cécile Fougeron. 2016. "Rôle des contextes lexical et post-lexical dans la réalisation du schwa : apports du traitement automatique de grands corpus". *Actes de la conférence conjointe JEP-TALN-RECITAL 2016*, volume 1 JEP: 633-641. Wu, Yaru, Martine Adda-Decker, Cécile Fougeron and Lori Lamel. 2017. "Schwa Realization in French: Using Automatic Speech Processing to Study Phonological and Socio-Linguistic Factors in Large Corpora". *Proceedings of Interspeech 2017*, 3782-3786

Zur Stabilität flexionsmorphologischer Variation: Die Dativformen des unbestimmten Artikels im Zürichdeutschen

Anja Hasse
Universität Zürich anja.hasse@ds.uzh.ch

Im Schweizerdeutschen variieren die Formen des unbestimmten Artikels im DAT.MASK/NEUTR (vgl. bereits Stalder 1819, 89). Die Variation über das Dialektgebiet hinweg betrachtet lässt sich teilweise diatopisch erklären, für variierende Formen innerhalb eines Dialekts werden teilweise verschiedene funktionale Belastungen angenommen (vgl. Nübling 1992, 230). Das Zürichdeutsche weist weiter variierende Formen auf, die vom gleichen Sprecher in den gleichen sprachlichen Kontexten verwendet werden. Konkret variieren hier im DAT.MASK/ NEUTR des unbestimmten Artikels Formen ohne, (1), und mit einem Suffix -*ne*, (2), wobei die Formen jeweils vokalisch, (1a und 2a), oder konsonantisch, (1b und 2b), anlauten können.

(1)   a.   *eme*, *mit=eme* 'einem, mit einem' DAT.MASK/NEUTR
         b.   *i=me* 'in einem' DAT.MASK/NEUTR
(2)   a.   *emene*, *mit=emene* id.
         b.   *i=mene* id.

In einer Korpusstudie wurden spontansprachliche Daten des modernen Zürichdeutschen und literarische zürichdeutsche Texte aus dem 19. und 20. Jahrhundert auf die Auftretens-bedingungen der Formen in (1–2) untersucht, wobei sich gezeigt hat, dass sich die Variation nicht nur synchron, sondern auch diachron nachweisen lässt. Die Formen (1–2) variieren aber nicht völlig frei. Ihre Auftretenshäufigkeit ist vielmehr durch phonologische und syntaktische Faktoren bedingt (vgl. Hasse 2018, 2019). Dabei handelt es sich um ein komplexes Zusammenspiel von Variation auf der morphologischen (Auftreten des Suffixes) und der phonologischen Ebene (Anlaut der Artikelform). Beim Nebeneinander verschiedener Formen innerhalb einer Zelle eines Flexionsparadigmas handelt es sich um *Overabundance* (vgl. Thornton 2011a), bei der phonologischen Variation des Anlauts um *Shape Conditioning* (vgl. Thornton 2011b), beides Abweichungen von einem kanonischen Paradigma, wie es im Rahmen der *Canonical Typology* beschrieben wird (vgl. Corbett 2007).

     In meinem Vortrag gehe ich auf die Struktur dieses Flexionsparadigmas, die Stabilität variativer Flexionsformen und auf die Entstehung der zürichdeutschen Formen im DAT.MASK/ NEUTR ein. Dazu beleuchte ich die diachronen Hintergründe, die zu den stark variativen Verhältnissen im Zürichdeutschen geführt haben. Ich skizziere, welche Faktoren die Verteilung der Formen beeinflussen und diskutiere, wie hoch letztlich der Grad an freier Variation in Diachronie und Synchronie ist.

References: Corbett, Greville G. 2007. "Canonical typology, suppletion and possible words". *Language* 83. 8-42. Hasse, Anja. 2018. *Morphologische Variation in der Indefinitartikelflexion des Zürichdeutschen*. Unveröffentlichte Doktorarbeit. Universität Zürich. Hasse, Anja. 2019. "The interaction of phonological and morphological variation in Zurich German". In: Dammel, Antje & Schallert, Oliver, eds. *Morphological Variation. Theoretical and empirical perspectives*. Amsterdam / Philadelphia: John Benjamins Publishing Company. 223-242. Nübling, Damaris. 1992. *Klitika im Deutschen. Schriftsprache, Umgangssprache, alemannische Dialekte*. Tübingen: Narr. Stalder, Franz Joseph. 1819. *Die Landessprachen der Schweiz*. Aarau: Heinrich Remigius Sauerländer. Thornton, Anna M. 2011a. "Overbundance (Multiple Forms Realizing the Same Cell)". In: Maiden, Martin, ed. Morphological Autonomy. *Perspectives from Romance Inflectional Morphology*. Oxford: Oxford University Press. 358-381. Thornton, Anna M. 2011b. *Shape conditions and paradigms*. Unveröffentlichter Vortrag, 8. April, University of Surrey, Guildford, UK.

Overabundance at the interface

Noah Diewald
Ohio State University
diewald.21@osu.edu

In English, there is a famous pattern exemplified by leap, where both leaped and leapt are legitimate past tense inflections (Haber, 1976). Recently, the term overabundance (Thornton, 2012) has been applied to similar patterns found throughout the world's languages where more than one inflected form of a lexeme is compatible with the same inflectional category.

Here I discuss two Native American instances, one of which I encountered during ongoing fieldwork, to demonstrate theoretical implications of overabundance. The first is a case documented in Cochabamba Quechua (Muysken, 2002), where *warmis*, *warmikuna*, *warmiskuna* and *warmikunas* are all valid plurals of *warmi*, 'woman', a combination of variable affix ordering and multiple exponence. The second is a pattern I observed in Wao Tededo, an isolate spoken in the Ecuadorian Amazon. There, a periphrastic future tense is in free variation with a merged form. Both *bekebo* and *beke kebo* mean 'I will drink.'

Overabundance is disruptive because popular theories were developed around mechanisms intended to preclude it. Since Robins (1959), there has been a trend to emphasize the role of syntactic information in determining morphological form. If there is only one inflected form of a lexeme per morphosyntactic feature set, while there are many morphosyntactic feature sets compatible with a form, there is an asymmetry that implies morphological dependence on syntactic information but not vice versa.

This asymmetric pattern is modeled using Paninian or elsewhere rule ordering, where only the most specific rule compatible with a set of morphosyntactic features applies. The issue is that overabundance requires more than one rule to be equally compatible with a set of features. Paninian rule ordering, and the interface it models, does not always hold.

Morphosyntactic information is important but phonological and purely morphological patterns are needed to explain a variety of phenomena, including overabundance. A fruitful strategy is to model inflection using two tiers, one form-oriented and the other distribution oriented (Sadler and Spencer, 2001). Roughly, categories of the former include inflection classes, while those of the latter include categories like past participle. Each tier is internally consistent but when the categories of the two are correlated, overabundant patterns may result.

References: Haber, Lyn R. (1976). "Leaped and leapt: A theoretical account of linguistic variation". In: *Foundations of Language* 14.2, pp. 211–238. Muysken, Pieter (2002). "La categoría del plural en el quechua boliviano". In: *La Romania Americana. Procesos Lingüísticos en Situaciones de Contacto*. Ed. by Norma Díaz, Ralph Ludwig, and Stefan Pfänder, pp. 209–217. Robins, Robert H. (1959). "In defence of WP". In: *Transactions of the Philological Society* 58.1, pp. 116–144. Sadler, Louisa and Andrew Spencer (2001). "Syntax as an exponent of morphological features". In: *Yearbook of Morphology 2000. YOMO (Yearbook of Morphology)*. Springer, pp. 71–96. Thornton, Anna M. (2012). "Reduction and maintenance of overabundance. A case study on Italian verb paradigms". In: *Word Structure* 5.2, pp. 183–207.

Paradigm splits across parts of speech

Tania Paciaroni
University of Zurich
paciaron@rom.uzh.ch

This paper explores how free variation in inflectional paradigms can differ across parts of speech. In most languages, given the morphosyntactic and morphological features and values, we arrive at a single set of forms. In the Italo-Romance dialect of Ripatransone (southern Marche) a further choice is necessary, since items belonging to almost all parts of speech have two sets of inflectional forms with the same specification of features: a full form, and a parallel reduced form whose selection depends on conditions operating at different levels. The phenomenon is exemplified in (1) with the adjective 'beautiful'.

(1) a. [NP nu   **bbiéllə/-a**   rusccia]

          INDF.M.SG   beautiful.nonF.SG   red(M).SG

     b. [NP nu   rusccia   **bbiéllu/*-ə/-a**]

          INDF.M.SG   red(M).SG   beautiful.M.SG/-nonF.SG

          'A beautiful red' (LuCa_RIP11_178)

In (1a) the adjective is in prenominal position and selects the reduced form in *-ə*, while in (1b) it is in postnominal position and selects the full form in *-u*, the suffix of the masculine singular.

A major point of interest is that reduced inflection differs across parts of speech both in terms of the persistence of the distinctions and the range and distribution of the suffixes. Consider the tables (2)-(3):

(2a) Adjective – Full inflection

(2b) Adjective – Reduced inflection

(3a) Verb – Full inflection

(3b) Verb – Reduced inflection

In the reduced paradigm of the adjective (2b), a stable opposition of two different forms is maintained (*-e*.F.SG vs. *-ə*.nonF.SG). In the finite forms of the verb (3b), however, we observe an on-going change towards neutralization with the collapse of any affixed distinction and the extension into all cells of two alternative default suffixes, *-e* -and *-ə*. This difference can be explained by invoking both an internal (morphosyntactic specification of the exponents) and an external reason (contact with Standard Italian and other Italo-romance varieties).

The paper aims to deepen the study of these different types of (free) variation through the analysis of new data collected during interviews with ten native speakers and annotated in *The Zurich Database of Agreement in Italo-romance* (DAI, 18,577 tokens).

References: DAI: Loporcaro, Michele, Tania Paciaroni, Diego Pescarini, Alice Idone, Serena Romagnoli, Chiara Zanini, and Taras Zakharko, 2020. *The Zurich Database of Agreement in Italo-Romance*. University of Zurich. http://www.dai.uzh.ch.

Eigenschaften der syntaktischen Allostruktionen: Am Beispiel des deutschen *je-desto*-Gefüges

Jirayu Tharincharoen
Universität Erlangen-Nürnberg
jirayu.tharincharoen@fau.de

Der Beitrag beschäftigt sich mit den deutschen komparativen Korrelativkonstruktionen (fortan: KK) oder des sogenannten *je-desto*-Gefüges wie in (1).

(1) [*Je länger ich dich kenne,*]ₐₙₜₑzₑdₑₙz [***desto** mehr liebe ich dich.*]ₖₒₙₛₑqᵤₑₙz

In (1) ist KK mit dem Konnektorpaar *je – desto* eingeleitet. Jedoch lassen sich deutsche KK auch mit anderen Konnektorpaaren wie *je – umso, umso – umso* usw. einleiten. Gemäß Reis (2009) sowie Fortmann (2016) unterscheidet sich das jeweilige Konnektorpaar durch Registerunterschiede. Beispielsweise ist *umso – umso* überwiegend umgangssprachlich.

Allerdings wurden sprachinterne Faktoren wie Teilsatzabfolge (präponierte Antezedenz/Konsequenz) usw. in der Forschungsliteratur kaum erwähnt. Demnach stellt sich die Frage, ob die lexikalische Variation der jeweiligen Konnektorpaare von sprachinternen sowie -externen Faktoren gesteuert wird. Und wenn nicht, dann stellt sich die weitere Frage, ob die freie lexikalische Variation als Allostruktion gemäß Cappelle (2009) klassifizierbar ist.

Zur Untersuchung wurde eine Korpusstudie durchgeführt. Gemäß der Korpusanalyse unterscheiden sich einige Konnektorpaare durch grammatische Kategorien wie Teilsatzabfolge unterscheiden. Beispielsweise präferiert das Konnektorpaar *je – immer* die präponierte Konsequenz.

Auch im Bereich der lexikalischen Kategorien wie Adjektivpaar ist zu beobachten, dass das Konnektorpaar *je – je* das Adjektivpaar *länger – mehr* wie in (2) präferiert. Damit lässt sich festhalten, dass sprachinterne Faktoren die lexikalische Variation der KK steuern.

(2) *Wo ich hingegen **je länger je mehr** Mühe kriege, ist mit dem "Absolutheitsanspruch" des missionalen Gemeindemodels*.
 http://blog.igw.edu/2009/06/19/hangen-geblieben-drei-monate-danach/ 19.06.2009

Jedoch weisen einige Konnektorpaare wie *umso – umso, umso – desto* usw. weder formale noch funktionale Unterschiede voneinander auf. Sie weisen ihre gemeinsamen Eigenschaften auf wie z.B. die niedrige Tokenfrequenz, die starke Präferenz für die präponierte Antezedenz, das Vorkommen in Textsorten mit konzeptioneller Mündlichkeit.
Dementsprechend lassen sie sich als Allostruktion betrachten.

Zusammenfassend lässt sich festhalten, dass einerseits sprachinterne Faktoren zur unfreien lexikalischen Variation der Konnektorpaare führen. Andererseits muss berücksichtigt werden, dass sprachexterne Faktoren die Allostruktionen bzw. freie Variation der Konnektorpaare ermöglichen können.

References: Cappelle, Bert. 2009. "Can we factor out free choice?" In Andreas Dufter, Jürg Fleischer and Guido Seiler, eds. *Describing and Modeling Variation in Grammar*. Berlin/New York: Mouton de Gruyter, 183–201. Reis, Marga. 2009. "Zur Struktur von Je-desto-Gefügen und Verwandtem im Deutschen." In Veronika Ehrich, Christian Fortmann, Ingo Reich and Marga Reis. *Koordination und Subordination im Deutschen*. Linguistische Berichte Sonderheft 16. Hamburg: Buske, 223–244.

Modelling free variation of linking elements after feminine noun stems in German

Jakob Maché
Universidade de Lisboa
jakob.mache@letras.ulisboa.pt

Assuming a hybrid moderate word-form lexicon, which contains both morpheme roots and complex word-forms, the present study outlines an analysis that models the free variation of the linking element -s- after feminine noun stems in German com- pounds and derivational suffixation.

As well known, there are a couple of feminine nouns in German which combine with the linking element (LE) -s- in German. Corpus data from the DTA and DeReKo show that there is a different degree obligatoriness depending on the feminine noun, whereas *Liebe* 'love' de facto requires the LE -s-, *Kraft* 'power' at the other and of the scale takes the LE only under very limited circumstances; other nouns like *Hilfe* 'help' occur in frequent formations with -s- (*Hilf-s-arbeiter* 'assistant worker' vs. *Hilfestellung* 'assistance' *hilf-Ø-los/*hilf-s-los* 'helpless'). This paper is investigating the free variation of LE -s- among different feminine nouns.

As will be shown here, none of these feminine nouns can be considered historically as a feminine simplex noun, rather the known cases can be divided into two big groups: Firstly, nouns which were subject to gender alternation (eg. *Hilfe*) or secondly derivates. As often confused with simplicia, we will take a closer look at two lesser investigated and more opaque classes of derivation: -t-derivates, which according to Wilmanns (1896: 327–345) are mainly derived from verb stems (§254–261) or noun stems (§262–264) and -j-derivates, which are all above derived from adjectives and occasionally from other nouns (cf. Wilmanns 1896: 248–251).

In order to explain the occurrence of non-paradigmatic LE -s- with feminine nouns, we are going to revive an assumption originally advocated by Nübling and Szczepaniak (2008, 2009), according to which a morphologically complex determinans makes the selection of LE -s- more likely. For a whole range of cases attested in DTA and DeReKo, in which the determinans is a feminine compound with -s-, and which become less acceptable once the determinans is no longer a compound: *Allmacht-s-phantasie* vs ??*Macht-s-phantasie*, *Eifersucht-s-drama* vs. ??*Sucht-s-drama*, *Völkerschlacht-s-denkmal* vs. ??*Schlacht-s- denkmal*, *Viehzucht-s-verein* vs. ??*Zucht-s-verein*. In doing so, it will be shown that the distribution of LE -s- after feminine nouns is not only conditioned phonological factors (cf. Kopf 2018a: 109; 2018b: 324–326), but also by morphological ones However, some extent of free variation are also caused by lexicalisation which conserves earlier stages (-s- after simple derivates like *liebe* or *Schamesröte*) and interspeaker variation. We conclude here that non-paradigmatic LE -s- can be analysed as a marker that can optionally attached to the class of morphologically complex determinans, for some selected classes and nouns however it is actually the only option due to phonological or individual lexical reasons.

References: Kopf, Kristin 2018a. "From genitive suffix to linking element". In Tanja Ackermann, Horst J. Simon, and Christian Zimmer, eds. *Germanic Genitives*. Amsterdam: Benjamins, 91– 114. 2018b. *Fugenelemente diachron*. Berlin: Mouton de Gruyter. Nübling, Damaris and Renata Szczepaniak 2008. "On the way from morphology to phonology". In: *Morphology* 18.1, 1–25. 2009. "*Religion+s+freiheit*, *Stabilität+s+pakt* und *Subjekt*(+*s*+)*pronomen*: Fugenelemente als Marker phonologischer Wortgrenzen". In: Peter O. Müller, ed. *Studien zur Fremdwortbildung*. Hildesheim: Olms, 195–222. Wilmanns, Wilhelm 1896. *Deutsche Grammatik Gotisch, Alt-, Mittel- und Neuhochdeutsch. Zweite Abteilung: Wortbildung*. Strassburg: Trübner.

Arbeitsgruppe 5

Encoding aspectuality in Germanic languages — empirical and theoretical approaches
Hanna Fischer, Melitta Gillmann & Mirjam Schmuck

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

The modal basis of progressive marking

Frank Brisard
University of Antwerp
frank.brisard@uantwerpen.be

My work on modal uses of markers of progressive aspect starts from an analysis of the simple vs progressive alternation in the English present-tense paradigm (Brisard 2002, De Wit & Brisard 2014). I contend that this alternation reflects not only a temporal contrast (roughly, perfective vs imperfective construals of a state of affairs coinciding with the time of speaking), but also, and arguably primarily, an epistemic one, corresponding to the basic cognitive distinction between "structural" vs "phenomenal" knowledge. This is in line with the claim in Cognitive Grammar that so-called grounding predications (e.g., tense and modals) convey epistemic meanings at the most schematic level, but it extends that claim to what I regard as periphrastic tense-aspect units, i.e., conventional collocations of aspect markers with certain tenses. One important implication of this analysis is that *all* central tense markers, including simple ones, have an aspectual value, even if that value is semantically underspecified (as in Germanic languages other than English for the present). In fact, the aspectually (non-)specific nature of a language's present tense turns out to be a good basis for predicting the more temporal (or, conversely, more modal) orientation of that language's progressive construction(s). I distinguish between languages, like English, in which progressive marking (at least in the present-tense paradigm), is grammatically obligatory to express ongoingness with certain types of verbs (i.e., dynamic ones), and those where it is optional and its use is, at least initially, more often than not motivated by non-temporal concerns, typically of an expressive/subjective nature (De Wit et al. 2020). This will be illustrated on the basis of existing case studies of Dutch, German, and Afrikaans (Anthonissen et al. 2016, 2019, Breed et al. 2017).

References: Anthonissen, Lynn, Astrid De Wit and Tanja Mortelmans. 2016. "Aspect meets modality: A semantic analysis of the German *am*-progressive." *Journal of Germanic Linguistics* 28: 1-30. Anthonissen, Lynn, Astrid De Wit and Tanja Mortelmans. 2019. "(Inter)subjective uses of the Dutch progressive constructions." *Linguistics* 57: 1111-1159. Breed, Adri, Frank Brisard and Ben Verhoeven. 2017. "Periphrastic progressive constructions in Dutch and Afrikaans: A contrastive analysis." *Journal of Germanic Linguistics* 29: 305-378. Brisard, Frank. 2002. "The English present." In Frank Brisard, ed. *Grounding: The Epistemic Footing of Deixis and Reference*. Berlin: De Gruyter Mouton. 251-297. De Wit, Astrid and Frank Brisard. 2014. "A Cognitive Grammar account of the semantics of the English progressive." *Journal of Linguistics* 50: 49-90. De Wit, Peter Petré and Frank Brisard. 2020. "Standing out with the progressive." *Journal of Linguistics* 56: 479-514.

Beyond progressive aspectuality: Aspectual *aan*-constructions in Dutch

Maarten Bogaards
Leiden University
m.p.m.bogaards@hum.leidenuniv.nl

Research into how Germanic languages encode aspectuality often proceeds in a top-down fashion. First a conceptual crosslinguistic category is taken as a point of departure, for instance 'progressivity'. Then that conceptual category is used to analyze and compare specific aspectual forms in language(s), for example the English '*ing*-progressive' (Boogaart 1999), the '*am*-Progressiv' in certain German varieties (Van Pottelberge 2004) and the Dutch '*aan het*-progressief' (Lemmens 2015).

This talk argues that a bottom-up, form-driven approach adds an important empirical dimension to this type of analysis. Specifically, it shows that the pattern widely analyzed as 'the Dutch progressive' — i.e. '*aan het* + infinitive' paired with *zijn* 'to be', cf. (1) — shares crucial formal and semantic properties with certain other phrases headed by the preposition *aan* 'on'. Instead of an infinitive, these may feature a verb stem, cf. (2), or a noun, cf. (3). Like (1), (2)-(3) appear to encode aspectual meaning.

(1)   Jan   is   aan   het   lezen
          John is   on   the   read.INF
          'John is reading.'
(2)   Jan   is   aan   de   wandel.
          John is   on   the   stroll.STEM
          'John is taking a stroll.'
(3)   Jan   is   aan   het   bier.
          John is   on   the   beer
          'John is drinking beer.'

Recognizing the formal and semantic connections between the aan-patterns in (1)-(3) complicates the idea of a 'standalone' Dutch progressive. Instead, a fine-grained analysis of their (semi-)aspectual similarities and differences contributes to a more empirically founded understanding of the Dutch aspectual inventory as such, and the place of conceptual categories like 'progressive aspectuality' within language-specific aspectual inventories.

References: Boogaart, Ronny.1999. *Aspect and temporal ordering. A contrastive analysis of Dutch and English*. Ph.D. thesis. Vrije Universiteit Amsterdam. Lemmens, Maarten. 2015. "Zit je te denken of ben je aan het piekeren? Persistentie in het synchrone gebruik van de PREP- en POS-progressiefconstructies in het Nederlands." *Nederlandse taalkunde* 20: 5-36. Van Pottelberge, Jeroen. 2004. *Der am-Progressiv. Struktur und parallelle Entwicklung in den kontinental-westgermanischen Sprachen*. Tübingen: Gunter Narr.

Diatopic and diachronic variations of the German *am*-progressive: A corpus-based investigation

Jianan Li
University of Göttingen
jianan.li01@stud.uni-goettingen.de

This talk concentrates on the *am*-progressive construction in German, such as *Ich bin am arbeiten* 'I am working'. The *am*-progressive has been regarded as typical of dialects in the Rhine region. However, it has evolved rapidly over the past 20 years and has been increasingly used nationwide. The *Atlas zur deutschen Alltagsprache* (Elspaß & Mölle 2003ff.) delivers a livelier image of the *am*-progressive in spoken German in the Rhine region as in other regions. While previous studies (e.g. Krause 2002; Flick & Kuhmichel 2013; Gárgyán 2014; Flick 2016) have explored mostly overall synchronic usage, my corpus-based investigation focuses on the latest diachronic development of the *am*-progressive considering its geographical spread. Using the Mannheim German Reference Corpus, eight local newspapers from eight dialect areas in Germany are selected to test the frequency and flexibility of the *am*-progressive in written German.

I will show that the frequency of use of the *am*-progressive increased overall from 2005 to 2015, with regional differences; that is, it increased in the West Central, West Upper and East Upper German regions but dropped significantly in the East Low German regions.

Furthermore, I will argue that in 2015, the *am*-progressive displays very large flexibility in use and less restrictions, without a clearly regional limit. It can be combined with most verb forms except passive and imperative. In conclusion, I will argue that the further advancement of the *am*-progressive from 2005 to 2015 endorses the view that this construction strongly contributes to the establishment of aspectuality in German.

References: Elspaß, S. und R. Möller. 2003ff. *Atlas zur deutschen Alltagssprache (AdA)*. http://www.atlas-alltagssprache.de. Flick, J. 2016. "Der am-Progressiv und parallel am V-en sein-Konstruktionen: Kompositionalität, Viabilität und Netzwerkbildung." *Beiträge zur Geschichte der deutschen Sprache und Literatur* 138(2): 163-196. Flick, J. und K. Kuhmichel (2013). "Der am-Progressiv in Dialekt und Standardsprache." *Jahrbuch für germanistische Sprachgeschichte* 4 (2013): 52-76. Gárgyán, G. 2014. *Der am-Progressiv im heutigen Deutsch: Neue Erkenntnisse mit besonderer Hinsicht auf die Sprachgeschichte, die Aspektualität und den kontrastiven Vergleich mit dem Ungarischen*. Frankfurt a. M.: Peter Lang. IDS Mannheim, *Das Deutsche Referenzkorpus DeReKo*. https://www.ids-mannheim.de/kl/projekte/korpora/. Krause, O. 2002. *Progressiv im Deutschen: Eine empirische Untersuchung im Kontrast mit Niederländisch und Englisch*. Tübingen: Niemeyer.

Grammaticalization in speech-islands: Possibilities and neglects

Adam Tomas
Ludwig-Maximilians-Universität München
adam.tomas@campus.lmu.de

The progressive markers of Germanic languages had being given a stepmotherly treatment and they were examined too hesitant for too long. The well-known explanation for this neglect was covered by the fact that the progressive is nowhere grammaticalized to the same extent as in English. There was also suspected a lack of grammatical morphemes in other Germanic languages to encode progressive aspect. Is this really true?

Within the linguistic diversity, however, it should be noted that some modern West Germanic languages have developed similar progressive constructions with structural similarities in the form of a prepositional progressive construction:

(1) German: *Ich bin am Lesen/lesen*.
(2) Dutch: *Ik ben aan het lezen*.
(3) Afrikaans: *Ek is aan die lese*.

It is surprising that the use of the extremely multifaceted German progressive *Verlaufsform* (*sein$_{Finitum}$+am+V$_{Infinitiv}$*) is avoided in the German standard written language. A sentence like *Egon ist ein Buch am lesen* opens the area of verbal aspectuality to German. The *am-* constructions are viewed with great scepticism by the general public because they are regarded as linguistic deviations from the norm. Such a view is no longer tenable from the author's point of view. The Pennsylvania German (PeD), the language of the Amish and Mennonite as German descendants in the USA, offers both expedient and surprising references. The PeD has so far not produced any prescriptive normalization, so that a complete morphological paradigm of the progressive markers has been established. The am- progressive paradigm in PeD is a simple proof of a grammatical unit, which does not exist in this form in any West Germanic language, except in English. In PeD in particular, am- progressive constructions represent a very frequent formal characteristic of incompleteness in the representation of a verbal situation. The am-progressive is used very often and with many additions or syntactical extensions, as shown in the data from my field-research 2014:

(4) *D Aenn is die Ebbel am schaela*. [Ann is pealing the apples.]
(5) *Ich bin sunndaags mei Guckbox am watscha*. [I am watching my TV sunndays.]

However, the progressive constructions in PeD have reached a previously unknown degree of grammaticalization with a functioning passive form (*sein$_{Finitum}$+am+PartizipII+werden$_{Infinitiv}$*):

(6) *Viele Haisa sind am gbaut werra do*. [Many houses are being built over there.]
(7) *Dei Pois sind am eingwrappt werra now*. [Your pies are now being wrapped.]

In my presentation I want to share some of my elicited data and explore these concrete questions about the use, acceptance and morpho-syntactic expandability of these grammatical forms in both PeD and standard German.

References: Ebert, Karin. 2000 . "The progressive markers in Germanic languages." In Dahl, Östen, eds. *Tense and aspect in the languages of Europe*. Berlin/New York: Morton de Gruyter. 42-74. Rödel, Michael. 2004. "Grammatikalisierung und die Folgen: Der Infinitiv in der deutschen Verlaufsform." *Muttersprache. Vierteljahresschrift für deutsche Sprache* 2: 138-150. Pottelberge, Jeroen van. 2004. *Der am-Progressiv. Struktur und parallele Entwicklung in den kontinentalwestgermanischen Sprachen*. Tübingen: Narr. Tomas, Adam. 2018. *Der am-Progressiv im Pennsylvaniadeutschen. Grammatikalisierung in normfernen Varietäten*. Tübingen: Narr Francke Attempto.

Periphrastic *tun* in Australian German: A past tense habitual marker?

Anna Saller
University of Regensburg
Anna.Saller@sprachlit.uni-regensburg.de

German varieties in Australia are shortly before language shift, and show many phenomena of language erosion, including analytical constructions such as *tun* + infinitive. In addition to its use (as shown below) in conditional clauses (1), in the subjunctive (2), and for emphasis (3), a habitual or imperfective character (4) is often discussed – or the tun periphrasis is discarded as a semantically empty, syntactic variant.

(1)   *wenn du sie nicht sprechen tust, vergisst du sie* ('if you do not speak it, you forget it')
(2)   *ich täte mich beeilen* ('I did (would) hurry up')
(3)   *sie tut sich ja Mühe geben* ('she does actually make an effort')
(4)   *wir tun immer montags Kaffee trinken* ('We do have coffee every Monday')

A diachronic study of Australian German based on spontaneous speech produced in interviews shows that periphrastic *tun* occurred in the 1960s/70s both in the present and in the preterite, while between 2009 and 2014, it was used almost exclusively in the preterite. The first data set is taken from the ›Monash Corpus of Australian German‹, by Prof. Dr. M. G. CLYNE, accessible via the database for spoken German (dgd.ids-mannheim.de). The second data set was provided by Prof. Dr. C. M. RIEHL from her project on ›Barossa German as a Relic Variety‹.

Given the fact that imperfective contexts predominate in the *tun* periphrases in both corpora, and that the present tense by nature produces an imperfective reading, it is reasonable to assume that the combination of past tense + imperfective reading is marked and that periphrastic tun is used for this semantic niche (KLEMOLA noted this for the South West English dialect in Somerset), as exemplified in the following sentences:

(5)   *er tat immer viele Briefe schreiben* ('He used to write a lot of letters')
(6)   *jeder hier tat Mandeln anbauen* ('Everyone here used to grow almonds')

An analysis of the situational contexts in which the *tun* periphrasis is used, as well as of morphological characteristics of the lexical verbs, helps to clarify whether this analytical construction has actually developed into a habitual or imperfective past tense marker over the course of 50 years, or into an analytical past tense only, and how this structure fits into the developmental course of dwindling varieties.

References: Bybee, J., R. Perkins and W. Pagliuca. 1994. *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World*. Chicago, London: University of Chicago Press. Comrie, B. 1976. *Aspect. An Introduction to the Study of Verbal Aspect and Related Problems*. London, New York, Melbourne: Cambridge University Press. Klemola, J. 1998. "Semantics of do in southwestern dialects of English English." In Tieken- Boon van Ostade, I., M. Van der Wal and A. Van Leuvensteijn, eds. *Do in English, Dutch and German. History and Present-Day Variation*. Münster: Nodus. 25-52. Riehl, C. M. 2015. "Language attrition, language contact and the concept of a relic variety: the case of Barossa German." *International Journal of the Sociology of Language* 236: 261-293.

Pseudo-coordinated *sitzen* ('sit') and *stehen* ('stand') in spoken German: A case of emergent progressive aspect?

Nadine Proske
Leibniz-Institut für Deutsche Sprache Mannheim
proske@ids-mannheim.de

In the Scandinavian languages, pseudo-coordinated posture verbs are in the process of becoming aspectual markers (e.g. Hesse 2009). In German, pseudo-coordination is generally said to be not as widely used, but it has been shown to exhibit possible starting points for the grammaticalization of motion and (change of) posture verbs into different directions (Proske 2017, 2019). Based on data from the Research and Teaching Corpus of Spoken German, FOLK (http://agd.ids-mannheim.de/folk.shtml), my study examines the pseudo-coordinated posture verbs *sitzen* ('sit') and *stehen* ('stand') in spoken German, which have developed subjective and aspectual meaning components and can mark the activity denoted by the verb in the second conjunct as temporally extended or in progress (see example 1 below).

(2)   und myrte STEHT dann da- und FÖHNT sich die hAAre als ich komme. ('And Myrte stands there and blow-dries [= is blow-drying] her hair when I come.')

The analysis shows that the degree of grammaticalization is low (the verbal semantics is not clearly bleached; the progressive reading largely relies on the co-presence of temporal adverbials; constituents may occur between the coordinated verbs and the conjunction; the locational adverbial of the posture verb is almost always realized). Nonetheless, the construction shows signs of fixedness (e.g., a preference for the locational adverbial to be realized by *da* ('there')) and a potential to extend to new contexts (e.g., stative verbs in the second conjunct, despite a preference for activity verbs). Its aspectual potential will be discussed in relation to its subjective meaning components (e.g. passivity, diligence and intentionality) and to the further grammaticalized German *am*-progressive (e.g. Flick 2016).

References: Flick, J. 2016. "Der *am*-Progressiv und parallele *am V-en sein*-Konstruktionen: Kompositionalität, Variabilität und Netzwerkbildung." *Beiträge zur Geschichte der deutschen Sprache und Literatur* (PBB) 138(2): 163-196. Hesse, A. 2009. *Zur Grammatikalisierung der Pseudokoordination im Norwegischen und in den anderen skandinavischen Sprachen*. Tübingen: A. Francke Verlag. Proske, N. 2017. "Perspektivierung von Handlungen und Zuschreibung von Intentionalität durch pseudokoordiniertes kommen." In Deppermann, A., Proske, N. und A. Zeschel, eds. *Verben im interaktiven Kontext. Bewegungsverben und mentale Verben im gesprochenen Deutsch*.Tübingen: 177-247. Proske, N. 2019. "Emergent pseudo-coordination in spoken German. A corpus-based exploration." *Yearbook of the German Cognitive Linguistics Association* (GCLA) 7: 115-136.

Is pseudocoordination an aspectual construction?

Torodd Kinn
University of Bergen
Torodd.Kinn@uib.no

Pseudocoordination is a frequent construction type in the Mainland Scandinavian languages (Danish, Norwegian, Swedish), and much research on pseudocoordination has focused on Germanic languages (Ross 2016). Early research typically concentrated on constructions involving posture verbs meaning 'lie', 'sit', and 'stand' as first verbs, e.g. as in *Ho **sit** og les dikt* [she **sit.PRS** and read. PRS poems] 'She's **(sitting)** reading poems'. Posture verbs are well-known grammaticalization sources of auxiliaries. It was observed that these contribute to the expression of aspectuality, along with a very few other verbs, e.g. 'go'. A number of syntactic and semantic properties distinguishing pseudocoordination from canonical coordination were identified. There developed a belief that constructions with these properties involve a closed class of aspectualizing first verbs.

However, research has gathered mounting evidence that there are number of variants of pseudocoordination which have the distinguishing properties. The paradigm of first verbs is far from small, including verbs of (assuming) posture, of (a)telic motion, of communication channel, and several other meaning types. Only very few of these first verbs express aspect. This has led to the realization that pseudocoordination does not fundamentally have to do with grammaticalization, but some types tend to become grammaticalization sources.

I will look more closely at pseudocoordination with two first-verb groups and one specific first verb:

- the verbs for 'lie', 'sit', and 'stand' in the Mainland Scandinavian languages (Kinn, Blensenius, and Andersson 2018). These constructions are highly conventionalized, but they continue to exhibit close ties between postures and concomitant activities or states, the former facilitating the latter. Still, there are some signs of bleaching and aspectualization.
- verbs of atelic motion (e.g., 'run around') in Norwegian (Kinn 2018). If supplied with an atelicizing adverbial ('around'), virtually any motion verb can be used in pseudocoordination. But absence of an adverbial correlates with bleaching and more prominent aspectuality.
- the verb *drive* 'carry on' (Hesse 2009, Kinn 2019). The Norwegian construction is the result of several reanalyses and involves the development from situational to viewpoint aspect.

References: Hesse, Andrea. 2009. *Zur Grammatikalisierung der Pseudokoordination im Norwegischen und in den anderen skandinavischen Sprachen*. Tübingen: Francke. Kinn, Torodd. 2018. "Pseudocoordination in Norwegian. Degrees of grammaticalization and constructional variants." In Evie Coussé, Peter Andersson and Joel Olofsson, eds. *Grammaticalization meets Construction Grammar*. Amsterdam: John Benjamins. 75-106**.** Kinn, Torodd. 2019. "Framveksten av pseudokoordinasjon med drive." *Norsk Lingvistisk Tidsskrift* 37: 207-236. Kinn, Torodd, Kristian Blensenius and Peter Andersson. 2018. *Posture, location, and activity in Mainland Scandinavian pseudocoordinations*. CogniTextes. Ross, Daniel. 2016. "Between coordination and subordination: Typological, structural and diachronic perspectives on pseudocoordination." In Fernanda Pratas, Sandra Pereira and Clara Pinto, eds. *Coordination and subordination*. Newcastle upon Tyne: Cambridge Scholars. 209-243.

Progressive and prospective in German dialects of Italy

Ermenegildo Bidese, Maria Rita Manzini
University of Trento, University of Firenze
e.bidese@lett.unitn.it, mariarita.manzini@unifi.it

German-based minority varieties in Italy in some cases seem to borrow constructions that can also be found in the surrounding Italo-romance varieties, such as the construction *nåsoin + inflected infinitive* ('to be after + inflected infinitive') in Cimbrian (cf. Ferraresi 2016):

(1)   I pin *nå   zo lesa*
        I am after   to read.INFL 'I'm reading'

There are other progressive constructions in Cimbrian, i.e. *dråsoin* + II (cf. Middle High German daran[e] 'thereby'), and *soin + drumauz* + inflected infinitive, specialized for inchoative and prospective meaning. We present new data collected in a large study about the progressive and prospective periphrasis with 34 speakers of Cimbrian. We compare the system of progressivity in Cimbrian with those of other German varieties spoken in Italy and other German-based minority languages, e.g. Pennsylvania Dutch (cf. Tomas 2018).

Theoretically, a question much debated within formal models regards the syntax of progressives. In the cartographic model of Cinque (2017), PROGR is a universal functional head in a monophrasal structure – though its overt realizations can stretch to apparently biclausal structures like (1). The alternative is that the meaning of the progressive is built from its component parts, very often biclausal structures with the embedded sentence introduced by a locative periphrasis (1). Recently Manzini et al (2017) argue that such structures are a good match to the Part-whole semantics of progressives proposed by Landman (1992). We argue that only such a constructivist perspective yields the required insights into microvariation, contact and change.

References: Ferraresi, Gisella. 2016. "Wandel im aspektuellen System des Zimbrischen." In Sergio Neri, Roland Schuhmann and Susanne Zeilfelder, eds. "*dat ihdir itnu bi huldi gibu*": *Linguistische, germanistische und indogermanistische Studien Rosemarie Lühr gewidmet*. Wiesbaden: Reichert Verlag. 101-112. Cinque, Guglielmo. 2017. "On the status of functional categories (heads and phrases)."**m***Language and Linguistics* 18(4): 521-576. Landman, Fred. 1992. "The Progressive." *Natural Language Semantics* 1: 1-32. Manzini, M. Rita, Paolo Lorusso and Leonardo M. Savoia. 2017. *a/bare finite complements in Southern Italian varieties: Mono-clausal or bi-clausal syntax?* Quaderni di Linguistica e Studi Orientali (QULSO)3. Tomas, Adam. 2018. *Der am-Progressiv in Pennsylvaniadeutschen*. Tübingen: Narr Verlag.

The syntactic expression of prospective aspect in German

Jens Fleischhauer
Heinrich-Heine Universität Düsseldorf
fleischhauer@phil.uni-duesseldorf.de

In his monograph on aspect, Comrie (1976: 64) briefly mentions an aspectual form he calls `prospective aspect´. Prospective aspect relates a state to a subsequent situation and "defines a temporal phase located close before the initial boundary of the situation" (Kuteva 2001: 92). Although German has not grammaticalized prospective aspect, it expresses prospective aspect by means of light verb constructions (LVCs) consisting of the light verb *stehen* and a PP headed by the preposition *vor* `in front of´ (1). The interpretation of the LVC in (1) is that the subject referent is close to the event denoted by the PP-internal NP.

(1) *Der Kessel steht   vor   der Explosion*.
     the boiler stands   in_front_of the explosion
      `The boiler is close to explosion.´

An essential question is which nouns are permitted within this construction? At first glance, it looks as if only eventive nouns denoting a change of state are permitted. However, actual language data show that other types of nouns are possible as well.

(2) *das geht jedem   so, der   vor   seinem   ersten   Wettkampf steht.*
     this goes everyone so, REL.PRON in_front_of his   first   competition stands
      `everyone is like this facing his first competition.´

In the talk, I present the results of a corpus study (based on the German reference corpus DeReKo) on the types of NPs admissible within the German prospective-LVCs. The talk presents the first corpus study on the expression of prospective aspect in German supplementing the compositional analysis presented in previous work (e.g. Fleischhauer & Gamerschlag 2019).

References: Comrie, B. 1976. *Aspect*. Cambridge: Cambridge University Press. Fleischhauer, J. and T. Gamerschlag. 2019. "Deriving the meaning of light verb constructions - a frame account of German stehen 'stand'." In C. Juchem-Grundmann, M. Pleyer and M. Pleyer, eds. *Yearbook of the German Cognitive Linguistics Association*, Vol. 7. Berlin/Boston: Mouton de Gruyter. 137-156. Kuteva, T. 2001. *Auxiliation. An Inquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press.

Go for ingressivity

Katharina Paul
University of Göttingen
Katharina.paul@uni-goettingen.de

Unlike many other languages, German does not have a grammaticalized morphological aspect system. Nonetheless, several strategies and periphrastic constructions (e.g. *am-* progressive) have recently developed to encode aspectual readings.

This presentation deals with the *gehen*+infinitive construction, which has hardly been discussed in terms of aspectuality yet. As illustrated in (1), this construction consists of an inflected form of *gehen* ('go') combined with the infinitive of another main verb:

(1) Max   geht   schlafen.
     Max   goes   sleep
     'Max goes to sleep.'

Based on an empirical investigation, Paul et al. (forthcoming) argue that this construction undergoes a currently observable grammaticalization process. Its output seems to encode aspectuality, more precisely ingressivity.

The aim of this talk is to show that the *gehen*+infinitive construction already underlies restrictions due to its obligatoriness in ingressive contexts and cannot be substituted by an aspectually underspecified or different expression. To test this, I employed a Likert-style questionnaire and tasked 24 participants to gauge the acceptability of items varying with regard to their aspectual interpretation (*gehen*+infinitive, *am*-progressive, underspecified) in ingressive contexts. In order to ensure a balanced distribution of the treatments, a Latin square design with three different lists was employed so that each of the 24 ingressive contexts were combined with only one of the three aspectual interpretations per list.

The results of the ANOVA and a priori t-tests show statistically significant preferences for the *gehen*+infinitive construction in ingressive contexts. In my analysis, I will discuss these results 1) in terms of grammaticalization parameters (cf. Lehmann 2015; *paradigmaticity*), and 2) in a broader context of emerging strategies of encoding aspectuality in Modern German.

References: Lehmann, C. 2015. *Thoughts on grammaticalization*. Berlin: Language Science Press. Paul, K. et al. (forthcoming). "*gehen* as a new auxiliary in German." In N. Catasso, M. Coniglio and C. De Bastiani, eds. *Language Change at the Interfaces. Intrasentential and intersentential Phenomena.*

Aspectual meanings of the present participle in Middle Low German

Sarah Ihden
University of Hamburg
sarah.ihden@uni-hamburg.de

In Middle Low German the combination of a finite auxiliary verb and a verb in the form of the present participle can be used for different aspectual meanings. According to Lasch (1974: 222, § 412, note 3) it can express an inchoative action as in *alse de sondach tôkomende was*, a durative action as in *se wêren sîner dar bêdende* or the simultaneity of two actions as in *He quam slîkende*. Sarauw (1924: 226–227) and Lübben (1882: 92–93) distinguish between the different types of auxiliary verbs used in this construction: While '*wēsen* + pres. part.' marks durative actions, '*wērden* + pres. part.' is used for inchoative actions.

Furthermore, Lübben (1882: 92–93) states the special function of a past tense form of wērden combined with the present participle, namely expressing the beginning as well as the duration of an action as in *he wart wenende* ('he began to cry and went on crying').

Besides the rather short descriptions of the phenomenon given by Lasch (1974), Sarauw (1924) and Lübben (1882) there is no further information on its development within the Middle Low German language period or on its use in different times, language areas or texts. The data of the recently published digital *Reference Corpus Middle Low German / Low Rhenish (1200–1650)* enables analyses on the construction '*wēsen/wērden* + pres. part.', its aspectual meanings and its use depending on various external factors. Within the presentation some of the first results on these issues will be shown.

References: Lasch, A. 1914 / 1974. *Mittelniederdeutsche Grammatik*. Tübingen (Sammlung kurzer Grammatiken germanischer Dialekte. A. Hauptreihe, 9). Lübben, A. 1882. *Mittelniederdeutsche Grammatik*. Nebst Chrestomathie und Glossar. Leipzig. Sarauw, C.1924. *Niederdeutsche Forschungen*. Vol. 2: Die Flexionen der mittelniederdeutschen Sprache. Kopenhagen (Historisk-filologiske Meddelelser, 10.1).

Temporal adverbs as aspecutality markers? On the grammaticalization of *als* and *viel* in German substandard varieties

Sophie Ellsäßer
University of Mainz/University of Münster
sophie.ellsaesser@uni-muenster.de

Research on German (and its dialects) has concentrated primarily on the verbal expression of aspectuality so far (Kuhmichel & Flick 2013, Weber 2017, Fischer 2018). Temporal adverbs are often classified as temporal rather than aspectual phenomena in German (Ebert 1996, Kuhmichel & Flick 2013). The corresponding works, however, primarily refer to the standard German temporal adverbs *gerade* 'just' and *jetzt* 'now'. Though, especially German substandard varieties have developed further temporal adverbs, which have an iterative or habitual meaning and thus could be *classified* as imperfective markers. Examples are *viel* and *als*, which can be traced back to the mass pronouns *viel* 'much' and *alles* 'everything' and mark temporal quantification in certain recent substandard varieties (see e.g. Grimm et al. 2008 and the following examples).

(1)   *Wir gehen viel in den Wald.*
        'We often go into the forest.'
(2)   *Wir gehen als in den Wald*.
        'We sometimes/often go into the forest.'

Though, the diachronic development, the diatopic distribution and the precise functional spectrum of these temporal quantification adverbs are still relatively unexplored. With reference to an initial analysis of different data types (historical dictionaries, dialectal corpus data from Zwirner-corpus), the talk will address the following questions:

- Can a grammaticalization path be reconstructed from the data?
- Are levels of this process reflected in different functions of the adverbs in geographical space?
- To what extent can this be defined as an extension of aspectuality?

References: Ebert, K. 1996. "Progressive aspect in German and Dutch." *Interdisciplinary journal for Germanic linguistics and semiotic analysis* 1(1): 41-62. Fischer, H. 2018. *Präteritumschwund im Deutschen. Dokumentation und Erklärung eines Verdrängungsprozesses*. Berlin & Boston: De Gruyter. Grimm, J.; W. Grimm und H.-W. Bartz. [6]2008. *Deutsches Wörterbuch*. Elektronische Ausgabe der Erstbearbeitung. Frankfurt am Main: Zweitausendeins. Kuhmichel, K. und J. Flick. 2013. "Der am-Progressiv in Dialekt und Standardsprache." In P. Vogel, ed. *Sprachwandel im Neuhochdeutschen*. Berlin: De Gruyter. 52-76. Weber, T. 2017. *Die TUN-Periphrase im Niederdeutschen. Funktionale und Formale Aspekte*. Tübingen: Stauffenburg.

"Wi wir am leben in alle plantation": The aspect system in Unserdeutsch (Rabaul Creole German)

Lena Schmidtkunz
Friedrich-Alexander-University Erlangen-Nuremberg
lena.schmidtkunz@fau.de

The aim of this paper is to describe the central verbal category of aspect in the German- based Creole language Unserdeutsch. The verbal categories in Unserdeutsch are—like in other languages—closely intertwined, so that it is hardly possible to consider them as isolated entities. The focus of the article is on the aspect category, which will be described and interpreted in relation to other verbal categories and in the context of Creolistic universal research.

Based on language data from semi-guided sociolinguistic interviews with basilectal speakers, it is first discussed which aspects Unserdeutsch possesses and to what extent they are grammaticalised or obligatory. It will be shown that Unserdeutsch has both a progressive and a habitual aspect, the two being realised through different verbal periphrases. In a further step, the grammatical means of aspect marking are shown, with two constructions in focus: (1) the *am*-Progressiv [copula + am + verb], which is similar to the German form but has a significantly higher grammaticalization level in Unserdeutsch and can express both progressive and habitual meaning, and (2) the constructions with *wid* [wid +verb], which is similar in use and function to the English *would*. It will be shown that wid is of particular interest insofar as it represents an integration of all central grammatical categories of the verb-constructions with *wid* can express (a) future tense, (b) habitual aspect and (c) irrealis mood:

> a. wi wid ni   charg-im du ein   ding   (…)
>   1PL FUT NEG   charge-TR 2SG ART.INDF thing
>   'We won't charge you anything.'
>
> b. du wid   afsteh-n am morgen vielleich so   sechs finf uhr
>   2SG HAB.PST get.up-V at morning maybe   around six   five o.clock
>   'We would get up at five or six o'clock in the morning.'
>
> c. du wid geht wo
>   2SG IRR go   whereto
>   'Where should we have gone to?'

From a comparative perspective, the findings show on the one hand that Unserdeutsch has a relatively elaborate aspect system compared to its lexifier language (German). On the other hand, the data also suggest that, in the context of creole languages (see Michaelis et al. 2013, Holm/Patrick 2007), the aspect system of Unserdeutsch belongs to the typological mainstream and features structural characteristics that are typical for these languages.

References: Holm, John and Patrick, Peter L.,eds. 2007. *Comparative Creole Syntax. Parallel Outlines of 18 Creole Grammars*. London: Battlebridge. Michaelis, Susanne Maria, Maurer, Philippe, Haspelmath, Martin and Huber, Magnus, eds. 2013. *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (available online at http://apics-online.info/, last accessed on 13.11.2020).

Non-encoding aspectuality in Old High German, or: Why are we failing?

Fabian Fleißner
Austrian Academy of Sciences
fabian.fleissner@unvie.ac.at

General linguistics has developed a variety of methods for the qualitative and quantitative study of aspectuality, which have been well proven in different individual languages. These include the disclosure of correlative patterns of perfectivity and anterior tense, deontic modality and narrative on the one hand, imperfectivity and temporal simultaneity, epistemic modality and non-narrativity on the other (cf. Abraham 1991, Hopper 1979, Smith 1997). So far, however, corresponding patterns could only be traced inadequately for various Old Germanic languages (cf. Heindl 2017). This explains the ongoing dispute about the general existence of a binary aspectual system being constructed via the respective linguistic successor to the **ga*-prefix. Nevertheless, with few exceptions, the existence of a verbal aspect category in Old High German is widely accepted in the German scientific literature of recent decades since Leiss (1992) and regarded as axiomatic within the field of German Studies and academic teaching as well. The divergent behaviour of one or the other morphological form that cannot be integrated into this scheme is often blamed on a supposed insufficiency of the Germanic system in the phase of its decline. I will use data from the ohg. *Evangelienbuch* by Otfrid of Weissenburg to show that previous approaches have largely failed to support this assumption. In addition, I will offer an alternative model that explains both a certain affinity of the prefix ohg. *gi-* to some contexts of perfectivity and the divergent behaviour in this respect. The hypothesis is that the functions of the element can be found in the explicit marking of an effect of verbal action on a particular actant, cf. ohg. *sehan/gisehan* ‚see‘ in (1) and (2):

(1) *Ságetun thaz sie gáhun stérron einan sáhun* (O, I, 17, 19)
    ‚They said that they recently saw a star‘ [− change of mental/physical state]

(2) *Sie blídtun sih es gáhun, sár sie nan gisáhun* (O, I, 17, 55)
    ‚They rejoiced immediately when they saw it‘ [+ change of mental/physical state]

All other readings of different grammatical categories such as aspect can derive from these functions.

References: Abraham, Werner. 1991. "Aktionsartensemantik und Auxiliarisierung im Deutschen." In Elisabeth Feldbusch, Reiner Pogarell, Cornelia Weiss, eds. *Neue Fragen der Linguistik.* Akten des 25. Linguistischen Kolloquiums, Paderborn 1990, Tübingen: Max Niemeyer. 125-133. Heindl, Olga. 2017. *Aspekt und Genitivobjekt. Eine kontrastiv-typologische Untersuchung zweier Phänomene der historischen germanischen Syntax*. Tübingen: Stauffenburg. Hopper, Paul J. 1979. "Aspect and foregrounding in discourse." In Givon, Talmy,eds. *Syntax and Semantics 12: Discourse and Syntax*. New York: Academic Press. 213-241. Leiss, Elisabeth. 1992. *Die Verbalkategorien des Deutschen. Ein Beitrag zur Theorie der sprachlichen Kategorisierung [Studia Linguistica Germanica 31]*. Berlin / New York: Walter de Gruyter. Smith, Carlota. 1997. *The parameter oft aspect [Studies in Linguistics and Philosophy 43]*. Dordrecht / Boston / London: Kluwer.

Perfect constructions in English and German varieties: Typologies and diachronic implications

Stephanie Hackert, Robert Mailhammer, Elena Smirnova
Ludwig-Maximilians-Universität München, University of Western Sydney, Université de Neuchâtel
Stephanie.Hackert@anglistik.uni-muenchen.de, R.Mailhammer@westernsydney.edu.au, elena.smirnova@unine.ch

This paper examines the expression of perfect meanings in varieties of two Germanic languages, English and German. Whereas the standard varieties of these two languages both possess grammaticalized perfect categories employing HAVE, this is not necessarily the case in non-standard varieties. A number of high-contact varieties of English and especially pidgins and creoles employ a range of forms to express the various meanings commonly associated with the category (cf. Dahl 1985: 132).

We will present data from a number of varieties of English and German and compare their expression of resultative, experiential, "hot news," and persistent perfect situations. We also look at the perfective, as the perfect's "anti-prototype" (Dahl 2014: 273). Our material includes the data on pidgins and creoles first presented in Hackert (2019), the data from a range of Australian Englishes (e.g. Aboriginal English, Mailhammer forthc.) and from varieties of German (Walser German, Amish German, Pennsylvanian German, Barossa German, and Russian-German Dialects). Specifically, we aim at answering the following research questions:

(1) What is the range of forms covering the semantic space of the perfect in varieties of English and German?

(2) Which varieties possess a grammaticalized perfect?

(3) Do marking patterns distinguish groups of varieties? Do these linguistically determined groups have geographical and/or sociohistorical correlates?

(4) How do the typological findings align with pathways of grammaticalisation (Smirnova et al. 2019) on micro and macro levels?

References: Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Blackwell. Dahl, Östen. 2014. "The perfect map: Investigating the crosslinguistic distribution of TAME categories in a parallel corpus." In Benedikt Szmrecsanyi and Bernhard Wälchli, eds. *Aggregating dialectology, typology, and register analysis. Linguistic variation in text and speech*. Berlin: de Gruyter. 268-289. Hackert, Stephanie. 2019. "The perfect in English-lexifier pidgins and creoles: A comparative study." *Journal of Pidgin and Creole Languages* 34: 195-242. Mailhammer, Robert. forthc. *English on Croker Island: the synchronic and diachronic dynamics of contact and variation*. Berlin/New York: De Gruyter Mouton. Smirnova, Elena, Robert Mailhammer and Susanne Flach. 2019. "The role of atypical constellations in the grammaticalization of German and English passives." *Diachronica* 36(3): 384-416.

Auxiliary variation in telic perfect constructions of Low German speakers

Kathrin Weber
University of Jena
weber.kathrin@uni-jena.de

Lexical aspect and telicity are major determinants of auxiliary selection in German (Sorace 2000; Keller and Sorace 2003; Gillmann 2011). Telic verbs are predominantly constructed with the auxiliary *sein* 'be', atelic verbs with the auxiliary *haben* 'have'. However, there are exceptions of the telicity principle in the German standard system as verbs like *anfangen* 'to begin' or *abnehmen* 'to decrease' are constructed with the auxiliary *haben* 'have'. From a regional perspective, telicity and lexical aspect have a stronger influence on auxiliary constructions in the Westphalian Low German area, especially with the verb *angefangen* 'to begin' (Weber 2020). Here, sentences like *ich bin angefangen* 'I be$_{AUX}$ begun$_{PP}$' are highly conventionalized. The main objective of the presentation is to apply a regional approach to the influence of telicity on auxiliary variation in Low German. Besides lexical aspect, the continuum between tense and aspect plays a crucial role in explaining auxiliary variation in the perfect tense with *angefangen* 'to begin'. Theoretically, the presentation follows a usage- based construction grammar approach, where constructions are understood as form- meaning-pairs. Methodologically, the talk pursues a mixed-methods approach by analyzing both authentic spoken interactions of Westphalian speakers (both dialect and regiolect data) and written data from the regional newspaper *Neue Westfälische*. The presentation shows that the auxiliaries *haben* 'have' and *sein* 'be' serve as grammatical markers in the tense- aspect-interface. Auxiliary constructions with *haben* 'have' and *angefangen* 'to begin' are mainly associated with tense meaning, while auxiliary constructions with *sein* 'be' mainly present current relevance meaning.

References: Gillmann, Melitta. 2011. "Die Grammatikalisierung des sein-Perfekts. Eine korpuslinguistische Untersuchung zur Hilfsverbselektion der Bewegungsverben im Deutschen." *Beiträge zur Geschichte der deutschen Sprache und Literatur* 133 (2): 203-234. Keller, Frank and Sorace, Antonella. 2003. "Gradient Auxiliary Selection and Impersonal Passivization in German: An Experimental Investigation." *Journal of Linguistics* 39 (1): 57-108. Sorace, Antonella. 2000. "Gradients in Auxiliary Selection with Intransitive Verbs." *Language* 76 (4): 859-890. Weber, Kathrin. 2020. *Regionale Auxiliarvariation: Interaktion, Schrift, Kognition*. Berlin, New York: De Gruyter.

The influence of grammatical and non-linguistic factors on motion event descriptions: A cross-linguistic study

Katharina Zaychenko
University of Kassel
zaychenko@uni-kassel.de

Cross-linguistic differences in the conceptualization of motion events have often been hypothesized to depend on the absence or presence of grammatical aspect in different languages. Empirical studies in this field focus in particular on presenting videoclips showing entities moving towards different endpoints in combination with the task to verbalize the event. Cross-linguistic differences become apparent, for instance, through the varying number of verbalized endpoints. An interesting observation concerns the result that speakers of languages with a grammaticized concept of temporality rather focus on the process of an event, whereas speakers of non-aspect languages have been shown to favor the motion endpoint (Stutterheim et al. 2012; Mertins 2018).

Recent investigations, however, speculate that cognitive factors such as the visual prominence (salience) of certain motion event components might influence motion event conceptualization, too (Bepperling & Härtl 2013; Georgakopoulos et al. 2019). To examine the interplay between non-linguistic and linguistic factors experimentally, an online survey was conducted which implements the non-linguistic factor 'endpoint salience' as an influence on motion event descriptions. Native speakers of German and English participated in two verbalization tasks and one non-linguistic memorization task. The results point at a main effect for 'endpoint salience' such that salient endpoints were verbalized more often than regular endpoints. While German speakers only show a tendency to verbalize more salient endpoints than regular ones, English speakers verbalized significantly more endpoints in the salient endpoint condition. Simultaneously, native speakers of English use significantly fewer progressive forms when they mention the motion endpoint in their descriptions. These results indicate that endpoint salience has a higher influence on speakers of English than German since German speakers focus on the endpoint in any case even if they do not verbalize it.

The results will be discussed in the light of an interdependency between linguistic and non- linguistic factors in motion event conceptualization.

References: Bepperling, S. and H. Härtl. 2013. "Ereigniskonzeptualisierung im Zweitspracherwerb - Thinking for Speaking im Vergleich von Muttersprachlern und Lernern." *Zeitschrift für Semiotik* 35 (1-2): 159-191. Georgakopoulos, T., H. Härtl and A. Sioupi. 2019. "Goal realization: An empirically based comparison between English, German, and Greek." *Languages in Contrast* 19.2: 280- 309. Mertins, B. 2018. *Sprache und Kognition: Ereigniskonzeptualisierung im Deutschen und Tschechischen* (Vol. 8). Berlin: Walter de Gruyter. Stutterheim, C. v., M. Andermann, M. Carroll, M. Flecken and B. Schmiedtova. 2012. "How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analysis, eye tracking data, and memory performance." *Linguistics* 50(4): 833-867.

Arbeitsgruppe 6

Empirical approaches to canonical and non-canonical uses of negation

Katharina Schaebbicke & Heiko Seeliger

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Processing accounts for negation in linguistic and non-linguistic domains

Carolin Dudschig
Eberhard-Karls Universität Tübingen
carolin.dudschig@uni-tuebingen.de

Negation is an omnipresent phenomenon in daily language. Nevertheless, producing and comprehending negation is often associated with increased processing costs. A core focus within psycholinguistic research on the topic of negation has been to understand when and why difficulties in comprehending negation arise and the circumstances that ease the processing of negation. In the present talk, I will address issues regarding both the comprehension and production of negation. First, findings from several experiments that use negation as an instructional command will be presented. Phrases such as "*Don't cross the street*" and "*Don't eat the chocolate*" are often inefficient to stop the negated behaviour, which is attributed to ironic effects of negation (Adriaanse et al. 2011). Specifically, the current studies investigate what produces ironic effects of negation, and how such effects can be integrated within theoretical accounts of negation processing (Dudschig & Kaup, 2018, 2020a). In this context, findings comparing processing difficulties by a negation instruction in the linguistic and pictorial domain are contrasted. Such comparisons between negation in different input domains are specifically interesting, given that very young children can already reject states in their environment via non-linguistic expressions (facial expression, gestures, vocal expressions, etc.), therefore negation in non-linguistic domain might ease processing. The second set of experiments investigate the time course of negation processing. Specifically, whether having longer to process the negation operator would be reflected in the N400 measures of negation integration, an ERP component that is often negation-blind (Dudschig et al. 2018, 2019; Palaz et al. 2020; cf. Nieuwland and Kuperberg, 2008). The final set of experiments investigate the circumstances under which we produce negation over an alternative affirmative possibility. Previous studies showed that negation is often particularly easy to process when used in a pragmatically licensed way. Here, we investigate whether other factors – such as cognitive effort – influence the use of negation in simple production tasks.

References: Adriaanse, Marieke, van Oosten, Johanne, de Ridder, Denise, de Wit, John, & Evers, Catharine. 2011. "Planning what not to eat: Ironic effects of implementation intentions negating unhealthy habits." *Personality and Social Psychology Bulletin 37*: 69-81. Dudschig, Carolin, Mackenzie, Ian Grant, Maienborn, Claudia, Kaup, Barbara, & Leuthold, Hartmut. 2019. "Negation and the N400: investigating temporal aspects of negation integration using semantic and world-knowledge violations." *Language, Cognition and Neuroscience 34*: 309- 319. Dudschig, Carolin, Mackenzie, Ian Grant, Leuthold, Hartmut, and Kaup, Barbara. 2018. "Environmental sound priming: Does negation modify N400 cross-modal priming effects?." *Psychonomic Bulletin & Review 25*: 1441-1448. Dudschig, Carolin and Kaup, Barbara. 2018. "How does "not left" become "right"? Electrophysiological evidence for a dynamic conflict-bound negation processing account." *Journal of Experimental Psychology: Human Perception and Performance, 44*: 716-728. Dudschig, Carolin and Kaup, Barbara. 2020a. "Negation as conflict: Conflict adaptation following negating vertical spatial words." *Brain and Language 210*: 104842. Dudschig, Carolin and Kaup, Barbara. 2020b. "Can We Prepare to Negate? Negation as a Reversal Operator." *Journal of Cognition 3*: 32.Fischler, Ira, Bloom, Paul, Childers, Donald, Roucos, Salim and Perry Jr, Nathan. 1983. "Brain potentials related to stages of sentence verification." *Psychophysiology 20*: 400-409.Nieuwland, Mante and Kuperberg, Gina. 2008. „When the truth is not too hard to handle: An event- rela ted potential study on the pragmatics of negation." *Psychological Science 19*: 1213-1218. Palaz, Bilge, Rhodes, Ryan and Hestvik, Arild. 2020. „Informative use of "not" is N400-blind." *Psychophysiology 57*: e13676. Wegner, Daniel. 2009. "How to think, say, or do precisely the worst thing for any occasion." *Science 325*: 48-50.

Validating the Performativity Hypothesis to Neg-Raising using corpus data: Evidence from Polish

Beata Trawiński
Leibniz-Institut für Deutsche Sprache, Mannheim
trawinski@ids-mannheim.de

It is known that with certain verbs (such as *think* or *believe*), a negation in the matrix clause can be understood as negating the embedded proposition. Such verbs are referred to as negation raising predicates (NR predicates) and are attested in many languages (Fillmore 1963, Bartsch 1973, Horn 1978, Gajewski 2007). Among many different approaches to Neg- Raising, the idea put forward by Prince (1976) is particularly interesting, since it explains a range of data which is difficult to account for within other, especially syntactic approaches.

Prince (1976) observed that NR predicates exhibit a marked preference to the first person present non-progressive form (in English) and in this regard, they behave in a way similar to performatives. In this paper, the performativity hypothesis is tested using corpus data based on evidence from Polish. The distribution of the negated NR predicate *sądzić* 'think' (Wierzbicka 1969, Modrzejewska 1981) was examined in the balanced version of the National Corpus of Polish (Przepiórkowski et. al. 2012). The negated verb *sądzić* exhibits two selectional patterns: It can select indicative complement clauses introduced by the complementizer *że* 'that' and subjunctive complement clauses introduced by the complementizer *żeby* or its variants. In contrast to sentences with *że*, where affirmative structures are possible, sentences with *żeby* obligatorily contain a negation in the matrix clause. Affirmative structures are excluded with *żeby*. However, the matrix negation in sentences with *żeby* can still be understood as negating the embedded proposition. The presence of a semantic negation in the embedded *żeby*-clause is evidenced, among others, by (strong) Negative Polarity Items (NPIs), which can be licensed within these clauses. Similar patterns can be observed in French. According to Prince (1976), French sentences with NR reading contain embedded clauses in the subjunctive rather than the indicative mood. This seems also to hold for Polish.

Based on the above observations and assuming the performativity hypothesis, the preference for the first person present form of NR predicates such as *sądzić* is expected to be stronger with *żeby*-clauses that with *że*-clauses. This is indeed evidenced by the corpus data. In particular, the results of the present corpus study show that the tense and person form of the negated verb *sądzić* correlate with its preferences for *że*- versus *żeby*- complements. The first person present form is associated with *żeby*-clauses significantly stronger that with *że*-clauses. These findings clearly support the performativity hypothesis. Pilot studies of other NP predicates in Polish, such as *wierzyć* 'believe' and *uważać* 'be of the opinion' yielded very similar results. Theoretical modeling of performativity effects in Neg- Raising still remains a challenge.

References: Bartsch, Renate (1973). 'Negative transportation' gibt es nicht. Linguistische *Berichte* 27(7). Fillmore, Charles J. (1963). The position of embedding transformations in a grammar. WORD 19(2), pp. 208–231. Gajewski, Jon R. (2007). Neg-raising and polarity. *Linguistics and Philosophy* 30(3), pp. 289– 328. Horn, Laurence Robert (1978). Remarks on neg-raising. *Syntax and Semantics* 9, pp. 129– 220. Modrzejewska, Ewa (1981). Neg-raising Predicates in English and Polish. *Papers and Studies in Contrastive Linguistics Poznan* 13, pp. 41–52. Prince, E. (1976). The Syntax and Semantics of NEG Raising, with Evidence from French. *Language* 52, pp. 404–426. Przepiórkowski, Adam et al. (2012). *Narodowy Korpus Języka Polskiego*. Warszawa: PWN. Wierzbicka, Anna (1969). *Dociekania semantyczne*. Wrocław: Ossolineum.

Affixal negation is not always negative: Evidence from Catalan and Spanish

Elisabeth Gibert-Sotelo
Universitat Rovira i Virgili
elisabeth.gibert@urv.cat

Outline. This study aims at offering a contrastive analysis of the two productive types of affixal negation available in Catalan and Spanish: *in-* prefixation and *des-* prefixation. The core proposal is that *in-* and *des-* are not prefixes of the same type: *in-* is a negative marker, but *des-* is a P (i.e., prepositional) element. In particular, it is argued that *in-* encodes negation and involves quantification as well as adjectival categorization, whereas *des-* lexicalizes a Source Path and accordingly expresses (physical or abstract) separation from an origin.

In- vs. des-. Both *in-* and *des-* may encode negative meaning (cf. *in-feliç* (Cat)/*in-feliz* (Sp) 'unhappy' and *des-honest* (Cat)/*des-honesto* (Sp) 'dishonest'), but their behaviour is different (Gibert-Sotelo 2017). First, *in-* licenses negative polarity items (1), whereas *des-* does not (2):

(1)   a. *Soc *(in)capaç de fer <u>res</u>*. [Google]   (Cat)
          'I am *(un)able to do <u>anything</u>'.
      b. *una realidad poética, *(in)traducible a <u>ninguna</u> clase de prosa*. [Google] (Sp)
          'a poetic reality, *(un)translatable to <u>any</u> kind of prose'.

(2)   a. **A la Marta li **(des)**agrada <u>res</u>*.   (Cat)
          '*Marta (dis)likes <u>anything</u>'.
      b. **María es **(des)**leal con <u>ninguno</u> de sus amigos*.   (Sp)
          '*María is (dis)loyal to <u>any</u> of her friends'.

Second, *in-* may encode both contrary (3a) and contradictory negation (3b). *Des-* only can encode contrary negation (4), since it always allows for a middle term in which both the prefixed and the non-prefixed predicate are simultaneously false (on the Aristotelian distinction between contrariety and contradiction, see Horn 1989):

(3)   a. *Él no es feliz, pero tampoco infeliz.* [Google]   (Sp)
          'He is not happy, but not unhappy either'.
      b. #*El sistema no és modificable, però tampoc immodificable*.   (Cat)
          '#The system is not modifiable, but it is not unmodifiable either'.

(4)   *Ni m'agrada ni em desagrada la ciència-ficció*. [Google]   (Cat)
    'I neither like nor dislike science fiction".

Finally, *in-* is only compatible with gradable adjectives (non-gradable adjectives systematically reject it: *in-solar '*unsolar'; cf. Scalise 1984), but *des-* is attested in verbs, nouns and adjectives (cf. Varela & Martín García 1999, Montero Curiel 1999, RAE & ASALE 2009, IEC 2016):

(5)   a. Nouns: *in-honor (Cat/Sp) vs. *des-honor* (Cat/Sp) 'dishonour'.
      b. Verbs: *in-conocer (Sp) vs. *des-conocer* (Sp) 'not know'.
      c. Gradable adjectives: *in-atent* (Cat) 'inattentive' vs. *des-atent* (Cat) 'rude'.

Analysis. Following a nanosyntactic approach to grammar (Starke 2009), I assume that morphemes spell out phrasal nodes encompassing multiple features (Phrasal Spell-Out). The proposal of the paper is that *in-* is a negative marker (NegP) that involves degree quantification (QP) (cf. De Clercq 2017) and adjectival categorization (AP) (cf. Newell 2008): as a negative marker, it licenses negative polarity items (cf. (1)) and allows for both contrary and contradictory readings (cf. (3)); as a degree quantifier, it is only compatible with gradable adjectives; and as an adjectival categorizer, it always gives rise to adjectives (cf. Cat. *color* 'colour' [noun] and *in- color* 'colourless' [adjective]) and is not attested in nouns and verbs (cf. (5)). By contrast, *des-* is not a negative marker: it does not license negative polarity items (cf. (2)) and only allows for contrary readings (cf. (4)). Rather, *des-* is a P

element that encodes separation, and accordingly I propose that it is the Spell-Out of a Source Path that, when embedded in dynamic verbs, gives rise to dynamic interpretations such as physical separation (cf. Sp. *desmoldar* 'remove from mould'), deprivation (cf. Cat. *desossar* 'bone'), or reversion (cf. Sp. *deshacer* 'undo'); but when embedded in non-dynamic predicates, gives rise to contrary opposition (cf. Cat. *deslleial* 'disloyal', Cat./Sp. *desagradar* 'dislike') (on the internal structure of Source Paths, see Pantcheva 2011):

(6)   a.   Structure of *in-*   b.   Structure of *des-*
             $_{\text{NegP}}$ Neg [$_{\text{QP}}$ Q [$_{\text{AP}}$ A]]]   [$_{\text{SourceP}}$ Source [$_{\text{PathP}}$ Path [$_{\text{PlaceP}}$ Place]]]

Negation scoping and focus in Mandarin biased questions: A VERUM account

Boer Fu
Massachusetts Institute of Technology
boerfu@mit.edu

The role of negation in biased yes/no questions has been under much discussion in the literature. I present evidence from the 4-way readings of Mandarin negative yes/no questions, which support Romero & Han's (2004) VERUM operator account. In English, VERUM is a silent operator that scopes relative to negation to derive two different biases in biased questions. I show that in Mandarin, VERUM can be pronounced at PF, and displays overt scoping with negation that confirms Romero & Han's prediction of the two readings. VERUM can also account for an unexpected assertion reading.

The 4 Readings A negative yes/no question constructed with the *ma*-particle have 4 readings, disambiguated by focus placement and boundary tone. The 2 readings with H% (a & b) are real questions. In (a), the content word 'eat' has prominence. It is a high NEG with epistemic bias for p (that he eats beef). When focus is placed on NEG in (b), it is a low NEG question in which the epistemic bias remains p, but there is an added evidential bias for $\neg p$. The 2 L% readings (c & d) are assertion-like, used when the speaker wants to assert something that she considers to be obvious.

(1)	ta	bu	chi	niurou	ma	he	NEG	eat	beef	ma

(a)	'Doesn't he eat beef?'	*Focus on 'eat',H%*	**High NEG question**
(b)	'Does he not eat beef?'	*Focus on NEG, H%*	**Low NEG question**
(c)	'He eats beef.'	*Focus on 'eat', L%*	**Ghost NEG assertion**
(d)	'He doesn't eat beef.'	*Focus on NEG, L%*	**Low NEG assertion**

VERUM and NEG The difference between the 2 H% readings are similar to the outside-NEG vs inside-NEG reading contrast observed by Ladd (1981). Ladd's intuition is fleshed out by Romero & Han (2004) with a silent VERUM operator, which creates an unbalanced partition that asks if it is for sure that the prejacent should be added to the CG. When NEG scopes above VERUM, the prejacent is p, where NEG contributes nothing to the prejacent itself. And when NEG scopes below VERUM, the prejacent is $\neg p$.

***shi* is VERUM** I argue that *shi*, homophonous to the copula, is the pronounced VERUM in Mandarin (per Hole 2012). When *shi* is added to the PF of a negative yes/no question, it disambiguates between the 2 H% readings by its scoping relative to NEG. In (2a), NEG scopes above *shi*, the sentence is a high NEG question with a bias for p. And the $\neg p$ bias in (2b) is derived by having *shi* scope above NEG.

(2a)	he	NEG (shi)	eat	beef	ma	H%

'Doesn't he eat beef?'   **High NEG question**: Q[$\neg$VERUM p]

(2b)	he	(shi) **NEG**	eat	beef	ma	H%

'Does he not eat beef?'   **Low NEG question**: Q[VERUM $\neg p$]

Ghost NEG assertion (1c) is not a straightforward assertion, for the presence of NEG indicates it should assert $\neg p$. Instead, it asserts p. And deleting the high NEG does not change its meaning. I argue that the ghost NEG assertion is actually a self-answering question. It is a composed of a question component and an answer component. In (3), an unbalanced partition is created by the question component. The '$\neg$FOR-SURE-CG$_x$ p', found in the partition, is at odds with the certainty conveyed by L%, which cues the presence of the elided answer 'FOR-SURE-CG$_x$ p'. VERUM is crucial in deriving a positive polarity for the ghost NEG assertion. Without it, the partition created by the

question is $\{p, \neg p\}$. The elided assertion is ambiguous between p and $\neg p$. Nothing can get asserted.

(3)   Question: Q[$\neg$VERUM p] = $\{\neg$FOR-SURE-CG$_x\, p$, FOR-SURE-CG$_x\, p\}$

Answer: [VERUM p]$_{elided}$ = FOR-SURE-CG$_x\, p$ **Ghost NEG assertion**

<u>Significance</u> The presence of VERUM operator is contentious due to its silent nature. The data from Mandarin negative yes/no question show that VERUM can be pronounced as *shi* and shows the scoping predicted by Romero & Han. By enlisting VERUM, the ghost NEG assertion can also be accounted for

Expletive negation and negative polarity: The view from Québec French

Aurore Gonzalez, Justin Royer
Harvard University, McGill University
auroregonzalez@g.harvard.edu, justin.royer@mail.mcgill.ca

Across languages, we find instances of sentential negation that do not always make a straightforward contribution to meaning, so-called "expletive negation" (ExN). In this paper, we investigate such an instance of negation in Québec French. We argue that ExN-*pas* is a dependent negative polarity item (NPI) that appears nowadays only in specific environments. **Data.** As in other variants of French, Québec French propositions are negated with *pas* (1). But a "nonnegative" use of *pas* is reported in sentences like (2) and (3) (Kemp, 1982, a.o.).

(1) J'aime **pas** les rats.
I.like NEG the rats
'I don't like rats.'

(2) J'ai fait tout ce que je pouvais **pas** faire.
I.have did all that C I could ExN do
'I did all I could.'

(3) C'est le pire livre {que tu peux **pas** lire} / {qu'il y a **pas**} / {#qu'il aime **pas**}.
it.is the worst book that you can ExN read / that.∃ ExN / that.he loves ExN
'It's the worst gift you could ever give me / there is / that he likes.'

In (2) and (3), ExN-*pas* does not negate the proposition expressed by the embedded clause.

Analysis. NPIs like English *any* have been argued to be semantically complex. On alternative-based accounts of negative polarity (Krikfa 1995; Lahiri 1998; Chierchia 2013), NPIs are existential items that obligatorily activate a set of domain alternatives (ALT) consisting of subsets of the relevant quantificational domain. We propose that ExN-*pas* is just one of the two ingredients in the composition of an NPI. Specifically, it requires the predicative existential expression it co-occurs with to involve a set of ALT. Unlike *any*, ExN-*pas* does not also contribute existential meaning. Our proposal is illustrated in (4). Sentence (2) has the LF in (4a) and asserts (4b). ExN-*pas* signals that the (existential) ability modal *pouvoir* triggers a set of ALT (4c). These alternatives are factored into meaning through the insertion of an exhaustification operator EXH, defined in (5). Because all the ALTs are entailed by the assertion, exhaustification is vacuous and simply returns the assertion (4d).

(4) $[\![(2)]\!] =$

a. EXH [tout [1 je pouvais pas faire t_1]] 2 [j'ai fait t_2]]

b. Assertion: $EXH \, \forall x [\exists w \in W [\text{I have done } x \text{ at } w] \rightarrow \text{I have done } x \text{ at } w_0]$
where $W = ACC(w_0, w)$

c. ALT: $\{\forall x [\exists w' \in W' [\text{I have done } x \text{ at } w'] \rightarrow \text{I have done } x \text{ at } w_0], W' \subseteq W\}$

d. After exhaustification: $\forall x [\exists w \in W [\text{I have done } x \text{ at } w] \rightarrow \text{I have done } x \text{ at } w_0]$

(5) $[\![EXH]\!]^{g,w}(\phi) = \phi_w \, \wedge \forall p \in ALT(\phi) \, [p_w \rightarrow \phi \subseteq p]$

Our analysis of ExN-*pas* correctly captures many of its distributional properties. (i) **Licensing environments**: ExN-*pas* occurs in the restrictor of *tout* and superlatives because these are downward-entailing environments that license NPIs (4). (ii) **Only with *tout***: ExN-*pas* cannot occur in RCs headed by *quelque chose* 'something'. Because RCs headed by *quelque chose* are upward entailing, the ALT are not entailed by the assertion. EXH thus negates them yielding an inference that contradicts the assertion. (iii) **Necessity of existential:** ExN-*pas* does not contribute existential meaning by itself (unlike *any*). As one part of a complex NPI, it requires the presence of an existential and triggers a set of ALT which will be used by EXH. This explains why ExN-*pas* must co-occur with an existential predicate (3).

Discussion. Our work suggests that at least some instances of ExN are NPIs. This questions recent work arguing for a uniform analysis of ExN (Makri 2013, Yoon 2011, a.o.) and supports work that views ExN as a non-unitary phenomenon (Greco, 2019).

Types of Negative Concord systems

Hedde Zeijlstra
Göttingen
hzeijls@uni-goettingen.de

It is a well-known fact that the landscape of polarity-sensitive elements is much more heterogeneous than has previously been assumed. However, many scholars still take Negative Concord to be a homogenous system, i.e., irrespective of what mechanism underlies NC, scholars have mostly assumed that this mechanism applies to all NC languages in the same way.

In this talk, I address the question as to whether Is the landscape of Negative Concord Items more uniform than the landscape of (other) NPIs or whether it is more pluriform than generally assumed? I will first discuss what kind of NC systems can be attested and how this variation must be constrained. I will show that the attested landscape of Negative Concord Items is then actually subject to two types of constraints: learnability diachronic constraints.

High negation questions are always polarity focused and sometimes contain VERUM

Cory Bill, Todor Koev

Leibniz-Center General Linguistics (ZAS), University of Konstanz

bill@leibniz-zas.de, todor.koev@uni-konstanz.de

High Negation Questions (HNQs) like (1-a) necessarily convey a *positive epistemic inference* (Romero & Han, 2004). However, a less noted property of HNQs (although see Krifka (2017)'s 'incredulity contour') is the fact that this epistemic bias comes in at least two strengths. That is, the epistemic bias in (1-a) can be strengthened by applying Höhle (1992)'s *verum accent* to the sentence (i.e. by applying a pitch accent to the finite auxiliary) to generate (1-b). This contrast in strength is made more evident by the contrast in the relative felicity of the follow-up phrases, when applied to each of these HNQs.

(1) a. Isn't Oliver from Australia? That is to say, I suspected / #was certain he was. (weak bias)

    b. ISN'T Oliver from Australia? That is to say, I #suspected / was certain he was. (strong bias)

We present a novel analysis which captures both the polarity and the strength of the epistemic inferences in (1-a) and (1-b). Essentially, we analyse the inference in (1-a) as being derived through polarity focus alone, whereas (1-b) involves both polarity focus and a covert epistemic operator (i.e. VERUM), which introduces a conflicting evidence presupposition.

    Core proposal: (i) HNQs, regardless of their pitch accent properties, always focus the polarity phrase, signaled by the high position of negation. (ii) This focus is always interpreted as contrastive. (iii) A pitch accent on the auxiliary signals the presence of a covert VERUM operator, which introduces no truth-conditional content but issues in the conflicting evidence presupposition in (2).

(2) $[\![$VERUM φ$]\!]$ = $[\![$φ$]\!]$, provided that there is conflicting contextual evidence about $[\![$φ$]\!]$

VERUM in declaratives: This simple semantics for VERUM derives the basic pattern in declaratives. The prototypical use of verum accent is in conflicting evidence contexts like (3-a), where the evidence for the prejacent of VERUM comes from prior discourse. For the same reason, our account rules out VERUM in answers to neutral polar questions. This use is possible only if there has already been some dispute regarding the issue raised by the polar question, i.e. (3- b).

(3) a. A: Oliver is from Australia.   B: Oliver is not from Australia.   A: He IS from Australia.

    b. A: Is Oliver from Australia? B: He IS from Australia. ⤳ *There has been a prior dispute about whether O. is from Australia.*

Bias in HNQs: Our account captures both the polarity and the strength of the epistemic bias conveyed by HNQs. A regular HNQ, like (4), is analysed as having the Logical Form (LF) in (4-a). That is, it contains polarity focus but no VERUM operator. Adopting a standard two- dimensional semantics for focus (Rooth, 1992), the LF in (4-a) results in the focus semantic value in (4-b). Polarity focus in HNQs (and polar questions generally) is interpreted as playing a contrastive function, because a presentational function would be redundant (i.e. the antecedent would be identical to the ordinary question meaning). The free variable C must, therefore, be resolved to p as this is the only element that satisfies Rooth (1992)'s condition on contrasting phrases. That is, C must differ from the ordinary value of its scope φ (i.e. $p \neq \neg p$) and C must be a member of the focus value of its scope φ (i.e. $p \in \{p, \neg p\}$). The speaker's choice to use a question form that generates this kind of contrastive focus rather than a neutral question form (i.e. a positive polarity question) necessarily conveys the inference that the speaker at least <u>suspected</u> that the affirmative alternative was true, so it generates the inference in (4). In this way, our account captures both the <u>affirmative</u> polarity and the <u>weak</u> strength of the epistemic inference in (4).

(4)   Isn't Oliver from Australia? ⤳ *The speaker suspected that Oliver was from Australia.*

a. *Logical Form*: [$_{CP}$ Q [$_{PolP}$ not$_F$ [$_{TP}$ Oliver from Australia]]$_\phi$ ~ C]

b. *Focus value of squiggle scope*:
$[\![\phi]\!]^f$ = {λw. from(w,oliver,australia), λw.¬from(w,oliver,australia)} = {p,¬p}

c. *Ordinary question meaning*: $[\![(4a)]\!]^o$ = {p,¬p}

When a verum accent is applied to a HNQ, as in (5), our account posits an LF containing both polarity focus and a covert VERUM operator, i.e. (5-a). The VERUM operator only serves to introduce the conflicting evidence presupposition in (5-c). As with the HNQ in (4), the polarity focus in (5) is expected to generate a weak, affirmative epistemic inference. However, the speaker's choice to use a question form that generates such an inference in a context where the conflicting evidence presupposition is satisfied necessarily increases the strength of the epistemic inference, i.e. it generates the <u>strong</u> epistemic inference in (5).

(5) ISN'T Oliver from Australia? ⤳ *The speaker was certain that Oliver was from* Australia.

a. *Logical Form*: [$_{CP}$ Q [$_{PolP}$ not$_F$ [$_{FP}$ VERUM [$_{TP}$ Oliver from Australia]]]$_\phi$ ~ C]

b. *Focus value of squiggle scope*:
$[\![\phi]\!]^f$ = {λw . from(w,oliver,australia), λw.¬from(w,oliver,australia)} = {p,¬p}

c. *Ordinary question meaning*: $[\![(5a)]\!]^o$ = {p,¬p}, provided there is conflicting evidence about p

Comparison with previous accounts and extensions: Several accounts of HNQs posit question partitions that contain some extra epistemic operators (e.g. {□p,¬□p}) (Goodhue, 2019; Repp, 2012; Romero & Han, 2004). One of the main issues with such accounts is the fact that these partitions do not appear to be realised in the answering patterns of HNQs, which seem to be closer to {p,¬p}. van Rooy and Šafárová (2003) attempt to capture HNQs using notions from decision theory, however they fail to account for the obligatory nature of the in- ference. Finally, Malamud and Stephenson (2015), Krifka (2017) and AnderBois (2019) develop accounts that are more focused on explaining the effects of HNQs on the unfolding discourse than on speaker bias. We would also note that our account is the first to identify and capture the noted variation in the strength of the epistemic inferences conveyed by HNQs. Moreover, we show that our analysis can be straightforwardly extended to provide a unified account of the epistemic inferences conveyed by stressed low negation questions (*Is Oliver NOT from Australia?*), reverse-polarity rising tag questions (*Oliver is from Australia, isn't he?*), and positive polar questions with a verum accent (*IS Oliver from Australia?*) or polar *really* (*Is Oliver REALLY from Australia?*).

Hungarian *nem-e* interrogatives: Marking the source of speaker bias

Beáta Gyuris
Hungarian Research Institute for Linguistics Budapest
gyuris.beata@nytud.hu

The paper analyses the interpretation of the (noncanonical) negative polar interrogative form type referred to as *nem-e interrogative* (*nem-eInt*) in Hungarian, which is sensitive to the source of the "original speaker bias" (cf. Domaneschi et al. 2017). (1) illustrates the form type *nem- eInt*, where the negative particle *nem* and the interrogative particle *-e* are merged into one unit. The latter is followed by the VM (verb modifier) + verb complex, lacking VM-verb inversion, characteristic of canonical positive interrogatives and declaratives, but not of canonical nega- tive ones. *Nem-eInt*s have ON readings (cf. compatibility with the PPI *valamikor*), but no IN readings (cf. incompatibility with the NPI *soha*). (Cf. Büring and Gunlogson 2000, Ladd 1981, Sudo 2013 for discussion of the IN/ON contrast, and Gyuris 2017, 2018 for a review of inter- rogative form types and their "bias profiles" in Hungarian.)

(1) Nem-e elrepült Pali   Freiburgba (valamikor   / * soha)?
    not-Q   VM.flew Paul Freiburg.to   at.some.time   never
    'Hasn't Paul (at some point / ever) flewn to Freiburg?'

*Nem-eInt*s are only compatible with epistemic bias by the speaker for *p,* cf. the infelicity of (2), which requires deontic bias. They are inappropriate as pure information questions, cf. (3):

(2) [M(other) knows that C(hild) did something inappropriate. M utters:]
    # Nem-e   szégyelled   magad?
     not-Q   be.ashamed.2SG yourself
    'Aren't you ashamed?'

(3) [A says to B:]
    What is the weather like?
    # Nem-e esik   az   eső?
     not-Q   falls   the   rain
    'Isn't it raining?'

*Nem-eInt*s are infelicitous as offers, and cannot be used to ask a question the hearer is known to be able to answer, cf. (4). *Nem-eInt*s are normally used to offer an explanation, cf. (5):

(4) # Nem-e éhes   vagy?   (5)   A: I can't find Pali.
     not-Q   hungry be.2SG   B: Nem-e elrepült Freiburgba?
    'Aren't you hungry?   not-Q   VM.flew Freiburg.to
                                          'Didn't he fly to Freiburg?

The analysis is based on the following assumptions. First, *nem-eInt* forms have originated from a biclausal structure, which contains a matrix existential clause with a copula (*van* 'be.3SG'), the demonstrative pronoun *az* ('that') in the preverbal (exhaustive) focus position, which marks the position of the subordinate clause within the matrix clause (cf. É. Kiss 2002), and where the constituent following *-e* originated as a subordinate clause. (6) shows the structure of (1):

(5) [NegP Nem [FocP ~~az~~   ~~van~~-e$_i$] [IP t$_i$ [CP ~~hogy~~ [IP elrepült Pali Freiburgba?]
         not   that is-Q   that   VM.flew P.   Freiburg.to
    'Isn't the (only) thing that that is the case identical to the following: P. has flewn to F.?'

Since the embedded clause is a positive declarative, the lack of inversion and compatibility with PPIs is explained. Second, *nem-eInt*s encode subquestions of a (possibly implicit) QUD, asking for the *single* interpretation, explanation, reason, cause, conclusion, or consequence of a state of affairs ⧄ in the context of the conversation. The talk will address the similarities between the behaviour of *nem-eInt*s and the "inferential construction" (*it's (just) that* or *it's not that*) in English,

cf. Delahunty (1995) and Remberger (2020), and derive the necessity of the speaker's epistemic bias for *p* from the fact that *nem-eInt*s indicate a rhetorical relation.

Yes and *no* in responses to negative (biased) questions: Russian vs. German

Ljudmila Geist, Sophie Repp
University of Cologne, *University of Cologne*
lgeist@uni-koeln.de, sophie.repp@uni-koeln.de

The cross-linguistic and inter-individual variation surrounding response particles has received increased interest in recent years, and experimental investigations have sharpened our understanding of the meaning and use of these particles in a number of languages. However, with respect to negative polar questions such as (i) serving as antecedents for response particles, there have been few systematic investigations that pay attention to the various meaning aspects that questions may express concerning speaker expectations and wishes, or the strength of the contextual evidence – the so-called *biases* of a question. There are different types of biases. For instance, questions may express that the speaker had a particular assumption about the truth of the proposition whose polarity is at issue, which is the so-called *epistemic bias*. Questions may also express that there is evidence in the context for or against the truth of that proposition, which is the so-called *evidential bias*. This is illustrated in (i) for a question with interrogative syntax and the negative marker cliticized onto the finite auxiliary.

(i)   A: Hasn't Ms Miller called him?   B: Yes/No, she has. // Yes/No, she hasn't.
   epistemic bias: Ms Miller has called him.
   evidential bias: Ms Miller hasn't called him. // no bias

Biases are quite important for theories of response particles because response particles are generally considered to be anaphoric devices that rely on antecedent propositions (or antecedent structures) in the discourse context, and biases essentially introduce propositions into the discourse context. Thus they influence what propositions will be available (and salient) as a potential antecedent for a response particle: In (i), the epistemic bias introduces a positive proposition, whereas the question form and the evidential bias seem to make a negative proposition salient. This might have consequences for the felicitous use of particles like *yes* or *no*. Importantly, what bias a question comes with depends on the syntactic form and lexical make-up of the question.

        In this talk we will present and discuss experimental results from an acceptability study testing three types of negative polar questions in Russian, and compare them to a similar study in German (Repp, Claus & Frühauf in prep.). A comparison between Russian and German is interesting for at least three reasons: First, Russian preferably uses declarative syntax for 'neutral' polar questions, which is relevant in this context because declarative V-second quest- ions in languages like German or English are not neutral and come with different biases than polar V-first questions (Gunlogson 2002; Trinh 2013). Second, Russian extensively uses different interrogative particles such as *razve* 'indeed', *neuželi* 'really' *ved'* 'but' and the particle *li*, which marks 'questioned constituents' to indicate different biases that have been described for negative polar questions. Third, Russian does not have a dedicated particle for rejections of negative antecedents comparable to German *doch*. We will analyze the results of our experiments in the semantic-pragmatic framework of Farkas & Roelofsen (2019), where response particles may on the one hand signal the polarity of the answer they express, and on the other hand, signal identity or complementarity of the response with the antecedent proposition. The role of the question bias will be analyzed in relation to current theories on the interaction between the (illocutionary) interpretation of lexical negation markers (cf. Romero, 2020) and different kinds of (illocutionary) particles (cf. Seeliger & Repp, 2018).

References: Farkas & Roelofsen (2019). "Response particles revisited." S&P 12(15). || Repp, Claus & Frühauf (in prep.). "Responses to biased questions." ms. || Seeliger & Repp (2018). "Biased declarative questions in Swedish and German: Negation meets modal particles (*väl*, and *doch wohl*)." In Dimroth &

Sudhoff (Eds.): *The grammatical realization of polarity contrast.Theoretical, empirical and typological approaches.* 129-172.

Is negation more difficult than affirmation?

Elena Albu, Oksana Tsaregorodtseva, Barbara Kaup
University of Tübingen
elena.albu@mnf.uni-tuebingen.de; oksana.tsaregorodtseva@uni-tuebingen.de; barbara.kaup@uni-tuebingen.de

Research question. In comparison with affirmation, the processing of negation is said to be more difficult when presented out of context (for an overview, see Kaup & Dudschig, 2020). When embedded in a supportive context, i.e. narrative stories where the proposition denied is either explicitly stated or strongly inferred (Lüdtke & Kaup, 2006) or the relevant attribute dimension is highlighted (Glenberg et al., 1999), the difficulty associated with negation is reduced or completely eliminated. In the present study, we investigated whether negation processing is also facilitated when presented in contexts provided by discourse connectives which deny contextual expectations (in the following: "denial contexts").

Experiment 1. We compared the response times (RT) of negative and affirmative sentences *([Contrary to expectations/ Surprisingly/ Unexpectedly/ Unpredictably], John has/hasn't eaten the soup*) in a sensibility-judgement-task. We expected an interaction between the factors *Context* and *Polarity* with (a.) significantly longer RTs for negative sentences in comparison with affirmation in the non-denial contexts and (b.) similar RTs for affirmative and negative sentences in the denial contexts.

Results. We analyzed the data of 79 participants (32 females; M_{age} = 38.13, SD_{age} = 11.32) by means of a repeated measures ANOVA with the factors *Polarity* (affirmative/negative) and *Context* (non- denial/ denial). There was a main effect of Polarity ($F(1,78)$ = 22.14, $p < .001$), with shorter RTs in the affirmative condition, and a main effect of Context ($F(1,78)$ = 145.1, $p < .001$), with longer RTs in the non- denial contexts. The interaction was not significant ($F(1,78)$ = 0.34, p = .512), invalidating our second prediction.

Experiment 2. Experiment 2 investigated the effect of context without the length confound present in Experiment 1: expressions with the same number of syllables were added to the non-denial contexts (*Everybody is convinced that/ Everyone thinks that/ We believe that/ Based on what we know, John has/hasn't eaten the soup*). The design and predictions were identical to those in Experiment 1.

Results. The data of 62 participants were analyzed (26 females; M_{age} = 39.96, SD_{age}=11.13). As in Exp. 1, the ANOVA revealed a main effect of Polarity ($F(1,61)$ = 21.02, $p < .001$) and a main effect of Context ($F(1,61)$ = 21.41, $p < .001$). This time, however, there were longer RTs in the non-denial contexts, possibly reflecting the complexity of the grammatical structures employed. Similarly to Exp. 1, there was no polarity-by-context interaction ($F(1,61)$ = 0.93, p = .339).

Experiment 3. To rule out that the previous results were an artefact of the task, as the RTs in the sensibility-judgement task included the time required for response decision and preparation, a self-paced reading paradigm was employed, where the participants read the sentences fragment by fragment (*Contrary to expectations // John has/hasn't eaten the soup*). Connectives with similar complexity were added to the non-denial context (*By all accounts/ Reportedly/ Apparently/ Supposedly, John has/hasn't eaten the soup*) The predictions were identical to those in Experiment 1.

Results. The analysis of the data (59 participants, 22 females; M_{age} = 39.76, SD_{age} = 13.11) revealed the same patterns: a main effect of Polarity ($F(1,58)$ = 56.31, $p < .001$), and a main effect of Context ($F(1,58)$ = 14.27, $p < .001$), but no significant interaction ($F(1,58)$ = 0.036, p = .851).

Conclusions. To sum up, this study aimed at investigating whether negation and affirmation behave similarly in denial contexts provided by discourse connectives. The discourse connectives were meant to provide the context of interpretation by activating, accommodating and rejecting contextual expectations. Both affirmative and negative sentences were designed around the mismatch between the polarities of contextual expectations and sentence meaning.

The results showed that the relevant interaction was not significant, indicating that polarity and context do not influence each other. In other words, the denial context provided by discourse connectives alone does not facilitate the processing of negation. In comparison with previous studies where negation was integrated in longer narrative stories, in our study, the context licenses negation but does not seem to provide any information about its relevance and informativeness, factors which appear to be crucial in reducing the processing difficulty associated with negation.

Slow and steady wins the race: Positive effects of the negated information on negative sentence comprehension in Italian dyslexic adults

Marta Tagliani
University of Verona & University of Göttingen
marta.tagliani@univr.it

State-of-the-art. In a previous work (in preparation) the author has provided compelling evidence that the visual prominence of the mentioned argument (i.e., the positive representation of the argument of negation) has a facilitating effect on the processing costs of negative sentences compared to the corresponding affirmatives. This evidence is in line with a non-incremental view of negation processing, suggesting that the higher processing difficulties traditionally reported for negative sentences must be attributed to the retrieval/activation of the negated information [1]. Previous literature on negation processing and developmental dyslexia [2,3] has reported an overall poor performance of dyslexic comprehenders in negative sentence interpretation compared to age-matched peers, as well as the classical interaction effect between truth-value and polarity. These results have been interpreted within a non-incremental theoretical framework of negation processing [4]: the simulation of the negated information and its momentary maintenance in the working memory is expensive in terms of processing resources, which are notoriously impaired in dyslexics due to limitations in their verbal working memory capacity [5,6].

 Aim of the present study. On one hand, experimental evidence with normally- developed comprehenders has shown that negation processing costs are hindered by the retrieval of the negated information; on the other hand, the retrieval of the negated information overloads limited working memory resources in dyslexics. The aim of the present follow-up work is to investigate whether the visual prominence of the mentioned argument might have a facilitating effect on dyslexics' negative sentence comprehension, avoiding the overload of working memory resources.

 The experiment. A picture identification task with a visual world set-up was administered. Following a two-second preview of the visual scenario, participants were auditorily presented with affirmative and negative sentences (Aladdin is (not) closing the door and Jasmine is cuddling a tiger), and they were asked to indicate the quadrant containing the referent of the verbal description while their eye movements were recorded. The visual prominence of the mentioned argument (i.e., Aladdin closing the door) was manipulated by parametrically varying the number of quadrants in which it appears from one to three (Fig. 1). Note that mentioned arguments constitute potential targets in positive sentences, whereas they are distractors to avoid in negative sentences. A group of 9 Italian adults diagnosed with developmental dyslexia participated in the study, and a second group of 9 typically developed Italian adults were included as control group. **Results and discussion.** We found an overall penalty for identifying the target in negative vs. positive conditions (p<.01).

 However, this penalty decreased as the number of mentioned arguments increased (p<.01). Furthermore, the fewer the mentioned arguments are the more steadily participants fixated the mentioned pictures in negative conditions: with one mentioned argument (Fig. 2A) participants' looks to the mentioned picture increase, regardless of the presence of negation, up to 1400ms from the onset of the critical word (Fig. 2A). However, as the number of mentioned arguments increases, participants shift their gaze away from the mentioned picture more rapidly (Fig. 2B-C). This indicates that the visual prominence of the mentioned argument facilitates the identification of the target in negative conditions. Moreover, dyslexics' negative sentence comprehension is enhanced by the visual representation of the negated information, as indicated by a good accuracy in target identification (65%).

 Conclusion. The findings of this follow-up study indicate that dyslexics benefit from a facilitating effect of the visual prominence of the mentioned argument on the interpretation

process, in line with a non-incremental view of negation processing. Working memory taxation is reduced by the easier retrieval/activation and maintenance of the negated information during sentence comprehension, allowing dyslexics to succeed in the task of target identification.

How negative concord licenses the acquisition of formal negation

Sumrah Arshad
University of Goettingen, Germany
sumrah.arshad@stud.uni-goettingen.de

Introduction: For decades Negative Concord (NC) has been a widely discussed topic within the generative framework. The main focus of the research on NC has been the syntactic and semantic status of negation and in particular negative elements involved in NC but little has been investigated about how young children acquire NC and use it as a clue to interpret negation/negative elements in their grammar. What determines how NC provides essential support for children in interpreting negation/negative elements is the main question of this paper.

    Theoretical background: While all languages express negation, Zeijlstra's proposed framework (2004, 2007, 2008a) predicts that languages may or may not require NegP, headed by a syntactic negative head, to express sentential negation. Double Negation (DN) languages in which a negative marker that is an adverb and serves as the negative operator and being able to be interpreted in the semantics directly, cannot have a NegP. On the other hand, languages where negative elements do not directly correspond to a negative operator are NC languages and may have NegP. Negation in NC languages is termed as syntactic/formal negation. For Zeijlstra (2004, 2008) NC is a syntactic agreement and negative elements are interpreted as carriers of formal features [i/uNEG]. Morphological evidence (MV) regarding the doubling effects credits negation as a formal syntactic category, and, this MV is provided in the form of NC sentences which contain multiple negative markers stipulating that one of them carries [uNEG] that must be checked in the syntactic component, this is what motivates a NegP.

    Proposal: I will use Zeijlstra's (2004, 2014) framework to explain the acquisition of negation and NC in SE. Zeijlstra proposes that children acquiring any language must settle the syntactic status of negative elements of their target language. SE is a DN language in which every negative form corresponds to a negative meaning (Labov 1972; Zeijlstra 2004, 2008, 2013). Consequently, it is predicted that children acquiring SE should start acquiring SE as a DN language where the negative marker is predominantly an adverb. At the same time, SE projects formal negation through NegP headed by a negative head *n't.* This means that children receive conflicting evidence. Since, the inclusion of *n't* must require the presence of negative formal features [iNEG]/[uNEG], because only the negative head co- occurs with a negative quantifier in NC, as shown in (3), we assume that children pass through a stage where they assume SE is full NC-language. Only after having reached this stage, will the child acquire that in SE only *n't* can be an NC item that agrees with a covert negative operator, and that all other negative elements are semantically negative as well.

    Data and results: This study provides empirical findings based on the analysis of Corpus data of SE retrieved from CHILDES database (McWhinney 2000). We predicted that children assume SE as a NC language and license NC. It confirms that children do produce NC but not before the age of 3 years, exactly the age when a negative head is projected in their grammar. Our analysis reflects that children's 'final' stage of negation passes through various stages of development. First, it shows that negation in early child English does not mirror the adults' most frequent form of negation. As predicted, it is also shown that children start acquiring SE as a DN language using extensively the adverbial forms of negation e.g. *no* and *not.* Their acquisition process goes through various stages until it reaches its final state, they place negation external to sentence. They gradually add negative expressions in their lexicon, i.e., *no > not > don't* etc., and so on (Thornton et al. 2013, 2016)*.* Our data findings confirm that children acquire the head *n't* after the adverbial forms which is the most frequent and throughout a stabilised form in adult speech. The acquisition of *n't* becomes considerably visible around the age of 36 months.

    Conclusion: Based on the finding, it is concluded that children's initial stages of acquisition of negation are instances of semantic negation but the adult speech contains consistent use of *n't*, the head form of negation. Children start from semantic negation and place negative adverb

no and *not* firstly in sentence external position and then internal. Given that the essential linguistic input (NC) required to acquire negation as a formal syntactic category is not readily available, their acquisition of *n't* is rather delayed. Based on the conflicting input (containing *a negative head* but not NC) children formalise negation and assume SE as an NC language and do produce and license NC. When the primary linguistic input violates children's hypothesis that SE is a NC language, only then the production of NC decreases in child speech and they realise SE as their target DN language.

Initial negation in Norwegian: A curious case of licensing

Henrik Torgersen
University of Oslo
hatorger@uio.no

The vast majority of Norwegian adverbs can front to the initial position quite freely. The negator *ikke*, 'not', may front to the initial position only under a narrow set of conditions. Sentential negation in the initial position of a V2 language has been discussed and argued to be marked (Zeijlstra 2014). Its interaction with syntactic hierarchy has been seen as problematic in how negation scopes over C0, creating a reading where the speaker negates their own illocutionary force (Frege 1892).Fronting of the negator is well documented for all the North Germanic languages except Danish. For Swedish, the phenomenon has been açnalyzed to a certain extent (Brandtler & Håkansson 2012; Seeliger 2018), but initial negation remains relatively unexplored in Norwegian.

At least 3 contexts allow for initial negation in Norwegian: enumeration, overt question answers and covert question answers. The enumerative usage (1) is shared with Swedish (Brandtler & Håkansson 2012) and Finland Swedish (Lindström 2009).

(1) Ikke   har han Vaska   og   ikke   har han   vært i butikken
     not   has he   washed   and not   has he   been in store.the
     "He has not washed and neither has he been to the store"

(Garbacz & Østbø 2012: 486)

Many Norwegian speakers may employ initial negation to answer overt questions (2a, b). Under this usage, the utterance marks the negative marker as a contrastive topic (CT; Büring 2013) and focuses another part of the sentence, *penger* - 'money' in this case. The speaker thus implies that they did receive something, but that this something was not money, as they might have expected.

(2) A. What   did   you   get   for   Christmas?
     B. Ikke$_{CT}$ fikk   jeg   penger$_F$ i hvert fall
         not   got   I   money   at least
         "I certainly didn't get any money at least"

The final usage context is in response to covert questions. These contexts reflect situations where there is a clear question under discussion (QUD - Roberts 2004) implicit to the discourse, but no direct question is posed. The idiomatic *ikke veit jeg* - 'I dunno' falls into this category

(3) [Context: Your co-host cannot find the receipt for a cake. They remark upon this fact out of frustration. You reply.]
     A. I   can't   find   that   damn Receipt
     B. Ikke$_{CT}$ husker   jeg$_F$ hva   Den   Kosta
         not   remember I   where   It   Is
         "I certainly don't remember where it is"

I answer what the licensing conditions of initial *ikke* are across the three contexts (1)-(3). To investigate the licensing conditions, I present data from 5 (LIA, CANS, TAUS, NOTA, NDC) spoken and 1 (NoWaC) written corpora of Norwegian that show when and how Norwegian speakers use initial negation, what constrains its usage and how we can account for the cases where fronting fails. The notion of contrastive topic (CT) captures the essence of (1) and (2), where CT-marking of the negative marker implicates the relevance of another sentences. This carries over straightforwardly to (3), where CT marking is used in response to a perceived rather than overt question. For the cases where initial negation fails, CT itself is discourse-structurally infelicitous.

The integration of acceptability tests into diachronic syntax: The case of presuppositional negation

Giuseppe Magistro
Ghent University
giuseppe.magistro@ugent.be

Research question: Can the gradual nature of the grammaticalization of an emerging negative polarity adverb be captured by an acceptability test experiment?

Theoretical background: In Italian, it is possible to mark the denial of an explicit presupposition through the negative polarity adverb *mica* (Cinque 1976, Zanuttini 1997). This element first originated as a minimizer, referring to a minimal quantity (literally *'crumb'*), and then became able to deny a presupposition via a grammaticalization process named Jespersen's cycle (Breitbarth et al. 2020 for an overview). It is often stated cross-linguistically that during this process, the emerging negative polarity adverb first starts by denying an explicitly mentioned presupposition and then becomes able to scope over implicit ones (Larrivée 2020). Eventually, at the end of the cline, it would be able to scope over completely new predicates, being promoted to the role of standard negator (Blaxter & Willis 2018). Although some quantitative studies have been undertaken, based on historical corpora (Hansen & Visconti 2009), the current stage of Jespersen's cycle in Italian has been mainly assumed with impressionistic intuitions and thus can be elusive. We believe that further empirical support to such theoretical claims may be given by formal experimentation in synchrony such as an acceptability test.

Material and methods: We programmed an experiment on Psytoolkit (Stoet 2010, 2017), where 56 speakers throughout Italy were asked to judge 50 target sentences and 14 fillers on a Likert Scale ranging from 1 (not acceptable at all) to 7 (completely acceptable). The set of target sentences was made up of 25 negative sentences where *mica* would deny an explicit presupposition and 25 non-presuppositional negative sentences that still had the emerging negator, as well. The sentences were also diversified pairwise by lexical domain and the possible position of *mica* in the clausal spine. In addition, reaction times were also measured for every item.

Results and Discussion: First of all, our results confirmed empirically the theoretical assumptions made so far in the previous literature: the presuppositional sentences had an average rating of 7 or 6, whereas those sentences with *mica* in non-presuppositional contexts had low ratings, namely 3 or 2 (the impact of semantic-pragmatic conditions was also tested with regression models: in standard contexts, the estimated rating would drop by -3.40 with a significance of $p < 0.001$). However, there were some sentences with median values corresponding to 4 and 5. These sentences, after a *post-hoc* exploration, revealed a possible accommodation of an implicit presupposition, inferable or retrievable by setting up plausible contextual scenarios, given the presence of presupposition triggers. More interestingly, in such a fuzzy area, where sentences had an average rating of 4 and 5, reaction times were also significantly higher than in those items that were rated as perfectly acceptable or completely unacceptable ($p < 0.001$). In light of the previous discussion, not only was this experiment useful in quantifying abstract and introspective judgments, but it also opened up to new experimental possibilities in capturing the change in progress, inasmuch as ratings occupy more medial position where the grammaticalization process is still beating its path (i.e. the denial of an implicit presupposition).

Further improvements and consequences: It would be insightful to test the methodology with strictly similar languages that are at different points in the grammaticalization cline. We may then expect a comparable shrinking of the fuzzy zone with a consequent decrease of RTs in more advanced varieties. Furthermore, it would be essential to come up with a factorial design to control rigidly the optional accommodation of an implicit presupposition.

Expletive negation: From embedded speech-acts to embedded propositions

Chloé Tahar
Institut Jean Nicod, DEC, ENS, EHESS, CNRS
chloe.tahar@gmail.com

Overview. This talk focuses on a non-canonical use of negation, characterised by a seemingly absence of meaning, expletive negation (ExN). In line with (Mari & Tahar, 2020), we argue that expletive negation in French continues the Latin prohibitive negation. We cast the analysis of prohibitive negation within (Krifka, 2014)'s model of embedded speech-act and propose that ExN is the 'fossilization' of prohibitive negation and it is what remains of a long gone embedded negative imperative in French. **Data.** ExN in French originates from the Latin prohibitive negation *ne*, a negation marker specialized for the construction of negative imperatives. In embedded clauses, *ne* occurs under directive verbs, like verbs of command (*impero*, 'order') and prohibition (*prohibeo*, 'forbid'), or desiderative verbs, such as verbs of wish (*vel*, 'wish') and fear (*metuo*, 'fear'). When embedded under prohibition or fear verbs, *ne* is said to receive a non-negative (or paratactic) reading.

(1) At   **ne** videas   velim.
     But *ne* see-2SG.SBJV wish-1SG
     `But I wish (that) you wouldn't see it.'
(2) Metuo, **ne** sero   veniam.
     fear-1SG *ne* late come-1SG.SBJV
     `I fear (that) I might be late.'

Analysis. Latin. We propose that in Latin, *ne*-clauses are embedded speech-acts (following (Krifka, 2014, see also (Jespersen, 1917); (Ageno, 1955); (Parry, 2013)). In this perspective, *ne* assumes a clause-typing function and heads the Force projection (Rizzi, 1997)).

(3) a. [$_{VP}$ velim [$_{ForceP}$ **ne** videas]]]]
     b. [$_{VP}$ metuo [$_{ForceP}$ **ne** sero veniam]]]]

In Latin, directives and desideratives are speech-act embedding predicates, (Krifka, 2014) serving a presentative function. Embedded under prohibition and fear verbs, *ne* does not receive an expletive reading, but a reading redundant with the meaning of the verb. Prohibitive verbs name the illocutionary act performed by the the embedded negative imperative ("X forbids: PROHIBITION") while fear verbs are meant to provide evidential motivation (see also (Krifka, 2017)) to the utterance of the embedded negative imperative ("X fears: PROHIBITION"). **French.** During the transition from Latin to French, the use of the complementizer *que* develops and systematizes**.** Directives and desideratives systematically subcategorize for *that*-clauses**.** Prohibitive *ne* loses its clause-typing function in embedded clauses and is reanalysed as a negation marker hosted by the MoodP (see (Cinque, 1999)).

(4) Je crains qu'il **ne** vienne (`I fear that he ExN might come.')
     [$_{VP}$ Je [$_{VP'}$ crains [$_{CP}$ que [$_{IP}$ il$_1$ [$_{MoodP'}$ **ne** [$_{Mood'}$ SBJV [$_{IP}$ t$_1$ vient]]]]

ExN conveys the speaker's preferential attitude towards the negation of the complement at the non-at-issue level (see also (Yoon, 2011); (Zovko-Dinkovic, 2017); (Liu, 2019)). We make use of the notion of 'Modal Concord' (à la (Huitink, 2012)) to explain why ExN yields a unitary semantic meaning of dispreference with the main verb. **Conclusion.** This talk brings historical evidence in favor of the hypothesis that languages can develop from parataxis to hypotaxis over time and argues that speech-act embedding is an intermediary step between these two stages. Our analysis of prohibitive negation as a sentential mood marker in Latin and as a verbal mood in French also brings a new piece of evidence to the claim that sentential and verbal mood marking are two intimately related phenomena.

Exploring the landscape of German polarity items and their licensing conditions

Katharina Schaebbicke, Heiko Seeliger
University of Cologne, *University of Cologne*
katharina.schaebbicke@uni-koeln.de, heiko.seeliger@uni-koeln.de

Positive and negative polarity items (PPIs and NPIs) are frequently used in linguistic research − either in research on polarity items themselves, or as diagnostics for other linguistic phenom- ena, such as propositional vs. non-propositional negation or question bias. However, we are faced with an empirical problem when it comes to the aforementioned research: research on and with polarity items often relies on the use and analysis of very few token NPIs and PPIs. There is, however, a vast and heterogeneous landscape of NPIs and PPIs, and it is an open question whether they can be theoretically accounted for by analysing just a few of them. A second problem is that often the precise licensing conditions of a given NPI or PPI are not known, which can lead to confounds when they are used in experiments. In this talk, we present three exploratory experiments that are aimed at providing a better empirical insight into a large set of German NPIs and PPIs. The goal of these experiments was to classify a set of German NPIs and PPIs into superstrong, strong, and weak negative and positive polarity items (van der Wouden 1994, Zwarts 1993), and also to explore the relationship of nonveridicality and the (anti-)licensing of those German NPIs and PPIs (Zwarts 1995, Giannakidou 1997, 1998, 2002, 2011). NPIs and PPIs were selected from the Collection of Distributionally Idiosyncratic Items (CoDII) (Sailer/Trawinski 2006a, Sailer/Trawinski 2006b, Trawinski/Soehn 2008, Trawinski et al. 2008). In **Experiment 1**, 60 German NPIs were selected. Participants were asked to judge the acceptability of sentences containing the NPIs in six different conditions that should allow the classification into superstrong, strong, weak and nonveridicality-licensed NPIs: antimorphic (*not*), anti-additive (*no*), downward entailing (*hardly*), nonveridical (*maybe*, question). Controls were positive assertions without any operator. Cluster analysis revealed seven clusters of NPIs, some of which confirm the licensing categorization from the literature (superstrong and weak NPIs). Other clusters show unclear patterns (overall good or medium ratings) and require further scrutiny in future research. One cluster showed an unexpected pattern, with high acceptability ratings only with the antimorphic and the question operator. **Experiment 2** tested whether the source of this unexpected distribution was a rhetorical interpretation of the questions. Results suggest that rhetoricity was not the sole source. **Experiment 3** tested 24 PPIs from the CoDII corpus in the same conditions as Experiment 1. Here, the reverse pattern from NPIs was expected. Cluster analysis revealed three major clusters: cluster 1 was rated acceptable only in positive assertions without an operator, cluster 2 was additionally rated acceptable in the nonveridical condition with *maybe*, and cluster 3 was rated acceptable in all non-negative conditions. Overall, the results of Experiment 1-3 show gradual rather than categorical differences in acceptability, with higher acceptability corresponding to stronger negativity for the NPIs, and vice versa for the PPIs. How this graded acceptability can be accounted for by theories of licensing remains an open question.

Arbeitsgruppe 7

On the nouniness of propositional arguments

Katrin Axel-Tober, Lutz Gunkel, Jutta M. Hartmann & Anke Holler

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

The evolving of nouny subordination in Hungarian: From parataxis or from correlatives?

Katalin É. Kiss
Research Institute for Linguistics
e.kiss.katalin@nytud.hu

The talk will analyze the evolution of Hungarian complement clauses and their complementizer hogy 'that', which is form-identical with the wh-phrase meaning 'how; as'. The following developmental path will be documented: Proto-Hungarian, similarly to present-day Khanty and Mansi, its conservative sister languages, only used non-finite subordination, and, in the case of verbs of communication, parataxis. The first sentence type with properties of finite subordination emerging in Khanty is the correlative construction, involving an indefinite/interrogative pronoun in the initial clause and an overt or dropped definite pronoun in the second clause, such as Who... he...; Where... there...; As... so.... The abundance of this construction in Old Hungarian suggests that correlatives represented the first type of subordination in Proto-Hungarian, as well. With the shift of word order from SOV to SVO, inverse correlative structures (He... who...; ...there where...) also started spreading. In ...so as... constructions, as-clauses containing an indicative verb functioned as clauses of manner, and those containing a subjunctive verb functioned as clauses of purpose. Verbs of communication, followed by a direct quotation, also contained the adverb so (He spoke so; He said so.).

The generalization of finite subordination resulted in the embedding of direct quotations. By analogy, the correlative [CP ...so [CP as ....]] pattern came to be extended to constructions involving a verb of communication complemented by an indirect quotation, as well. Hogy 'as' eventually developed into a general complementizer. Later, the adverb so associated with indirect quotations came to be replaced by the pronoun az-t 'that-ACC', the strong version of the 3rd singular [-human] pronoun. Object clauses coindexed with an (overt or pro-dropped) pronoun trigger agreement on the verb, which is evidence of their nouniness – as object-verb agreement is only elicited by DPs (Bartos 2000).

In late Middle Hungarian, free relative complement clauses underwent a further change: their pronominal associate in the main clause came to be reanalyzed as part of the relative wh-pronoun, as a result of which relative wh- pronouns assumed an a-prefix, and the pronoun had to be spelled out again (i.e., az-t, mi-t 'that-ACC, what-ACC' > azmi-t > ami-t > az-t, ami-t).

The developmental paths to be presented may be relevant for the controversy concerning the evolution of Germanic that-clauses (cf. Lenerz 1984; Hopper & Traugott 1993; Axel-Tober 2017; etc.).

Japanese nominalizations and the copula

Paul Poirier
University of Toronto
paul.poirier@mail.utoronto.ca

There is a vast literature focused on nominalizations embedding different major clausal projections, such as TP and CP (e.g. Kornfilt and Whitman 2011). Increasing research argues the need for a more fine-grained ap- proach to clausal structure, such as the articulated C-domain espoused by Rizzi (1997). I will make the case here that such an approach is also re- quired for Japanese nominalizations headed by *koto* and *no*, providing evi- dence from the copular paradigm. While adjectival and nominal predicates surface with non-past copula *da* in matrix contexts, the special adnominal form *na* appears in the nominalizations, patterning with other complex NPs:

(1)   a. Sarah-wa shinsetsu da/*na
            Sarah-Top kind   Cop/NA
            'Sarah is kind.'

      b. [John-ga shinsetsu *da/na   koto/no]-ni   odoroi-ta
          John-Nom kind   Cop/NA   koto/no-Dat surprise-pst
           'It surprised me that John is kind.'

Following Rizzi's [Force>Topic*>Focus>Topic*>Fin] C-domain hierarchy, I propose that these nominalizations embed only the lowermost head of the C- domain– Fin. I will argue that non-past *da* is best analyzed as being depen- dent on the Focus head, as its presence in matrix clauses forces an exhaus- tive reading for the subject, which has been tied to raising to Spec/FocP (Watanabe 2003). This suggests that *da* should also find its locus on the Focus head. Meanwhile, the FinP analysis of *koto*/*no* clauses entails the absence of Focus from their structure, and explains why *da* cannot surface. Other analytic and past-tense copular forms do not show the same sensitivi- ty to C-domain structure, and surface uniformly across clause types.

      Since the different forms of the copula depend on different heads in the clausal spine, we can extend the use of the copular distribution to determine how much functional structure is present in other embedded contexts, includ-ing interrogatives and conditionals. Moreover, the distribution of the copular forms in *koto* and *no* clauses present further evidence to support the need for an articulated C-domain. This would suggest that nominalizing heads are sensitive to this more fine-grained clausal structure, and consequently, that CP-nominalizations across languages do not constitute a uniform class.

References: Kornfilt, Jacklin, and John Whitman. 2011. Afterword: Nominalizations in syn-tactic theory. *Lingua* 121:1297-1313. Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, ed. Liliane Haegeman, 281-337. Dordrecht: Kluwer. Watanabe, Akira. 2003. Wh and operator constructions in Japanese. *Lingua* 113:519-558.

Nouny propositions and their individual correlates: The view from Japanese

Elizabeth Bogal-Allbritten, Keir Moulton, Junko Shimoyama
Götesborg universitet, University of Toronto, McGill University
Elizabeth.bogal-allbritten@gu.se, keir.moulton@utoronto.ca, junko.shimoyama@mcgill.ca

In this paper, we examine languages with morphosyntactically nominal propositional arguments (NomProps). Based on evidence from Korean and Japanese, we propose that NomProps can denote either individual assertion events (Krifka 2014, Bogal-Allbritten and Moulton 2017) or ordinary individuals with propositional content (Kratzer 2006, Moulton 2015).

In Japanese, finite clauses can be nominalized by the element *-no*. These can complement attitude verbs like *shinji-* 'believe' (1).

(1)   Watashi-wa [Johnny-ga shukudai-o   zembu shi-ta   no]-o   shinji-teir-u.
        I-TOP   J.-NOM   homework-ACC all   do-PST no-acc believe-ASP-NONPAST
         'I believe that Johnny finished his homework.'

As with Korean (Bogal-Allbritten and Moulton 2017), these Japanese NomProps require a discourse-familiar assertion. (1) can follow a discourse such as "*Johnny finished his homework. Can he play?"* but not one where the proposition is not asserted, e.g. "*Did Johnny finish his homework?".* We argue this shows NomProps at least sometimes denote assertion events.

We then show via an ambiguity in memory-reports that Japanese NomProps can also simply describe contentful individuals. Like English, the verb *oboe-* 'remember', can describe a direct/vivid memory (*I remember him winning*) or an indirect memory (*I remember that he won*) (Stephenson 2010). In Japanese, however, a plain NomProp cannot express the indirect version; instead the element *toyuu,* which contains a grammaticalized verb of saying, is required.

 (2)   [1703-nen-ni Kuranosuke-ga nakunat-ta *(**to-yuu) no**]-o   oboeteiru
         1703-year-in K.nom   passed-PST TO-YUU NO-ACC remember
         'I remember that K. passed away in 1703.'

We analyze the difference between bare NomProps in (1) and those in (2) with *toyuu* in terms of how the propositional meaning arises: in (1), the NomProp describes a familiar assertion event whereas in (2) it describes an abstract individual with propositional content, building on analyses of *toyuu* by H. Saito (2018). We then show that the meaning differences independently follow from selectional properties of the embedding verbs.

References. Bogal-Allbritten, E. & K. Moulton. 2017. Nominalized clauses and referent to propositional content. SuB 21. Kratzer, A. 2006. Decomposing attitude verbs. UMass. Krifka, M. 2014. Embedding illocutionary acts. Recursion: Complexity in cognition. Saito, H. 2018. (De)categorizing speech. UConn. Stepehnson, T. 2010. Vivid attitudes. SALT 20.

Relatively nouny?

Carlos de Cuba
Kingsborough Community College, City University of New York
carlos.decuba@kbcc.cuny.edu

In an effort to reconcile Kayne's (1994) Linear Correspondence Axiom with Chomsky's (1995) Bare Phrase Structure, Kayne (2008), following Guimarães (2000), proposed that a head *x* can merge with itself, yielding the singleton set {x}. This solved a projection problem that occurs when merging two heads that would otherwise be in a symmetrical c-command relationship, causing a linearization problem (with mutual c-command it would be unclear which head should project). An upshot of the proposed analysis in the paper is the claim that nouns do not project, meaning that they do not take complements. Given that nouns have traditionally been analyzed as taking a number of different types of complements, the onus was then on Kayne to show that what we have been calling complements of nouns are not in fact complements. Kayne's solution was to propose that instead of complements, we were dealing with relative clause structures, which are adjuncts (see Arsenijević 2009 for a related proposal). This relative clause analysis has gained a lot of traction over the years, enough so to be featured prominently in the call for papers for this workshop. However, in this talk I will attempt to throw some cold water on the relative clause analysis. I will present a number of problematic issues that arise with the proposal that all complement clauses can be analyzed as RCs. I will show that the evidence that has been put forth in favor of the RC analysis in the literature is weak, and that cross-linguistic evidence points strongly away from a uniform treatment of complement clauses as RCs.

References. Arsenijević, Boban. (2009). Clausal complementation as relativization. *Lingua* 119(1). 39–50. DOI: https://doi.org/10.1016/j.lingua.2008.08.003. Chomsky, Noam (1995). *The Minimalist Program*, The MIT Press, Cambridge, Mass. Guimaraes, M. (2000). In Defense of Vacuous Projections in Bare Phrase Structure. In M. Guimaraes, L. Meroni, C. Rodrigues & I. San Martin (Eds), *University of Maryland Working Papers in Linguistics*, 9, 90–115. Kayne, Richard. (2008). Antisymmetry and the lexicon. *Linguistic Variation Yearbook* 8. 1–31. DOI: https://doi.org/10.1075/livy.8.01kay. Kayne, Richard. (2010). Why isn't *this* a complementizer? In Richard Kayne 2010. *Comparisons and contrasts*, 190–227. Oxford: Oxford University Press.

CP-complementation and selection

Ellen Brandner
University of Stuttgart
eleonore.brandner@ling.uni-stuttgart.de

In this talk, I will provide further evidence for the idea that clausal complements should be analyzed as modifiers of a (possibly null) pronominal argument of the verb instead of being selected as such, Kratzer (2006). I will approach the issue by considering cross-clausal dependencies, i.e. 'long extraction' in languages where the type of the complementizer changes, i.e. as it is the case in Celtic languages and in Alemannic, see Brandner & Bucheli-Berger (2018):

(1)　a. des isch des buech [**wo**　de　Peter glese hät]　　RC
　　　　 this is 　the 　book 　　RC 　the Peter read 　has

　　 b. %wer 　hesch 　　gseet [**wo** d'　Marie moant [**wo** (*er)　en Unfall 　　gha hät]]
　　　　 who 　 have.2sg 　said 　RC the Maria thinks 　RC 　　 　an accident 　had had

　　 c. wer hesch 　　gseet [**dass** d'Marie 　moant [**dass** *(er)　en Unfall 　　gha hät]]
　　　　 who have (you) said 　that 　the Maria 　thinks 　that 　　　an accident 　had had
　　　　 b. and c. 'Who did you say that M. thinks that had an accident.'

Whereas a propositional complement is (usually) realized with a complementizer of the d-pronoun series (*dass*), this may change when extraction has taken place. In this case, the relative clause (RC) particle, exemplified in (1a), shows up as in (1b) – although the dass-complementation is a possibility as well, (1c). Note that (1b) does not allow a resumptive pronoun whereas this is nearly obligatory in the *dass*-case. The crucial point is that the length of both constructions is identical, i.e. the insertion of the resumptive cannot be due to complexity/parsing problems. The first question arising is how the matrix verb can tolerate a relative clause as its complement, since a relative clause can hardly be taken as being c-selected by a verb. If we do not want to give up the widely established analyses of RCs as being modifiers of nominal expressions, the answer can only be that the CCs in these cases are introduced into the structure as an RC (with an inherent gap) and not as a (selected) complement of the verb in form of an embedded clause with successive cyclic movement. The next question is whether this analysis can be transferred to the cases in (1c) with the 'usual' complementizer. Following the analyses by Kayne (2014) and Axel-Tober (2017), it will be argued that this type of clausal complements are indeed 'explicative relative clauses' that occur without a gap, cf. (1c). It will be argued that the long distance dependency in this case is established via a proleptic construction, see Salzmann (2006). This analysis is transferred to clausal embedding in general with a (possibly null) nominal correlate in the matrix clause. If it were true that verbs directly select for their clausal complements, the differences in shape of the complementizer nor the varying distribution of resumptives could not be captured.

References: Axel-Tober, K. (2017). The development of the declarative complementizer in German. *Language*, *93*(2), 29-65. Brandner, E. & C. Bucheli (2018). Über lange W-Extraktion im Alemannischen. In: Sardis aus Saarbrücker Sicht 2 (ZDL Beihefte, 170). Kayne, R. S. (2014). Why isn't this a complementizer. *Functional structure from top to* toe, 188-231. Salzmann, M. (2006). Resumptive Prolepsis: A study in Indirect A'-dependencies. Utrecht: LOT Publications.

Where propositional arguments and participial relative clauses meet

Éva Dékány, Ekaterina Georgieva

Research Institute for Linguistics, Budapest

dekany.eva@nytud.hu, ekaterina.georgieva@nytud.hu

Claims: In this talk we look at non-finite propositional arguments which have nominal properties and involve the same suffix as participial relative clauses (pRCs). This is observed in the Uralic, Altaic, Quechua and Tibeto-Burman languages (Koptjevskaya-Tamm 1993; Serdobolskaya & Paperno 2006; Shagal 2018). We investigate this phenomenon in Udmurt and Khanty (Uralic) as well as Kazakh, Modern Standard Turkish, Uyghur and Korean (Altaic). We argue that in the relevant cases non-finite propositional arguments with nominal properties structurally contain the projection of pRCs; however, the polysemy arises from different underlying structures. Languages exhibiting the polysemy fall into 3 types, parametrically differing in the structure of pRCs and propositional arguments.

Analysis: We argue that the shared suffix of pRCs and propositional arguments with nominal properties expones an aspectual head in the extended VP (Collins 2005; Baker 2011; *pace* Doron & Reintges 2005). Variation is observed in the structure of pRCs and nominalizations. We propose that:

1) pRCs fall into two types: they are either bare AspPs or they are nominalized before they are merged with the head noun. The nominalized status of the pRC is reflected in the obligatory possessive agreement of the clause and the genitive marking of the subject.
2) Nominalizations comprise the Asp of pRCs and an additional (covert) element that gives the external nominal distribution to the phrase. The additional element may be: (i) a nominal functional head, e.g., *n* or D, that takes AspP as its complement (mixed extended projections, cf. Borer 1997; Borsley & Kornfilt 2000; Alexiadou 2001); or (ii) a covert N with the meaning of 'event' or 'fact' taking the AspP as an RC modifier/complement. Empirically, the latter type is manifested by the alternation of overt and covert nouns or the presence of overt light Ns.

Cross-linguistic variation: The languages under consideration fall into the following types: (i) Udmurt and Kazakh have bare pRCs and mixed extended projections as propositional arguments; (ii) Modern Standard Turkish employs mixed extended projections in both RCs and propositional arguments; (iii) propositional arguments in Korean, Uyghur and Kazym Khanty involve covert/overt light Ns.

References: Asarina & Hartman. 2011 Uyghur Genitive Subjects and the Phase Impenetrability Condition. Proceedings of WAFL7 Kim 2009. E-type anaphora and three types of kes-construction in Korean. NLLT. Kornfilt 2000. Some Syntactic and Morphological Properties of Relative Clauses in Turkish.

That relatives! and the relativization of *dass*-clauses in German

Kalle Müller
University of Tübingen
kalle.mueller@uni-tuebingen.de

This talk is concerned with the question whether *dass* can introduce relative structures in German. It has been argued recently that many instances of *that-* clauses involve relative structures rather than complement structures (e.g. Aboh 2005, Kayne 2008, Arsenijevic 2009, Haegeman & Ürögdi 2010). This concerns especially cases like *the fact/claim that* what will be referred to by the neutral term 'noun related clauses' (= NRCs). However, this claim has been challenged by de Cuba (2017) who argues that crosslinguistically, languages which, unlike English, have separate forms for declarative complementisers and relative particles always employ the latter to introduce relative clauses. Distinguishing between internal and external issues of syntax, I argue that *dass*-clauses are compatible with operator movement but are replaced by different relativisers in the case of more accessible constituents. NRCs can be either complements or adjuncts of the noun, hence the form of non-complement NRCs (i.e. V-final vs. V2) cannot be selected syntactically. Evidence is drawn from various phenomena: Clefted temporal adverbials like *Es war zu dieser Zeit, dass ich müde wurde* 'It was at this time that I became tired' show how the lack of an adverbial relative pronoun *wann* 'when' in German sparks the competition between *als* and *dass* to relativise the time argument. Furthermore, I apply Fabricius-Hansen & von Stechow (1989) test for implicative vs. explicative constituents to show the NRC behaves like an adjunct for some nouns but like a complement for others. Finally, following Reis (1997), who has argued that embedded V2-clauses are not syntactically licensed as complements, the same is argued for NRCs in the form of V2-clauses: they can only be semantically licensed as adjuncts.

References. Aboh, E. 2005. Deriving relative and factive clauses. In Brugè et al. (eds.) Contributions to the Thirtieth Incontro di Grammatica Generativa. 265–285. Venezia: Cafoscarina. Arsenijević, B. 2009. Clausal complementation as relativization. Lingua 119(1). 39–50. de Cuba, C. 2017. Noun complement clauses as referential modifiers. Glossa: a journal of general linguistics 2(1): 3. 1–46. Fabricius-Hansen, C. & A. von Stechow. 1989. Explikative und implikative Norminalerweiterungen im Deutschen. Zeitschrift für Sprachwissenschaft 8. 173–205. Haegeman, L. & B. Ürögdi. 2010. Referential CPs and DPs: An operator movement account. Theoretical Linguistics 36(2–3). 111–152. • Kayne, R. S. 1994. The antisymmetry of syntax. MIT Press. Kayne, R. 2008. Antisymmetry and the lexicon. Linguistic Variation Yearbook 8. 1–31. Reis, M. 1997. Zum syntaktischen Status unselbständiger Verbzweit-Sätze. In Dürscheid et al.(eds.) Sprache im Fokus. 121–144. Tübingen: Niemeyer.

Reconsidering the syntax of correlates and propositional arguments

Andreas Blümel, Nobu Goto

University of Göttingen, Toyo University

ablueme@gwdg.de & ngoto@toyo.jp

This talk reconsiders well-known instances of nominal correlates and their associated propositional arguments primarily in German and English, and suggests a novel descriptive generalization with respect to their syntax: The association of a correlate with a sentential argument to the matrix predicate requires a category label on that sentential argument.

It is well-known that in numerous cases, German and English exhibit asymmetries between subordinate clausal arguments and root clauses fulfilling the same function. In German, such contrasts are manifested by V-final clauses introduced by complementizers and V2-clauses. In English, such contrasts are manifested e.g. by clauses headed by C°=*that* and such clauses headed by what is commonly analyzed as a null-C-head, C°=∅. One such said asymmetry is that *that/dass*-clauses can function as complements to nouns, while null-C- headed clauses and V2 cannot.

We seek to derive these patterns as follows: Drawing on the labeling algorithm LA suggested by Chomsky (2013, 2015), Blümel & Goto (2020) propse that root clauses are syntactically characterized by the obligatory absence of a category label. Assuming that a category label is syntactically required for the ongoing (Set Merge) computation, a label becomes superfluous when the derivation comes to an end – which is the case at the root node.

Given this much, we propose that unlabeled syntactic objects cannot associate syntactically with nominal elements, such as correlates and nominals that can take clausal arguments. For the sake of this talk, we stipulate (1):

(1) An unlabeled syntactic object must not be co-indexed with a nominal.

We show how (1) derives the attested empirical patterns. As to the question how English meet the requirement to leave root-clauses unlabeled. We suggest the following: Based on Chomsky's (2015) idea that the phase head **v* is a "syntactic affix" which is invisible to the LA, Obata (2018) proposed that his C-deletion analysis can be recast: C° and T° form a complex head amalgam in which C°'s phasehood is cancelled out (cf. also EKS 2016, Sugimoto 2016). Her proposal opens up the possibility that English root clauses are analyzed as C°=∅ which can undergo Set Merge with TP, yielding {C°=∅, TP}. Assuming with Obata that the C-head is invisible to the LA, {C°=∅, TP} is the structure of English root clauses – an exocentric structure, as desired. Based on this, we show how the attested asymmetries between null-C-headed clauses and *that*-CPs derive.

References: Blümel, A. & N. Goto (2020) Head Hiding. *Proceedings of NELS 50.* Obata, M. (2018) Eliminating C-deletion in the syntax: structure- building by Merge. *Koganei Journal of the Humanities* 14:21–34.

Nouny clauses: The clausal prolepsis strategy

Nikos Angelopoulos
KU Leuven
n.angelopouloss1@gmail.com

1. **Introduction.** This abstract proposes a new analysis of object clausal prolepsis in Dutch (cf. 1) inspired by the BigDP configuration that has been proposed for clitic doubling in Romance (cf. Uriagereka 1995 i.a.).

   (1) Jan betreurt het dat Marie onstlagen is].
        Jan regrets   it   that Marie fired   is

2. **The analysis.** Under the proposed analysis, (2), *het* 'it', the proleptic form used in clausal prolepsis, is a D head that selects a silent pronoun (as complement) and in prolepsis, a CP (as specifier).

   (2) [$_{DP}$ CP [$_{D'}$ het$_D$ pro]]

3. **No accidental homophony.** A first advantage of the proposed analysis is that it can capture the different meanings we will *het* 'it' can have in Dutch without postulating different accidentally homophonous lexical entries. In this analysis, *het* is a D head that (c-/s) selects an NP complement in DPs like *het boek* 'the book' or, a *pro* which can be individual or propositional denoting.

4. **The internal structure of the proleptic proform.** The analysis in (2) can also explain why in contrast to *het*, other DPs e.g. *dat* 'that', cannot double an embedded clause in Standard Dutch:

   (3) Jan betreurt (het/*dat/) [dat Marie onstlagen is.]
        Jan regrets   (it/ *that)   that Marie fired   is

   Concretely, it is well known that demonstratives, e.g. *dat*, occupy the Spec DP (cf. Leu 2007 i.a.). This makes them incompatible with a doubled CP in the same position thus, blocking prolepsis.

5. **The distribution of clausal prolepsis.** (2) can account for a hitherto unobserved generalization:

   (4) *The Prop-Prolepsis Generalization*: Clausal prolepsis can occur in all and only those contexts that allow for propositional het.

   Under (2), (4) is accounted for: a proleptic clause is simply a propDP with a specifier. Clausal prolepsis is available, if propDP is selected by a verb.

6. **Clause type restrictions on clausal prolepsis.** Lastly, (2) can capture the fact that in clausal prolepsis, *het* in prolepsis imposes stricter restrictions on the kind of clause that it doubles, e.g. it cannot double a wh-question, (5):

   (5) *Jan vroeg het wie   mijn lievelingsdichter   was
        Jan   asked it   who   my   favorite poet   was

   This restriction follows from the fact that the CP is selected in (2) by *het* and that this element only selects for familiar clauses.

Reference: Uriagereka, J. (1995). Aspects of the syntax of clitic placement in Western Romance. Linguistic inquiry, 26(1), 79-123.

A syntactic account of clausal complementation in Jula

Alassane Kiemtoré
University of Stuttgart
akiemtor04@yahoo.fr

This paper attempts a unified syntactic derivation for complement clauses constructions in the West African Language Jula (Manding-Niger-Congo, SOV), using two mechanisms: predication (cf. Bowers 1993, Den Dikken 2006, Citko 2011) and Case assignment à la Koopman (1992). Two types of constructions are considered: (i) the complement clauses associated with correlate, (ii) the complement clauses without correlate. Despite their difference in the surface, I propose to derive the two constructions from the same underlying structure. In practice, the relation between correlate and complement clause (CP) is analyzed as an instance of predication. In this respect, they are both base-generated within a predication phrase to the right of the hosting matrix clause. The position of the correlate within the matrix clause results from a SpecX to SpecX movement triggered by Case assignment, in accord with the SOV word-order of the language. The absence of correlate is due to a principle active in Jula grammar according to which the specifier position of covert case assigning head must remain covert. Overall, the proposed analysis has at least two theoretical implications. First, it supports the observation that complement clauses can be base-generated in a non-argument position (Postal and Pullum 1988, Haider 1995, Moulton 2009 Frey 2016, i.a.,). Therefore, their position does not result from movement out of the matrix clause, aka extraposition (Culicover and Rochemont 1990, Schwabe 2013, i.a.,). Second, by treating the relation between correlate and complement clause as a case of predication, the analysis, in some way, goes against approaches that view complement clauses as complements to nominal heads or treat them on a par with relative clauses (cf. Aboh 2005, Arsenijevic 2009, Kayne 2014, i.a.,). As the analysis predicts, complement clauses in Jula are not involved in such a relation.

References: Aboh, Enoch (2005). "Deriving relative and factive clauses". In: Brugè, Laura et al. (eds.), Contributions to the thirtieth Incontro di Grammatica Generativa, Venezia: Libreria Editrice Cafoscarina, 2005, pp. 265-285. Venezia, Libreria Editrice Cafoscarina. Arsenijevic, Boban (2009). "Clausal complementation as relativization". In: Lingua 119.1, pp. 39–50. Bowers, John (1993). "The syntax of predication". In: Linguistic inquiry 24.4, pp. 591– 656. Citko, Barbara (2011). "Small clauses". In: Language and Linguistics Compass 5.10, pp. 748–763. Culicover, Peter W and Michael S Rochemont (1990). "Extraposition and the complement principle". In: Linguistic Inquiry 21.1, pp. 23–47. Den Dikken, Marcel (2006). Relators and linkers: The syntax of predication, predicate inversion, and copulas. Vol. 47. MIT press. Frey, Werner (2016). "On properties differentiating constructions with inner-sentential pro-forms for clauses". In: Inner-sentential propositional proforms: Syntactic properties and interpretative effects, pp. 1– 21. Haider, Hubert (1995). "Downright down to the right". In: On extraction and extraposition in German 245271. Kayne, Richard S (2014). "Why isn't this a complementizer". In: Functional structure from top to toe: A Festschrift for Tarald Taraldsen. Koopman, Hilda (1992). "On the absence of case chains in Bambara". In: Natural Language & Linguistic Theory 10.4, pp. 555–594. Moulton, Keir (2009). "Natural selection and the syntax of clausal complementation". Postal, Paul M and Geoffrey K Pullum (1988). "Expletive noun phrases in subcategorized positions". In: Linguistic Inquiry 19.4, pp. 635–670. Schwabe, Kerstin (2013). "Eine uniforme Analyse sententialer Proformen im Deutschen". In: Deutsche Sprache 41, pp. 142–164.

Verb-y and noun-y complementation in Kipsigis

Imke Driemel, Maria Kouneli
Universität Leipzig
imke.driemel@uni-leipzig.de, maria.kouneli@uni-leipzig.de

In this talk, we present novel data from Kipsigis (Nilotic, Kenya) that reveal two types of CP complements. The starting point of our investigation is what has been described as upwards-oriented complementizer agreement with a matrix subject (Diercks & Rao 2019, Diercks et al. 2020): the complementizer consists of the root of the verb *le* 'say' and an agreement prefix. We argue that what has been described as an agreeing 'say'-based complementizer in Kipsigis is in fact the lexical verb 'say'. We offer the following arguments in favor of this position: i) *le* 'say' can be used as a matrix verb, ii) *le* is inflected in the indicative in matrix uses, but in the subjunctive in complementation uses, and iii) applicative and reflexive morphology – usually associated with verbs – is possible on *le*, even when used in complementation contexts. However, we also find a non-agreeing form that contrasts with the agreeing forms in creating noun-y clausal complements. For example, complements headed by the non-agreeing form, unlike those headed by the agreeing forms, can appear in a pre-verbal position that is generally restricted to noun phrases in the language. Based on these (and other) observations, we analyze the agreeing forms of *le* as heads of <v,t> type complements and the non-agreeing form as heads of <e,t> type complements. Our analysis thus supports the claim that the semantic type of CPs varies cross- linguistically: CP complements are not propositional, but rather constitute properties of individuals or properties of eventualities depending on the language (Kratzer 2013, Özyıldız et al. 2018, Moulton 2019, Demirok et al. 2020 a.o.). Kipsigis is also added to a list of languages whose 'say'-based complementizers are analyzed as verbs (Koopman 1984, Koopman & Sportiche 1989 a.o.). Different 'say'-based complementizers with verb-y and noun-y properties have also been described for Zulu (Halpert 2018). It is an open question at this point whether both types are attested in all languages with 'say'- based complementation.

References: Demirok, O., Özyıldız, D., and Öztürk, B. 2020. Complementizers with attitude. In Baird, M. and Pesetsky, J., (eds.), *Proceedings of the NELS 49*. GLSA, Amherst. Diercks, M., van Koppen, M., and Putnam, M. 2020. Agree Probes Down: Anaphoric Feature Valuation and Phase Reference. In Smith, P.W. et al., (eds.), *Agree to Agree: Agreement in the Minimalist Programme*, p. 347-389. Language Science Press, Berlin. Diercks, M., and Rao, M. 2020. Upward-oriented complementizer agreement with subjects and objects in Kipsigis. In Clem, E. et al., (eds.), *Theory and description in African Linguistics: Selected papers from the ACAL*, p. 369-393. Language Science Press, Berlin. Koopman, H**.**, and D. Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry*, p.555-588**.** Kratzer, A. 2013. Modality for the 21st century. In Stephen R. Anderson, J.M. and Reboul, F., (eds.), *L'interface Langage-Cognition/ The Language- Cognition Interface: Actes du 19e Congrès International des Linguistes Genève*, p. 179-199. Librarie Droz.

Definitely factive

Vesela Simeonova
University of Tübingen
vesela-simeonova@uni-tuebingen.de

The relation between factivity and definiteness has been the subject of copious works, starting with 'Fact' by Kiparsky and Kiparsky (1970). The tradition senses a parallel between the clausal and nominal domains and often implements it as factive clauses headed by a covert nominal layer and a determiner. The present paper sympathizes with the intuition that nominals and factive clauses share the property of definiteness, but models this theoretically without a mediating nominal layer or treating factive clauses as noun-modifying clauses (*pace* Kiparsky and Kiparsky, 1970). I show that noun modifying clauses cannot be the solution of the puzzle of 'nouny' factive *vs* less so non-factive clauses because noun modifying clauses exhibit **the same puzzle** (cf. 1-a-ii *vs* 1-b-ii)—an observation that has escaped the literature so far. The parallel and the theoretical solution proposed here are summarized below. The proposal: non- factive verbs and nouns select for CONTENT CLAUSES (cP_{cont}) (in the sense of Moulton, 2009, label de Cuba, 2017 extended here to factive clauses as well), while factive verbs and nouns select for DEFINITE CLAUSES (cP_ι) (building on ideas by Melvold, 1991).

(1) a. **factive domain**:
   (i)   Factive verbs: He regrets/resents/is happy [$_{cP\iota}$ OPι [$_{CP}$ that Edna is a thief]]
   (ii)   Fact(ive) nouns: the fact/realization/regret [$_{cP\iota}$ OPι [$_{CP}$ that Edna is a thief]]
  b. **non-factive domain**:
   (i)   Content verbs: He believes/said [$_{cP cont}$ OP$_{cont}$ [$_{CP}$ that Edna is a thief]]
   (ii)   Content nouns: the idea/rumor/belief [$_{cP cont}$ OP$_{cont}$ [$_{CP}$ that Edna is a thief]]

The cP layer allows differentiating C- and *c*-complementizers. This captures cross-linguistic facts: Greek uses *oti* for content clauses and *pou* for factive ones, while languages like English have an all-purpose complementizer. I propose that Greek-type are *c*-complementizers and English *that* is a C-complementizer. In both types of languages, the *cont/ι* distinction is in the *c* head, not in C.

(2)   Greek: [$_{cP}$ *oti*$_{cont}$/*pou*$_\iota$ [$_{CP}$ Ø [...]]]
(3)   English: [$_{cP}$ Ø$_{cont}$/Ø$_\iota$ [$_{CP}$ that [...]]]

To summarize, this paper upgrades decompositional semantics with a unified account of factivity in the nominal and verbal domains, and addresses the thorny issue of the syntactic size of factive and non-factive clauses and the interpretation of complementizers.

References: de Cuba, Carlos. 2017. Noun complement clauses as referential modifiers. Glossa 2(1). doi:10.5334/gjgl.53. Kiparsky, P. and C. Kiparsky. 1970. Fact. In Progress in linguistics, 143–173. Mouton, The Hague. Melvold, J. 1991. Factivity and definiteness. In MIT WPL, vol. 15. 97–117. Moulton, K. 2009. Natural selection and the syntax of clausal complementation. UMass Dissertation.

From D to N, CPs as nominals in Greek

Richard Faure
Université Côte d'Azur, CNRS, BCL, France
richard.faure@univ-cotedazur.fr

The talk aims to account for the distribution of Classical Greek (CG) *hóti* fi- nite complement clauses (CC), provide fresh arguments in favor of the CP- as-nominal hypothesis (e.g., Baunaz & Lander 2017), arguing that the C is a D itself (but not a probe, as in Angelopoulos 2019), and explain the change from D to N of these clauses between CG and Modern Greek (MG).

Hóti-clauses have syntactic properties supposedly mutually exclusive, thus offering a contradictory picture in two respects. First, they seem to both (A) be low within VP (binding in and A-movement out of them are possible) and (B) extrapose (they always are rightmost in their clause, contrary to argumental DPs); second, like DPs, (C) they can be coordinated with DPs, but unlike DPs, (D) they cannot occupy focus (preverbal, Dik 1995) and subject (Spec,TP) positions.

I first show that CCs are actually *in situ* by means of coordination data and partial topicalization. I then address the question of the satisfaction of the selection of their embedding verbs, which otherwise select for DP. The com- plementary distribution of the C with articles, the association with demonstra- tives, and their agreement properties (showing that they carry φ-features) indicates that *hóti*-clauses are DPs. However, restriction (D) is still to be clari-fied. It points towards an inability for *hóti*-clauses to be case-marked (see Stowell 1981). Although they are DPs of type e and carry φ-features, which makes them suitable for θ-marking *in situ*, they are not allowed to be A-moved to a case-position (Spec, TP or Spec, vP, Chomsky 2001, the latter being both an A and Ā position in CG). Note that a dem. or a DP extracted out of them does not undergo such restrictions and goes to such a position, an operation that involves θ-marked *hóti*-clauses in a chain headed by a case-marked DP, as required in Θ Theory. Elsewhere, a silent expletive is present. Typologically, languages like English that do not have such proxies resort to other repair means like movement (which is moreover needed if they are not DPs, Moulton 2015).

Finally, the status of *hóti*-clauses changed with time: from CG to MG, (*h*)*óti* gained the ability to be nominalized with an article, which made it more N-like than D-like. At the same time, (*h*)*óti*-clauses spread to believe verbs. A feature impoverishment made possible this extension.

References: Angelopoulos N., 2019, Complementizers and Prepositions as Probes: In-sights from Greek, Phd diss. UCLA. Baunaz L.&E. Lander, 2017, "Syncre-tisms with the nominal complementizer", *SLinguistica* 72, p.537-570. Chomsky N., 2001, "Derivation by phase", in M. Kenstowicz, *Ken Hale : A life in language*, MIT Press, p.1-52. Dik H., 1995, *Word order in ancient Greek*, Gieben. Moulton, K. 2015. CPs: Copies and compositionality. Linguistic Inquiry 46.305-42. Stowell T., 1981, Origins of Phrase Structure, PhD dis. MIT.

Argument clauses and definite descriptions

Jürgen Pafel
Universität Stuttgart
juergen.pafel@ling.uni-stuttgart.de

Argument clauses aren't noun phrases, nevertheless they have several interesting similarities with noun phrases. Long time ago, Herling called the complementizer *daß* a sentential article (›Satzartikel‹) in his *Syntax der deutschen Sprache* (1832). Depending on one's syntactic framework, argument clauses and noun phrases are similar to a certain extent. The most startling affinity, however, can be detected looking at the semantics. Argument clauses and several types of noun phrases can be analyzed as definite descriptions. It seems that we can distinguish four types of definite descriptions which are instantiated by noun phrases as well as argument clauses. These four types will be introduced in my talk.

– Type I denotes a maximal plurality
Examples:
  (1) (Joe knows) the politician(s) representing his county.
  (2) (Joe knows) that Mary is a gifted politician.
  (3) (Joe knows) whether Mary is a gifted politician.

– Type II denotes a minimal plurality
Examples:
  (4) The speakers of all factions (met in parliament yesterday). [scope: all>the speakers]
  (5) (We know) which politicians met in parliament yesterday.
  (6) (We know) which politicians each of them met in parliament. [each>which]

– Type III denotes a kind
Examples:
  (7) Dinosaurs (are not mammels).
  (8) (We wonder) which politicians met yesterday in parliament.
  (9) (We know) where to meet a democrat. [mention-some reading]

– Type IV denotes a smallest kind
Examples:
  (10) Politicians from all factions (agreed on a declaration). [all>politicians]
  (11) (We wonder) which politicians each of them met in parliament yesterday.
                                              [each>which]

On the nouniness of V2-clauses under preference predicates

Frank Sode
Goethe University Frankfurt
sode@em.uni-frankfurt.de

This talk addresses embedded V2-clauses under "preference predicates" that typically are marked by subjunctive mood, cf. Frank (1998); Meinunger (2007).

(1) a. *Ich { wollte / wünschte }, ich wäre schon zu Hause.*
    I { want.SUBJ / wish } I be.SUBJ already at home

  b. *Maria wäre froh, sie wäre schon zu Hause.*
    Maria be.SUBJ glad she be.SUBJ already at home

  c. *Es wäre gut, ich wäre schon zu Hause.*
    It be.SUBJ good I be.SUBJ already at home

These embedded V2-clauses dont't really fit into the standard picture of embedded V2-clauses since predicates like "wollen" ('want'), "gut (sein)" ('be good') and "froh (sein)" ('be glad') are not assertive and don't license embedded root phenomena.

  In a first step, I bring together syntactic evidence, Williams (1974); Pesetsky (1991), semantic evidence Heim (1992) and cross-linguistic evidence, Iatridou (2000); von Fintel & Iatridou (2017), that suggest that the predicates in (1) form a natural class across languages: At their core they are evaluative predicates that take conditional clauses as their arguments which share semantic and syntactic properties with nominal arguments. Second, I argue that embedded V2-clauses under preference predicates have the same distribution as "complement fulfilling conditionals" suggesting that they are argument-conditionals.

  The overall consequence of this view is that embedded V2-clauses under preference predicates restrict a hidden conditional operator in the same way as "if"-clauses do. This solves the puzzle of V2-clauses under preference predicates since V2-clauses under preference predicates are not embedded root phenomena and don't pose any threat to the standard picture that relates embedded V2 to assertivity or at-issueness (Antomo (2012)).

References: Frank, N. (1998). Präferenzprädikate und abhängige Verbzweitsätze. Magisterarbeit Universität Tübingen. Arbeitspapiere des SFB 340. Bericht Nr. 128. Heim, I. (1992). Presupposition Projection and the Semantics of Attitude Verbs. JoS, 9.3:183–221. Iatridou, S. (2000). The grammatical ingredients of counterfactuals. LI, 31.2:231-270. Meinunger, A. (2007). In the mood of desire and hope. In Tense, Mood and Aspect, p. 155–176. Rodopi. Pesetsky, D. (1991). Zero syntax. Vol. 2: Infinitives. Ms. Williams, E. S. (1974). Rule Ordering in Syntax. PhD thesis, MIT.

Objects of attitude ascriptions

Patrick D. Elliott
Massachusetts Institute of Technology
pdell@mit.edu

An embedded declarative of the form "that *p*", and a content nominal of the form "the proposition that *p*" are not intersubstitutable *salva veritate* - this is a special case of Prior's (1971) *substitution problem* (see also Moltmann 2003 on the "objectivization effect"). It's tempting to conclude that *syntactic category* is responsible for failure of substitution (see, e.g., Forbes 2018). In previous work (Elliott 2017), I argued that this position is untenable, on the basis of evidence from Moltmann's "special quantifiers". Rather, there is reason to believe that embedded declaratives and content nominals play distinct compositional roles in attitude ascriptions - embedded declaratives are eventuality modifiers, whereas content nominals are *bona fide* arguments, a distinction which I argued cross-cuts the complement/adjunct distinction. In this talk, I reassess the empirical landscape, paying specific attention to cases where declarative clauses appear to take on certain characteristics of nominals, such as clausal subjects, and the relationship between "nouniness" and factivity.

References: Elliott, Patrick D. 2017. *Elements of Clausal Embedding.* Ph.D. thesis, University College London. Forbes, Graeme. 2018. Content and theme in attitude ascriptions. In A. Grzankowski and M. Montague (Eds.), *Non-Propositional Intentionality*, pp. 114–133. Oxford: Oxford University Press. Moltmann, Friedrike. 2003. Propositional attitudes without propositions. Synthese 135(1), 77–118. Moltmann, Friedrike. 2013. *Abstract Objects and the Semantics of Natural Language*. Oxford: Oxford University Press. Prior, Arthur N. 1971. *Objects of Thought.* Oxford, Clarendon Press.

The transfer of nominal (ordinary individual) to propositional (phenomenal individual) properties in German particle verb constructions

Patrick Brandt
Institut für Deutsche Sprache
brandt@ids-mannheim.de

We argue that properties that are presumably nominal in origin get transferred regularly in certain German Particle Verb constructions to properties that are propositional insofar as they concern the internal structuring of eventualities as understood to be described, by and large, by propositional (= truth-assessable) representations.

According to our analysis, the oft-noted perfectivizing function of certain verbal particles like *ein-* in *einfahren* (cf. e.g. Kühnhold 1972) is the effect of redressing a conflict at the syntax-semantics interface: On the one hand, constructions like *in* [*die Grube*]$_{AKK}$ *einfahren* exhibit transitive syntax (Gehrke 2008), requiring that the syntactic arguments are mapped onto well-distinguished or DIFFERENT referents in the semantics. On the other hand, *in/ein* codes a spatiotemporal inclusion relation between its relata, contradicting the requirement imposed by the transitive syntax.

We follow Brandt (2019) in assuming that the interface executes a maneuver that delays the interpretation of part of the contradiction-inducing DIFFERENCE feature. It is not locally interpreted (semantically represented) *in toto* but in part passed on to the next syntactic-semantic computational cycle. Here, the passed-on meaning is interpreted in the locally custom terms: there are times where the state of affairs that defines the Givonian post-state of the depicted eventuality does not hold. No hidden element codes the superficially surprising meaning, nor ambiguity. Instead, part of an actually coded but locally unrealizable semantics in terms of ordinary individuals spills over to the phenomenal domain (using Husserl's term) and yields the interpretive effect observed.

References: Brandt, Patrick (2019): *Discomposition Redressed. Hidden Change, Modality, and Comparison in German*. Tübingen: Narr. Gehrke, Berit (2008): *Ps in motion: on the semantics and syntax of P elements and motion events*. PhD thesis, Utrecht: LOT. Givón, Talmy (1972): Forward implications, backward presuppositions and time axis verbs. In Kimball, John P., editor, *Syntax and Semantics*. New York: Seminar Press, pages 29—50. Husserl, Edmund (1928): *Phänomenologie des inneren Zeitbewusstseins*. In Heidegger, Martin, editor, Edmund Husserls Vorlesungen zur Phänomenologie des Inneren Zeitbewusstseins. Jahrbuch für Philosophie und Phänomenologische Forschung. Halle: Niemeyer. Kühnhold, Ingeburg (1972): Präfixverben. In Moser, Hugo, editor, *Deutsche Wortbildung*. Erster Hauptteil: Das Verb. Düsseldorf: Schwann, pages 141-363.

S-selection and presupposition in quotational complementation

Jan Wiślicki
University of Warsaw
j.wislicki@uw.edu.pl

1 Observations. There are four observations underlying the present proposal.
First, though quotation does not involve TMA and phasal relations, allowing also gibberish as in (1), it is not blind for relations with root verbs, as in (2):

(1)   He didn't say 'He is gwlch'.
(2)   #He asked 'He is smart'

Second, relations with verbs are based on presupposition, rather than standard selection. Though (2) is clearly odd, it is not simply ungrammatical. Rather, it is at odds with the presupposition involved by the verb that quotatotion represents a question. This is supported by the fact that quotation passes the *hey-wait-a- minute* test, e.g. in the context of code-switching quotation:

(3)   A: He asked 'Suzuki ga kita'. [*Suzuki ga kita* is a declarative in Japanese]
      B: Hey, wait a minute, I didn't know that's a question!

Third, quotation enters a relation typical for complementation (with attitude verbs, e.g. *say*, *claim*) or adjunction (with irrealis verbs, e.g. *agree*, *decide*):

(4)   He didn't agree 'He is smart'.

While in (1) the quoted speaker did not utter the quoted string, in (4) he expressed his disagreement by uttering the quoted string. **Fourth**, tenseless verbs (e.g. *avoid*, *finish*) treat direct quotation as purely phonological strings, without entering formal relations with their content:

(5)   He finished 'He is smart'.   [i.e. finished writing, uttering, …]

2 **Analysis.** These effects align with the hierarchy of verbs developed by Wurmbrand & Lohninger (2019). Most of attitude verbs allow quotations, treating them as complements. Irrealis verbs allow quotations less often and combine with them via adjunction. Finally, only some tenseless verbs allow quotation, treating them as strings of symbols. Still, the above effects require a more fine-grained semantic account, which is secured by Cooper's (2005) TTR framework. First, rather than simple types, like e or et, it provides dependent types, e.g. $f(e)$. Second, except single formulas like $\lambda x.dog(x)$, it allows various pieces of information encoded in separate fields within a bigger record of formulas. This goes in hand with the above observations. The root verb *ask* selects not an object of type QUEST, but *presup*(QUEST), presupposing that it is a question; hence the odd, but not ungrammatical, character of (2) and the effect in (3). Tenseless verbs take arguments of type STRING, as shown in (5). Finally, adjunction is interpreted as providing a separate field (not an argument of verb), so that negation in (4) scopes over the verb *agree*, but not quotation.

References: Wurmbrand S. & M. Lohninger. 2019. An implicational universal in complementation–Theoretical insights and empirical progress. Cooper R. 2005. Records and record types in semantic theory.

Arbeitsgruppe 8

Ditransitives across languages and frameworks
Cherlon Ussery, Jóhannes Gísli Jónsson & Nicole Dehé

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Universal patterns in case and agreement alignment

András Bárány
Universität Bielefeld

It has been argued that no language has NOM–ACC case alignment but ERG–ABS agreement alignment (Moravcsik 1978, Bobaljik 2008). In other words, there is no language in which the transitive verb (always) indexes an ACC object rather than a NOM subject. In contrast, languages with ERG–ABS case alignment allow both NOM–ACC agreement (Shipibo, Nepali) or ERG–ABS agreement (Hindi, Tsez); NOM–ACC alignment in both also exists (English, Finnish).

Ditransitives show a similar gap (Haspelmath 2005, 2013, Bárány 2017): **no language has secundative or neutral case-marking and only indirective agreement**. In Hungarian, (1), the single object of a monotransitive (P, ACC; not shown) is coded like the theme (T, ACC) in a ditransitive, while the recipient (R, DAT) is coded differently. The verb indexes P and T, not R.

(1)   Hungarian: Indirective case and **indirective agreement**

    [$_R$ *Neked*   ] *ad-ja*   [$_T$ *a*   *kutyá-t*].
    you.SG.DAT   give-3SG.SBJ>3.OBJ   the dog-ACC
    "S/he gives you the dog."
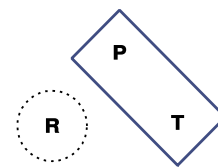
Amharic, (2), has indirective case-marking (like Hungarian) but secundative agreement: the verb indexes the DAT R, not ACC T, while in a monotransitive, the verb indexes the ACC P.

(2)   Amharic: Indirective case and **secundative agreement**

    Ləmma   [$_R$ *l-Almaz*   ] [$_T$ *tarik-u-n*   ] *nəggər-at.*
    Lemma.M   DAT-Almaz.F   story.M-DEF-ACC   tell.3.M.SBJ-3.F.OBJ
    "Lemma told Almaz the story."   (Baker 2012: 261)
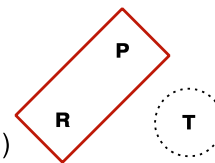
Languages can have secundative case-marking and agreement (e.g. Nez Perce and Ka- laallisut), but among those reported in Dryer (1986), Haspelmath (2005, 2013), Malchukov *et al.* (2010) and Bárány (2017), **no language has secundative case and only indirective agreement**.

Absence of alignment types as a universal. Evidence for this universal claim comes from different sources. **First**, grammars beyond those discussed in the current literature ('grammar-mining'). **Second**, standard assumptions about the syntactic structure of ditransitives, the locality of agreement and a case hierarchy predict the absence of the missing pattern (Bárány 2017), providing a theoretical explanation for the typological gap: in secundative (or neutral) alignment, the verb must be able to index R because of its case. **Third**, there are close parallels between the gaps in monotransitive and ditransitive alignment types. Analogous structural explanations account for both, and analogous apparent counterexamples are found to both generalisations: in so-called 'symmetric' languages (e.g. Bembe, Bantu), either R or T can be indexed on the verb, in seeming violation of the ditransitive generalisation. In languages with inverse agreement, the verb can index the object rather than the subject, in seeming violation of the monotransitive generalisation. I show that these agreement patterns are always options, never the only pos-sibility, and therefore not real counterexamples. In both mono- and ditransitives, they show the sensitivity of agreement to information structure and/or person.

In sum, a broad empirical basis, independently motivated aspects of linguistic theory and analogous syntactic domains provide converging evidence for universals in alignment types.

References: Baker, Mark C. 2012. On the relationship of object agreement and accusative case: Evidence from Amharic. Linguistic Inquiry 43(2). 255–274. https://doi.org/10.1162/LING_a_ 00085. Bárány, András. 2017. Person, case, and agreement: The morphosyntax of inverse agreement and global case splits. Oxford: OUP. Bobaljik, Jonathan David. 2008. Where's phi? In Daniel Harbour, David Adger & Susana Béjar (eds.), Phi theory: Phi-features across modules and interfaces, 295–328. Oxford: OUP. Dryer, Matthew S. 1986. Primary objects, secondary objects, and antidative. Language 62(4).808–845.

https://doi.org/10.2307/415173. Haspelmath, Martin. 2005. Argument marking in ditransitive alignment types. Linguistic Discov- ery 3(1). https://doi.org/10.1349/PS1.1537-0852.A.280.Haspelmath, Martin. 2013. Ditransitive constructions: The verb 'give'. In Matthew S. Dryer & Martin Haspelmath (eds.), WALS. Leipzig. http://wals.info/chapter/105 (29 August, 2019). Malchukov, Andrej L., Martin Haspelmath & Bernard Comrie (eds.). 2010. Studies in ditransitive constructions: A comparative handbook. Berlin: De Gruyter. Moravcsik, Edith A. 1978. On the distribution of ergative and accusative patterns. Lingua 45(3– 4). 233–279. https://doi.org/10.1016/0024-3841(78)90026-8.

Dative case assignment and ditransitives in Lithuanian

Milena Šereikaite˙,
Yale University

Introduction: This study investigates dative case assignment and the structure of ditransitives in Lithuanian. In some languages e.g., Icelandic (Zaenen et al. 1985), dative experiencers behave like grammatical subjects, while in other languages e.g., German, dative DPs behave like non-subjects (Fanselow 2002). I show that this dichotomy between languages is not universal since Lithuanian has two types of datives. The dative indirect object (IO) in distransitives (1) bears a non-structural case: it is retained in the derivation and is not visible for A-movement – a property of an inert case (McGinnis 1998). *Lack* verbs (*trūkti* 'to lack', *reikėti* 'need', etc) (2) have a dative quirky subject, which is also retained in the derivation, but, unlike the inert dative, it is visible for A-movement.

(1) Mama davė vaikui$_i$ obuolį
    mother.NOM gave child.DAT apple.ACC
    jo$_i$ namuose.
    his house
    'Mother gave the child$_i$ an apple in his$_i$ house.'

(2) Jonui$_i$ trūko pinigų savo$_i$
    Jonas.DAT lacked money.GEN self
    reikmėms.
    needs
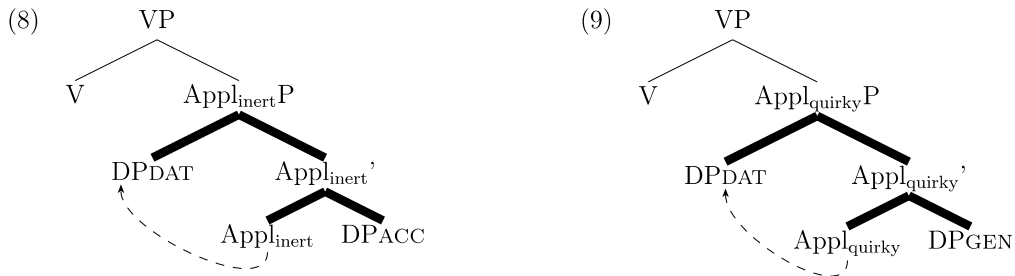   'Jonas lacked money for his own needs.'

To capture this dichotomy, I propose that Lithuanian has two types of low ApplPs which assign two distinct datives: a quirky dative and an inert dative. I further show that ditransitives are not restricted to one type of ApplP. In some ditransitives, the applied argument (IO) in SpecApplP receives a structural accusative case from the verbal phrase as well suggesting that some low ApplPs do not assign case to IOs at all. This study expands McGinnis' (1998) classification of ApplPs showing that low ApplPs can vary within a single language in terms of their case assignment properties.

Two types of datives: Datives in (1-2) are non-structural cases because they are retained in the derivation. **I)** The dative IO is retained in passives, it cannot be nominative (3-4). **II)** In evidentials, a nominative structural subject becomes genitive (Lavine 2005); however, the dative of *lack* is retained (5), it cannot be genitive, which is expected if dative is a non-structural case. Furthermore, the dative DP of *lack* is a subject which has undergone A-movement, whereas the dative IO is invisible of A-movement. **III)** The dative IO retains its original binding relations in the passive: it binds the anti-subject oriented anaphor (3), and thus behaves like the dative object (1). **IV)** The dative DP of *lack* is visible for A-movement: it binds the subject-oriented anaphor *savo*, thus behaves like a subject (2). If XP is relativized in reduced relatives, then that XP is a subject (Poole 2016). **V)** The dative IO cannot be a relativized element (6), whereas the dative of *lack* can at least for some speakers behave like a subject.

(3) Vaikui$_i$ buvo duotas
    child.DAT was give.PASS.NOM
    obuolys jo$_i$/*savo$_i$ namuose.
    apple.NOM his/self house
    'The child$_i$ was given an apple in his$_i$ house.'

(4) *Vaikas buvo duotas
    child.NOM was give.PASS.NOM
    obuolį.
    apple.ACC
    'The child was given an apple.'

(5) Berniuk-ams/*-ų trūkta žinių.
    boys-DAT/-GEN lacked.PASS.N knowledge.GEN
    'The boys must have lacked knowledge.'

(6) vaikams$_i$, *[tėv-ai duodantiems
    children.DAT parents.NOM giving
    t$_i$ klaunus]
    clowns.ACC
    'children giving parents clown toys'

(7) %pirkėjams$_i$, [t$_i$ reikiantiems kapitalo greitai]
    buyers.DAT needing capital.GEN quickly
    'buyers needing capital quickly'(Internet)

Analysis: Lithuanian disallows symmetric passives with ditransitives (3-4) meaning that it lacks high applicatives (Pylkkännen 2008). I propose that ditransitives and lack constructions have a low

ApplP. I rule out the possibility that the dative is assigned by a preposition (cf. Alexiadou et al. 2014) since Lithuanian has no prepositions that assign dative. The two Appl heads in (8-9) assign dative to their applied argument in SpecApplP along with a θ-role. Nevertheless, the two non-structural cases are distinct. The dative of IO in SpecApplP (8) is an inert case in the sense of McGinnis (1998): it isn't visible for A-movement and it doesn't block agreement relation between T and the theme as evidenced by the passive in (3). The dative in (9) is also assigned thematically like an non-structural case, but DPs with this case are quirky subjects visible for A-movement.

(8)

$$\begin{array}{c} \text{VP} \\ \diagup \diagdown \\ \text{V} \quad \text{Appl}_{inert}\text{P} \\ \diagup \diagdown \\ \text{DP}_{DAT} \quad \text{Appl}_{inert}' \\ \diagup \diagdown \\ \text{Appl}_{inert} \quad \text{DP}_{ACC} \end{array}$$

(9)

$$\begin{array}{c} \text{VP} \\ \diagup \diagdown \\ \text{V} \quad \text{Appl}_{quirky}\text{P} \\ \diagup \diagdown \\ \text{DP}_{DAT} \quad \text{Appl}_{quirky}' \\ \diagup \diagdown \\ \text{Appl}_{quirky} \quad \text{DP}_{GEN} \end{array}$$

Two types of IOs: In addition to the two types of low ApplPs in (8-9), I propose that Lithuanian has a third type of ApplP which does not assign dative case to its applied argument. This is evidenced by ditransitives in (10) which take an accusative IO followed by an instrumental theme. The accusative IO bears a structural case. **I)** The accusative IO becomes genitive under negation, which is a property of DPs assigned structural object case, (Sigurðsson&Šereikaitė 2020). The dative IO cannot become genitive when negation is present, which is predicted if a DP bears a non-structural case. **II)** Unlike the dative IO (3-4), the accusative IO advances to nominative in the passive and becomes a grammatical subject in that it binds the subject-oriented anaphor *savo* (13). Thus, ditransitives like (10) contain an ApplP whose applied argument IO receives a structural accusative case from v.

(10) Draugai apkaltino Joną
friends.NOM blamed Jonas.**ACC**
vagyste.
robbery.INS
'Friends blamed Jonas for robbery.'

(11) Draugai neapkaltino Jon-o/*-ą
friends.NOM NEG.blamed Jonas-GEN/-ACC
vagyste.
robbery.INS
'Friends didn't blame Jonas for robbery.'

(12) Mama nedavė vaik-ui/*-o obuolio.
mother.NOM NEG.give.PST.3 child-DAT/-GEN apple.GEN
'Mother didn't give the child an apple.'

(13) Jonas$_i$ buvo apkaltintas
Jonas.**NOM** was blame.PASS.NOM
vagyste dėl savo$_i$ kaltės.
robbery.INS because self fault
'Jonas was blamed for the robbery because of his own fault.'

(14)

$$\begin{array}{c} \text{VP} \\ \diagup \diagdown \\ \text{V} \quad \text{ApplP} \\ \diagup \diagdown \\ \text{DP}_{ACC} \quad \text{Appl'} \\ \diagup \diagdown \\ \text{Appl} \quad \text{DP}_{INS} \end{array}$$

References: Fanselow 2002. Quirky Subjects and other specifiers. Lavine 2006. Is there a passive evidential strategy in Lithuanians? McGinnis 1998. Locality and inert case. Poole 2016. Deconstructing quirky subjects.

German double-accusative verbs: Different solutions for avoiding a marked construction

Gabriele Diewald, Vera Lee-Schoenfeld, Maud Kelly
Leibniz Universität Hannover, University of Georgia, University of Georgia
gabriele.diewald@germanistik.uni-hannover.de, vleesch@uga.edu, maud.m.kelly@gmail.com

As discussed in Lang 2007 and Lee-Schoenfeld & Diewald (L-S&D) 2017, the ditransitive verb *lehren* 'teach' is undergoing a change from the exceptional double-accusative (ACC>ACC) pattern to the prototypical ditransitive dative-accusative (DAT>ACC) pattern. This is evident from its acceptable use in the *kriegen*-passive construction, which targets DAT arguments:

(1) *Er   kriegt   den Seiltrick   gelehrt.*
     he.NOM   gets   the.ACC   taught
     'He's getting taught the rope trick.'
     [L-S&D 2017: 1]

In the active equivalent of (1), the person being taught (*er* 'he') must be DAT-marked (*ihm* 'him'). Here, we extend L-S&D's corpus investigation to the four other ACC>ACC verbs: *abfragen/abhören* 'test, quiz', *kosten* 'cost', and *fragen* 'ask'.

The hypothesis is that each verb follows its own individual path to overriding the highly marked ACC>ACC pattern, with the latest usage data revealing notably different results as to the verbs' most typical syntactic patterns, meaning variants, and contextual features. Specifically, we propose that this small group of verbs makes use of three different "strategies" for avoiding the ACC>ACC pattern: (i) change of major valency frame from ditransitive to monotransitive pattern (*abhören/abfragen*), (ii) limiting the second object to primarily a clausal or prepositional one (*fragen*), and (iii) semantic diversification / polysemy combined with different preferences as to the valency pattern per meaning (*kosten*). We aim to confirm this hypothesis by comparison of the verbs' usage patterns at the beginning of the 20th century and at the beginning of the 21st century (via corpus analyses using DWDS). Our synchronic corpus search (of deWaC and DeTenTen), initiated by consultation of the Duden (1995, 2006, 2016), is complete and has yielded the results and analysis laid out in what follows.

To give a broad overview of our findings, *abfragen/abhören*, though mainly used mono-transitively now, come close to patterning like *lehren* in that they allow for the DAT>ACC pattern:

(2) *Ich weigere mich auch, ihr   die Vokabeln   zum   5. Mal   abzuhören…*
     I   refuse   REFL also her.DAT   the.ACC vocabulary   for-the 5th time quiz
     'I also refuse to quiz her on the vocabulary for the fifth time…'
     [From grundschultreff.de via German Web Corpus (DeTenTen) 2013]

The verb *kosten*, on the other hand, patterns more robustly as ACC>ACC (see (3) and (4)), and it disallows passivization. It does exhibit DAT>ACC when particular affectedness is expressed (see (5)).

(3) Number of ACC>ACC vs. DAT>ACC constructions with *kosten* in a partial search of SpiderLing's German Web Corpus (deWaC) 2009:

Total ditransitives	ACC>ACC	DAT>ACC	Ambiguous ditransitives (ACC>ACC or DAT>ACC)
137	77 (56.2%)	12 (8.76%)	48 (35.03%)

(4) *Siebenhundert Galleonen   haben mich   die   Karten   gekostet.*
     seven-hundred   galleons.ACC have me.ACC the.NOM tickets cost
     'Seven hundred galleons is what the tickets cost me.'
     [From akbi.de via German Web Corpus (DeTenTen) 2013]

(5) *Die   letzte Runde   hätte   <u>ihm</u>   beinahe   <u>das Leben</u>   gekostet*.
the.NOM last round   would-have   him.DAT   nearly   the.ACClife   cost
'The last round nearly cost him his life / caused him to lose his life.'
[From swnation.de via German Web Corpus (DeTenTen) 2013]

The pattern of *fragen* is again different, with no DAT-object being allowed and the second object being limited to a PP, clause, or indefinite generic:

(6) *Der   Ranger fragt <u>ihn</u>   noch so <u>Einiges</u>   und…*
the.NOM ranger asks   him.ACC   still   like things.ACC and
'The ranger was still asking him this and that and…'
[From womo-abenteuer.de via German Web Corpus (DeTenTen) 2013]

In line with L-S&D, we analyze the first, animate object of ditransitive *lehren, abfragen*, and *abhören* as structurally ACC-marked and the second, inanimate object as lexically (idiosyncratically) case-marked. The first object corresponds to the indirect (DAT) object of the normal ditransitive pattern, typically a Recipient, and the second object was originally an "inner object" (already included in the meaning of the verb itself) but can often be interpreted as a more Patient-like (ACC) object. This causes the prototypical ditransitive schema, a scene of transfer, with an inherently DAT-marked Recipient/Source (Affectee) and a structurally ACC-marked Patient/Theme, to become available: for *lehren*, 'give someone(DAT) something(ACC) to learn', and for *abfragen* and *abhören*, 'elicit from someone(DAT) some(ACC) info'.

Unlike *lehren*, *abfragen* and *abhören* are hardly used ditransitively anymore. The newer, monotransitive meaning of *abfragen* is 'request, check' in the context of a computer query, and that of *abhören* is 'spy on, listen in' in the context of bugging a phone line or a doctor using a stethoscope. In three of the few ditransitive examples of *abhören* we found (9 out of 7,992 attestations of the verb, i.e., 0.112%), the animate object was ACC-marked; in another three, it was a name or the pronoun *uns* 'us', which can be ACC or DAT; and in the remaining three, it was DAT-marked. This is evidence of a *lehren*-like development from ACC>ACC to DAT>ACC (again, see (2)). The Duden (2016) mentions this ditransitive development but says nothing about the monotransitive meanings.

Kosten* can be shown to be polysemous, whereby the three meanings relevant here, namely 'be worth a certain amount', 'require effort by someone', and 'cause someone to lose something', tend to be expressed by distinct syntactic patterns. Given the two more eventive meanings of the verb rather than the stative 'be worth a certain amount' in connection with a particularly gravely affected animate object, this animate object can again be interpreted as Affectee, i.e., be DAT-marked, as shown in (5). According to Burzio's (1986) Generalization, if the cost-incurring participant isn't analyzable as an external argument, like *die Karten* in (4), the impossibility of passivization is unsurprising. If it is analyzable as a cause and thus external argument, like *die letzte Runde* in (5), on the other hand, we expect passivization to be possible. We speculate that it isn't (yet) either because the eventive meaning of *kosten* is still establishing itself or because the construction is semi-frozen in that it necessarily involves inalienable possession (*Leben* 'life', *Verstand* 'mind', *Freiheit* 'freedom', etc.).

Finally, *fragen* sets itself apart from the other verbs in that no DAT-marking of the animate object is allowed. Our explanation here is straightforward. It's the person being asked that is interpreted as Patient, not the clausal, prepositional, or indefinite generic second object (again, see (6)), so no scene-of-transfer interpretation involving Affectee and Patient/Theme becomes available, and thus no development from ACC>ACC to DAT>ACC is expected.

References: Burzio, Luigi. 1986. Italian Syntax: A Government-Binding Approach. Dordrecht: Reidel. Duden. 1995. Drosdowski, Günther (ed.) *Grammatik der deutschen Gegenwarts-sprache*, Volume 4. 5th edition. Mannheim: Dudenverlag. Duden. 2006. *Die Grammatik.* Der Duden in zwölf Bänden, Volume 4. 7th edition. Mannheim: Bibliographisches Institut. Duden. 2016. *Die Grammatik.* Der Duden in zwölf Bänden, Volume 4. 9th edition. Mannheim: Bibliographisches Institut. Lang, Peter. 2007. Grammatik und Norm: Direktes Objekt, indirektes Objekt und der doppelte Akkusativ. Ms. Universität Zürich. Lee-Schoenfeld, Vera, and

Gabriele Diewald. 2017. Passivization possibilities in double-accusative constructions. *Proceedings of the Linguistic Society of America* 2, 9:1-14. http://dx.doi.org/10.3765/plsa.v2i0.4050.

Nominalizations of ditransitives in Icelandic

Jim Wood

Yale University jim.wood@yale.edu

According to the applicative theory of ditransitives (Pylkkänen 2002, Cuervo 2003, et seq.), ditransitive verbs do not, contrary to appearances, select two internal argument DPs. Instead, such verbs select one argument, and the other internal argument is introduced by a special Appl(icative) head. For the purposes of this talk, I assume that a 'verb' is really a category- neutral root adjoined to a little v head, and that Icelandic applicative constructions are uniformly low applicatives. Another way that a verb can appear to select two internal argument DPs is if one of the DPs is actually introduced by a silent preposition.

(1)   a. $[_{vP} [_{v} \text{√ROOT v}] [_{ApplP} DP_{\text{indirect argument}} [_{Appl'} \text{Appl } DP_{\text{direct argument}}]]]$

      b. $[_{vP} [_{v} \text{√ROOT v}] [_{PP} DP_{\text{direct argument}} [_{P'} \text{P } DP_{\text{indirect argument}}]]]$

Either way, at most one DP is actually the *direct argument* of the verb, and any other DP internal argument present is related to the verb only indirectly; I will refer to such arguments as *indirect arguments*. In this talk, I will argue that nominalizations in Icelandic can diagnose the status of an argument as *direct* or *indirect,* and discuss some conclusions that this leads us to in the analysis of non-obligatory arguments, using the verb *kenna* 'teach' as a case study.

      When a ditransitive verb is nominalized, only the direct argument can be expressed as a genitive DP. For example, consider (2) (adapted from Jóhannsdóttir 1995:65).

(2)   Pósturinn   afhenti   frúnni   pakkann.
      postman.the.NOM   delivered   lady.the.DAT   package.the.ACC
      'The postman delivered the lady the package.'

      a.   *   afhend-ing   frúarinnar   b.   afhend-ing   pakkans
             deliver-NMLZ   lady.the.GEN   deliver-NMLZ   package.the.GEN
                                            'the delivery of the package'

By hypothesis, the dative DP *frúnni* is introduced in the specifier of an Appl head, and it is typically assumed that the Appl head assigns dative to its specifier (Cuervo 2003, McFadden 2004, Wood 2015, E.F. Sigurðsson 2017); it is an indirect argument of the verb. Thus, it cannot correspond to the genitive of the nominalization in (2a). The accusative DP *pakkann*, however, is a direct argument of the verb, not introduced by any Appl head or silent P. Thus, it can correspond to the genitive of the nominalization in (2b). Consistent with this conclusion, the dative argument of *afhenda* 'deliver' is optional, while the accusative argument is obligatory.

(2)   c.   *   Pósturinn   afhenti   frúnni.
             postman.the.NOM   delivered   lady.the.DAT
             INTENDED: 'The postman delivered to the lady.'

     d.   Pósturinn   afhenti   pakkann.
             postman.the.NOM   delivered   package.the.ACC
             'The postman delivered the package.'

The verb *kenna* 'teach' may also appear to take two internal arguments. By hypothesis, one would assume that the verb phrase in this case has the same basic structure as with *afhenda*, with the dative being introduced in the specifier of an Appl head.

(3)   María   kenndi   börnunum   tungumálið.
      Mary.NOM   taught   children.the.DAT   language.the.ACC
      'Mary taught the children the language.'

However, with the nominalization of *kenna*, *kennsla*, the genitive can correspond to either the dative or the accusative argument.

(4) a. kenn-sla   barnanna   b. kenn-sla   tungumálsins
       teach-NMLZ   children.the.GEN   teach-NMLZ   language.the.GEN
       'the teaching of the children'   'the teaching of the language'

Given the reasoning above, this suggests that the dative argument of *kenna* can be a direct argument. This conclusion is supported by the fact that either the dative or the accusative can appear as the lone argument -- neither is obligatory.

(5) María   kenndi tungumálið.   (6) María   kenndi börnunum.
     Mary.NOM   taught language.the.ACC   Mary.NOM   taught children.the.DAT
     'Mary taught the language.'   'Mary taught the children.'

Presumably, 'the children' is an applied argument in when there are two internal arguments as in (3), but a direct argument (e.g. complement) of the verb in (6), and it is the structure in (6) that explains the availability of (4a).

Perhaps somewhat surprisingly, however, even when 'the children' is the direct argument, it is still marked with dative case. This suggests that the verb root √KENNA may occur with a dative assigning v_{DAT} head (Svenonius 2006; H.Á. Sigurðsson 2012; Wood 2015; E.F. Sigurðsson 2017). But then we seem to end up with a strange picture of this verb: √KENNA attaches to ordinary v, and the 'learner' is optionally introduced by an Appl head that assigns it dative, or √KENNA attaches to v_{DAT}, and the 'learner' is introduced as a complement. Instead, I suggest that √KENNA attaches to v_{DAT} in all cases, and that the dative case in (3), like (6), comes from v_{DAT} and not from an Appl head; v_{DAT} simply assigns dative case to the nearest DP, which is the applied argument in (3) and the complement in (6). This would imply that (5) is deceiving: the dative is not actually missing, it is just silent. Returning to nominalizations, (4a) is possible because 'the children' can be a direct argument of the verb when it is a sole complement, and (4b) is possible because 'the language' is the direct argument of the verb when the verb takes an ApplP complement. Support for this comes from sentences like (7). There, the direct object learner is still dative, but it is unlikely to be the specifier of an ApplP because the object is a PP rather than a DP.

(7) Ég   skal kenna þér   á þvottavélina.   (example from snara.is)
     I.NOM will teach   you.DAT on washing.machine.the.ACC
     'I will teach you how to use the washing machine.'

The analysis implies that Appl does not always assign dative to its specifier, and moreover that even when we see ditransitives with a NOM-DAT-ACC case frame, we cannot necessarily assume that the dative comes from an Appl head.

References: McFadden, Thomas. 2004. The position of morphological case in the derivation: A study on the syntax-morphology interface. Doctoral dissertation, University of Pennsylvania. Sigurðsson, Einar Freyr. 2017. Deriving case, agreement and voice phenomena in syntax. Doctoral dissertation, University of Pennsylvania. Sigurðsson, Halldór Ármann. 2012. Minimalist C/case. *Linguistic Inquiry* 43:191–227. Svenonius, Peter. 2006. Case alternations and the Icelandic passive and middle. Ms., University of Tromsø. https://ling.auf.net/lingbuzz/000124. Wood, Jim. 2015. *Icelandic morphosyntax and argument structure.* Dordrecht: Springer.

Why no double objective construction in Shupamem

Xiaomeng Ma
The Graduate Center, CUNY
xma3@gradcenter.cuny.edu

Introduction: Shupamem (Grassfields Bantu, Cameroon; SVO) doesn't allow double object construction (DOC). Only prepositional dative construction (PDC) is available. Recent studies have concluded that PDC and DOC are not related to each other derivationally (see, Rappaport Hovav and Levin, 2008), although the underlying structures are still under debate. The most widely accepted theories can be divided into two categories: the semantic decomposition approach, which proposes that the structures are derived from the se- mantic decomposition of the verb 'give' (Harley, 2002), and the applicative approach, which argues that the indirect object in DOC is introduced by an applicative head (e.g. Marantz, 1993; Bruening, 2010). This study applies both approaches to explain why Shupamem doesn't allow DOC.

The Semantic Decomposition Approach: Harley (2002) revised Pesetsky (1995) and pro- posed two structures for PDC and DOC. The verb 'give' is decomposed into a CAUSE component and an abstract preposition head either encodes location (P_{LOC}), with direct object (DO) as its specifier, or encodes possession (P_{HAVE}), with indirect object (IO) as its specifier. Based on these structures, Harley (2002) hypothesized that if a language doesn't have P_{HAVE}, it doesn't allow DOC. Furthermore, she made three predictions about languages without P_{HAVE}: (a) DOC is not allowed; (b) There is no verbal 'have' to express possession; In a sentence that denotes possession, the possessor does not always c-command the possessee. This study shows that Shupamem data follows all the predictions.

The Applicative Approach: Marantz (1993) proposed two structures for PDC and DOC based on the Voice theory. In DOC, the indirect object (IO) is introduced by an applicative head that takes the VP as complement. Marantz (1993) didn't make any hypothesis about why languages don't allow DOC. In order to derive the word order for DOC, the applicative head has to be empty or null. It's reasonable to assume that a language doesn't allow DOC because it doesn't have an empty or null applicative head. Unlike other Bantu languages, Shupamem doesn't have a typical applicative construction, since most of the applicatives are introduced by a preposition, such as the instrumental and the locative applicative. In addition, the benafactive applicative is expressed through a serial verb construction with the phrase 'fá nə' ('give to').Therefore, the lack of empty or null applicative head could account for why DOC is not allowed in Shupamem.

References: Bruening, B. (2010).Double object constructions disguised as prepositional datives. Linguistic inquiry, 41(2):287–305. Freeze, R. (1992). Existentials and other locatives. Language, pages 553–595. Guéron, J. (1995). On have and be. In Proceedings-NELS, volume 25, pages 191–206. University of Massachusetts. Harley, H. (2002). Possession and the double object construction. Linguistic variation yearbook, 2(1):31–70. Kayne, R. S. (1993). Toward a modular theory of auxiliary selection. Studia linguistica, 47(1):3–31. Kittilä, S. (2005). Recipient-prominence vs. beneficiary-prominence. Walter de Gruyter. Marantz, A. (1993). Implications of asymmetries in double object constructions. In Theoretical aspects of Bantu grammar, pages 113–150. CSLI. Nchare, A. L. (2012). The grammar of Shupamem. New York University. Pesetsky, D. M. (1995). Zero syntax: Experiencers and cascades. Number 27. MIT press. Rappaport Hovav, M. and Levin, B. (2008). The English dative alternation: The case for verb sensitivityl. Journal of linguistics, 44(1):129–167.

Scope in Faroese ditransitives

Cherlon Ussery, Hjalmar P. Petersen

Carleton College, University of the Faroe Islands

cussery@carleton.edu, Hjalmar@setur.fo

This talk examines quantifier scope in double object (DO) and prepositional (PP) constructions in Faroese. Work in this area is of particular interest because Faroese is undergoing a shift in which the PP construction is increasingly available alongside DO sentences (Henriksen 2000; Petersen 2010; Ussery & Petersen, to appear). Further, with the exception of Sandøy (1992, 2014) and Lindstad (2009), little has been documented about Faroese quantifiers, with the most robust set of examples found in Petersen (2020). We illustrate that both the syntactic structure and the ordering of quantifiers affects the interpretation: DO constructions favor surface scope more so than PP constructions do and the *hvør…ein* 'every…a' ordering is more amenable to inverse scope than the *ein…hvør* 'a…every' ordering is. These observations are consistent with both theoretical and experimental work conducted on other languages.

Bruening (2001, 2010a) argues that the asymmetry in scope interpretations in English ditransitives is attributed to differences in the syntactic structure. In the DO construction, the theme direct object is an argument of the verb while the goal indirect object is an argument of the Appl head, which is higher in the structure, as shown in (1a). As such, the indirect object moves first in scope-taking operations, delivering surface scope. In the PP construction, both objects are arguments of the preposition and are, consequently, equidistant to higher positions, as shown in (1b). Either object can move in scope-taking operations, delivering ambiguous scope. This proposal accounts for the contrast in the *ein…hvør* 'a…every' Faroese examples listed in (2a) and (2b).

However, as is well-known, the ordering of the quantifiers also affects scope interpretations, and we show that this is also the case in Faroese. Inverse scope is more readily available in the *hvør…ein* 'every...a' DO construction in (3a), than in the *ein…hvør* 'a...every' DO construction in (2a). Through a series of experiments which examined scope in transitive constructions in English, Anderson (2004) reports that while speakers have an overall preference for surface scope, this preference is stronger in *a...every* sentences. Since in *every...a* sentences, the inverse scope interpretation is a subset of the surface scope, speakers may not tease the two readings apart. In the Faroese examples in (3a) and (3b), it could be coincidental that every coworker is lent the same screwdriver, for instance. The effect of quantifier order in Faroese is corroborated by the fact that inverse scope is more available in the *hvør…ein* 'every...a' PP construction in (3b) than in the *ein…hvør* 'a...every' PP construction in (2b).
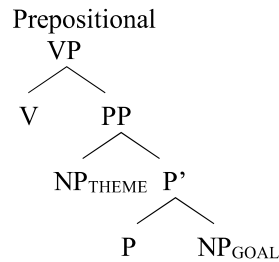
The DO construction in (3a), however, illustrates that the syntactic structure does indeed privilege certain interpretations. Even though the *hvør…ein* 'every...a' ordering allows for ambiguity in a way that the *ein…hvør* 'a...every' DO construction in (2a) does not, surface scope is still preferred in (3a). Scope in Faroese is, therefore, computed via an interaction of syntactic structure and quantifier order. Surface scope is most available in *ein…hvør* 'a...every' DO constructions, while inverse scope is most available in *hvør…ein* 'every...a' PP constructions.

Examples

(1)

| Double Object | (1b) | Prepositional |

Double Object   (1b)   Prepositional

ApplP   VP

NP_GOAL   Appl'   V   PP

      Appl   VP   NP_THEME   P'

           V   NP_THEME   P   NP_GOAL
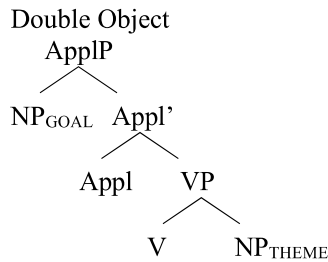
(Bruening 2010a, EX 6-7)

(2a)   **DO: *ein…hvør* 'a...every'**
       Timburmaðurin lænti einum   starvsfelaga hvørt (einasta) skrúvublað.
       the carpenter   lent   a   coworker   every (single) screwdriver
               Surface: a > every

(2b)   **PP: *ein…hvør* 'a...every'**
       Timburmaðurin lænti eitt skrúvublað til (ein og)   hvønn starvsfelaga.
       the carpenter   lent   a   screwdriver to (each and) every   coworker
           Ambiguous: a > every/every > a

(3a)   **DO: *hvør…ein* 'every...a'**
       Timburmaðurin lænti hvørjum starvsfelaga eitt skrúvublað.
       the carpenter   lent   every   coworker   a   screwdriver
          Ambiguous: every > a/ a > every (preference for surface)

(3b)   **PP: *hvør…ein* 'every...a'**
       Timburmaðurin lænti hvørt (tað einasta) skrúvublað til ein starvsfelaga.
       the carpenter   lent   every (single)   screwdriver to a   coworker
          Ambiguous: every > a/ a > every (preference for inverse)

References: Anderson, Catherine. 2004. *The structure and real-time comprehension of quantifier scope ambiguity.* Doctoral Dissertation. Evanston: Northwestern University. Bruening, Benjamin. 2001. "QR obeys superiority: frozen scope and ACD." *LI* 32: 233-274. Bruening, Benjamin. 2010a. "Double object constructions disguised as prepositional datives." *LI* 41: 287- 305. Henriksen, Jeffrei. 2000. *Orðalagslæra*. Vestmanna: Sprotin. Lindstad, Arne Martinus. 2009. "On the negative polarity sensitive indefinite determiner *nakar* 'any' in Faroese." *Nordlyd* 36(2): 208-230. Sandøy, Helge. 1992. "Indefinite pronouns in Faroese." Í J. Louis-Jensen & J. H. W. Poulsen (ritstj.) *The Nordic Languages and Modern Linguistics*, bls. 547-554. Føroya Fróðskaparfelag: Tórshavn. Sandøy, Helge. 2014. "Somme har nok høyrt eitkvart om kvantoren nokon før." In: Janne Bondi Johannessen & Kristin Hagen (red.), Språk i Norge og nabolanda. Ny forskning om talespråk, s. 261–289, Novus 2014. ISBN978-82-7099-795-4. Petersen, Hjalmar P. 2010. *The Dynamics of Faroese-Danish Language Contact*. Winter: Heidelberg. Petersen, Hjalmar P. 2020. *Grunddrøg í føroyskari syntaks*. Nám: Tórshavn. Ussery, Cherlon and Hjalmar P. Petersen. To appear. Ditransitives in Faroese: The Distribution of IO/DO and PP. Submitted to T. Colleman, M. Röthlisberger, and E. Zehentner eds. *Ditransitive Constructions in the Germanic Languages*. Amsterdam: John Benjamins.

[Person] intervention effects with Romanian ditransitive constructions

Alina Tigău & Klaus von Heusinger
University of Bucharest, University of Cologne
alina.tigau@lls.unibuc.ro, kheusing@uni-koeln.de

1.   Introduction This paper discusses Romanian (Rom) ditransitive constructions (DC) focusing on some intervention effects arising when a differentially marked direct object (DOMed DO) co-occurs with an indirect object (IO). Our study rests on three grammaticality judgement experiments involving 480 Romanian natives and testing quantificational binding relations between the two internal arguments. One of the most comprehensive studies on Rom DCs is Diaconescu & Rivero (2007)`s alternative projection account. They argue that the two interpretations of *give*-verbs, *caused movement* and *caused possession* mirror configurations (1) and (2) respectively:

(1)   *Theme c-commands Goal*: $[_{VoiceP} DP_{Agent} Voice[_{vP} v [_{PP} DP_{Theme} P DP_{Goal}]]]$

(2)   *Goal c-commands Theme:* $[_{VoiceP} DP_{Agent} Voice[_{vP} v [_{ApplP} DP_{Goal} [cl_{Appl}] [_{VP} V DP_{Theme}]]]]$

In (1) the dative is a PP argument c-commanded by the Theme while in (2) the Goal is introduced by a low Appl°, it is interpreted as a Possessor and it c-commands the Theme, determining the asymmetries in Barss&Lasnik (1987) for the English Double Object Construction (DOC). D&R claim that the DOC properties obtain *only if the Goal is clitic doubled (CD)* with Appl° spelling out as a dative *clitic*. In order to test these predictions, we conducted a series of three grammaticality judgement tasks manipulating i) the surface order of DO and IO; ii) direction of binding between DO and IO; iii) the presence of a dative clitic doubling IO and iv) the DO form (unmarked vs. DOMed vs. CDed+DOMed). As the first three parameters already yield a 2x2x2 design, we decided to design three similar questionnaires with 32 test items each and which differed through the DO type used. The results (in 3) show that: a) binding dependencies *do not depend on CD*; b) the two internal arguments show *symmetric c- command;* c) surface word order is an important factor for acceptability: items where the surface word order matches binding directions obtain higher scores than instances where reverse binding obtains. d)The presence of a dative clitic doubling the IO significantly lowers the acceptability of the items in comparison to their undoubled counterparts. e) Finally, sequences where a DOMed DO co-occurs with a CDed IO are highly problematic for surface word order *DO precedes IO*. The analysis in (1) & (2) is thus severely *incomplete*, excluding many grammatical patterns. These findings suggest that the alternative projection account in its present form cannot be maintained.

(3)   Mean values for acceptability judgements of a Likert scale (1 very bad – 7 very good) for DCs showing quantificational binding dependencies between the two internal arguments

	DO before IO				IO before DO			
	DO binds into IO		IO binds into DO		DO binds into IO		IO binds into DO	
	-cl	+cl	-cl	+cl	-cl	+cl	-cl	+cl
Unmarked DO	4.57	3.64	3.08	3.00	3.40	3.22	4.56	4.35
DOMed DO	4.43	**2.64**	3.05	**2.36**	3.36	3.27	4.32	3.68
CDed+DOMed DO	4,51	3,52	3,47	3,18	3,67	3,42	5,52	3,51

2.   Aim The paper has two aims: a) to provide a derivational analysis of Romanian DCs; b) to account for the difference in the evaluation of the DOMed DO > cl-IO pattern. The co- occurrence of DOMed DOs with CDed IOs was granted very low acceptability scores by the respondents in both directions of binding for the *DO before IO* surface order (4). This effect was not noticed with the counterparts of these patterns in the experiments featuring unmarked DOs (5) and CDed+DOMed DOs, which prompt us to hypothesize that the lower acceptability has to do with the internal structure of the DPs involved.

(4) **DOMed DO + CDed IO (low acceptability)**

   DO before IO; DO binds into IO; +cl

Poliția   *le*-a   înapoiat   *pe fiecare copil$_i$*   pierdut pe plajă *părinților lui$_i$*   *diperați.*
Police.the them.Dat-has returned *pe* every   child   lost on beach   parents.Dat his desperate
´The police returned every child lost on the beach to his desperate parents.´

   DO before IO; IO binds into DO; +cl

Poliția   *i*-a   înapoiat   *pe copilul său$_i$* pierdut peplajă   *fiecărui părinte$_i$,*
după îndelungi căutări
Police.the him.Dat-has returned   *pe* child.the his lost on beach   every.Dat parent
after long search
'The police returned his child lost on the beach to every parent, after a long search.´

(5) **Unmarked DO + CDed IO (high acceptability)**

   DO before IO; DO binds into IO; + cl

Editorii   i-au   trimis *fiecare carte$_i$ autorului   ei$_i$* pentru corecturile finale.
Editors.the him.Dat-have   sent   every book   author.Dat. its for corrections final
'The editors sent every book to its author for final corrections.'

   DO before IO; IO binds into DO; + cl

Editorii   i-au   trimis *cartea   sa$_i$ fiecărui   autor$_i$* pentru corecturile finale.
Editors.the him.Dat.-have sent   book.the his   every.Dat author for corrections final
'The editors sent his book to every author for final corrections.'

3.   A derivational account. Given the experimental findings, we argue in favour of a derivational account for DCs. The symmetric binding potential of the two internal arguments obtains as a consequence of their relative hierarchical order in the VP (6) combined with subsequent movement for reasons of case assignment and [Person] checking (the feature valuation system used is that from Pesetsky & Torrego 2007)

(6)   ApplP/ αP

                    Appl/ α

      Appl^0/ α VP

                  IO   V′

                          V   DO
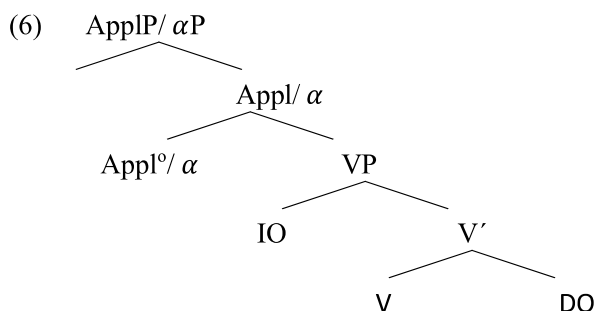
We further posit some priority criteria with respect to feature valuation between the two objects: DO has general priority over IO, but this may change function of the feature specification of the two objects. The proposed system allows us to account for all the patterns assessed as acceptable, and to explain the problematic cases where a DOMed DO interacts with a CDed IO. In the latter case, the analysis draws on the internal featural make-up of the two internal arguments and shows that the problem amounts to a locality issue: DOMed DOs carry [iPerson] and only need to check case. The CDed IO needs to check both case and its [iPerson:] feature. Since IO has more features to verify it gains priority over DO (closer proximity to Appl is also important). The IO enters an Agreement relation with Appl^0 (specified as [uPerson: val]) and checks both case and [iPers:]. The [uPerson: val] feature of Appl is EPP and the IO moves to SpecAppP. As such, it acts as an intervener for the DO, which may no longer move to a Spec of Appl to get its case feature valued by *v*.

References: Barss, A. and H. Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17: 347–35. Diaconescu, C. & Rivero M. L. 2007. An Applicative Analysis of DOC in Romanian. *Probus*. 19 (2): 209-23. Pesetsky, David and Torrego, Esther. 2007. The Syntax of Valuation and the Interpretability of Features. In S. Karimi, V. Samiian, W.K. Wilkins, eds. *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*. Amsterdam: John Benjamins. 262-194.

Verb class and differential object marking in Spanish ditransitive constructions

Klaus von Heusinger, Marco García García & Diego Romero Heredero
Universität zu Köln
Klaus.vonHeusinger@uni-koeln.de; marco.garcia@uni-koeln.de; d.romeroheredero@uni- koeln.de

In Spanish, a human definite direct object in a transitive construction must be marked by the differential object marker (DOM) *a* (which merges with the definite article to become *al*); see (1) (Pensado 1995, Fábregas 2013). However, for ditransitive sentences, the literature (Comrie 2013, Real Academia Española 2009) assumes that *a*-marking is generally blocked or disfavoured by the concurrence of an indirect object noun phrase. The latter is obligatorily marked by the dative marker *a*, which is homonymous with the direct object marker; see (2). Blocking effects for DOM in ditransitive contexts seem to be influenced by different factors, such as clitic doubling of the indirect object; see (3). However, it is controversial whether clitic doubling strengthens the blocking effect for DOM (Rodríguez-Mondoñedo 2007, Fábregas 2013) or mitigates it (Company Company 2001). Moreover, Rodríguez-Mondoñedo (2007: 215) assumes that unmarked direct objects must precede indirect objects.

In order to test the parameters that enhance or block DOM of direct objects in ditransitive sentences, we conducted two forced choice experiments. We presented short contexts with three sentences, the last of which had two continuations – i) with DOM for the definite direct object and ii) without; see Table 1 for one example. In both experiments we manipulated verb class and clitic doubling of the indirect object (CD vs. noCD). With respect to verb class, we compared verbs of change of location (or verbs of caused motion) such as mandar 'to send' with verbs of no change of location (or verbs entailing a secondary experiencer), such as mostrar 'to show'. In Exp1 we also tested word order (DO > IO; IO > DO), and in Exp2 we selected direct objects and indirect objects with different number marking, one in the plural and the other in the singular, in order to prevent a potential ambiguity of the clitic pronoun *le*. We used the online platform Google Forms to present the questionnaires that each experiment consisted of. In the first questionnaire, 96 native Spanish speakers participated, mainly from Spain and Mexico, while in the second one, which was restricted to Spain, we had 157 participants.

The results from both experiments (Table 2) show that (i) participants used DOM in half of the instances across conditions; (ii) there was a significant effect of verb class: verbs of change of location show a higher rate of DOM than verbs with no change of location (estimate: 1,02; $p = 0,02$); (iii) clitic doubling of the indirect object has a clear blocking effect on DOM (estimate: -1,36; $p < 0,001$), but, more interestingly, (iv) this effect is much higher for verbs of change of location (estimate: -2,51; $p < 0,001$) than for the other verb class (estimate: -0,6; $p < 0,001$); (v) in Exp. 1 there was no significant effect of word order (estimate: -0,002; $p = 0,991$).

Taking the results together, DOM alternations are not a "stylistic variation" (against Comrie 2013, Real Academia Española 2009), but a grammatical effect depending on verb class and clitic doubling of the indirect object, but not on word order (against Rodríguez- Mondoñedo 2007). In the final part of the paper we will discuss the contribution of these results to our understanding of the structure of ditransitive sentences and the function of clitic doubling and DOM in such constructions.

(1) *Busc-o   *el/al   médico.*
     look_for-PRS.1SG.   *the /DOM.the doctor
     'I am looking for the doctor.'

(2) *El   profesor envi-ó   el/al   chico a-l   jefe de estudios.*
     The   teacher send-PST.3SG the /DOM.the boy   to-the head of studies
     'The teacher sent the boy to the head of studies'.

(3) *El   profesor le_i   envi-ó el./(al)   chico a-l   jefe de estudios*
    The teacher   CLIT.DAT.3SG send-   PST.3SG the /DOM.the boy   to-the   head of studies
    'The teacher sent the boy to the head of studies.'

Table 1: Experimental item for Exp1 with the condition noCD and DO > IO

Context: *El espectáculo iba a comenzar y el actor protagonista ya se encontraba preparado. Aquella noche entre el público había un cazatalentos en busca de algún actor para su nueva película. Al cazatalentos le estaba gustando tanto la actuación del protagonista que se puso en seguida en contacto con su mánager. Este le dijo que esperase a que terminara la función. Y de esta forma…*
i)   *el mánager presentó <u>al actor</u> al cazatalentos en cuanto terminó lafunción*
ii)   *el mánager presentó e<u>l actor</u> al cazatalentos en cuanto terminó lafunción*
'The show was about to start and the leading actor was already ready. That night there was a talent scout in the audience looking for an actor for his new film. The talent scout was so fond of the protagonist's performance that he immediately contacted his manager. He told him to wait until the show was over. And so…' the manager introduced the actor to the talent scout as soon as the show was over'

Table 2: Percentages of DOM for verb class and clitic doubling of the indirect object

DOM	Exp01			Exp02		
verb class / clitic doubling	all	CD	noCD	all	CD	noCD
verbs of change of location	59,07%	36,81%	81,32%	61,94%	41,72%	82,17%
verbs of no change of location	37,09%	31,32%	42,86%	44,37%	38,64%	50,11%

References: Company Company, C. (2001). Multiple dative-marking grammaticalization: Spanish as a special kind of primary object language. *Studies in Language 25*(1):1–47. Comrie, B. (2013). Human themes in Spanish ditransitive constructions. In Bakker, D. and Haspelmath, M., eds, *Languages Across Boundaries. Studies in Memory of Anna Siewierska*, pages 37–52. de Gruyter, Berlin. Fábregas, A. (2013). Differential object marking in Spanish: State of the art. *Borealis 2*(2):1– 80. von Heusinger, K. and Kaiser, G. A. (2011). Affectedness and differential object marking in Spanish. *Morphology*, 21(3–4):593–617. Leonetti, M. (2004). Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics*, 3:75–114. Pensado, C. (1995). El complemento directo preposicional. Estado de la cuestión y bibliografía comentada. In Pensado, C., eds, *El complemento directo preposicional*, 11–60. Visor, Madrid. Real Academia Española. (2009). *Nueva gramática de la lengua Española*. Madrid, Espasa- Calpe. Rodríguez-Mondoñedo, M. (2007). *The syntax of objects: Agree and differential object marking*. University of Connecticut dissertation.

Topicalization: The IO/DO asymmetry in Icelandic

Elena Callegari, Anton Karl Ingason
University of Iceland
ecallegari@hi.is, antoni@hi.is

Synopsis: *We investigated differences in the frequency of direct-object versus indirect-object topicalization in Icelandic using the Icelandic Parsed Historical Corpus (IcePaHC, Wallenberg et al. 2011). Specifically, we queried for double-objects constructions to determine which type of object (IO or DO) was more likely to be topicalized. Our queries revealed that the overall incidence of DO topicalization is double that of IO topicalization. We argue that this is a consequence of the cross-linguistic preference towards having topical information appear before focal one: while DO topicalization can help ensure that this configuration is obtained when the IO is in focus, there is nothing to gain from topicalizing the IO when the DO is in focus, as IO > DO is already the unmarked order in Icelandic.*

In Icelandic double-objects constructions, the accusative DO can either precede or follow the dative IO, with IO > DO being the unmarked order:

(1) Ég   gaf   Elínu   bókina
     I   gave   Eileen(DAT) book.the(ACC)

While the inverse order (DO > IO) is certainly possible, its availability is dependent on factors such as stress patterns, definiteness and idiomaticity (Collins & Thráinsson 1996, Falk 1990, Ottósson 1991). According to Dehé (2004), the IO > DO order is in fact preferred even in those cases where the inverse order would be licensed, e.g. when the IO is in focus. This overall preference for the IO>DO order raises the question of whether this preference is maintained when the two objects no longer appear in the same local domain, i.e. when either object is fronted to a pre-verbal position through topicalization. Both IO and DO object topicalization (as exemplified in ex. 2 - 3) are possible in Icelandic; our goal was to determine whether these are equally frequent, or whether one is more frequent than the other. In particular, if the preference towards having the DO precede the IO is a linear type of constraint (e.g. it applies regardless of the relative distance and structure between DO and IO), we would expect IO topicalization to be more frequent than DO topicalization.

(2) Elínu   gaf   ég   bókina   IO topicalization Eileen(DAT)
     gave   I   book.the(ACC)

(3) Bókina   gaf   ég   Elínu   DO topicalization
     book.the(ACC) gave   I   Eileen(DAT)

We searched the IcePaHC using PaCQL (Parsed Corpus Query Language, Ingason 2016) through the freely available online platform treebankstudio.org. The IcePaHC (1,002,390 words) is a collection of texts from the 12th to the 21st century. We searched for all instances of matrix-clause double-object constructions, and coded these depending on whether (i) IO topicalization had occurred, (ii) DO topicalization had occurred, (iii) no topicalization had occurred. Our search returned a total of 1110 hits. Out of these, 89 were instances of DO topicalization (incidence: 8%) and 39 were instances of IO topicalization (incidence: 3,5%); we provide an example for each structure below.

(4) Reykelsi   færum   vér   honum
     Incense(ACC)   bring   we   he(DAT) …
     'We bring him incense'
   (From *Íslensk Hómilíubók*, late 12th century or early 13th century)

(5)   … og [öllum ríkismönnum þeim   sem   þar   voru]
        … and [all   powerful-men those that   there   were](DAT)

        gaf   hann   nokkura   góða   gjöf   og   sæmilega
        gave   he   some   good   gift   and   respectable

'and he gave all the powerful men that were present some good and respectable gift'
(From *Finnboga Saga Ramma*, 1330-1370)

Of the 39 cases of IO topicalization, only six consisted in the fronting of a pronominal element. Many of the topicalized datives were rather quite heavy (Indriðadóttir & Ingason 2019), as it is the case for (5) above; the average length of topicalized DOs was 2.6 words.

DO topicalization was thus more than twice as frequent as IO topicalization. As tables 1. and 2. show, while the relative frequency of DO and IO topicalization differs depending on the century, the trend which sees DO topicalization being more frequent than IO topicalization is robust and attested for almost all centuries represented in the IcePaHC.
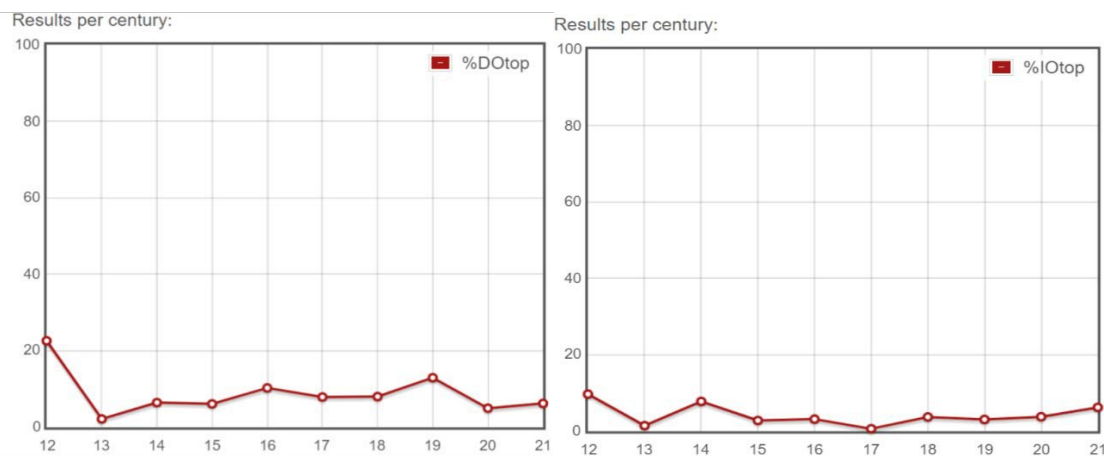
We thus observe an asymmetry between topicalized and non-topicalized structures: while the order IO>DO is preferred when both objects appear VP-internally, the order DO>IO order has a higher incidence the moment the two objects no longer appear in the same local domain. If we assume there is a general, cross-linguistic preference towards having topical constituents precede focal ones, the overall higher incidence of DO topicalization finds a ready explanation: in those instances where the IO is a topic and the DO is in focus, no word order permutation is needed to ensure that the constituent in topic precedes what is in focus, as this order is already the unmarked one in Icelandic ditransitive constructions. If the DO is topical and the IO is in focus, on the other hand, the fronting of the DO to a preverbal position is warranted in order to restore the Topic > Focus precedence relation. This alone would explain why DO topicalization is more frequent than IO topicalization: while there is nothing to be gained (information-structure-wise) from fronting a topical IO in Icelandic, fronting a topical DO helps in ensuring that old/topical information appears before new/focal one.

References: Collins, C., & Thráinsson, H. (1996). VP-internal structure and object shift in Icelandic. *Linguistic inquiry*, 391-444. Dehé, N. (2004). On the order of objects in Icelandic double object constructions. *UCL working papers in linguistics*, 16, 85-108. Falk, C. (1990). On double object constructions. *Working Papers in Scandinavian Syntax,* 46, 53-100. Ingason, A. K. (2016). PaCQL: A new type of treebank search for the digital humanities. *Italian Journal of Computational Linguistics*, 2(2), 51-66. Indriðadóttir, I. H., & Ingason, A. K. (2019). Linguistic end-weight is really edge-weight. Observing heaviness in a parsed corpus. *DHN*, 240-249. Ottóson, K.G. (1991). Icelandic double objects as Small Clauses. *Working Papers in Scandinavian Syntax*, 48, 77-97. Wallenberg, J. C., Ingason, A. K., Sigurðsson, E. F., & Rögnvaldsson, E. (2011). *Icelandic Parsed Historical Corpus (IcePaHC)*. Version 0.9. Available at http://www.linguist.is/icelandic_treebank

Investigating person-case effects in Standard German and Swabian

Johannes Rothert
University of Potsdam
rothert@uni-potsdam.de

Claim: We report on a small-scale acceptability rating study that was conducted to test Anagnostopoulou's (2008) claims about the emergence of the Person-Case Constraint (PCC) in German. She argues (i) that German exhibits visible person-case effects in ditransitive constructions when the subject follows a weak object cluster, but not when it precedes the cluster, and (ii) that this effect is present in Standard German, but not in the Southern German dialect Swabian. The results of our investigation indicate that Standard German and Swabian do not exhibit any visible person-case effects, regardless of the position of the subject.

Background: In languages subject to the PCC, the grammaticality of a ditransitive construction with phonologically weak objects depends on the person value of both objects. The constraint applies to a large number of typologically unrelated languages and comes in several versions that differ regarding which combinations of weak objects are prohibited (see, among many others, Nevins 2007). Anagnostopoulou (2008) claims that German is subject to the PCC. According to her, the combination of a local (1st or 2nd) person direct object and a 3rd person indirect object is ungrammatical, but only when the subject follows the weak object cluster. This is illustrated with the contrast in (1) (examples taken from ibid., p. 26). In order to verify the reliability of these judgements, we carried out a small-scale acceptability rating study.

(1) a. *weil   dich ihm   irgendwer vorgestellt hat   b. weil   sie dich ihm   vorgestellt hat
       because you to him someone   introduced has   because she you to him introduced has
       'because someone has introduced you to him'   'because she has introduced you to him'

Design and Method: The study had a fully crossed 2 2 2 design with the factors DIRECT OBJECT PERSON (local vs. 3rd person), INDIRECT OBJECT PERSON (local vs. 3rd person), and SUBJECT POSITION (subject > object vs. object > subject). For the construction of the items, we used ten ditransitive verbs that allow both their objects to be animate / local person. An example showing the four subject > object conditions for the verb *vorstellen* 'introduce' is given in (2).

(2)   Man erwatet, dass jemand   ihn mir   / ihn ihr   / mich dir   / mich ihr   vorstellt.
      one expects   that someone him to me / him to her / me   to you / me   to her introduces
      'One expects that someone introduces him to me / him to her / me to you / me to her.'

The ten target items and 24 filler items were merged into an online questionnaire that was completed by eight native speakers each of Standard German and Swabian. The participants saw all experimental sentences in randomized order and their task was to rate the acceptability of the sentences on a scale ranging from 1 (= very unacceptable) to 7 (= fully acceptable). If Anagnostopoulou's (2008) claims are correct, the proportion of ratings in the lower rating categories in the local person direct object > 3rd person indirect object > subject condition should be large for the Standard German speakers, but not for the Swabian speakers. In all other conditions, the proportion of ratings in the lower rating categories should be very small for both speaker groups.

Results: We ran ordinal regressions in the Bayesian framework to analyze the acceptability ratings. For both speaker groups, we carried out independent analyses on the data from the subject > object and the object > subject conditions. In the discussion of the results, we focus on interpreting the distribution of the ratings per condition as predicted by the respective model. The results are shown in Figure (1) and point towards the following conclusion. (i) The four object combinations differ in their acceptability, but not in their grammaticality, indicating that German does not exhibit any visible person-case effects. Across subject positions and speaker groups, the proportion of ratings in the lower rating categories is very small for all four object combinations. (ii) The emergence of the PCC does not depend on the position of the subject. Irrespective of whether it precedes or follows the object cluster, the pattern of acceptability ratings is very

similar. (iii) The emergence of the PCC does not depend on the dialect. The Standard German and Swabian speakers show a very similar pattern of acceptability ratings.

Conclusion: Our investigation has important implications for the emergence of person-case effects in weak pronoun languages as well as the importance of acceptability rating studies. On the one hand, they provide experimental evidence that weak pronouns are not affected by the PCC in the same way as clitics or agreement markers (see also Doliana 2013). On the other hand, they show that judgements from trained linguists do not always match the intuition of linguistically naive speakers and can therefore lead to wrong generalizations (see e.g., Featherston 2007)
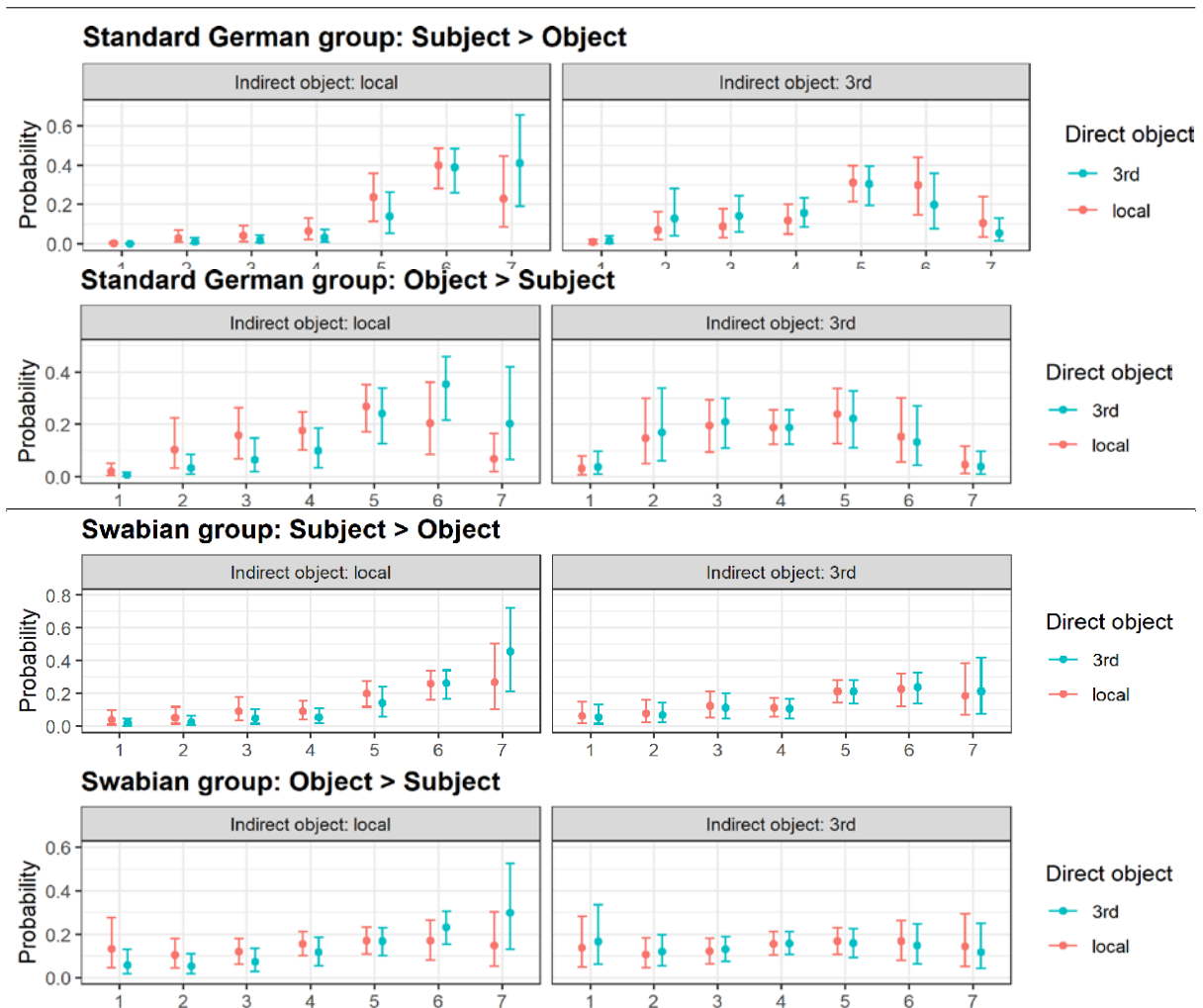
Figure 1: Standard German (top) and Swabian (bottom) speakers' acceptability ratings for the sentences of the experiment. Points indicate posterior mean estimates for the probability that the ratings fall into the seven rating categories and error bars indicate 95% credible intervals.

References: Anagnostopoulou, Elena. 2008. Notes on the Person Case Constraint in Germanic (with special reference to German). In Roberta D'Alessandro, Susann Fischer & Gunnar Hrafn Hrafnbjar- garson (eds.), *Agreement restrictions*, 15–48. Berlin, Germany: Mouton de Gruyter. Doliana, Aaron. 2013. *On the Person-Case Constraint: From the giga to the zero version with copy impoverishment and check*. Leipzig, Germany: Universität Leipzig BA thesis. Featherston, Sam. 2007. Data in generative grammar: The stick and the carrot. *Theoretical Linguistics* 33(3). 269–318. Nevins, Andrew I. 2007. The representation of third person and its consequences for person-case effects. *Natural Language & Linguistic Theory* 25(2). 273–313.

Inverting objects in Icelandic: Report from a corpus study

Jóhannes Gísli Jónsson
University of Iceland
jj@hi.is

Although the neutral order of two objects in Icelandic is clearly IO-DO, inverting the order of the objects is also possible under certain circumstances:

(1)   a.   Ég   gaf   nemandanum   bókina
                I   gave   student.the-DAT   book.the-ACC

       b.   Ég   gaf   bókina   nemandanum
                I   gave   bool.the-ACC   student.the-DAT

            'I gave the student the book.'

Object Inversion (OI) in Icelandic has been discussed in the theoretical literature for some time (Falk 1990, Ottósson 1991, Holmberg & Platzack 1995, Collins & Thráinsson 1996, Dehé 2004 and Ussery 2017, 2018) but still many questions remain. This paper reports the findings of an extensive corpus study of OI, based on the Risamálheild Corpus (Steingrímsson et al. 2018). While confirming some earlier claims about OI, the results also yield new insights that will undoubtedly play an important role for theoretically oriented work on Icelandic ditransitives in the near future.

The most surprising result is that OI is incredibly rare in Icelandic. With most ditransitive verbs, the rate is only about 1% of all examples where both objects appear. For a handful of ditransitive verbs, the rate is significantly higher, e.g. *afhenda* 'deliver, hand over' (11%), *selja* 'sell' (17%) and *tilkynna* 'announce' (28%). The study also reveals that OI strongly favors examples where the DO is phonologically lighter than the following IO. Moreover, the DO must encode old information and the verb must belong to the biggest class of ditransitive verbs in Icelandic, the so called *gefa*-verbs (give-verbs), which take a dative IO and an accusative DO. (The other classes display the case patterns DAT-DAT, DAT-GEN, ACC-DAT and ACC-GEN.) To be sure, some potential counterexamples with ditransitive verbs outside the *gefa*-class are attested in the Risamálheild Corpus but nearly all of them can be argued to involve Heavy Object Shift moving the IO to the right of the DO rather than OI, as in (2):

(2)   að   undanþiggja   skyldunni   þessar   litlu   stofnanir
        to   exempt   the.duty-DAT   these-ACC   small-ACC   institutions-ACC

        sem   ég   hef   verið   að   tala   um
        which   I   have   been   to   talk   about

        'to exempt from this duty these small institutions that I have been talking about'

DO-IO orders have been claimed to be base-generated in Icelandic (Falk 1990, Holmberg & Platzack 1995), but the facts discussed above suggest that OI is the result of movement (see also Ottósson 1991). Thus, OI is clearly very different in many ways from Prepositional Datives in languages like English where the IO is expressed in a PP following the DO. To the best of my knowledge, a movement account of OI has never been developed, but it is plausible to derive OI by movement of the DO to a specifier position in a projection just above the ApplP hosting the IO. In this derived position, the DO checks accusative case from v. As a result, this position is not available to DOs from other classes of ditransitive verbs in Icelandic.

References: Collins, Chris & Höskuldur Thráinsson. 1996. VP-internal Structure and Object Shift in Icelandic. *Linguistic Inquiry* 27. 391-444. Dehé, Nicole. 2004. On the Order of Objects in Icelandic Double Object Constructions. *UCL Working Papers in Linguistics* 16. 85–108. Falk, Cecilia. 1990. On Double Object Constructions. *Working Papers in Scandinavian Syntax* 46. 53–100. Holmberg, Anders & Christer Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. Oxford University Press, Oxford. Ottósson, Kjartan. 1991. Icelandic Double Objects as Small Clauses. *Working Papers in Scandinavian Syntax* 48. 77–97. Steingrímsson, Steinþór, Sigrún Helgadóttir, Eiríkur Rögnvaldsson, Starkaður Barkarson & Jón Guðnason. 2018. Risamálheild: A Very Large Icelandic Text Corpus. *Proceedings of LREC 2018*, 4361-4366. Myazaki,

Japan. Ussery, Cherlon. 2017. Double Objects Again. . . but in Icelandic. In Nicholas LaCara, Keir Moulton & Anne-Michelle Tessier (eds.), *A schrift to fest Kyle Johnson*, 375–388. Linguistics Open Access Publications. Ussery, Cherlon. 2018. Inversion as Rightward-Dative Shift in Icelandic Ditransitives. In Sherry Hucklebridge and Max Nelson (eds.), *Proceedings of the 48th Meeting of the North East Linguistic Society*. University of Massachusetts, Amherst.

Preposition reanalyzes and ditransitive sentences in Brazilian Portuguese

Ana Regina Calindro, Maria Aparecida Torres Morais
Universidade Federal de Rio de Janeiro, Universidade de São Paulo
anacalindro@letras.ufrj.br, torres.mariacida@gmail.com

1. Introduction: According to recent diachronic and synchronic studies, Brazilian Portuguese (BP) presents different strategies of encoding the Indirect Object (IO) in ditransitives sentences. We begin our presentation with facts of European Portuguese (EP), because modern BP and modern EP share a common historical background. According to Torres Morais and Salles (2010, 2016, 2019), there is strong syntactic and semantic evidence to corroborate the hypothesis that the IO in EP is morphologically expressed with dative case. As a DP, it is introduced by the preposition *a,* a dative Case marker. As a pronominal form, it is expressed by the 3rd person dative clitic *lhe/lhes* (cf. 1-3). BP behaves differently. As examples (4) to (6) show, BP has reanalyzed the IO as a prepositional complement, introduced by transitive prepositions, namely *a* and *para* (cf. Calindro, 2015, 2016). Additionally, in its pronominal expression, the IO is not a dative clitic, but a strong/full pronoum. Our main goal in this presentation is to propose an analysis of the innovative BP prepositional system. Based on Svenonius (2004) ideas and Cuervo (2010), we argue that prepositional phrases (pP) can account for the relation between direct objects (DO) and indirect objects (IO) in the context of BP ditransitive structures

2. Theoretical background: Based on the discussion on English and Bantu languages (Marantz 1993, Pylkkänen 2002), Cuervo (2003) for Spanish and Diaconescu and Rivero (2007) for Romanian proposed applicative heads for ditransitive sentences, based on obligatory dative clitic doubling. Hence, in Spanish, for example, the dative clitic (*le*) co- occurs with the DP introduced by a functional preposition *a*. The clitic is the *Spell-out* of the ApplP, because it is responsible for lexicalizing the number and person features of the DP in SpecApplP. Therefore, the main argument for applicative heads in Romance languages is the dative morphological expression. According to Torres Morais & Salles (2010, 2019), EP also has applicative constructions, because the main characteristic for this assumption is that the IO can always alternate with the dative clitic *lhe(s).* Hence, the clitic can only be used when the IO is introduced by *a,* so the *a*-DP has dative Case. When the IO is introduced with *para* it does not alternate with *lhe,* Even though EP and BP shared the same structure in the past, BP has undergone a diachronic change for encoding IOs, as mentioned in the introduction. Therefore, in this presentation, we intend to show that this change in BP has affected its argumental structure, separating this variety from the European one.

3. Proposal for BP: As mentioned before, we are assuming BP has lost the functional preposition *a* and pronominal evidences for an applicative grammar. Consequently, we will assume with Svenonius (2004) and Cuervo (2010) that prepositions may have semantic content (cf. Calindro 2015, 2016, 2020). They are *transitive elements*, because they can project *complemen*t and *specifier*. Hence, in languages with transitive prepositions in ditransitive sentences, as we are assuming for BP, a pP projection is adequate to represent the DO and IO relation. The prepositions which introduce IOs in BP, *para* and also *a* (cf. 4- 5) have been reanalyzed as lexical elements. Hence, transitive prepositions determine selection restrictions to its complement IO, but not for the DO.. As prepositions can project Spec and complement positions, a pP can introduce a thematic relation between the DO and the IO (cf. Wood 2012). Otherwise, if there is only a PP projection in the structure, the DO *theme* would be in SpecPP - being subcategorized by the preposition, when actually its relation is with the verb (cf. 7a), what can be accounted for using little *p*, similar to little *v*'s ability to introduce external arguments in transitive structures (cf. 7b).

The representation in (7a) shows the relation between the IO and P, which can be related to the asymmetry between the verb and both its complements in ditransitive constructions. Thus, in a pP configuration, the preposition remains inside the PP, as it imposes restrictions to the IO, not the DO. This means that *p* can introduce a thematic relation, in Cuervo's (2010) terms. Hence, a ditransitive sentence, as (4), in BP, can be represented as illustrated in (8).

European Portuguese

(1) A   Maria   <u>enviou</u> uma   carta   **ao João**   /enviou-**lhe**   **uma carta.**
The Maria   sent   a   letter   P$_{a(to)}$ the João.DAT   /sent-3SG.DAT   letter.
‘Maria sent a letter to João/sent him a letter.’

(2) A   Maria   <u>atirou</u> a   bola   **ao João**   /<u>atirou</u>-**lhe**   a bola.
The Maria   threw   the   ball   P$_{a(to)}$ the João   /threw-CL.3rd.DAT   the ball
‘Maria threw the ball to João / threw him a ball.’

(3) A   Maria   <u>preparou</u> o jantar   **ao João**   /<u>preparou</u>-**lhe**   o jantar. (*BP)
The Maria   prepared   the dinner P$_{a(to)}$ the João   / prepared-CL.3rd.DAT   the dinner
‘Maria prepared the dinner for João’

Brazilian Portuguese

(4) Maria <u>enviou</u> uma carta   **para/a**   **o João**   /**para ele**.
Maria sent   a letter   P$_{para(to)/a(to)}$ the João.OBL   /to   him.3SG
Maria sent a letter to João/to him.’

(5) Maria <u>atirou</u>   a bola   **para/a**   **o João/ele**.
Maria threw   the ball   P$_{para(to)/a(to)}$ the João.OBL/him
‘Maria threw the ball to João.

(6) Maria <u>preparou</u> o jantar   **para o João**   /**para ele**.
Maria prepared the dinner   P$_{para(to)}$the João.OBL   /for him.3SG
‘Maria prepared dinner for João/for him.’

(7) a. [$_{p}$P *DO Figure*   [$_{p}$’   [$_{p}$ [PP [P [*IO Ground*]]]]]]

      b. [$_{v}$P *Subject* [$_{v}$’   [$_{v}$ [VP [V [*DO*]]]]]]

(8) [$_{v}$P Maria [$_{v}$’ [$_{v}$ [$_{VP}$ enviou [$_{p}$P uma carta [$_{p}$’ Ø [$_{PP}$ *para /a*   o João]]]]]]].

References: Calindro, Ana. 2015. *Introduzindo Argumentos: uma proposta para as sentenças ditransitivas do português brasileiro*. [Introducing arguments: the case of ditransitives in Brazilian Portuguese]. São Paulo: University of São Paulo. (Doctoral Dissertation.) Calindro, Ana. 2016. "Introducing indirect arguments: the locus of a diachronic change. *Rivista di Grammatica Generativa,* Perugia, Itália, v. 38. Calindro, Ana. 2020. "Ditransitive constructions: what sets Brazilian Portuguese apart from other Romance languages?" In Anna Pineda and Jaume Mateu (eds.). *Dative constructions in Romance and beyond* (Open Generative Syntax 7). Berlin: Language Science Press, v. 7, 2020. p. 75-95. Cuervo, Cristina. 2003. *Datives at large*. MIT. (Doctoral Dissertation). Cuervo, Cristina. 2010. "Against Ditransitivity." *Probus,* 22. 151-180. Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In Mchombo, Sam (ed.), *Theoretical aspects of Bantu grammar*. 113-151. Pylkkänen, Liina. 2002. *Introducing Arguments*. MIT. (Doctoral Dissertation). Svenonius, Peter. 2004. "Adpositions, particles and the arguments they introduce." In Eric Reuland, Tammoy Bhattacharya, and Giorgos Spathas (eds.), *Argument Structure.* Philadelphia: John Benjamins. 63–103. Torres Morais, Maria Aparecida and Heloísa Salles. 2010. "Parametric change in the grammatical encoding of indirect objects in Brazilian Portuguese." *Probus* 22. 181- 209. Torres Morais, Maria Aparecida and Heloísa Salles. 2016. The external possessor construction in European Portuguese and Brazilian Portuguese. In KATO, M. & ORDÓNEZ, F. (eds.), The morphosyntax of Portuguese and Spanish in Latin America. Oxford, Oxford University Press: 204-235. Torres Morais, Maria Aparecida and Heloísa Salles. 2019. "Possessivos de 3ª. pessoa na história do Português Brasileiro." In Charlotte Galves, Mary Kato, and Ian Roberts. Português Brasileiro. *Uma segunda viagem diacrônica*, Unicamp, VII: 174-201.Wood, Jim. 2012. *Icelandic Morphosyntax and Argument Structure.* New York: New York University. (Doctoral Dissertation).

On the derivation of prepositional dative constructions in Irish and Gaelic

Gary Thoms
New York University

Collins (2017) proposes an approach to the dative alternation where the double object construction (DOC) is the base structure and the prepositional dative construction (PDC) is derived by 'smuggling' the theme over the goal within a lower VP constituent. This is supported by c-command asymmetries in English: the recipient always asymmetrically c-commands the goal in DOCs, (1)-(2), but theme and goal may c-command each other, (3)-(4) (the 'backward' c-command being derived via reconstruction). (5) gives a partial tree for the PDC derivation.

1. I sent [every worker]$_i$ her$_i$ apprentice.
2. *I sent her$_i$ boss [every apprentice]$_i$.
3. (?)I sent her$_i$ apprentice to [every worker]$_i$
4. I sent [every apprentice]$_i$ to her$_i$ boss.
5.
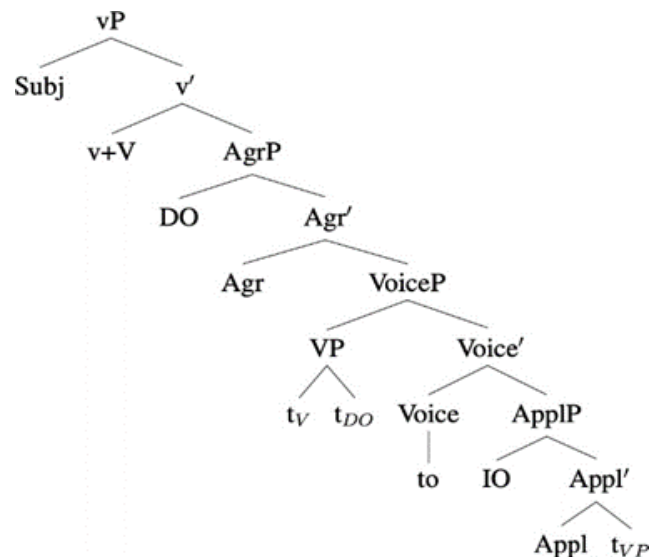
I argue that a version of approach to PDCs receives strong support from Irish and Scottish Gaelic, which do not allow for DOCs (Jung et al 2012). The first source of evidence for VP-fronting comes from c-command facts: the goal always c-commands the theme for binding purposes, (6), while for some speakers the theme may not c-command the goal, (7) (Jung et al 2012). (6) follows if there is VP-fronting plus reconstruction, and the variation with (7) would follow from variation in whether the DO undergoes some additional step of object shift following VP-fronting (as in Collins' derivation for English).

6. Thug Seán a$_i$ pheann-fhéin do   chuile   bhuachaill$_i$
   gave J   his pen-self   to   every   boy
   "John gave every boy his own pen"   (Ir)

7. *Sheall   Máiri a h-uile   leabhar$_i$ dha$_i$ úghdar$_i$
   showed M   every   book   to-its author
   "Mary showed every book to its author"   (SG)

The second source of evidence comes from word order variation in Irish PDCs noted by Maki & Ó Baoill (2008). They note that while DO-IO order is the default, IO-DO order is possible in cases where the IO is a quantifier (8) or a definite but not if it is a non-heavy bare indefinite, (9), an apparent case of rightward object shift of DOs. Similar facts are noted by Pearson (2004) for Malagasy (where VP-fronting derivations for basic clause structure have much independent plausibility), and so I adopt a version of his analysis: the DO-IO order is derived by applying

definiteness-driven object shift of the DO out of the lower VP to a low position below the IO's base-generated position prior to VP-fronting, (10). I offer some speculative remarks on how this account could be extended to account for pronoun postposing in Irish and SG, the properties of which are more involved (Bennett et al 2018).

8. Thug Máire do Sheán úllaí   go leor
   gave M   to S   apples plenty
   "M gave many apples to John" (Ir)

9. Thug Milo do Bhinclí   {*caiserbhán / an caiserbhán}
   gave M   to B   a-dandelion   the dandelion
   "Milo gave a/the dandelion to Binkley" (Ir)

10. [$_{vP}$ subj [$_{XP}$ [$_{VP}$ … V t$_{DO}$] [$_{X'}$ X [$_{ApplP}$ PP$_{IO}$ [$_{Appl'}$ Appl [$_{YP}$ DP$_{DO}$ [$_{Y'}$ Y … t$_{VP}$]]]]]]]

A third argument for the VP-fronting analysis comes from Irish quantifier float facts, also from Maki & Ó Baoill (2008). (11)-(12) shows that the quantifier *uilig* 'all' can be associated with an IO with the IO-DO order (as in 8) but not a DO with the DO-IO order. (12) can be understood on the VP-fronting analysis if the DO-IO order is derived by VP-fronting, if QF is derived by A- movement (Sportiche 1988 on French, McCloskey 2001 on Irish), and would be mysterious if that order (and the c-command facts) were derived by A-movement of the DO over the IO. The fact that QF is possible with the IO, which is a PP in (11), is remarkable in the landscape of QF facts, but I argue it can be understood if the IO's PP-syntax is derived by raising-to-complement-of- preposition, which is known to be a feature of the syntax of these languages (McCloskey 1984, McCloskey and Sells 1988). Revising (10), I argue that the IO is base-generated as a nominal subject of the embedded small clause and then raised, via at least one stopping off point where it may strand its quantifier, on the way to the complement of P, which is introduced in a position which is itself below where the VP (containing the DO) is moved to. I provide a detailed articulation of the VP structure required for this in the talk.

11. Thug Maire do *na mic leinn* ullai go leor   *uilig* inne.
    gave M   to   the students apples plenty   all   yesterday
    "Mary gave plenty apples to all the students yesterday"

12. *Thug Maire *na lebhair*   do Sean *uilig* inne.
    gave   M   the books   to Sean all   yesterday
    Intended: "Mary gave all the books to Sean yesterday"

References: Collins, C. 2017. A smuggling approach to the dative alternation. Ms., NYU (lingbuzz). Jung, H., A. Carnie, H. Harley. 2012 On the lack of double object constructions in Scottish Gaelic (and Modern Irish). Handout from a talk at *New Perspectives on Celtic Syntax*, Berkeley. Pearson, M. 2004. Two types of VO languages. In P. Svenonius (ed)., *The derivation of VO and OV*. John Benjamins: Amsterdam, p.327-363. Maki, H. and D. Ó Baoill. 2008. The theme goal construction in Modern Irish. *English Linguistics* 25:2, 439-451.

Thematic role and movement to subject position: Muskogean evidence for a 'deactivation'-based account

Matthew Tyler
Christ's College, University of Cambridge
mt516@cam.ac.uk

Introduction. The structure in (1) could schematize an applicative of an unaccusative, or a passive of a ditransitive—two configurations where there is (a) a subject position that needs filling, (b) no external argument, and (c) two internal arguments that are each potential subjects.

(1)   Derivation before movement to subject position (Spec-TP)

$[_{TP}$ T^0 $[_{VoiceP}$ Voice $[_{ApplP}$ NP_{ApplO} $Appl^0$ $[_{VP}$ V NP_{DO}]]]]

In such configurations, we find three attested patterns of syntactic behavior. In the asymmetric pattern in (2), only the higher of the two internal arguments—usually the indirect object (IO) or applied object (ApplO)—may become the subject. In the **symmetric** pattern in (3), either of the two internal arguments may become the subject, with the choice between them determined by factors outside of thematic structure, such as information structure. And in the 'reverse-asymmetric' pattern in (4), only the lower internal argument — usually the direct object (DO) — may become the subject (this pattern is less discussed than (2/3) but is attested, e.g. McGinnis 1998:53f. on Albanian).

(2) Asymmetric

$[_{TP}$ NP T^0 $[_{VoiceP}$ Voice $[_{ApplP}$ NP_{ApplO}) $Appl^0$ $[_{VP}$ V NP_{DO}]]]]

(3) Symmetric

$[_{TP}$ NP T^0 $[_{VoiceP}$ Voice $[_{ApplP}$ NP_{ApplO}) $Appl^0$ $[_{VP}$ V NP_{DO})]]]]

(4) Reverse-asymmetric

$[_{TP}$ NP T^0 $[_{VoiceP}$ Voice $[_{ApplP}$ NP_{ApplO} $Appl^0$ $[_{VP}$ V NP_{DO})]]]]

Locality vs. deactivation. The choice between (2) and (3) can vary by language (Bresnan & Moshi 1990), by verb (Van der Wal 2018), and by the thematic role of the ApplO (Alsina & Mchombo 1993), among other factors. Various proposals exist to account for the range of attested patterns. As Haddican & Holmberg (2015, 2019) note, the theories can be broadly classified into two groups. On the one hand, *locality*-based theories hold that the behavioral difference stems from whether or not the DO moves to an intermediate position (e.g. the outer specifier of ApplP), from where it can be targeted for movement to the subject position (McGinnis 1998, Anagnostopoulou 2003). On the other hand, *deactivation*-based theories hold that the difference stems from whether or not the ApplO/IO may be 'deactivated' in some way, rendering it ineligible for movement to the subject position and allowing the DO to move instead (Baker 1988, Woolford 1993). Deactivation of ApplO is typically via Case-{assignment/licensing/ valuation}, Agreement, or encasing ApplO within a PP 'shell'.

Today. I look to an understudied language family to shed light on this debate. In Choctaw and Chickasaw (Muskogean), applicatives of unaccusatives can show patterns (2), (3) or (4), with the choice of pattern depending on the *thematic role* of ApplO. I argue that this is more elegantly explained in a deactivation-based analysis than a locality-based analysis.

Muskogean. All ApplOs are indexed on the verb with an agreeing DAT/ABS prefix and sometimes an additional applicative prefix. Subjecthood can be diagnosed by (a) word order (subjects precede objects), (b) obligatory nominative case-marking, and (c) ability to serve as a pivot for switch-reference (Tyler 2020).

Choctaw (data from original fieldwork). ApplOs added to non-active verbs (which include unaccusatives and 'lexical pas- sives') may become the subject or object of their clause. (5) shows

that ApplOs become the subject if they are interpreted as affected experiencers, 'engineers' (in the sense of Myler 2016), external possessors or predicative possessors. (6) shows that ApplOs become objects if they are beneficiaries or locations. That is, we find asymmetric (5) and reverse-asymmetric configurations (6). Note also that the type of applicative is not determined by the host verb (cf. (5a) vs. (6a)).

(5) Choctaw: ApplO = subject

   a. Katie-at   tali   i-kochoofa-tok.   (ApplO = affected experiencer)
        Katie-NOM metal 3.DAT-bend.NACT-PST

        'The metal bent on Katie.'

   b. Miko-yat aapisa móyyoma-k-a i-tiwa-t   taha-tok.   (ApplO = engineer)
        chief-NOM window all.YG-COMP -OBL 3.DAT-open.NACT-PTCP finish.NACT-PST

        'The boss had all of the windows opened.'

   c. Alíkchi-yat ókfochoosh im-illi-tok.   (ApplO = external possessor)
        doctor-NOM duck 3.DAT-die-PST

        'The doctor's duck died.'

   d. Alíkchi-m-at   ofi i-lawa-h.   (ApplO = predicative possessor)
        doctor-DEM-NOM dog 3.DAT-many-TNS

        'That doctor has a lot of dogs.'

(6) Choctaw: ApplO = object

   a. Tali-t   Katie-ano i-kochoofa-tok.   (ApplO = beneficiary)
        metal-NOM Katie-OBL 3.DAT-bend.NACT-PST

        'The metal bent for Katie.'

   b. Miko i-katos-at   pro₃ im-ittola-tok.   (ApplO = source/location)
        chief DAT-cat-NOM   3.DAT-fall-PST

        'The chief's cat fell from her.'

Chickasaw (data from published work by Pam Munro). Applicatives of unaccusatives in Chickasaw behave much the same as in Choctaw. But in addition to asymmetric and reverse-asymmetric configurations, (7) shows that some applicatives show fully symmetric behavior—the choice between (7a) and (7b) is determined by non-thematic, information-structural factors (Munro & Gordon 1982, Munro 1999, 2016).

(7) Chickasaw: ApplO = subject *or* object

   a. Anaakoot nampanaa'-at a-sa-shiiyalhchi-taha.
        1SG.NOM   string-NOM   APPL-1SG.ABS-be.tied-be.done

        'I have the string tied on me.'

   b. Nampanaa'-at anaako a-sa-shiiyalhchi-taha.
        string-NOM   1SG.A CC APPL-1SG.ABS-be.tied-be.done

        'The string is tied onto me.'   (Munro 1999:263)

Analysis. Of the competing locality vs. deactivation accounts, a deactivation-based account provides the simplest way of stating the reverse-asymmetric pattern in (4): this Appl obligatorily deactivates its specifier. Other Appls either will *never* deactivate their specifier, leading to the asymmetric pattern in (2), or they will *optionally* deactivate their specifier, leading to the symmetric pattern in (3). The correlation between thematic role and deactivation behavior is also easily accounted for: Appls with different thematic behaviors are different functional heads, so they may have different syntactic properties.

By contrast, locality-based accounts need to say something extra about how the DO becomes the *only* eligible target for movement to subject position in (6)—movement of the DO to Spec-ApplP, as proposed in McGinnis (1998) and much subsequent work, only derives symmetry, but not reverse-asymmetry. Locality-based accounts also have difficulty with the correlation between thematic role and syntactic behavior: if reverse-asymmetry is derived by moving the DO to a projection *above* ApplP, then the head that attracts the DO, while ignoring the ApplO, must 'know' the thematic behavior of Appl.

Two more arguments for deactivation. (i) Choctaw ditransitives freely permit the DO to move to the left of the ApplO—a necessary prerequisite for symmetric and reverse-asymmetric derivations. Contra the predictions of the locality account, this movement is not restricted or mandated according to the thematic interpretation of the ApplO (data omitted here for space). (ii) The syntactic implementation of deactivation may vary across languages (e.g. Case vs. Agreement vs. PP shells), allowing for an account of 'partial symmetries' in double-object configurations cross-linguistically (cf. Van der Wal 2018).

The nature of deactivation in Choctaw. Deactivation cannot be tied to verb agreement, since *all* applied arguments are indexed on the verb (cf. (5-7)). Deactivation probably does not involve encasing the NP within an FP shell either, as we would expect this to render an NP un-targetable by agreement, as well as rendering it immovable. Deactivation *may* be related to Case, though at a fairly abstract level: deactivated ApplOs show the same morphological case-marking (non-obligatory oblique case) as any other non-subject NP.

References: Alsina, Alex & Sam Mchombo. 1993. Object asymmetries and the Chichewˆ a applicative construction. In Sam Mchombo ed. Theoretical aspects of Bantu grammar. Stanford, CA: CSLI Publications. 17–45. Anagnosto–poulou, Elena. 2003. The syntax of ditransitives: Evidence from clitics. Berlin: de Gruyter. Baker, Mark C. 1988. Incorporation: A theory of grammatical function changing. Chicago: University of Chicago Press. Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. Linguistic Inquiry 21, 147–185. Haddican, Bill & Anders Holmberg. 2015. Four kinds of object symmetry. In Complex visibles out there: Proceedings of the Olomouc Linguistics Colloquium 2014, 145–162. Haddican, Bill & Anders Holmberg. 2019. Object symmetry effects in Germanic. Natural Language & Linguistic Theory 37, 91–122. McGinnis, Martha. 1998. Locality and case. MIT Dissertation. Munro, Pamela. 1999. Chickasaw subjecthood. In Doris Payne & Immanuel Barshi eds. External possession. Amsterdam: John Benjamins. 251–289. Munro, Pamela. 2016. Chickasaw switch-reference revisited. In Rik van Gijn & Jeremy Hammond eds. Switch reference 2.0 Amsterdam: John Benjamins. 377–424. Munro, Pamela & Lynn Gordon. 1982. Syntactic relations in Western Muskogean: A typological perspective. Language 58, 81–115. Myler, Neil. 2016. Building and interpreting possession sentences. Cambridge, MA: MIT Press. Van der Wal, Jenneke. 2018. Flexibility in symmetry: An implicational relation in Bantu double object constructions. In Michelle Sheehan & Laura Bailey eds. Order and structure in syntax II: Subjecthood and argument structure. Berlin: Language Science Press. 115–152. Tyler, Matthew. 2020. Argument structure and argument-marking in Choctaw. Yale Dissertation.

High applicatives in Icelandic adjectival constructions

Einar Freyr Sigurðsson & Jim Wood
*The Árni Magnússon Institute for Icelandic Studies*

1. Introduction

Icelandic prima facie poses a problem to the assumption sometimes made within generative grammar that adjectives cannot assign case (Chomsky 1981), as pointed out by, e.g., Jónsson & Pétursdóttir (2012).

(1) Fundurinn   er opinn (öllum).
    the.meeting   is open everyone.dat
    'The meeting is open to everyone.'

(2) Þessi náttúruauðlind   er (okkur) að eilífu glötuð.
    this natural.resource is us.dat forever   lost

(3) Ég er   ekki líkur *(þér).
    I am   not like you.dat
    'I am not like you.'

Icelandic applicatives are not limited to verbs (Ingason 2016). Here I focus on adjectives and adjectival passives that take, e.g., dative benefactive arguments and argue that these are high applicatives. That is a somewhat surprising result as Icelandic high DP applicatives are very restricted (e.g., Maling 2002, Wood 2013). Furthermore, I argue that dative case is assigned in the same way that indirect objects of verbs are assigned their case.

2. High Applicatives

I set aside adjectives as in (3), where both arguments are obligatory. Instead, I focus on adjectives and adjectival passive participles like *opinn* 'open' and *glataður* 'lost', respectively, whose structures only require a nominative argument, here taken to be generated in SpecPredP. The optional dative argument on top of PredP, which in turn consists of the adjective *opinn* in (4), is interpreted as a benefactive, which is generally taken to be a part of a high applicative structure; it relates an entity to an event or state, unlike low applicatives, which relate two entities.

(4) $v_{BE}P$
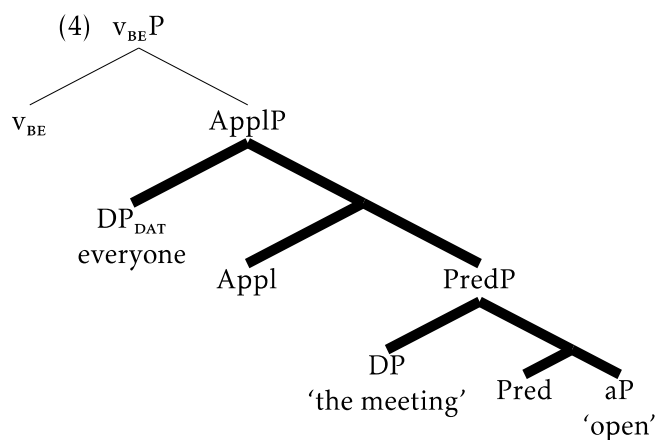
Adjectival applicatives seem to be more varied than verbal applicatives. (5)–(6) demonstrate; the dative argument *mér* is only allowed in the adjectival passive with the root *hylja* 'hide'.

(5) Þetta er (mér) hulið.
     this is me.dat hidden

(6) Hún huldi (*mér)   þetta.
     she hid   (*me.dat) this

Appl selects a small clause PredP in the structure shown in (4) and is in turn selected by the verb *vera* 'be'. This is a key to understanding why adjectival applied arguments is more varied than verbal applied arguments. Note, however, that according to our structure, "adjectival applicatives" is sort of a misnomer.

3. Order of arguments

It is noteworthy that the dative does in general not move to subject position whereas the nominative case argument usually does. This may indicate that high applicative arguments, at least those that Appl relates to a PredP, are not eligible as subjects.

4. Against a Silent PP Analysis

A natural question to ask is whether the benefactives discussed above are concealed PPs; many high applicatives are in fact overt PPs (with prepositions like *fyrir* 'for' and *handa* 'to the hands of'). However, such PPs are right-adjoined and cannot easily move in front of the adjective, see (8), unlike high DP applicatives, see (7):

(7) Þetta er henni gagnlegt.
     this is her.dat useful
     'This is useful for her.'

(8) *Þetta er fyrir hana gagnlegt.
     *this is for her.acc useful
     Intended: 'This is useful for her.'

In (7) and (8), the DPs in question do not even get assigned the same case (although that is not always so for high DP vs. PP applicatives). In addition to this, it is not always possible to spell out an overt PP instead of the high applicative DP.

5. Case Assignment

Finally, even though it looks like the adjectives and adjectival passives I discuss in this presentation assign dative case, it is really Appl that does that when the two merge. This is the same mechanism as with indirect objects of ditransitive verbs, where Appl has a dative case feature assigned via Merge (Sigurðsson 2017).

References: Chomsky (1981): *Lectures on Government and Binding* Ingason (2016): Realizing morphemes in the Icelandic noun phrase. UPenn diss. Jónsson & Pétursdóttir (2012): Þágufallsandlög með lýsingarorðum í íslensku og færeysku. Snæ- dal & Sigurðardóttir (eds.): *Frændafundur 7* Maling (2002): Það rignir þágufalli á Íslandi: Verbs with dative objects in Icelandic. *Íslenskt mál* Sigurðsson (2017): Deriving case, agreement and Voice phenomena in syntax. UPenn diss. Wood (2013): The unintentional causer in Icelandic. *Proceedings of NELS* 41

Prepositions and "double object constructions" in Cantonese

Kevin Kwong

Cornell University kk936@cornell.edu

Claim: Analyzing understudied patterns of relativization and parasitic gaps, I argue that all Cantonese dative constructions, which express events where agents cause themes to be possessed by affectees, contain a potentially null preposition selecting the indirect object. Ostensible double object constructions (DOCs) really involve null prepositions at Logical Form (LF) or prepositional deletion at Phonological Form (PF).

Data: The three traditionally recognized dative constructions in Cantonese are:

(1) **Dative A** (S V DO P$_{bei}$ IO)
'Overt-*bei* prepositional'
ngo^5 bei^2/deng^3 [go^2 di^1 syu^1]
1SG give/throw [DEM CL book]
bei^2 [go^2 di^1 jan^2]
DAT [DEM CL person]
'I give/throw the books to the people.'

(2) **Dative B** (S V DO IO)
'Inverse double object'
ngo^5 bei^2 [go^2 di^1 syu^1]
1SG give [DEM CL book]
[go^2 di^1 jan^2]
[DEM CL person]
'I give the books to the people.'

(3) **Dative C** (S V IO DO)
'Regular double object'
ngo^5 gaau^3 [go^2 di^1 jan^2]
1SG teach [DEM CL person]
[go^2 di^1 zi^6]
[DEM CL letter]
'I teach the letters to the people.'

The theme-DO precedes the affectee-IO in A/B, but follows it in C. The DO is always bare. The overt dative preposition *bei^2* introduces IO in A, but not in B/C. Verbs select these constructions: *bei^2* 'give', only A/B; *gaau^3* 'teach', *zuk^1* 'wish', only C; other verbs (*deng^3* 'throw', *waan^4* 'return' etc.), only A.

Analysis: Relativization confirms Tang's (1998) claim that A & B are underlyingly prepositional and derivationally related; prepositional *bei^2* occurs in A at LF & PF, but is optionally deleted in B at PF under haplology with verbal *bei^2* when the DO is phonologically light. B thus does not exemplify the typologically rare Inverse DOC found in Manchester English (Haddican 2010). Furthermore, contra Tang, I argue that C is prepositional rather than a Regular DOC, but C is unrelated to A & B, since its preposition is always null:

(4) **Dative B** (S V DO P̶$_{bei}$ IO) 'Covert-*bei* prepositional'
ngo^5 bei^2 [go^2 di^1 syu^1] **bei^2** [go^2 di^1 jan^2]
1SG give [DEM CL book] DAT [DEM CL person]
'I give the books to the people.'   cf. (2)

(5) **Dative C** (S V P$_\emptyset$ IO DO) 'Null prepositional'
ngo^5 gaau^3 Ø [go^2 di^1 jan^2] [go^2 di^1 zi^6]
1SG teach DAT [DEM CL person] [DEM CL letter]
'I teach the letters to the people.'   cf. (3)

In all constructions, the bare DO is relativizable by gapping. In contrast, the IO triggers obligatory resumption, indicating its dependence on a preposition, whether overt in A, deleted at PF in B, or always null in C:

(6) ngo^5 bei^2/deng^3 [go^2 di^1 syu^1]
1SG give/throw [DEM CL book]
bei^2 *(keoi^{5}dei^2_i) [go^2 di^1 jan^2_i]
DAT 3PL [DEM CL person]
A: 'the people that I give/throw the books to'

(7) ngo^5 bei^2 [go^2 di^1 syu^1] **bei^2**
1SG give [DEM CL book] DAT
*(keoi^{5}dei^2_i) [go^2 di^1 jan^2_i]
3PL [DEM CL person]
B: 'the people that I give the books to'

(8) ngo^5 gaau^3 Ø *(keoi^{5}dei^2_i) [go^2 di^1 zi^6] [go^2 di^1 jan^2_i]
1SG teach DAT 3PL [DEM CL letter] [DEM CL person]
C: 'the people that I teach the letters to'

Similarly, bare DOs license parasitic gaps (Cheung 2015), but prepositional IOs do not, requiring resumption:

(9) a. hai⁶ mat¹je⁵-je⁵$_i$, ming⁴zai² ᴬ[mou⁵ waan⁴-dou³ **PG**$_i$ bei² keoi⁵] / ᴮ[mou⁵ bei²-dou³ **PG**$_i$ bei² keoi⁵] / ᶜ[mou⁵
         is   what-thing Mingzai [NEG return-COMPL   DAT 3SG] / [NEG give-COMPL   DAT 3SG] / [NEG
         gaau³-dou³ Ø   keoi⁵ **PG**$_i$] ji⁴gaa¹ zung⁶ m⁴gin³-zo² e$_i$ tim¹ aa⁴
         teach-COMPL DAT 3SG   ] now   even lose-PFV   PRT Q
         'What is it that Mingzai has now lost without ᴬreturning / ᴮgiving / ᶜteaching to him?'   DO
     b. hai⁶ bin¹go³$_i$, ming⁴zai² ᴬ[mou⁵ waan⁴-dou³ go² di¹ zi¹liu⁶ **bei²** *(**keoi⁵**$_i$)] / ᴮ[mou⁵ bei²-dou³ go² di¹ zi¹liu⁶
         is   who Mingzai [NEG return-COMPL DEM CL info DAT 3SG   ] / [NEG give-COMPL DEM CL info
         bei² *(**keoi⁵**$_i$)] / ᶜ[mou⁵ gaau³-dou³ Ø   *(**keoi⁵**$_i$) go² di¹ zi¹liu⁶] zau⁶ caau²-zo² e$_i$ aa⁴
         DAT 3SG   ] / [NEG teach-COMPL DAT 3SG   DEM CL info] then fire-PFV   Q
         Intended: 'Who is it that Mingzai has fired without ᴬreturning / ᴮgiving / ᶜteaching the info to?'   IO

Conclusion: The variety of ditransitive constructions in Cantonese can be reduced to two prepositional constructions (Dative A/B vs. C).

References: Cheung, Candace Chi-Hang. 2015. "On the fine structure of the left periphery: the positions of topic and focus in Cantonese." In Wei-Tien Dylan Tsai, ed. Cartography of Chinese syntax. Oxford: Oxford University Press, 75–130. Haddican, W. 2010. "Theme-goal ditransitives and theme passivisation in British English dialects." Lingua 120: 2424–2443. Tang, S.-W. 1998. "On the 'inverted' double object construction." In Stephen Matthews (ed.), Studies in Cantonese linguistics. Hong Kong: Linguistic Society of Hong Kong. 35–52.

Can English idioms undergo the dative alternation? A priming investigation

Breanna Pratley, Philip Monahan
University of Toronto
bre.pratley@mail.utoronto.ca, philip.monahan@utoronto.ca

Theoretical Motivation. While the dative alternation in English has two structural options, i.e., Double Object (DO) and Prepositional Dative (PD), idioms with verbs that should alternate are cited as being restricted to the DO (Bruening 2010). This restriction is often used as evidence to support theories in which the DO and PD are construed as entirely distinct structurally (Harley, 1997; Richards, 2001; among others), and evidence against theories that analyse the dative alternation as derivationally related (Larson, 1988; among others); however, these same idioms appear to take on the PD form when the sentence involves Ā extraction, such as heavy NP shift (1b-1c; Bresnan and Nikitina 2007).

(1)   a.   The lighting here gives me a headache.
        b.   *The lighting here gives a headache to me.
        c.   The lighting here gives a headache to everyone in the room.

Bresnan and Nikitina (2007) take idioms in the PD form as evidence for dative alternation theories in which the two structures are derivationally related (e.g., Larson, 1988). There is debate, however, about whether idioms like (1c) are truly PDs, as this would contradict longstanding observations that idioms do not alternate. An alternative hypothesis is that idioms with this surface order are a type of DO that has undergone a process called Rightward Dative Shift ((Figure 1); Bruening, 2010). Crucially, this construction is structurally a DO, but with the goal projected to the right. This results in a surface order similar to the PD.

Current Experiment. To adjudicate between these hypotheses, we conducted a syntactic priming experiment. Participants read prime sentences that were either DO, PD, or Rightward Dative Shift. Then, they made a forced choice between DO and PD target sentences that described a picture. If idioms like (1c) share structure with PD, then the results of the Rightward Dative Shift Condition should pattern like the results of the PD Condition. Syntactic priming targets structural relations (Bock and Loebell, 1990; Pickering et al. 2002). As such, this method can be used to test whether sentences like (1c) are structurally like PD or DO. If a Rightward Dative Shift prime (1c) results in fewer PD responses in test trials than a standard PD prime (see Table 1), idiomatic sentences in this form are not likely to have a PD structure. Our results suggest that these idioms are not structurally similar to Prepositional Datives, contrary to Bresnan and Nikitina's (2007) proposal.

Methods. Native English-speaking participants (n=40) completed 144 trials. In each trial, they were shown a sentential prime, followed by a forced-choice picture description task. Primes were displayed in one of four conditions: Prepositional Dative, Double Object, Rightward Dative Shift, and a Control Condition (Table 1). We created four lists in a Latin Square design. In each trial, participants read the prime aloud, then chose which of two sentences better described a line drawing. The test sentences were presented in the lower portion of the screen, differed only in structure, and were counter-balanced for side of presentation. Forty-eight trials tested the dative alternation, and ninety-six filler trials tested active/passive priming. The active/passive trials were included to ensure that the forced-choice priming task was effective.

Results. Results were submitted to a linear mixed effects model with a logistic regression function (Jaeger 2008), including a fixed effect of Condition, and a maximal random effects structure. Significant priming effects were found in the active/passive condition, (Δ=19% between Active and Passive Conditions), confirming the validity of the forced-choice task. Figure 2 illustrates the dative alternation results in terms of proportion PD response. Test trials after PD primes resulted in significantly more PD responses than after DO (Δ=8%, β=0.36, SE=0.14, z=2.58, p<0.01) or Control primes (Δ=6%, β=-0.29, SE=0.14, z=-1.99, p<.05). There was no difference between the Rightward Dative Shift condition and any other prime condition.

Implications. The rate of PD responses following a Rightward Dative Shift prime is not different from a PD prime; however, unlike PD primes, it is also not different from a DO prime. These results point to many factors influencing syntactic preferences in priming, including perhaps lexical

overlap of *to* in both the Rightward Dative Shift and PD Conditions. If these idioms were truly PD, however, the rate of PD responses in the Rightward Dative Shift Condition should be different from the DO Condition. Therefore, while it is unclear whether the structure in (2) is the correct hypothesis to account for our findings, it is clear that idioms like (1c) are not true Prepositional Datives (cf. Bresnan and Nikitina 2007), which ultimately lends some support to theories which construe the dative alternation as distinct structures (Harley, 1997; among others).

Figure 1
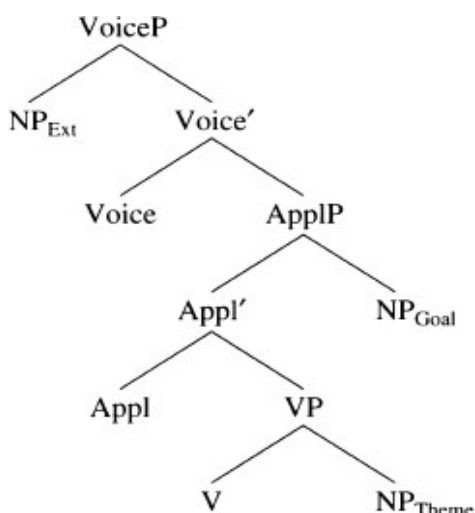
Figure 2
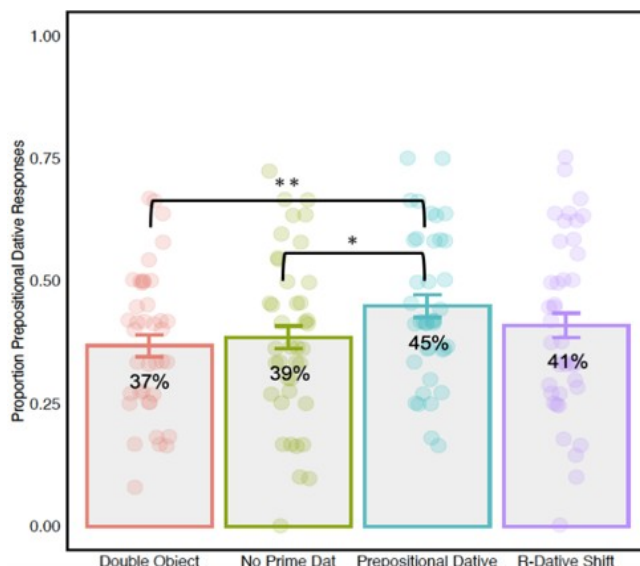
Table 1

Prime Condition	Example Prime	Test Trial
Double Object	The conductor gave the quiet girl on the evening train the ticket	
Prepositional Dative	The conductor gave the ticket to the quiet girl on the evening train	
Rightward Dative Shift	The conductor gave the creeps to the quiet girl on the evening train	
Control	Fully flowery and intricately patterned	The man gave the child a cookie. The man gave a cookie to the child.

References: Bock, Kathryn, and Helga Loebell. 1990. "Framing Sentences." *Cognition*, vol. 35, no. 1, pp. 1-39., doi: 0010-0277/90 Bresnan, Joan, and Tatiana Nikitina. 2007. "The gradience of the dative alternation." *Reality exploration and discovery: Pattern interaction in language and life*, ed. By Linda Uyechi and Lian Hee Wee. Stanford, CA: CSLI Publications. Bruening, Benjamin. 2010. "Double Object Constructions Disguised as Prepositional Datives." *Linguistic Inquiry* 41(2): 287-305. Harley, H. 1997. "If you have, you can give." In *Proceedings of the 15th West Coast Conference on Formal Linguistics*, ed. by Brian Agbayani and Sze-Wing Tang, 193-207. Stanford, CA: CSLI Publications. Jaeger, T. F. 2008. "Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models." *Journal of Memory and Language*, 59(4), 434-446. Larson, R. 1988. "On the Double Object Construction." *Linguistic Inquiry*, vol. 19 no.3, pp. 335-391. www.jstor.org/stable/25164901 Pickering, Martin J., Holly P. Branigan, and Janet F. McLean. 2002. "Constituent Structure is Formulated in One Stage." *Journal of Memory and Language*, 46, 586-605, doi:10.1006/jmla.2001.2824 Richards, N. 2001. "An Idiomatic Argument for Lexical Decomposition." *Linguistic Inquiry*, 32(1), 183-192, doi:10.1162/002438901554649

Arbeitsgruppe 9

The semantics and pragmatics of conditional connectives

Mingya Liu & Mathias Barthel

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

An experimental approach to the semantics and pragmatics of conditional connectives: German *wenn/nur, wenn/wenn und nur wenn*

Mingya Liu, Mathias Barthel
Humboldt University Berlin, Humboldt University Berlin
mingya.liu@hu-berlin.de, mathias.barthel@hu-berlin.de

This paper focuses on the semantics, pragmatics and processing of the lexically related German conditional connectives (CCs) *wenn 'if', nur wenn 'only if' and wenn und nur wenn 'if and only if'*. While in logic *if* is treated as a binary truth-functional connective of material implication, Kratzer (1986) proposes to treat the natural language *if* as a restrictor with no conditional meaning. The interpretation of conditionals has been shown to be subject to semantic/pragmatic modulation, but the modulating role of CCs remains unclear. Logically, modus ponens (MP) should be valid for all conditional sentences. Based on the semantics proposed for *only* in (Horn 2002), *nur-wenn* sentences should also entail the affirmation of the consequent inference (AC). The biconditional status of *nur wenn* is doubted by (Herburger 2015), however. We investigated the meaning of the respective CCs in four experiments.

In **Experiment 1**, participants created two sentences each using the three CCs above. Qualitative analyses of the produced sentences show that both MP and CP are valid for all sentences using *wenn und nur wenn*, indicating its biconditionality. In terms of MP, the same general pattern holds for *wenn* sentences (90% valid) but not for *nur wenn* sentences (63% valid). In contrast, the inference not p → not q was qualified as valid for all sentences with *nur wenn* but only for 35% of *wenn* sentences.

In **Experiment 2**, participants read short scenarios including a conditional sentence with *wenn* or *nur wenn* and a sentence containing the affirmative or negated antecedent proposition (p / not p). An incomplete final sentence had to be completed by participants. Both in the CCs, <1% of responses after a confirmed antecedent (if p, q; p) contained a negative (or downward entailing) consequent (not q). After a negated antecedent (if p, q; not p), however, 11-15% of sentence completions contained a negated consequent, suggesting that for both CCs MP is valid but neither of them was strictly treated as biconditional.

In **Experiment 3**, participants were presented with a conditional sentence (if p, q.) with either *wenn* or *nur wenn* and another sentence containing either the true or the negated antecedent (p/not p). A final sentence contained a question about the truth of the consequent (MP/AC), which participants had to rate on a 5-point Likert scale. A Bayesian ordinal mixed model with CC and Inference (MP/AC) plus their interaction revealed that the biconditional interpretation is most prominent overall. However, *nur wenn* lent itself more to a biconditional reading than *wenn*. Acceptance rates in MP were at ceiling for *wenn*, as expected, but lower for *nur wenn,* casting doubt on the strict biconditionality of the latter. An analysis of decision times for ratings mirrored the results obtained in the ratings.

Structurally similar scenarios used in a self-paced reading task in **Experiment 4** a conditional sentence, a negated antecedent in a follow-up sentence and a negated or non-negated consequent in a final sentence (q / not q). A Bayesian mixed effects regression model with CC (*wenn/nur wenn*) and q (positive/negative) plus their interaction revealed that reading times (RTs) of the positive quantifier in the final sentence were statistically equivalent, but the negative quantifier was read decisively faster in $CC_{nur\ wenn}$ than in the CC_{wenn}, indicating that the meaning not p → not q is activated more strongly by *nur wenn* p, q than by *wenn* p, q.

In conclusion, *wenn und nur wenn* is semantically biconditional. *Nur wenn* and *wenn*, on the other hand, are not biconditional connectives. While for *wenn* all p-cases are q-cases, only some not-p-cases are not-q-cases. For *nur wenn* all not-p-cases are not-q-cases and only some p-cases are q-cases – a novel empirical finding calling for further analytic description.

References: Kratzer, Angelika. 1986. "Conditionals." *Chicago Linguistics Society* 22(2), 1-15. Horn, Laurence R. 2002. "Assertoric inertia and NPI-licensing." *Chicago Linguistics Society 38: Parasession on Negation and Polarity*, 55-82. Herburger, Elena (2015). "*Only if*: If only we understood it." *Proceedings of Sinn Und Bedeutung* 19, 304-321.

Manipulation of nonveridical equilibrium produces negative bias in conditionals

Anastasia Giannakidou
University of Chicago
giannaki@uchicago.edu

Since Kratzer's seminal work, conditional antecedents (or, *protasis*, i.e., *if-clauses*) are taken to merely function as restrictors of (implicit or explicit) modal quantifiers (i.e. quantifiers over worlds or situations). The conditional clause itself is assumed to contain no modal force of its own. Yet it is also known that not all antecedents are created equal, and that complementizer choice correlates with higher or lower degree of commitment to the potential veridicality of the antecedent. The difference between *falls* and *wenn* in German is often discussed in this connection (Reis and Wollstein 2010, and Liu's 2019 notion of *elastic veridicality).* The empirical generalization for *falls* and *wenn* is that while *wenn* appears to have no constraints in its use, the choice to use *falls* implies that the speaker considers *not p* a more likely outcome. I will call this *negative bias in conditionals*. In my presentation, I argue that negative bias in conditionals arises from (a) the existence of a modal layer in the antecedent which (b) manipulates the equilibrium via a meta-evaluating *ranking* function O (Giannakidou 2013, Giannakidou and Mari 2018, 2021, GM) in favor of *not p.* Nonveridical equilibrium reveals that that the speaker considers *p* and *not p* as equal possibilities:

(1)   *Nonveridical equilibrium* (Giannakidou 2003: (14)): An information state M is in nonveridical equilibrium iff M is partitioned into p and ¬p, and there is no bias towards p or ¬p.

Giannakidou calls nonveridical equilibrium 'prototypical inquisitiveness', or 'true uncertainty': it characterizes also questions and possibility modals. A conditional antecedent is a state of equilibrium par excellence: the speaker hypothesizes, and has no (epistemic or doxastic) reason to favor *p* over *not p*. With a modal such as MUST, the function *O* compares *p*-worlds (called Ideal in GM) to *not p* worlds, and positively biases towards *p* by ranking Ideal as better possibilities:

(2)   *Positive* bias of epistemic necessity modals (weak necessity):

Ideal$_S$ is a better possibility than ¬ Ideal$_S$, relative to M(*i*) and *O*; *i* is the speaker, S ordering source

According to GM, the bias producing function *O* is always present in a nonveridical space, and is often realized as an adverb (*She must probably be a genius, She might perhaps be late*). Modal adverbs can also be used in conditionals (including *really*) as well as modal verbs and subjunctives as we see in (3), and of course distinct complementizers. The function of modal elements in conditional antecedents reveals the presence of the modal function *O* which now manipulates in favor of *not p*, thus producing negative bias. Following GM, I argue that the modal elements are realizations of the ranking function *O*. When applied to a state characterized by equilibrium, *O* will bias towards *not p*:

(3)   a. English: If/In case it really /perhaps rains/should rain, I will stay here.
      b. German: Wenn/Falls es wirklich/vielleicht regnet/regnen sollte, bleibe ich hier.

(4)   Negative bias in nonveridical equilibrium contexts:

¬Ideal$_S$ is a better possibility with respect to Ideal$_S$, relative to M(*i*) and *O*.

This manipulation, I will show, characterizes the use of *O* also in questions thus explaining why in both conditionals and questions positive bias modals such as MUST are excluded.

References: Giannakidou, Anastasia. 2013. Inquisitive Assertions and Nonveridicality. In Maria Aloni, Michael Franke, F. Roelofsen (eds.): 115-26.Giannakidou, Anastasia and Alda Mari. 2018. The Semantic Roots of Positive Polarity: Epistemic Modal Verbs and Adverbs. *Linguistics and Philosophy* 30(4): 461-87. Liu, Mingya. (2019). The Elastic Nonveridicality Property of Indicative Conditionals. *Linguistic Vanguard* 5(3).Reis, Marga and Angelika Wöllstein. 2010. Zur Grammatik (vor allem) konditionaler V1- Gefüge im Deutschen. *Zeitschrift für Sprachwissenschaft* 29: 111-179

All that in conditionals

Juliane Schwab, Mingya Liu
Osnabrück University, Humboldt University of Berlin
jschwab@uni-osnabrueck.de, mingya.liu@hu-berlin.de

Conditional antecedents are known as a typical licensing environment for negative polarity items (NPIs), something that is commonly attributed to either their non-veridical nature (Giannakidou 1998) or their scalar properties (von Fintel 1999). We make the novel observation that there are some NPIs, namely understating ones like *all that* (1a/b), which are degraded in indicative compared to subjunctive conditionals. In this talk, we will show how this challenges existing accounts of NPI licensing. Focusing on English *all that,* we will propose a scalar licensing mechanism that builds on Krifka (1995), Israel (1996), and Condoravdi (2010). Then, we will discuss the degradation in indicative conditionals as pragmatic phenomenon.

(1) a. ?If Mary is all that$_{NPI}$ intelligent, she will pass the test.
    b. If Mary were all that$_{NPI}$ intelligent, she would pass the test.

***All that:* An understating NPI (uNPI)**. In much of the existing work on NPIs, the focus has been on indefinites like *any* and *ever*, or on minimizer NPIs like *to lift a finger*. Scalar ap- proaches to polarity sensitivity (Israel 1996; Krifka 1995; *inter alia*) assume that the usage of these expressions is acceptable only if their presence in the sentence makes the assertion stronger than its alternatives. There are, however, also NPIs that have the opposite effect, including, for instance, English *all that*. These NPIs have sometimes been called understating or attenuating (Israel 1996). Within Israel's approach to polarity sensitivity, they are assumed to be restricted to scalar contexts where the proposition with the uNPI is less informative than its alternative, that is, where the proposition is entailed by a contextually provided alternative proposition.

Proposal. Our proposal builds on Israel's work, but formalizes the licensing of *all that* in terms of lexically triggered ordered alternatives (Krifka 1995). We assume that, as degree modifying expression, *all that* lexically evokes (weaker) lower degree alternatives. We then put forward a licensing mechanism (see (2)) that employs and adapts Condoravdi's (2010) revised version of scalar assertion: We argue that for the NPI *all that* to be licensed there must be an alternative P' such that there is a world compatible with the context where P' is true and P' is informationally stronger than P.

(2) **Proposed licensing condition:** $\lambda w.\{w \in c \mid w \in [\![P]\!]_c \wedge \exists P' \in Alt(P)\ (\exists w' \in c \mid w' \in [\![P']\!]_c \wedge c + P +_{str} P' \neq c + P)\}$

Critically, both indicative and subjunctive conditionals satisfy this licensing condition. In our talk, we will show that the degradation of *all that* in indicative conditionals can instead be at- tributed to a conflict between the licensing condition and (strengthening) implicatures present in conditionals.

References: Condoravdi, Cleo. 2010. "NPI licensing in temporal clauses." *Natural Language and Linguistic Theory 28*(4): 877–910. Giannakidou, Anastasia. 1998. "*Polarity sensitivity as (non) veridical dependency*." *Linguistik Aktuell/Linguistics Today*. Amsterdam/Philadelphia: John Benjamins. Israel, Michael. 1996. "Polarity sensitivity as lexical semantics." *Linguistics and Philosophy 19*(6): 619–666. Krifka, Manfred. 1995. "The semantics and pragmatics of polarity items." *Linguistic Analysis 25*(3–4): 209–257. von Fintel, Kai. 1999. "NPI licensing, Strawson entailment, and context dependency." *Journal of Semantics 16*(2): 97–14

Conditional conjunctions informed by Japanese and Korean

Magdalena Kaufmann, John Whitman
University of Connecticut, *Cornell University*
magdalena.kaufmann@uconn.edu, jbw2@cornell.edu

Many languages attest conditional readings for sentential conjunctions (CCs) like *Mary starts singing and John leaves the bar* for `If Mary starts singing, John leaves the bar', e.g. Keshet 2013). Recent accounts argue that thanks to lexical or construction-specific prosodic clues the first conjunct of a CC introduces a hypothetical state of affairs as the topic, respective to which the second conjunct is evaluated (Starr 2018, Kaufmann 2018). We investigate CCs in Japanese and Korean, two languages with morphological topic marking and a rich inventory of conjunctive affixes. Garden variety clausal conjunctions (Japanese –*te*, Korean –*mye/-ko*) in these languages receive only a Boolean interpretation. But CCs result when the first conjunct is topicalized in various ways: (i) overtly marked as a topic (Japanese –*te wa*), (ii) the affix on the first clause is derived from the grammaticalized combination of coordinate affix and topic marker (Korean –*myen*: conjunctive -*mye* plus topic marker =*(nu)n*; Martin 1992), (iii) for Japanese =*to*, normally reserved for NP coordination, we argue that conditionals marked with =*to* are instances of a sentential coordination with a topicalized first conjunct (building on Hasegawa 2017, Koizumi 2000). The Korean and Japanese data suggest that topicalization of the first conjunct in a sentential conjunction itself permits hypotheticality (*pace* Starr 2018 and Kaufmann 2018). Like CCs ambiguous with Boolean conjunctions (as in English), the conjunction-derived hypothetical conditionals in Japanese and Korean can express notions of immediate consequence, causation, or result but not epistemic conditionals (Bolinger 1969). The typological link between topics and conditionals is well known (Haiman 1978), in fact, almost all conditional markers in Modern Japanese (-(*r)eba, -te wa, -tara -to, nara*, Takubo 2020) involve a topic marker (-*ba*, as in -*reba*, -*tara* < *raraba*, *nara* < *naraba*, from Old Japanese topic marker =*pa*, Ono 1974). This poses the question of why only some conditional markers share the semantic restriction characteristic of CCs. We assume that the difference arises because not all connectives derive from fully symmetric coordinations. While markers like –*reba, -tara* and *nara* combine with finite clauses that can, for instance, contain modals, -*tewa* and =*to* conditionals involve syntactically smaller antecedents that express properties of non-maximal situations and cannot describe epistemic possibilities (sets of possible worlds). This receives support from English CCs (Bjorkman 2013, Keshet 2013), and offers novel connections to non-conditional temporal modifier readings of =*to*-clauses when appearing with past tense matrix clauses.

References: Bjorkman, Bronwyn. 2010. "A syntactic correlate of semantic asymmetries in clausal coordination". Proceedings of NELS 41, UPenn. Bolinger, Dwight. 1967. "The imperative in English." To honor Roman Jakobson: Essays on the occasion of his seventieth birthday, Janua Linguarium, Vol. 1. The Hague, Paris: Moulton. 335–362. Haiman, John .1978. "Conditionals are topics." Language 54: 565–589. Hasegawa, Nobuko. 2017. "Modals." In Masayoshi Shibatani, Shigeru Miyagawa, and Hisashi Noda, eds. Handbook of Japanese Syntax. de Gruyter Mouton: 371-402. Kaufmann, Magdalena. 2018. "Topics in Conditional Conjunctions." Talk at NELS 49. Keshet, Ezra. 2013. "Focus on conditional conjunction." Journal of Semantics 30: 211–256. Koizumi, Masatoshi. 2000. "String vacuous overt verb raising." *Journal of East Asian Linguistics* 9: 227-285.Martin, Samuel E. 1992. *A Reference Grammar of Korean.* Rutland, Vt: Charles E. Tuttle. Ono, Susumu. 1974. *Kogo jiten*. [Dictionary of Premodern Japanese]. Tokyo: Iwanami. Starr, Will. 2018. "Conjoining imperatives and declaratives." *Proceedings of Sinn und Bedeutung* 21: 1159-1176. Takubo, Yukinori. 2020. "Conditionals in Japanese." In Wes Jacobsen and Yukinori Takubo, eds. *Handbook of Japanese Semantics and Pragmatics*. de Gruyter Mouton.

Iffy discourse: Japanese *moshi* in conditionals and nominal topics

Muyi Yang
University of Connecticut
muyi.yang@uconn.edu

Japanese conditionals are obligatorily marked by conditional connectives that appear in the end of connectives, but can also be accompanied by *moshi*, a marker that appears at the beginning of antecedents, cf. (1). In addition, there is a previously unnoticed usage of *moshi* in topics-marked nominals, cf. (2). I observe that *moshi* is subject to an "iffiness" requirement in both conditionals and nominal topics, and present an analysis that captures this requirement.

(1) (moshi)   Mary-ga   ki-tara,   John-mo kuru   darou.
     MOSHI   M-NOM   come-COND J-ADD   come   MOD
     "If Mary comes, John will probably also come."

(2) (moshi)   tameshi-ta koto nai kata-wa taiken shi-ta hou-ga   ii   desu yo!
     MOSHI   try-PST   thing NEG ppl-TOP try   do-PST way-NOM good COP   SFP
     lit. "People who haven't tried are s.t. they should." ≈ "If you haven't tried, you should."

Iffiness: In conditionals, *moshi* is disallowed if antecedent proposition the speaker is not "iffy" about the antecedent proposition. One such case is factual conditionals, whose antecedents are presupposed to be true. *moshi* is not allowed in factual conditionals (Arita 2007), as in (3).

(3) A: This curry tastes terrible
    B: (#moshi) karee-ga   sonnnani mazui nara,   nan-mo   oishiku-nai   darou.
         MOSHI curry-NOM so.much   awful COND naan-ADD delicious-NEG   MOD
         "If the curry is so awful, the naan is probably also not tasty."

Iffiness holds for *moshi* in nominal topics, too. Consider (4) for two contexts of (2), uttered by a salesman. As shown by the felicity of (2) in context (a) vs. (b), *moshi* is odd if the speaker has full knowledge about whether the property described by the topic holds for each individual.

(4) a. Some customers said they haven't tried the product, others didn't say anything.
    b. The customers are separated into two groups – one with people who have tried the product, standing on the salesman's left, another with people who haven't, standing on his right.
    (2) with *moshi*: ok in (a), # in (b); (2) without *moshi*: ok in both (a) and (b)

Basic set-up: I assume that a context c is a tuple $\langle w, CS, \Pi \rangle$. w is the world of c. CS represents the set of worlds compatible with the mutual joint beliefs of all participants at w (Stalnaker 1978). Π is the Question Under Discussion (Roberts 2012). Following Gronendijk & Stokhof (1982), I assume a question denotes a equivalence relation between worlds (type $\langle s, st \rangle$). Π_c thus induces a partition of CS_c into sets of worlds agreeing on its (strongly exhaustive) answers.

Proposal: Drawing on the idea that a question makes salient an *n*-place property, where *n* is the number of *wh*-elements (Groenendijk & Stokhof 1982 a.o.), I analyze *moshi* as follows:

(5) $\llbracket$ *moshi* $\rrbracket^c = \lambda p_{\langle \sigma, t \rangle}. p$, defined iff
    (a) (i) if σ = s, Π_c maps a world to a 0-place property of *p,*
          (ii) if σ = e, Π_c maps a world to a 1-place property of *p* over the domain of *e*; and
    (b) The speaker doesn't know the answer of Π_c.

So, in conditionals, *moshi* presupposes a polar question about the antecedent proposition (cf. Starr 2014). In nominal topics, *moshi* presupposes a single *wh*-question regarding the property expressed by the topic (e.g. in (2), Π_c is the question "Which human haven't tried?"). (5) predicts the iffiness requirement. Factual conditionals have antecedents *p* that are presupposed to be true; hence, the partition induced by the presupposed Π_c "Whether *p*?" on *CS* is trivial. Subsequently, the speaker knows the answer of Π_c already, leaving the presupposition in (5b)

unsatisfied. Likewise, in nominal topics, if the speaker knows for each salient individual whether the property named by the topic holds (cf. Context (4b) for (2)), the speaker knows the strongly exhaustive answer of the presupposed Π_c already, thus, again, rendering the pre- supposition in (5b) unsatisfied.

References: Groenendijk, Joroen and Martin Stokhof. 1982. "Semantic analysis of wh-complements." *Linguistics and Philosophy* 5: 172-233

Non-Boolean Conditionals

Paolo Santorio, Alexis Wellwood

University of Maryland, College Park, (USC)

santorio@umd.edu, wellwood@usc.edu

Overview. Standard theories predict that indicative conditionals (ICs) behave in a Boolean fashion when interacting with *and* and *or*. We test this prediction by investigating probability judgments about sentences of the form "$a{\to}b\{$ and/or $\}c{\to}d$". Our findings are incompatible with a Boolean picture.This is challenging for classical theories, while trivalent theories may account for our data.

1. Background. Boolean interpretations of *and* and *or* entail constraints about probabilities of compounds (see e.g. Adams 1998). The following two are relevant here:

***and*-drop.** If A doesn't entail B, $Pr(A) > Pr(A{\wedge}B)$.

***or*-drop.** If A doesn't entail B, $Pr(A{\vee}B) > Pr(A)$.

These constraints apply to all sentences of natural language that express propositions. Thus, if truth-conditional theories of ICs are correct (see a.o. Stalnaker 1968, Kratzer 2012), the sentences in (1) are predicted to conform to the constraints on the right below.

(1) a. If Lea danced, Mia danced, or, If Lea didn't dance, Nina danced.
     b. If Lea danced, Mia danced.
     c. If Lea danced, Mia danced, and, If Lea didn't dance, Nina danced.

$$Pr(1a)>Pr(1b)$$
$$Pr(1b)>Pr(1c)$$

***and/or*-drop** also hold on semantics where ICs don't express propositions, but (i) ICs have probability and (ii) connectives are Boolean (e.g. Van Fraassen 1976, Kaufmann 2009, Bradley 2012).

2. Experiment. Our experiment tests ***and*-drop** and ***or*-drop** for natural language ICs. Subjects were presented with several sentences and asked to perform a likelihood estimation task. Three main variables were m`anipulated: presence and type of connective (And vs Or vs None; within); compatibility of the two antecedents, when sentences involved two ICs (Compatible vs Incompatible; between); and frequency of the event described in the consequent, given the antecedent (50/50 vs 75/25; between). In a training phase, participants viewed 24 animations of 1 shape (Incompatible conditions) or 1-2 shapes (Compatible) traveling by "car" into a "tunnel", whereupon they changed into 1 of 2 colors. Then, participants viewed two sets of 4 "mystery car" an``imations, and gave likelihood estimates for (i) the simple ICs in (2) and (ii) the compounds schematized in (3).`

(2) a.   If the car was carrying the sqare, the sqare turned { red / yellow }. $s{\to}r$, $s{\to}y$
     b.   If the car was carrying the circle, the circle turned { green / blue }. $c{\to}g$, $c{\to}b$
(3) a. $s{\to}r$ { and / or } $c{\to}g$ b. $s{\to}y$ { and / or } $c{\to}b$

Finding. Likelihood estimates were not impacted by the factors Compatibility or Connective, ps > .53.

Discussion. ***and*-drop** or ***or*-drop** predict lower probability estimates for "$s{\to}r$ or $c{\to}g$" over $s{\to}r$, and for $s{\to}r$ over "$s{\to}r$ and $c{\to}g$". This was not observed, revealing non-Boolean behavior.

3. Analysis. Our findings are challenging for standard theories, but can be vindicated by some trivalent theories. In particular, we consider a semantics with the following features. Every clause A has definedness conditions D(A) and truth conditions T(A). A${\to}$B is defined iff A is true and B is defined, and true iff A and B are true. A${\wedge}$B (A${\vee}$B) is defined iff at least one of A and B is defined, and true iff all (at least one of) the defined conjuncts (disjuncts) are true. Combined with a notion of trivalent probability (see Cantwell 2006), this semantics predicts failures of *and*-drop and *or*-drop.

Signalling conditional relations

Markus Egg, Debopam Das
HU Berlin, *HU Berlin*
{markus.egg, dasdebop}@hu-berlin.de

Many discourse relations (DRs) are signalled by connectives like *and* or lexical cue phrases like *as a result*. Recent work (e.g., Redeker et al. 2012, Webber 2013, Das & Taboada 2018, 2019) extends the range of DR signals, e.g., to lexical chains, subject-verb inversion, and lexical sense relations. We investigate subgroups of CONDITIONAL DRs (CONDITION, CONTINGENCY, HYPOTHETICAL OTHERWISE) to test three predictions w hich link the amount of DR signalling to expectedness (Asr & Demberg 2012), but with the full range of discourse signals.

(1) The causality-by-default hypothesis (CBD; Sanders 2005) says that the expected linking of discourse units is causal, so causal DRs should be little marked. (2) The claim that discourse is continuous by preference (Segal et al. 1991, Murray 1997) entails that the temporal order of events or situations is kept in presenting them, i.e., in continuous CONDITIONAL DRs, antecedents are expected to precede consequents; thus, discontinuous CONDITIONAL DRs should be more marked. (3) The hypothesis of uniform information density (UID; Frank & Jaeger 2008) claims that information is spread out evenly across a discourse; this suggests more marking for more informative and hence less expected DRs.

Our data is from the RST Signalling Corpus (RST-SC; Das & Taboada 2018), which includes a wide range of discourse signals. Here the least informative subtype of CONDITIONAL is CONDITION. E.g., HYPOTHETICAL is more informative in that the antecedent must be believed.

CBD is relevant for CONDITIONAL DRs since they are related to CONDITIONAL DRs: both introduce a causal link between antecedent and consequent. This is reflected in classifications, e.g., both groups form the CONTINGENCY group in the PDTB (Webber et al. 2018). They differ in that only CAUSAL DRs have factive arguments. The hypothetical status of arguments in CONDITIONAL DRs is taken to add semantic complexity. Thus, CBD should extend to CONDITIONAL DRs, but in a weaker form than for CONDITIONAL DRs. This is confirmed by the RST-SC (81.4% and 85.7% signalled CAUSAL and CONDITIONAL DRs, with an average 92.7% for all DRs).

Next, the continuity hypothesis suggests that continuous CONDITIONAL DRs are less marked, but in our data, the marking for this subgroup is stronger than for the discontinuous one. Finally, UID predicts the CONDITION subtype to be the least marked in the CONDITIONAL group, as the other subtypes are more informative, but the RST-DT data confirm this claim only for OTHERWISE. CONDITION is more marked than HYPOTHETICAL and CONTINGENCY, with HYPOTHETICAL being the most unmarked subtype, in spite of its high degree of informativity.

Still, these results need not be evidence against the hypotheses, rather, they suggest other interacting factors, e.g., the difference between inter- and intra-sentential relations: Intra-sentential DRs most often call for signalling, mostly by connectives. The high percentage of intra-sentential CONDITION DRs is in our view responsible for its consistent signalling, even though its informativity is lower than the one of HYPOTHETICAL DRs. This subtype occurs predominantly inter-sententially, hence, shows less marking despite its higher informativity.

References: Feresteh Asr & Vera Demberg. 2012. Implicitness of discourse relations. *COLING 2012*. Debopam Das & Maite Taboada. 2018. RST Signalling Corpus. *Lang. Resources and Evaluation.* 52. Debopam Das & Maite Taboada. 2019. Multiple signals of coherence relations. *Discours*. Austin Frank & Florian Jaeger. 2008. Speaking rationally: Uniform information density as an optimal strategy for language production. *30th Annual Meeting of the Cognitive Science Society*. John Murray. 1997. Connectives and narrative text. *Memory and Cognition*, 25. Gisela Redeker, Ildikó Berzlánovich, Nynke van der Vliet, Gosse Bouma, & Markus Egg. 2012. Multi-Layer Discourse Annotation of a Dutch Text Corpus. *LREC 2012.* Ted Sanders. 2005. Coherence, causality and cognitive complexity in discourse. *Proceedings of the Symposium on the Exploration and Modelling of Meaning*. Erwin Segal, Judith Duchan & Paula Scott. 1991. The role of interclausal connectives in narrative structuring, *Discourse Processes*, 14. Bonnie Webber. 2013. What excludes an alternative in coherence relations? *IWCS 2013*. Bonnie Webber, Rashmi Prasad, Alan Lee, & Aravind Joshi. 2018. *The PDTB 3.0 Annotation Manual.*

A causal relevance analysis of (hidden) conditionals

Robert van Rooij, Katrin Schulz
ILLC, University of Amsterdam, *ILLC, University of Amsterdam*
r.a.m.vanrooij@uva.nl, k.schulz@uva.nl

The following sentence is inappropriate and misleading: `If it was sunny in Berlin yesterday, there are COVID-19 casualties in Brazil today.' Why? Because this conditional sentence strongly suggests that what happened in Berlin (the antecedent) is *relevant* to what happens in Brazil (the consequent). The sentence is misleading because we know no such relevance `relationship exists.

Although what makes this conditional sentence misleading is clear, the standard semantic analyses of conditionals do not predict that anything is wrong with it, because they do not include relevance as part of the meaning of conditionals. According to one such theory (e.g. Adams, 1976), the only thing that counts for the meaning and use of conditional is that the consequent is likely, or probable, given the antecedent. Given that the consequent of the above conditionals is likely, or even certain, it is falsely predicted that the whole conditional sentences is appropriate to use. The same wrong prediction follows from the other popular analysis of conditional sentences (e.g., Kratzer, 2012), which demands, instead of relevance, that the consequent is true in all most similar/normal antecedent worlds.

We will discuss two ways to tackle this problem: according to the first *pragmatic* proposal, the semantics is just like the above, but relevance comes out because of the *implicature* that the consequent is not believed. On the contrasting *semantic* solution (Douven, 2008), relevance is built in into the meaning of the conditional. We discuss two probabilistic ways to work out such a semantic approach. On the first semantic analysis, 'If A, then C' is assertable if $\Delta P_A^C = P(C|A) - P(C|\neg A) > 0$. This analysis seems natural, because the notion ΔP is used frequently to measure the learned association between A and C. Unfortunately, the use of this notion gives rise to various empirical problems: for instance, the inferences of 'contraposition' and 'denying the antecedent' are (falsely) predicted to be valid, and this analysis predicts far too many counterexamples to transitivity. According to the second semantic solution, the conditional expresses a *causal relation* between A and C. Mostly, this is the causal power of A to generate C, which can normally be captured by $\Delta * P_A^C = \frac{P(C|A, B) - P(C|\neg A, B)}{1 - P(C|\neg A, B)}$, with B the relevant causal background. At other times, causal power reduces to the conditional probability $P(C|A)$. This second relevance-based semantic analysis gives much better predictions than the first. The relation with the pragmatic analysis is much more controversial.

We will also argue that this causal analysis is natural for many types of sentences that are normally analysed as *hidden conditionals,* such as generic and habitual sentences and disposition attributions. It can explain, for instance, why `Primary school teachers are female' is normally considered false, although `Primary school teachers are *ususally/normally* female' is not. In addition, we will argue that our causal analysis can naturally explain the *inherence bias* for generic and habitual sentences: the assumption that behavior is caused primarily by the agent's inherent disposition instead of the external situational characteristics to which the agent respond.

References: Adams (1976), 'Probability and the Logic of Conditionals'; Cimpian & Salomon (2014), 'The inherence heuristic'; Douven (2008), 'The evidential support theory of conditionals'; Kratzer (2012), Modals and Conditionals ; Skovgaard-Olsen, Singmann & Klauer (2016), 'The relevance effect and conditionals'; van Rooij & Schulz (2019), 'Conditionals, causality and conditional probability'.

Indicatives, subjunctives, and the falsity of the antecedent

Niels Skovgaard-Olsen, Peter Collins
University of Göttingen, *University of Greenwich*
niels.skovgaard-olsen@psych.uni-goettingen.de, p.j.collins@greenwich.ac.uk

In this paper, several hypotheses concerning the doxastic status of the antecedent in conditionals are investigated.

In a first strand, conversational implicature and presupposition accounts of the falsity of the antecedent of subjunctive conditionals are tested via experimental tasks that apply the family of sentences test (Kadmon, 2001) to investigate whether belief-state assumptions concerning the antecedent of conditionals project under operators like negations, possibility modals, and interrogatives (Experiments 1, 1a) and tasks that compare the cancellability of belief-state assumptions concerning the antecedent of indicative and subjunctive conditionals with cancellation of entailments and scalar implicatures (Experiment 2).

In our experiments, it is found that the results across studies are most consistent with a conversational implicature hypothesis of the belief-state assumptions. These results have a bearing on various discussions at the interface of psychology and linguistics. In psychology, it has, for instance, been common to speak within mental model theory of the falsity of antecedent and consequent as part of the default meaning (e.g. Khemlani, Byrne, & Johnson-Laird, 2018) but also the "presupposed facts" (see, e.g., Byrne, 2005, 2016, 2017). In linguistics, implicature-based accounts of the falsity of the antecedent have proved popular (see, e.g., Iatridou, 2000; Ippolito, 2003; Leahy, 2011). But they are yet to be tested experimentally.

In a second strand, a novel experimental task is developed for testing the highly influential, but experimentally underexplored, possible worlds account of subjunctive conditionals (Stalnaker, 1968; Lewis, 1973). As a novelty of this study, it is found that a possible worlds semantics is capable of accounting for participants' truth value assignments in this task. Moreover, a new finding concerning accommodation in truth value assignments to indicative conditionals is reported.

References: Byrne, R. M. J. (2005). The Rational Imagination: How People Create Alternatives to Reality. MIT Press. Byrne, R. M. J. (2016). Counterfactual Thought. Annual Review of Psychology, 67(1), 135– 157. Byrne, R. M. J. (2017). Counterfactual Thinking: From Logic to Morality. Current Directions in Psychological Science, 26(4), 314–322. Iatridou, S. (2000). The Grammatical Ingredients of Counterfactuality. Linguistic Inquiry, 31(2), 231–270. Ippolito, M. (2003). Presuppositions and Implicatures in Counterfactuals. Natural Language Semantics, 11(2), 145–186. Kadmon, N. (2001). *Formal Pragmatics*. Malden, MA: Blackwell Publishers. Khemlani, S. S., Byrne, R. M. J., & Johnson-Laird, P. N. (2018). Facts and Possibilities: A Model-Based Theory of Sentential Reasoning. *Cognitive Science*, *42*(6), 1887–1924. Leahy, B. (2011). Presuppositions and Antipresuppositions in Conditionals. *Semantics and Linguistic Theory*, *21*, 257. Lewis, D. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press. Stalnaker, R. C. (1968). A Theory of Conditionals. In: Rescher, N. (Eds.), *Studies in Logical Theory (pp. 98-112)*. Oxford: Basil Blackwell.

Jiu-conditionals in Mandarin Chinese

Mingya Liu, Yuting Wang
Humboldt University of Berlin, *Nanjing University*
mingya.liu@hu-berlin.de, wangyuting0205@hotmail.com

The particle jiu in Mandarin shows different distributional properties Some researchers argue for two basic kinds: it can have left or right associates, see (1), with jiu1 unstressed and jiu2 stressed (Lu 1984). In addition, jiu can also function as a conditional particle, see (2).

(1) *Zhangsan jiu   chi-le   yi-zhi   ji.*
    Zhangsan jiu1/jiu2   ate-ASP   one-CL chicken
    a. [**Zhangsan**]$_F$ jiu1:   Zhangsan alone ate a chicken. (rough translation)
    b. jiu2…[**yi zhi ji**]$_F$:   Zhangsan only ate a chicken.

(2) *ni   qu,   ta   jiu   kai   che.*
    you   go   she   jiu   drive car
    a. Assertion: If you go, she will drive.
    b. SI: It is less likely that she will drive with **you going** than with other contextual alternatives (e.g. someone else / more than one person going).

We propose a uniform analysis for *jiu1* and *jiu* in conditionals, henceforth, simply jiu1. The antecedent CP (p) in *jiu*-conditionals is base-generated in the consequent TP (q), which undergoes movement to the left periphery, possibly, yielding a (contrastive) topic (see Haiman 1978, Pan and Paul 2018). Furthermore, we propose that *jiu1* triggers a scalar inference (SI), which is computed based on the alternatives of its left associates, see (2b), cf. Kellert's (2018) treatment of Italian *già* 'already'. In entailment-cancelling contexts such as questions, the SI survives, indicating that it is non-at-issue. While we do not have a fully developed answer as to how the conditionality is derived in *jiu*-conditionals, any account needs to take into account the fact that not only *jiu* but also other particles (e.g. *cai* 'only' or *ye* 'also') trigger conditionality without conditional connectives. Instead of treating these as cases of conditional conjunction as Liu (2017), in the spirit of Klinedinst and Rothschild (2012), we propose a pragmatic account: the (non)veridicality property of the first clause p (Giannakidou 1998) is decided by the context. The parser interprets the entire sentence, based on the context and/or intonational and stress patterns of p and *jiu.* If p in (3) is presupposed, the entire sentence gets a conjunctive reading; otherwise, it gets a conditional reading. Without contextual support, the sentence is ambiguous. We assume that the basis for the selection of p as a condition (when p is nonveridical) from a list of possible conditions - for q, is "the extra-linguistic principle of relevance" (Haiman 1978).

(3) *Ta shi yisheng jiu hui zheyang xiang.* (he is doctor *jiu* can this-way think)
    a. jiu1: If he is a doctor, he already can think so.
    b. jiu2: He is a doctor. Thus, he only can think so.

References: Lu, Bingfu. 1984. Fuci *jiu* de yixiang fen he wenti (The diversity and uniformity question of the adverb "jiu"), *Chinese Language Learning* 1: 31-34. Haiman, John. 1978. Conditionals are topics. *Language* 54: 564-589. Giannakidou, Anastasia. 1998. Polarity Sensitivity as (Non)veridical Dependency. John Benjamins, Amsterdam-Philadelphia. Liu, Mingming. 2017. Mandarin conditional conjunctions and *only*. *Studies in Logic* 02: 45-61. Kellert, Olga. 2018. PPIs under negation: A case study of Italian *già*. *Linguistics* 56(2): 333-359. Pan, Victor Junnan, and Waltraud Paul. 2018. The syntax of complex sentences in Mandarin Chinese. *Linguistic Analysis* 42 (1-2): 63-161.

Past tense morphology and the choice of connectives in Korean counterfactual conditionals

Jiyeong Kim, Sung-Eun Lee
Seoul National University
kimjy04@snu.ac.kr, cristlo5@snu.ac.kr

Counterfactual conditionals convey the speaker's belief that the proposition in the antecedent is contrary to fact. As Iatridou (2000) claimed, past tense morphology is the hallmark of subjunctive conditionals cross-linguistically, which is also the case for Korean (Han,2006). Besides, Korean has two types of conditional connectives*, -myen* and *-tamyen*. The connective *-tamyen* has been related with the speaker's hypothetical or irrealis attitude while *-myen* is a conditional connective in any type of attitude. However, some debates have arisen regarding the hypothetical property of *-tamyen* since it is freely interchangeable with *-myen* and it can also be used in a situation where the antecedent clause is quoting another's utterance (Bak, 2003; Noh, 2009). This paper examined how past tense morphology and two different types of connectives in Korean conditionals contribute to counterfactual interpretation using an experimental method.

Thirty Korean native speakers participated in a naturalness judgement task (7-point Likert scale). A conditional sentence, differing in two levels (Past vs. Non-past, *-myen* vs*. -tamyen* in antecedent) was first presented to the participants as a context. The participants then judged the following sentence indicating realization or non-realization of consequence in the prior conditional. If the antecedent of a conditional sentence is interpreted as having a strong counterfactual meaning, realization of a consequence condition will be more likely to be judged as unnatural while the unrealized condition will be natural.

The results show that counterfactual meaning in Korean conditionals is affected not only by the use of past tense morphology but also by the choice of connectives. Significantly lower naturalness was observed in the realized consequence condition when past tense was marked in the antecedent. Furthermore, when *-tamyen* was used, counterfactual interpretation was reinforced when past tense was marked, supported by lower naturalness in the realized consequence condition.

The effect of using connective *-tamyen* over *-myen* can be explained with scalar implicature based on a relation of asymmetric entailment between two connectives (Ippolito, 2003). Also, the mechanism how *-tamyen* contributes to counterfactual meaning was discussed with evidential property it has.

References: Bak, Sung-Yun. 2003. "Conditionals in Korean revisited." *Discourse and Cognition* 10.2:25–52.Han, Chung-hye. 2006. "Variation in form-meaning mapping between Korean and English counterfactuals." *Journal of East Asian Linguistics* 15.2: 167-193. Iatridou, Sabine. 2000. "The grammatical ingredients of counterfactuality." *Linguistic inquiry* 31.2: 231-270. Ippolito, Michela. 2003. "Presuppositions and implicatures in counterfactuals." *Natural language semantics* 11.2: 145-186. Noh, Eun-Ju. 2009. "The Korean conditional markers *myen* and *tamyen*: epistemicity vs. modes of language use." *Journal of East Asian Linguistics* 18.1: 21-39.

Conditional perfection in causal and conventional conditionals

Natalia Zevakhina, Veronika Prigorkina
HSE University
nzevakhina@hse.ru, vaprigorkina@edu.hse.ru

Conditional Perfection (CP) introduced in Geis and Zwicky (1971) has been acknowledged as a quantity inference: *If you mow the lawn, I'll give you 5$ ~> If you don't mow the lawn, I won't give you 5$*. This paper studies the role of negation and clause order in derivation of CP in causals and contingent universals. Following Fillenbaum (1975), Prediction A is that contingent universals yield derivation of CP to a lesser extent than causals. Following Evans and Newstead (1977) and Schroyens et al (1999), Prediction B is formulated as follows.

Reading times for conditionals with no negation are faster than reading times for conditionals with single negation (negation in the consequent is processed faster than negation in the antecedent), and single negation is less time-consuming than negation in both the antecedent and consequent. Prediction C is that reaction times to inferential questions with no negation is faster than conditionals with single negation, and single negation is less time-consuming than negation in both the antecedent and consequent of a conditional. Following Plogmann (2011), Prediction D is that the amount of yes-responses for direct (if p, q) order won't differ from the amount of responses for inverse (q, if p) order. To verify the predictions, two studies were conducted in Russian via IbexFarm. The 1st study tested derivation of CP and had a 2 x 2 x 2 design: Conditionals (causals vs. contingent universals) x Negation in antecedent (negation vs. no negation) x Negation in consequent (negation vs. no negation). Control items were contexts expected to receive *yes* or *no* answers. The critical and control items were designed as a combination of reading, reaction and inference tasks. For each critical/control context, 36 people (21 fem., m.a.=27) had to press a gap after reading one sentence, so that the following sentence appeared on the screen. Participants' reading times were recorded.

The last sentence of each context was formulated as an inference with negated conditionals and followed with *yes* (key "J") and *no* (key "G"). Participants' reaction times were recorded. *Yes*-responses for critical items indicated derivation of CP. The 2nd study also tested derivation of CP in causals and contingent universals. However, instead of negation, clause order (direct: if p, q vs. inverse: q, if p) was taken as a factor (2 x 2 design). The materials and procedure were similar to the ones of the 1st study. In overall, 33 people (18 fem., m.a.=30) took part in the 2nd study. Mixed-effects logistic regression for yes/no answers and Linear mixed- effects model for reaction/reading times confirmed prediction A and partially confirmed predictions B, C and D. In overall, the results showed that the two types of conditionals are pro- cessed quite differently and are affected by negation and order in a different way.

References: Evans, Jonathan, and Stephen Newstead. 1977. "Language and reasoning: A study of temporal factors." *Cognition* 5: 265-283. Fillenbaum, Samuel. 1975. "If: some uses." *Psychological Research* 37: 245-260. Geis, Michael and Arnold Zwicky. 1971. "On invited inferences." *Linguistic inquiry* 2: 561-566. Plogmann, Marie-Christine. 2010. "Focus affects conditional perfection: experimental evidence." PhD thesis. Goethe University Frankfurt. Schroyens, Walter, Walter Schaeken, Niki Verschueren, and Gery d'Ydewalle. 1999. "Conditional reasoning with negations: matching bias and implicit versus explicit affirmation or denial." *Psychologica Belgica* 39: 235-258.

Conditional connection explored: The case of Sicilian *cusà*

Luisa Brucale, Egle Mocciaro, Maria Cristina Lo Baido
University of Palermo, University of Brno**,** *University of Cagliari*
luisa.brucale@unipa.it, egle.mocciaro@gmail.com, mariac.lobaido@unica.it

The paper explores the discourse profile of *cusà* (wh-+know:3ps; see also Spanish *quiza(s)*, ancient Portuguese *quiga*, English *who knows*, Dutch *wie weet*), which in Sicilian was described as adverb, interjection or conjunction (Piccitto & Tropea 1977). After conventionalising as a modal adverb with epistemic functions (Fortuna 2002), *cusà* acquired a set of functions ranging from near-connective in conditional contexts to adverbial related to the speaker's epistemic stance, to contrastive-corrective and, finally, to pragmatic functions.

We will analyse the grammaticalisation path that led the original wh-question to the encoding of (inter)subjective functions (Traugott 1995), taking into account both the semantic and the syntactic characteristics of the marker, as well as the mechanisms of formal reduction it underwent over time (Hopper & Traugott 2003; Heine, Claudi & Hünnemeyer 1991).

The analysis is based on a survey conducted on CMC ('computer-mediated communication') data and on diachronic data. In addition, a questionnaire has been administered to a sample of 200 Sicilian speakers.
The analysis brought to light the following functions for <u>cusà</u>:

(1)   it heads *wh-* and *if*-clauses encoding the speaker's epistemic stance
(2)   it specifies the dubitative value expressed in a clause introduced by *si* ('if')
(3)   it acts as a near conditional connective
(4)   it acts as a full conditional connective thus introducing non-predictive and meta-discursive conditionals (Dancygier 1998: 141; Lombardi Vallauri 1999)
(5)   it behaves like a complex epistemic adverb (Mauri & Sansò 2014)
(6)   it plays the pragmatic function of mitigation
(7)   it acts as a dubitative corrective adverb
(9)   it constitutes a conversational turn.

We will show that *cusà* acquired a complex conditional function often in co-occurrence with meta-discursive and meta-communicative conditionals. The relation of conditionality may reach the realm of dubitative corrective functions, where *cusà* challenges a given presupposed statement and introduces the correct consequence. We will also argue that the reference to conditionality may be exploited to convey pragmatic meanings of attenuation and of insinuation.

References: Dancygier, Barbara. 1998. Conditionals and prediction. Cambridge: Cambridge University Press. Fortuna, Antonella. 2002. Grammatica siciliana: principali regole grammaticali, fonetiche e grafiche (comparate tra i vari dialetti siciliani). Caltanissetta: Terzo Millennio. Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. Grammaticalization: A conceptual framework. Chicago: The University of Chicago Press. Hopper, Paul, and Elisabeth C. Traugott. 2003. Grammaticalization, 2 ed. Cambridge: Cambridge University Press.Lombardi Vallauri, Edoardo. 1999. Grammatica funzionale delle avverbiali italiane. Roma: Carocci. Mauri, Caterina, and Andrea Sansò. 2014. "Pathways to conditionality: two case studies from Italian." *Archivio Glottologico Italiano* 99: 97-121. Piccitto, Giorgio, and Giovanni Tropea. 1977. *Vocabolario Siciliano*. 5 Vol. Catania: Centro di studi filologici e linguistici siciliani, Opera del Vocabolario siciliano.Traugott, Elisabeth C. 1995. Subjectification in grammaticalisation. In Dieter Stein, and Susan Wright, eds. *Subjectivity and subjectivisation*. Cambridge: CUP. 31-54.

Conditional constructions in Spanish: Overt connectives, ellipsis, and juxtaposition

Laura Margarita Merino Hernández
Indiana University
lmerino@iu.edu

Cross-linguistically speakers use a wide variety of morphosyntactic structures to express conditionality including juxtaposition, non-finite protases, and subordination including the prototypical structure *if p, (then) q* (e.g., Elder 2019; Montolío 1999). Furthermore, conditional constructions (CCs) intersect with other types of constructions such as adversativity and causality (e.g., Kortmann 1997; van Rooij & Schulz 2019). The purpose of this talk is to present an empirically based inventory of CCs in Spanish and the sociolinguistic factors that condition their use.

The data comes from 32 speakers of Mexican Spanish who were presented with an opinion interview and a contextualized situations task designed to elicit CCs. Based on previous research (e.g., Sweetser 1990; Elder 2019) a CC had to meet the following criteria: (a) have a protasis and an apodosis implicitly or explicitly realized, (b) the protasis is a sufficient, but not necessary, condition that causes or enables fulfilment of the apodosis, (c) The antecedent is uncertain (not known to be true) by the speaker, and (d) the construction can be replaced by a canonical *if*-clause (regardless of tense and mood shift). Each CC was coded for five linguistic variables: marking of the protasis and marking of the apodosis (e.g., modal adverb or *then*), verbal form of the protasis, verbal form of the apodosis, and surrounding discourse (e.g., appearing embedded in an adversative clause *but if*), and three social variables: sex, age, and level of education. The data was then submitted to a linear mixed effects logistic regression in SAS 9.4.

A total of 977 CCs were identified, which were divided into 35 types grouped into three major categories: overt connective, elliptical, and juxtaposition. Overt connectives constitute 43% of the data (N = 418), followed by elliptical constructions (those without an overt protasis) 34% (N=337) and juxtaposed clauses without a connective 23% (N = 222). Thus, the preferred method for marking a conditional relation is overtly via a connective. The results from the statistical analysis show that juxtaposed clauses have a tendency of not marking the protasis (p = 0.02), but the odds of marking the apodosis are stronger (p = .0001). Juxtaposed conditional protases (p = .0001) and apodoses (p = .0001) have a higher probability of occurring with a verb in the present tense (p = .0001) when compared to nonfinite/verbless protases and other conjugated forms like the subjunctive. Finally, neither the surrounding discourse nor any of the social variables were significant.

These results show the heterogenous nature of CCs and highlight the importance of analysing conditionality beyond overt markers like *if* or *in case that*. Furthermore, research on CCs have usually focused on the protasis, however, this work also underscores the importance of the apodosis's marking as when there is no overt connective it is usually the independent clause the one who carries the load of triggering the conditional inference (e.g., via a modal adverb like *tal vez* 'maybe').

References: Elder, Chi-Hé. 2019. Context, cognition and conditionals. Switzerland: Springer. Kortmann, Bernd. 1997. Adverbial subordination: A typology and history of adverbial subordinators based on European languages. Berlin: de Gruyter. Montolío Durán, Estrella. 1999. Las construcciones condicionales. Gramática descriptiva de la lengua española, eds. Ignacio Bosque and Violeta. Demonte, 3643-738. Madrid: Espasa Calpe. van Rooij, Robert and Katrin Schulz. 2019. Conditionals, causality and conditional probability. Journal of Logic, Language and Information, 28(1): 55-71.

Emojis and conditionals: Exploring the super linguistic interplay of expressive modifiers and conditional meaning

Patrick Georg Grosz
University of Oslo
p.g.grosz@iln.uio.no
(based on joint work with Elsi Kaiser, Gabriel Greenberg and Christian De Leon)

Face emojis can be analyzed as expressive modifiers (Grosz, Kaiser & Pierini fc.; Maier 2020), similar to expressives such as *f*ing* and the emotive marker *alas* (Rett 2020). As formal semantics is venturing into the integration of extra-linguistic devices such as emojis with the text that they accompany, our focus turns on examples such as (1) and (2), where the 'worried face' emoji (🙁) gives rise to different inferences for the counterfactual conditional (1) and the indicative conditional (2). Intuitively, the emoji targets the implicature (1b) of the counterfactual (1a), whereas it targets the implicature (2b) of the indicative (2a), giving rise to the emotive inferences in (1c) and (2c).

(1)   a.   if the movie weren't violent, Sam would love it 🙁
       b.   *implicature:*   the movie is violent and Sam does not love it
       c.   *inference:*   **I am sad that** Sam does not love the movie.

(2)   a.   if the movie is violent, Sam will hate it 🙁
       b.   *implicature:*   the movie may be violent and Sam may hate it
       c.   *inference:*   **I am worried that** Sam will hate the movie.

In this talk, I provide a first exploration of the interactions of face emojis with the presuppositions and implicatures that arise from conditional constructions. I also explore implications of the findings for expressives and emotive markers in conditionals, as illustrated by (3a) and (3b).

(3)   a.   if the movie weren't violent, Sam would love it, **alas**!
       b.   if this wasn't **f*ing** true, I would laugh

References: Grosz, Patrick Georg, Elsi Kaiser, and Francesco Pierini. forthcoming. "Discourse Anaphoricity and First-Person Indexicality in Emoji Resolution." *Proceedings of Sinn und Bedeutung 25.* URL of SuB25 presentation: https://osf.io/rdhc9/ Maier, Emar. 2020. The Semantics of Smiles and Smileys. Talk at *Sinn und Bedeutung 25*. URL: https://osf.io/mhejw/ Rett, Jessica. 2020. The semantics of emotive markers and other illocutionary content. Manuscript, UCLA
URL: https://linguistics.ucla.edu/people/Rett/Rett_2020_emotive_markers.pdf

What embedded counterfactuals tell us about the semantics of attitudes

Nina Haslinger, Viola Schmitt
Universität Göttingen, Humboldt-Universität Berlin
nina.haslinger@uni-goettingen.de, viola.schmitt@hu-berlin.de

Background: Semantic analyses of *believe* often rely on a subject-dependent set DOX of doxastic alternatives (Hintikka 1969). Counterfactuals (CF) embedded under *believe,* like (1), raise a *prima facie* problem for this view: Just as unembedded CFs convey that their antecedent is false, (1) suggests Ada does *not* believe Bert won the election. So to evaluate counterfactuals in *believe*-contexts, we must go beyond the subject's doxastic alternatives.

(1)   *Ada believes that if Bert had won the election, there would have been a revolt.*

Lewis (1973) a.o. assumes that CFs are sensitive to a partial ordering $\leq_w$ on possible worlds relativized to an evaluation world w, which encodes a 'closer to w' relation essentially determined by the generalizations holding in w: Roughly, a CF is true iff the consequent is true in all the most plausible worlds (given the generalizations of w) where the antecedent holds.

One could combine this with Hintikka's (1969) analysis of *believe* by evaluating the CF 'distributively' for every world in DOX: (2) is then true iff for each of Ada's doxastic alternatives *w',* there is a revolt in all of the most plausible worlds *relative to w'* where Bert won the election. **Problem**: This treatment runs into a problem very similar to that raised by Yalcin (2007) for epistemic modals under attitudes: The domain of quantification for a modal embedded under *believe* should be determined separately for each doxastic alternative. Consider (4a) with the German existential modal *möglicherweise* in an embedded CF.

(2a)   *Die Ada glaubt, dass wir jetzt möglicherweise Kaffee trinken würden, wenn das Paket nicht gestohlen worden wäre.*   **true** in (2b)
'Ada believes that we would possibly be drinking coffee now if the package hadn't been stolen.'

(2b)   SCENARIO: Ada had intended to order two things: a toaster and a coffee maker. However, she ultimately ordered only one thing and forgot which. Today, she was told that a package had been delivered. She cannot find the package. Her neighbors say it was stolen.

A 'distributive' analysis for CFs predicts (2a) to be false in (2b): In some doxastic alternative *w',* Ada only ordered a toaster, so in all of the closest worlds to *w'* making the antecedent true, no coffee maker was delivered. Hence, it is *not* the case that for every doxastic alternative *w',* Ada gets a coffee maker in some of the closest worlds to *w'* where the package wasn't stolen. **Proposal**: To avoid the incorrect 'double quantification', we combine Lewisian ordering semantics with domain semantics (Yalcin 2007). Yalcin accounts for analogous examples without CFs by positing complex indices that consist of a world parameter and a domain parameter. The latter provides the quantificational domain for epistemic modals and is shifted by attitude predicates. We extend this by making the domain parameter a partially ordered set, where the ordering reflects a belief state. The minimal worlds in the ordering are the doxastic alternatives. In addition to quantifying over the world component of an index, attitude verbs like believe shift the ordering component to the ordering $\leq_{w,x}$ reflecting the attitude subject x's belief state in w. For simple belief ascriptions with a subject x and no further modals, this won't have any effect: They only quantify over indices involving $\leq_{w,x}$ whose world component is one of x's doxastic alternatives. But embedded CFs are sensitive to the other worlds in the ordering: (1) quantifies over the minimal worlds wrt. $\leq_{w,\text{Ada}}$ in which Bert won the election. Similarly, the CF in (2a) picks out the $\leq_{w,\text{Ada}}$-minimal worlds in which the package wasn't stolen, which makes (2a) true in scenario (2b). **Consequences**: The proposal predicts that there should be other expressions that are sensitive to counterfactual beliefs when embedded under attitudes. This is borne out by 'Hob-Nob' sentences (Geach 1967) and cumulative belief sentences, which involve intentional identity and distinctness relations that can be stated in counterfactual terms.

References: Geach, Peter. 1967. "Intentional identity." *Journal of Philosophy* 64: 627-632, 1967. Hintikka, Jaakko. 1969. "Semantics for Propositional Attitudes." In D. Hockney and W. Wil- son, eds. *Philosophical Logic.* Dordrecht: Reidel. 21-45. Lewis, David. 1973. *Counterfactuals*. Cambridge: Harvard University Press. Yalcin, Seth. 2007." Epistemic modals." *Mind* 116: 983-1026.

Modals as a diagnostic for biconditional vs. material interpretations of conditionals

Magdalena Sztencel, Sarah E. Duffy
University of Gdańsk, Northumbria University
magdalena.sztencel@ug.edu.pl; sarah.duffy@northumbria.ac.uk

Consider the use of the '*q, only if p*' conditional and the modal verb *should* in the following example of practical reasoning taken from Kratzer (2012):

> a. I want to become a mayor.
> b. (*q*) I will become a mayor only if (*p*) I go to the pub.
> c. Therefore, I should go to the pub.

Given what the cogniser wants (a) and the relevant circumstances (b), the conclusion that the cogniser goes to the pub follows necessarily. Hence, the use of the necessity modal in (c). Indeed, given the context of (a), the necessity modal in (c) is simply a reflection of the necessity of *p* for *q*, which is lexicalised in the '*q, only if p*' construction. In this talk, we look into whether modalized indirect reports of conditionals (i.e. reports which involve the use of a modal verb) lexicalise the necessity of *p* for *q* in cases where necessity is contextually available but not lexicalised in a conditional, as in '*if p, q*' formulations.

We report on two online experiments into the relation between (i) contextually available necessity or sufficiency of the truth of a conditional antecedent for the truth of the consequent, and (ii) the formulation of an indirect report of a conditional with necessity or possibility modals (*have to, should* or *could*, respectively). Following Politzer (2004), we assumed that the conditional comes with an implicit guarantee of normality. Accordingly, the conditionals used in Experiment 1 were believable in the sense that the lack of satisfaction of relevant complementary necessary conditions (CNCs) or the presence of alternative consequents was not suggested in the co-text. We hypothesised that, if a conditional is believable, the presence of alternative antecedents in the co-text (Condition 1) should result in the sufficiency of *p* for *q* interpretation of the conditional and trigger an indirect report with *could*. On the other hand, the existence of no reasonable or salient alternative antecedents(Condition 2) should result in the necessity of *p* for *q* interpretation and trigger an indirect report with *should* or *have to*. 139 native English speakers worked under Condition 1 or 2. For each condition two scenarios were created, one involving conditional advice and the other a conditional inducement. The experiment revealed a reliable difference in the use of modalized indirect reports between participants working under Conditions 1 and 2 (*p* < 0.0001), with *could* being preferred in Condition 1 and *have to* in Condition 2. This suggests that modals used in indirect reports of '*if p, q*' conditionals may be a diagnostic for biconditional versus material interpretations of conditionals.

The aim of Experiment 2 was to find out whether these results could be replicated in contexts which lower/eliminate the believability of the major premise, i.e. in contexts which evoke a denial of or doubt in the satisfaction of CNCs and/or introduce alternative causes. It was found that manipulating the believability variable has no reliable effect on the results, indicating that indirect reports may not be definable in terms of belief attribution by the hearer of the indirect report to the reported speaker.

References: Cummins, Denise D., Todd Lubart, Olaf Alksnis, and Robert Ris. 1991. "Conditional reasoning and causation." Memory and Cognition 19: 274–282. Kratzer, Angelika. 2012. Modals and Conditionals. Oxford: Oxford University Press. Politzer, Guy. 2003. "Premise interpretation in conditional reasoning." In D. Hardman and L. Macchi, eds. Thinking: Psychological perspectives on reasoning, judgment, and decision making.Wiley. 79-93. Politzer, Guy. 2004. "Reasoning, Judgment, and Pragmatics." In I. Noveck and D. Sperber, eds. *Experimental Pragmatics*. London: Palgrave. 94-115.Thompson, Valerie A. 1994. "Interpretational factors in conditional reasoning." *Memory & Cognition* 22(6): 742-758.

Arbeitsgruppe 10 (Kurz-AG)

Prosodic boundary phenomena

Fabian Schubö, Sabine Zerbian, Sandra Hanne & Isabell Wartenburger

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Prosodic phrasing and syllable prominence in spoken prose: Prediction from text and validation

Isabelle Franz[1,2], Christine Knoop[1], Gerrit Kentner[1,2], Sascha Rothbart[1], Vanessa Kegel[1], Julia Vasilieva[1], Sanja Methner[1] & Winfried Menninghaus[1]

Max-Planck-Institute for Empirical Aesthetics[1], Goethe University Frankfurt[2]
isabelle.franz@ae.mpg.de

Current systems for prosodic boundary prediction from text focus on syntax/semantic-based automatic decoding of sentences that need to be annotated syntactically (Atterer & Klein 2002; Windmann et al., 2011). A replicable system for manually coding prosodic boundaries and syllable prominence in longer sentences or even texts is lacking so far, let alone its validation with the phonetic realization. Based on work in the fields of metrical phonology (Liberman & Prince 1977) and existing pause coding systems (Gee and Grosjean 1983; Windmann et al. 2011), we developed a manual for coding prosodic boundaries with 6 degrees of juncture (and syllable prominence).

The manual consists of a set of rules that are to be applied in a prescribed order. These rules mainly refer to the number of pre-annotated potential phrasal accents according to POS and neighboring words (Hayes 1989; Windmann et al. 2011), word and syllable count (Siebenhaar et al. 2004), syntactic structure (Selkirk 1984) and punctuation (Kalbertodt et al. 2015). Based on theories on cliticization (Selkirk 1984) and algorithms for phrase formation (Hayes 1989) and chunks (Schmid & Schulte im Walde 2000), the rules include markers for lexical word boundaries where a prosodic boundary is unlikely (see Gee & Grosjean 1983 for English). In order to develop a broadly usable manual for annotators that are not necessarily trained linguists, the wording in the rules refers to simple cues from the text, like POS, neighboring POS, word count and punctuation.

Three independent annotators applied the coding system to the beginning pages of four different German novels (~90 000 syllables). With an inter-annotator agreement close to 1 (Cohen's κ .90 - .96), the conflicting cases were discussed and solved between the annotators resulting in a final consensus coding. We used the consensus coding to predict prosodic boundary strength and relative syllable prominence in the phonetic realization. As for prosodic boundaries, we predicted a positive correlation between annotated boundary strength and pause duration in the phonetic realization.

For the validation of the coding system, eight professional speakers read the texts aloud. We annotated the speech signal automatically, using MAUS (Schiel 1999), matching the spoken syllables with citation form syllables. Using PRAAT (Boersma & Weenink 2019), we extracted duration and F0 range for each syllable. These parameters were compared to predicted syllable prominence and prosodic boundary strength. The validation with the speech signal and the high interrater agreement show that our annotation system reliably predicts syllable prominence and prosodic boundaries.

In comparison to Gee and Grosjean (1983) who developed a system to predict pauses from text with an infinite number of boundary degrees, our system generates six degrees of boundaries. This is comparable to GToBI (Baumann et al. 2000) where the speech signal is annotated. Since our annotation works with plain text, there are additional potential applications of the coding system, covering author profiling and style recognition, synthetic speech, and (psycho) linguistic research on prosody.

Final and pre-final lengthening in 13 languages

Ludger Paschen, Susanne Fuchs, Frank Seifart
Leibniz-Zentrum Allgemeine Sprachwissenschaft
{paschen, fuchs, seifart}@leibniz-zas.de

Final lengthening (FL) is a process whereby segments preceding a prosodic boundary, e.g. a pause, are lengthened compared to other positions. FL has been observed in a number of languages and is often considered a universal prosodic property of languages (Fletcher 2010). It has also been suggested that FL increases progressively from the penultimate syllable to the final syllable (Schubö & Zerbian 2020). However, most insights on FL are based on experimental studies with speakers of Indo-European languages using a small set of stimuli.

In this paper, we present findings from a corpus of spontaneous speech in a diverse set of 13 languages. The data come from language documentation collections and have been enriched by forced phone alignments as part of the DoReCo project (doreco.info). As the UNESCO has declared the decade of indigenous languages, DoReCo presents an effort to mobilize fieldwork data from lesser studied languages for cross-linguistic research. The languages used here are Arapaho, Beja, Bora, Fanbyak, Kamas, Lower Sorbian, Movima, Sadu, Sanzhi Dargwa, Svan, Urum, Yali, and Yongning Na. Each corpus contained ~10k word tokens, adding up to a total of ~250,000 segments. The data were manually checked for misalignment at the level of word boundaries, disfluencies, code-switching, and gaps in the transcription. We consider vowels before a silent pause (final position), in penultimate (pre-final) position and all other positions without surrounding pauses (non-final).
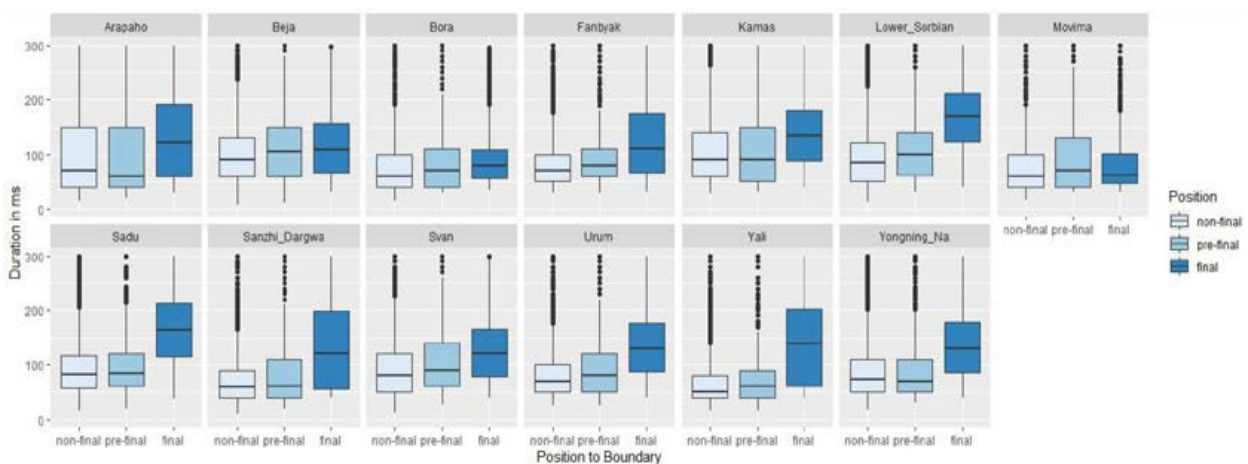
Figure 1: Vowel durations in final, pre-final and non-final positions.

Our results (Figure 1) show a continuum ranging from languages with massive FL (e.g. Sadu) to languages with no FL at all (Movima). Moderate lengthening of pre-final Vs is less common but attested in a subset of languages (e.g. Beja). These results will be discussed in light of language-specific processes such as word-level penultimate lengthening in Movima, but also with respect to factors such as presence of a phonemic vowel quantity contrast.

References: Fletcher, Janett. 2010. "The prosody of speech: Timing and rhythm". In William J. Hardcastle,John Laver, and Fiona E. Gibbon, eds. *The handbook of phonetic sciences*. Wiley-Blackwell. 521 602. Schubö, Fabian, and Zerbian, Sabine. 2020. "Phonetic content and phonological structure affect pre-boundary lengthening in German". *Proceedings Speech Prosody 2020*, Tokyo. 111-115.

Phonetic cues to IP-initial boundaries: Acoustic data from English, Spanish, and Portuguese

Ricardo Napoleão de Souza
University of Helsinki
ricardo.napoleaodesouza@helsinki.fi

Research has yet to determine the ways through which phonetic properties mark the beginnings of prosodic domains (but see Cho 2016), especially regarding its relationship with lexical prominence (cf. Turk & Shattuck-Hufnagel 2007 for domain-final edges). Additionally, the hypothesis that language-specific phonological contrasts modulate the phonetic marking of initial boundaries (e.g. Cho & McQueen 2005), needs direct testing using comparable methods in a cross-linguistic sample. Using similarly constructed stimuli, the present study seeks to address these issues by looking at acoustic differences in the phonetic expression of domain-initial boundaries on unstressed syllables in three languages that differ in how lexical stress manifests itself: English, Spanish, and Portuguese.

Method. Trisyllabic words with penultimate stress were selected in each language (e.g. 'tequila'). Target syllables consisted of a CV sequence of /p t k/ plus a monophthong in word-initial position. Target words were embedded in pre-nuclear position in carrier sentences controlled for length under two prosodic conditions: IP-initial or IP-medial (where IP = Intonational Phrase). Fourteen speakers of American English, Mexican Spanish, and Brazilian Portuguese each read the language-specific stimuli three times in pseudo- randomized order among filler items (N= ~1,260 tokens). Acoustic measures included VOT, and occurrence of a burst for /p t k/; and duration, f0 and f1-f2 ratio for vowels. Pre-boundary pause duration, duration of the stressed syllable, and articulation rate were measured for control. Praat scripts segmented and extracted the data, which were hand-corrected and normalized for comparison. It was expected that post-boundary, word-initial consonants in target CV syllables would manifest phonetic properties linked to domain-initial strengthening, whereas target vowels would be subject to the language-specific effects of lexical stress.

Results. The acoustic analyses of the three datasets suggest that the phonetic marking of the domain-initial boundary occurs both in the vicinity of the boundary, and on the non-adjacent stressed syllable. Target syllables were affected differently between the languages: in English, only /p t k/ showed longer VOTs in IP-initial position, with the following vowel showing no difference in the variables investigated. In both Portuguese and Spanish, consonants in the target CV syllables showed no difference between conditions, whereas the vowel in the target syllables were significantly longer IP-initially. Additionally, Spanish vowels showed higher f0 values near the domain-boundary, whereas Portuguese vowels showed less centralization. Contra predictions, the stressed syllable in the three languages showed significant differences in duration, with Portuguese being longer in the IP-initial condition. In English and Spanish, the stressed syllable was shorter near the IP.

Discussion. In the three languages investigated, phonetic properties marked the domain-initial boundary in the IP. In all three, differences in the duration of the stressed syllable, which was not adjacent to the boundary, suggest lexical stress plays a role in how the initial boundary is phonetically expressed. Furthermore, the hypothesis that the marking of domain-initial is language-specific receives support from the current results. Not only did Spanish and Portuguese show an effect on the vowel in target CV syllables, unlike English, but they also differed in how the effect was translated. Mirroring the phonological patterns of these languages, Spanish target vowels showed an effect of f0, whereas Portuguese vowels showed less centralization near the IP boundary. Put together, these results suggest the variables investigated may also constitute important markers of domain-initial prosodic boundaries that should be included in future studies of both production and perception.

The effect of predictability on the duration of phrase-final syllables

Bistra Andreeva, Bernd Möbius, Omnia Ibrahim, Ivan Yuen
Universität des Saarlandes
[andreeva, moebius, omnia, ivyuen]@lst.uni-saarland.de

There is a growing body of evidence suggesting that speakers and listeners have access to probability distributions over linguistic units (e.g., Jaeger 2010). This entails that speakers' choices and listeners' preferences are affected by the probability and frequency of occurrence of realizations of such units in a variety of contexts. According to the strong version of the Smooth Signal Redundancy (SSR) hypothesis (Aylett and Turk 2004, 2006), there is an inverse relationship between predictability and acoustic realization of phonetic structures moderated through prosodic structure. The weak version of SSR accepts that another major factor, viz. pre-boundary lengthening, modifies the relationship between prosodic structure, acoustic realization, and predictability.

The present study is an extension of our previous work and examines the impact and interaction of information density and prosodic structure on phrase-final syllable duration in a subset of the DIRNDL Radio News Database. DIRNDL is manually annotated for pitch accents and prosodic boundaries following the autosegmental intonation model. Each data point in our analysis is the last syllable before an intermediate phrase boundary (ip) or an intonational phrase boundary (IP). We analysed a total of 2382 ip and 2393 IP final syllables. Information density (ID) is defined as contextual predictability or surprisal of a syllable and estimated from language models based on DeWaC by means of a statistical language model as the inverse log probability of a syllable to occur in the context of two preceding syllables, that is, in a syllable trigram context. Our prediction is that (a) at final boundaries of prosodic constituents of the same type, the syllables with higher surprisal values will have longer duration compared to syllables with lower surprisal, and (b) since the magnitude of phrase- final lengthening varies systematically with the level of the constituent in the prosodic hierarchy, the increase in duration due to surprisal will be significantly greater before ip than before IP boundaries.

Regarding our first hypothesis, we calculated Pearson's r correlations between final syllable durations per boundary type (ip and IP) and the corresponding surprisal values. Syllable duration and surprisal were significantly positively correlated, which confirms our first hypothesis. To test our second hypothesis, we calculated linear mixed-effects models with predictors trigram surprisal, accent (factor levels: accented, unaccented), boundary type (factor levels: ip, IP) as well as their interactions. The continuous dependent variable syllable duration was log-transformed due to positive skewness. All categorical variables were treatment coded. The random structure included random intercepts for speaker, syllable identity (which reflects the segmental make-up of the syllable), and word identity. As expected, trigram surprisal, accent, and boundary type significantly lengthen the phrase-final syllable duration. Surprisal and accent as well as surprisal and boundary type interact in explaining syllable variability. Accent and surprisal complement each other in their positive effects on syllable duration. As for the interaction of surprisal and boundary type we found that syllable durations become longer with increasing surprisal, but this increase in duration is significantly greater before ip than before IP boundaries, which is generally compatible with the weak version of the SSR hypothesis.

References: Jaeger, Florian T. 2010. "Redundancy and reduction: speakers manage syntactic information density." *Cognitive psychology*, 61 (1): 23-62.Aylett, Matthew and Alice Turk. 2004. "The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech." *Language and Speech*, 47(1): 31-56. Aylett, Matthew and Alice Turk. 2006. "Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei." *Journal of the Acoustical Society of America*, 119(5): 3048-3058.

Pause duration and other prosodic boundary cues are not monotonically correlated

Gerrit Kentner, Isabelle Franz, Christine Knoop, Winfried Menninghaus
Max Planck Institute for Empirical Aesthetics, Frankfurt
firstname.lastname@ae.mpg.de

Prosodic boundary cues like phrase-final lengthening, pitch excursion and pause duration are generally thought to be positively correlated, i.e., the stronger the boundary, the longer the duration of phrase-final syllables, the greater the pitch excursion, and the more likely or longer a pause is predicted to be (Wightman et al 1992, Krivokapic 2007, among others). However, most studies we are aware of are concerned with only a single boundary cue (e.g., final lengthening) and/or with only a limited set of prosodic boundary strengths (e.g., comparing only two levels). Exploring a large sample of read prose, we consider pause duration, final lengthening, and pitch excursion and compare these across five predicted levels of boundary strength (0: no break predicted; 1: phrase break, no comma; 2: short comma phrase; 3: long comma phrase; 4: sentence boundary). The results show that these phonetic cues are not correlated with boundary strength in a simple monotonic fashion.

Methods: Eight professional speakers read aloud four prose text samples each (~1500-1800 words). The read texts (~6h of speech) were automatically segmented for words and syllables using WebMAUS (Kisler et al. 2017). We applied a manual coding scheme (Franz et al., forthcoming) to predict 5 degrees of boundary strength on the basis of textual features.

Results: While pause length monotonically increases with predicted boundary strength, this is not the case for the other boundary cues. Compared to the no-boundary condition (break index 0), pre-boundary syllables are lengthened and have higher pitch excursion, but final lengthening and pitch excursion are strongest for break index 2 and show a significant decrease through break indices 3 and 4.

Discussion: The monotonic increase of pause duration along the predicted scale, and the non-monotonic increase of the other prosodic boundary cues indicates that these phonetic signals reflect different processes in speech production. The increasing pause durations probably reflect the closure of units of increasing size and planning of upcoming clauses (Ferreira 1991). We tentatively suggest that the phonetic cues on the pre-boundary syllables reflect current planning complexity: Breaks provide a time window for speech planning, and planning complexity is high as long as clauses and sentences are not finalized. Breaks with commas (index 2 and 3) offer more planning time than breaks without comma (index 1). However, index 3 breaks are more likely to close off a clause, while at index 2, planning for the current clause is likely to be still ongoing. Therefore, planning complexity at break index 2 will be higher than at break index 3. Finalized clauses (break index 3) and sentences (index 4) require less or no time for current phrase planning, with syllable duration and, concomitantly, pitch excursion consequently decreasing.

In sum, this research shows that the suggested monotonic correlation between pause duration and other phonetic boundary cues is not valid.

References: Ferreira, Fernanda. 1991. "Effects of length and syntactic complexity on initiation times for prepared utterances" *Journal of Memory and Language* 30: 210-233. Franz, Isabelle et al. forthcoming. "Prosodic phrasing and syllable prominence in spoken prose – Prediction from text and validation." Ms. MPI for Empirical Aesthetics. Krivokapic, Jelena. 2007. "*The planning, production, and perception of prosodic structure*." University of Southern California Dissertation. Kisler, Thomas, Reichel, Uwe, & Schiel, Florian. 2017. "Multilingual processing of speech via web services." *Computer Speech & Language* 45: 326-347. Wightman, Colin et al. 1992. "Segmental durations in the vicinity of prosodic phrase boundaries." *JASA* 91.3: 1707-1717.

Temporal prediction in speech segmentation is modulated by foregoing utterance length

Laurence White
Newcastle University
laurence.white@newcastle.ac.uk

Listeners' generation of predictions about the duration of upcoming vowels and consonants supports both perception of speech sounds and interpretation of prosodic structure (Dilley and Pitt, 2010; Reinisch et al., 2011). Such temporal prediction is potentially crucial for word segmentation given that lengthening of speech sounds reliably marks prosodic boundaries: thus, for various languages, lengthening of word-final syllable rhymes is a cue to subsequent boundaries (Price et al., 1991) and word-onset consonant lengthening cues preceding boundaries (White et al., 2020).

We used a novel paradigm – nonword segmentation – to investigate how native English-speaking listeners' exploitation of timing cues to boundaries is affected by the length of the foregoing utterance. Participants heard 12-syllable nonsense utterances (e.g., *dumipako_libeku_binudafolu*) followed by trisyllabic nonword probes (e.g., *libeku*), and had to indicate whether the probed target had been in the nonsense utterance, 45 of 90 trials being target-present. In the Flat timing baseline, all utterance segments were the same duration. In four other timing conditions, specific segments were lengthened in the embedded target nonword: Syl1-C: first syllable onset consonant; Syl1-V: first syllable vowel rhyme; Syl2-C: second syllable onset consonant; Syl3-V: third syllable vowel rhyme. We also varied the placement of the target – Early, Medial, Late – within the carrier utterance on target-present trials.

Listeners performed at chance in all timing conditions for Early target detection, probably due to memory demands, whilst Medial target detection was above chance, but showed no variation between timing conditions. A differential timing effect emerged in Late targets, however: thus, detection was highest for Syl1-C lengthening, whilst all other Late timing conditions were equivalent. We also found shorter response latencies for Late Syl1-C targets; furthermore, similar mediation of timing cue use by utterance position also emerged in replications and extensions of this original study.

We interpret these findings as: a) support for the critical role of word onsets in segmentation and recognition; b) evidence of temporal prediction modulated by utterance length. To make useful predictions about segment duration, and thus detect lengthening as a boundary cue, listeners require sufficient prior utterance context. Finally, we tentatively suggest that such temporal prediction may be associated with entrainment of auditory cortex theta oscillations to the speech amplitude envelope (Luo and Poeppel, 2007).

References: Dilley, L.C., Pitt, M.A., 2010. Altering context speech rate can cause words to appear or disappear. *Psychological Science* 21, 1664–1670. Luo, H., Poeppel, D., 2007. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. Neuron 54, 1001–1010.Price, P.J., Ostendorf, M., Shattuck-Hufnagel, S., Fong, C., 1991. The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 90, 2956– 2970. Reinisch, E., Jesse, A., McQueen, J.M., 2011. Speaking rate affects the perception of duration as a suprasegmental lexical-stress cue. *Language and Speech* 54, 147–165. White, L., Benavides-Varela, S., Mády, K., 2020. Are initial-consonant lengthening and final- vowel lengthening both universal word segmentation cues? *Journal of Phonetics* 81, 100982.

An ideal-observer approach to structured talker variability in prosodic productions

Xin Xie[1], Andres Buxó-Lugo[2], Chigusa Kurumada[1]

University of Rochester[1], University of Maryland[2]
xxie13@ur.rochester.edu

Talker variability is one of the fundamental challenges for perceiving prosodic boundaries: Each talker can mark locations and types of a boundary with different acoustic details, which demands that a listener detect and adjust to different talkers' way of speaking (Buxó-Lugo, 2017; Arvaniti, 2019). On the other hand, talker variability is rarely random, which means that the knowledge of the variation can *facilitate* recognition of a boundary in continuous and often ambiguous speech. The structure of talker variability in prosody – how prosodic distributions vary within and across talkers – is not yet well characterized, however. The current study investigated the distributional structure of acoustic cues (e.g., mean fundamental frequency (F0) and duration of a syllable) and their variability across 65 native speakers of American English.

The main phenomenon examined was utterance-final pitch movements in marking a question vs. a statement. That is, we look at production variability in encoding phrasal boundary tones (Pierrehumbert, 1980). Each talker produced 48 tokens of the English construction "It's X-ing" (e.g., *It's raining*) to encode question vs. statement meanings, resulting in a total of 2974 tokens (after excluding speech errors). Recorded utterances were segmented into three sections 1) *it's,* 2) *X* (the stressed syllable), and 3)-*ing*. F0 and duration of each syllable were extracted (**Fig. A,** and examined with respect to the structure of variability in the cue distributions (**Fig. C**).
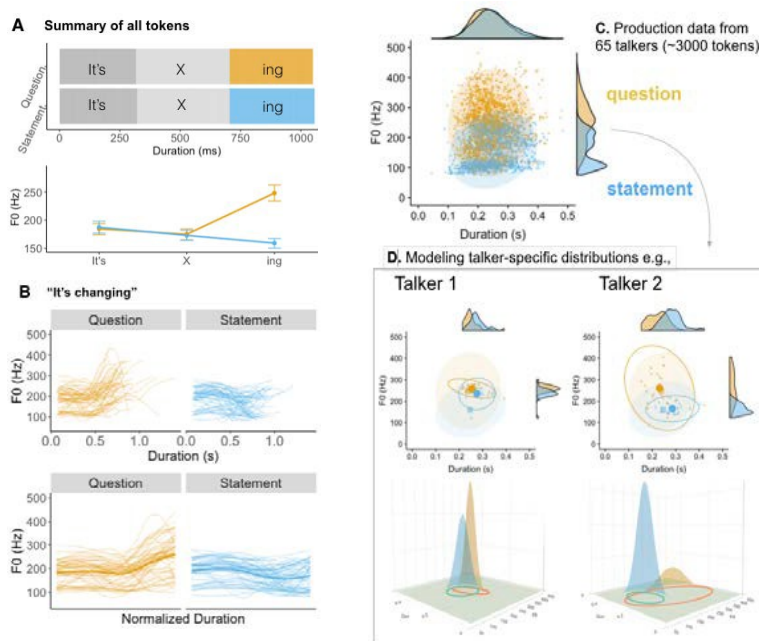
Figure A-D.
A. Summary statistics of duration (top) and fundamental frequency (F0, bottom) in the intonation contours for "It's X-ing" utterances produced by 65 native English speakers.
B. F0 values of individual tokens of "It's changing" to illustrate the magnitude of talker variability seen for each item type.
C. Group-level variations of syllable mean F0 (y-axis) and duration (x-axis) in the ~3000 tokens collected;
D. Talker-specific ideal observer models of productions for two example talkers (Talker 1 and Talker 2).

We then used the data to investigate whether the knowledge of talker-variability in production would "in-principle" benefit recognition of underlying categories (e.g., Is this a question or a statement?). We trained a set of Bayesian ideal-observer models on the production data, either with or without the capacity to index *talker-specific* information (**Fig. D**) (extending Kleinschmidt, 2019). Model predictions were then tested against human listeners' (N = 240) judgments on two new talkers' speech in a categorization task. Models with talker-specific information consistently outperformed those without it, supporting the hypothesis that **listeners draw on the structure and amount of talker-specific, acoustic, variability** in perceiving utterance final boundary tones. We discuss how this analysis approach can be extended to

AG 10: *Prosodic boundary*

quantify the extent and structure of prosodic variability at various syntactic boundaries. The workflow, assumptions, and limitations of the Bayesian approach will be examined (e.g., Schad, Betancourt & Vasishth, 2019) to obtain robust interpretations of models to be used in relevant future investigations.

Comparing cues: a mixed methods study of intonation unit boundaries in three typologically diverse languages

Naomi Peck[1], Kirsten Culhane[1], Maria Vollmer[1, 2]

University of Freiburg[1], ARC Centre of Excellence for the Dynamics of Language[2]
naomi.peck@linguistik.uni-freiburg.de, kirsten.culhane@linguistik.uni-freiburg.de, maria.vollmer@linguistik.uni-freiburg.de

In this presentation, we compare intonational boundary cues in three understudied languages which have three distinct typological profiles: Kera'a (Tibeto-Burman, NE India), which has lexical tone and possibly stress, Waima'a (Austronesian, Timor-Leste), which has no lexical tone or stress, and Warlpiri (Pama–Nyungan, Northern Australia), which has lexical stress and no tone. We take a mixed methods approach to investigate which cues are the most salient, which cues co-occur most frequently, and which factors motivate cue choices. The relative importance of intonational boundary cues has been said to vary on a language- specific basis (Izre'el and Mettouchi 2015: 23), and even the most common cues are realised differently across languages (Himmelmann et al. 2018: 239). We hypothesise, firstly, that speakers of each selected language will differ in what cues they use most frequently and secondly, that the choice of cues will be affected by the information structure of an utterance, as well as the phonological and intonational profile of each language.

Recent work by Himmelmann et al. (2018) argues for the universality of intonation unit boundary cues, like pauses, pitch resets, final lengthening and initial rushes. The authors also note the use of less systematic language-specific boundary cues such as creaky voice and unit-initial glottal stops to reinforce the perception of prosodic boundaries (also see Wagner and Watson 2010). Using naturalistic field data from Kera'a, Waima'a and Warlpiri, we segmented our data into intonation units and then annotated each one in Praat for the afore-mentioned phenomena. We also annotate for other phenomena we find in these languages, such as tonal parallelism and final devoicing. Our text selection is controlled for speaker and genre effects, as these could affect cue choices.

Our initial results show intra-language diversity in how intonation unit boundary cues are employed. In each language, speakers use different strategies depending on genre. For example, we find differences between procedural and narrative texts. We also observe information structural effects on cue choice, such as the prosodic encoding of new versus old information. Our initial results also demonstrate inter-language diversity in how speakers of each language encode intonation unit boundaries differently, even if the chosen cue is the same. Our data support the claim that pitch reset and pauses are the most important boundary cues, but we find that these cues differ as to their relative importance.

Furthermore, we contend that creaky voice is a more consistent boundary cue than the literature suggests. Additionally, traditional concepts of anacrusis and final lengthening proved insufficient for annotating our data, as speakers also make use of final rushes and initial lengthening. We suggest that the cross-linguistic diversity we observe may be linked to the specific typological profiles of each language, indicating a relationship between phonological and intonational inventories.

References: Himmelmann, Nikolaus P., Meytal Sandler, Jan Strunk, and Volker Unterladstetter. 2018. On the universality of intonational phrases: A cross-linguistic interrater study. *Phonology* 207–245.Izre'el, Shlomo, and Amina Mettouchi. 2015. Representation of speech in CorpAfroAs: Transcriptional strategies and prosodic units. In Amina Mettouchi, Martine Vanhove, and Dominique Caubet, eds. *Corpus-based Studies of Lesser-described Languages: The CorpAfroAs Corpus of Spoken AfroAsiatic Languages.* Amsterdam: Benjamins. 13–41.Wagner, Michael, and Duane G. Watson. 2010. Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes* 25:905–945.

Prosodic boundaries in phrase processing, a click-detection study

Sandrien van Ommen, Natalie Boll-Avetisyan, Barbara Höhle, Thierry Nazzi
CNRS – *Université de Paris, Universität Potsdam, Universität Potsdam, CNRS - Université de Paris*
sandrienvo@gmail.com, nboll@uni-potsdam.de, hoehle@uni-potsdam.de,
thierry.nazzi@parisdescartes.fr

In this study we test the influence of the native prosodic system on perceptual chunking. We use a click detection paradigm, building on Abrams & Bever (1969) who established slower detection of clicks at phrase-final boundaries, especially in unfamiliar phrases. Coupled with the evidence that clicks `migrate' to phrase boundaries in memory, even when they were presented in the middle of phrases (Cutler et al., 1997), this evidence suggests that cognitive resources are engaged in speech processing at phrase boundaries, and that clicks are registered separately. We use this separation of resources to assess the influence of top-down knowledge on phrase-prosodic processing.

It is unclear to what extent perceptual chunking is influenced by language-specific prosody; is it an automatic memory/phonetics-driven event, or is it influenced by top-down phonological experience? With a fully crossed design of prosodic structure (see below), stimulus language (French vs. German), stimulus register (infant-directed and adult-directed speech; IDS & ADS) and participant language (French vs German), we investigate different top-down effects on phrase processing.

Forty German- and forty French-speaking adult listeners listened to phrases of the structures   (A) *[[Name 1]$_{AP}$[and Name 2]$_{AP}$[and Name 3]$_{AP}$]$_{IP}$*   and

(B) *[[Name 1 and Name 2]$_{iP}$[and Name 3]$_{AP}$]$_{IP}$*.

Each stimulus contained a click at one of six possible locations, i.e. after the first or second syllable of one of the three names. Participants were requested to push a button as soon as they heard a click. Data were analyzed with linear mixed-effects models. Reaction times (RT) to clicks were the dependent measure, a covariate of real time controlled for faster responses later in the phrase.

Models for ADS stimuli revealed slower responses to higher-level prosodic positions with a three-way interaction of Position (within or between words), Part (Name 1 or 2) and Structure (A or B) (t = 4.41; p < .0001). Testing language-specific effects, models show that the above three-way interaction is strong in French listeners hearing German (t = 3.44; p <.005) and French (t = 3.55; p < .001); but weaker in German listeners hearing German (t = 1.91; p < .1) and French (t = 2.01; p < .05), suggesting a stronger effect of prosody in French listeners. Models for IDS stimuli, lastly, showed that the three-way interaction was not significant in either of the stimulus-participant language combinations.

These results confirm our prediction that there is a role for top-down processing on perceptual chunking. The same stimuli elicited different results in listeners with different native languages. Effects of prosodic phrase structure showed slower responses to clicks at stronger phonological boundaries, and faster responses inside phrases. The difference between ADS and IDS was as expected, with stronger top-down effects in the former. This expectation stems from the presumed top-down knowledge in adults but not infants, moving the IDS speaker to adjust their speech to stronger bottom- up cues. We discuss our results in the context of anticipatory resource allocation in line with theories on entrainment such as dynamic attending theory (Jones & Boltz, 1989), predictive timing (Friston & Buzsáki, 2016), and earlier accounts on the role of rhythmicity in speech processing (Martin, 1972).

References: Abrams, K., & Bever, T. G. (1969). Syntactic structure modifies attention during speech perception and recognition. *Quarterly J. of Exp. Psychology.,21*(3), 280–290. Cutler, A., Dahan, D., & Donselaar, W. van. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*, 141–201.Friston, K. J., & Buzsáki, G. (2016). The functional anatomy of time: What and when in the brain.*Trends in Cognitive Sciences*, *20*(7), 500–511.Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 459–491. Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, *79*(6), 487–509.

Effects of prosody and collocation frequency on language chunking

Nele Ots, Piia Taremaa
Goethe-University of Frankfurt am Main, University of Tartu
ots@em.uni-frankfurt.de, piia.taremaa@ut.ee

Language chunking is a strategy to process language by grouping smaller linguistic units together into larger ones in order to expand the limits of working memory and reduce the processing difficulty (Christiansen and Charter, 2016). Since the structural properties of language chunks are still poorly understood, the aim of this study is to investigate the perception of prosodic phrasing in spontaneous speech. Prosodic phrases as pause-internal units have a great potential to function as language chunks because acoustic breaks and intonational movements have proved to aid memory for verbal information (e.g. Baumann and Trouvain, 2001). In addition, memory for words is strongly supported by lexical frequencies (e.g. Stuart and Hulme, 2000), and the effect of word frequency has been found in perception of prosodic prominence (Baumann and Winter, 2018) as well. In a similar vein, we propose that the frequency of lexical collocations modulates the perception of prosodic phrase boundaries.

Native speakers of Estonian (n=51) listened to spontaneously spoken Estonian utterances (n=396) and were instructed to mark between the words where they heard some sort of juncture. No additional explanations about the nature of junctures was provided. The prosodic strength of word boundaries of the excerpts was estimated with the help of continuous wavelet analysis (Suni et al., 2017). In this analysis, the greater boundary values indicate offsets of higher-level prosodic units (e.g. intonation phrases) while lower values associate with word boundaries. In addition, each word was coupled with a likelihood of another word (bigram frequencies) or two other words (trigrams) to follow it based on the n- gram frequencies from fictional literature (Raudvere and Uiboaed, 2018). We predict that (i) the offsets of higher-level prosodic units boost boundary perception, and that (ii) the lower likelihood of a collocation relates to higher probability of boundary marking.

As predicted, the prosodic boundary strength affected prosodic boundary perception such that the probability of boundary marking increased together with the prosodic boundary strength. The examination of collocations indicated that the wavelet determined boundary strength correlated well with the bigram but not with trigram frequencies. More importantly, the bigram frequencies modulated the probability of boundary marking such that the likelier a word was followed by another word, the less likely a boundary was perceived. The trigram frequencies did not affect boundary perception. These results suggest that lexical collocations (esp. bigrams) constitute mentally coherent units that interact with prosodic boundary strength and in addition to prosody affect the perception of prosodic phrases. In overall, the results support the notion of language chunks to be determined in terms of higher-level prosodic phrases (i.e. intonation phrases) and lexical collocations.

References: Baumann, Stefan and Bodo Winter. 2018. "What makes a word prominent? Predicting untrained German listeners' perceptual judgments". *Journal of Phonetics* 70: 20-38. Baumann, Stefan and Jürgen Trouvain. 2001. "On the prosody of German telephone numbers". In *Interspeech*. Christiansen, Morten H., and Nick Chater. 2016. "The now-or-never bottleneck: A fundamental constraint on language." *Behavioral and Brain Sciences* 39: 1-72. Raudvere, Uku and Kristel Uiboaed. 2018. „Uuema eesti ilukirjanduse mitmikute loendid." http://dx.doi.org/10.15155/re-8; http://datadoi.ee/handle/33/41.Stuart, Georg and Charles Hulme. 2000. "The effects of word co-occurrence on short-term memory: associative links in long-term memory affect short-term memory performance." *Journal of experimental psychology. Learning, memory, and cognition, 26,* 796-802. Suni, Anti, Juraj Simko, Daniel Aalto and Martti Vainio. 2017. "Hierarchical representation and estimation of prosody using continuous wavelet transform." *Computer Speech and Language,* 45,123-136.

Arbeitsgruppe 11 (Kurz-AG)

Edge-asymmetries in morphophonology

Daniel Gleim & Marie-Luise Popp

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Morphological symmetry, prosodic asymmetry: The case of Huave mobile affixes

Yuni Kim
University of Essex
y.kim@essex.ac.uk

Overview: Huave (isolate: Mexico) has two phonological processes that apply within the stem-plus-suffix domain, but which fail to apply across a prefix-stem boundary. This prefix-suffix asymmetry is particularly compelling because Huave has mobile affixes that alternate between prefixal and suffixal realizations (Noyer 1994, Kim 2010), and they participate in the phonological processes in question only when appearing as suffixes. I analyze the stem-plus-suffix domain as a prosodic word, and argue that the Huave case supports proposals that deal with morphosyntax-prosody mapping in terms of constraints referring to individual constituent edges (e.g. Cheng & Downing 2012), rather than in terms of whole constituents (e.g. Match Theory; Selkirk 2011). Furthermore, prosodification constraints in Huave are violable: within a restricted morphosyntactic domain, prefixes and stems may be prosodified together if the stem would otherwise be subminimal, which is striking given that a repair strategy of prosodic augmentation is attested for prefixes that are structurally further away from the stem. Together, the behavior of mobile affixes and the violability of prosodification constraints locate the Huave prefix-suffix asymmetry firmly in the synchronic grammar, rather than as an epiphenomenon arising from the lexical entries of individual morphemes.

Key data: In Huave vowel harmony, the quality of epenthetic suffix vowels is determined by both segments in the preceding VC sequence (Kim 2008: ch.4); both stems and preceding suffixes qualify. However, the example in (1) shows that when the preceding C is a monoconsonantal stem, the prefix vowel is unexpectedly invisible to the process. In the context of preceding /ewʲ/, the regular outcome of harmony should be /e/, diphthongized to *[ja]. Instead, only the palatalized C matters, and the default palatal vowel /i/ (diphthongized to [jə]) results.

(1) /t-e-wʲ-n/ [[$_{L1}$t-e-[$_{Stem}$w]]-**jən**$_{L4}$] 'you (pl.) borrowed it'
     CPL-2-borrow-1/2PL

Attempting a representational analysis, though, we cannot simply state that prefixes are always prosodified separately from the stem. In (2), a non-suffixed form shows that the resulting prosodic constituent would consist of a single consonant, violating general minimality conditions.

(2) /t-e-wʲ/ [t-e-[$_{Stem}$w]] 'you (sg.) borrowed it'
     CPL-2-borrow

Also failing to apply across prefix-stem boundaries is Fricative Dissimilation, which deletes postvocalic /h/ in the context of a voiceless fricative within a one-syllable distance. In (3), prefixal [s] fails to trigger deletion of a [h] or [+s.g.] feature in the stem.

(3) [$_{L3}$**s**-[$_{Stem}$a-**h**tʃ]]] 'I give (it)' *satʃ 1-tv-give

Because this process also applies between suffixes and within derived stems, its non- application in (3) does not readily submit to a root-faithfulness or stem-control analysis.

The suffixing preference and the edge-asymmetry reversal in reduplication

Thomas Schwaiger
University of Graz
thomas.schwaiger@uni-graz.at

This paper investigates reduplicative edge asymmetries in relation to the suffixing preference in the world's languages. Morphological reduplication, whereby formally a whole word or some circumscribed subpart is repeated mainly for functions like plurality, intensity, diminution or lexical enrichment, is exemplified for its cross-linguistic variation in (1)-(4):

(1) Indonesian *gula* 'sugar' – *gula~gula* 'sweets' (Sneddon 1996: 16)
(2) Ngiyambaa *giːdjan* 'green' – *giːdja~giːdjan* 'greenish' (Donaldson 1980: 73)
(3) Mangarayi *galŋbam* 'spouse' – *galŋbam~bam~yi* 'spouses' (Merlan 1982: 215)
(4) Daga *baraen* 'he put' – *ba~ra~raen* 'he put and put until full' (Murane 1974: 73)

The examples show the basic divide between full and partial reduplication. The former pertains to the repetition of morphological units like roots, stems or words (1) (rarely also affixes), while the latter comprises phonological units like feet (2), complex syllables (3) or the unmarked CV syllable (4). Moreover, a reduplicative exponent (the reduplicant) can occur in front (2), in the middle (3) or at the end (4) of its unreduplicated counterpart (the base).

Much of the pertinent theoretical literature has viewed reduplication as a special kind of affixation, e.g. Marantz (1982), who was very influential in treating reduplicants like segmentally underspecified affix templates to be filled by phonological copying from their bases. However, such an approach cannot explain why, in contrast to the well-known suffixing preference in the languages of the world, (partial) reduplication is preferably of the initial type like (2) (see also Inkelas 2012: 358).

The present contribution argues for a non-affixational approach to reduplication, recognizing the latter's independence as a process of morphology that is heavily driven by phonological considerations as well. Based on typological data, reduplicants are analyzed as essentially derivational morphological devices, a property connected to their phonological and morpho-semantic features as interpreted from a word-based stance on the form-meaning relationship (see Bybee 1985). Accordingly, the different positional preferences for affixes and reduplication arise from their fundamentally different status. Additional support for this reversal comes from the so-called root privilege: "No matter what the specific morphological and phonological conditions on reduplication may be, reduplication ends up copying at least a portion of the morphological root" (Inkelas 2012: 358). Together, these characteristics are claimed to explain the behavioral asymmetry of affixation and reduplicants, including the fact that even in reduplicating languages which otherwise exclusively display suffixes in their morphology, (partial) reduplication tends to be initial and thus at the opposite edge of the word. From a psycholinguistic perspective, this difference is linked up with Berg's (2015) structural-processual account of temporal asymmetries between prefixes and suffixes and logical asymmetries between stems and affixes.

References: Berg, Thomas. 2015. "Locating affixes on the lexicon-grammar continuum." *Cognitive Linguistic Studies* 2: 150-180. Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: Benjamins. Donaldson, Tamsin. 1980. *Ngiyambaa: The language of the Wangaaybuwan*. Cambridge: Cambridge University Press. Inkelas, Sharon. 2012. "Reduplication." In Jochen Trommer, ed. *The morphology and phonology of exponence*. Oxford: Oxford University Press. 355-378. Marantz, Alec. 1982. "Re reduplication." *Linguistic Inquiry* 13: 435-482. Merlan, Francesca. 1982. *Mangarayi*. Amsterdam: North-Holland. Murane, Elizabeth. 1974. *Daga grammar: From morpheme to discourse*. Norman: SIL. Sneddon, James Neil. 1996. *Indonesian: A comprehensive grammar*. London: Routledge.

Prefix independence as root-initial percept maximization

Noah Elkins
UCLA
noahelkins1@g.ucla.edu

The prefix-suffix asymmetry in phonology is an imbalance in the application of phonological processes by which prefixes are far less likely than suffixes to cohere to their roots. In this talk, I refer to this exclusionary characteristic of prefixes as prefix independence. For example, prefixes may fall outside of the domain of stress assignment, hiatus may not be repaired at the prefix-root juncture, and prefix- controlled vowel harmony systems are virtually unattested.

This talk has two main goals. First, I present the results of a typological survey of languages and language processes for which prefix independence is observed. This typology includes 85 languages from a diverse range of families and regions for which prefixes are less cohering than suffixes with respect to some process. At present, this is the largest survey of prefix independence ever compiled. The following major processes have been shown to demonstrate the prefix-suffix asymmetry (n = number of observed cases): affix control (12), vowel harmony (13), consonant harmony (6), tone spread (6), footing/stress assignment (29), syllabification (8), hiatus resolution and other phonotactics (11). These results suggest that this asymmetry is widespread and robust; the one attested example of non-cohesion targeting suffixes instead (Kabardian, NW Caucasian) can actually be explained via other mechanisms.

The second goal of this talk is to discuss a proposed theoretical mechanism which can account for the typological results. My proposal rests on the fact that root-initial syllables constitute a privileged position: initial segments are articulated with greater magnitude cross-linguistically (Keating et al. 1997), and initial syllables are more resistant to alternation than non-initial ones (Beckman 1998, Becker et al.2012). Prefixes, while constituting the first syllables of entire morphological words, are not nearly as salient: they tend to contain fewer phonemic contrasts than root morphemes (Bybee 2004), and have been argued to be processed only after their roots are (Taft & Forester 1975, Taft 1994). This root-initial strengthening can act as a boundary signal, aiding in lexical access (Fougeron & Keating 1997). In order to maximize the efficacy of the root-initial percept, then, initial segments are hesitant to share their features leftward to target prefixes, as this would blur the strong root-initial boundary; prefixes are unwilling to affect roots for the same reason. I argue that a highly-ranked CRISPEDGE constraint (Itô & Mester 1999) relativized to the left edge of initial syllables can account for this pattern.

For phenomena like stress assignment which are not governed by feature spreading, I argue that prosodic words, which are the domain for such processes, are preferentially aligned to the edges of roots, as opposed to whole morphological words. This has a similar effect, namely preserving the environments in which root-initial segments are articulated most robustly. A preferential ranking of ALIGN-L(Root, PrWd) ≫ ALIGN-R(Root, PrWd) is proposed; an alternate ranking, while possible, should be more marked, as left edges are more informationally beneficial for lexical access. This approach is shown to avoid issues encountered by Transderivational OO-Correspondence (Benua 1997, Bakovic 2000) and Stratal OT (Kiparsky 1982) which have also been invoked to address the prefix-suffix asymmetry.

References: Bakovic, Eric (2000). Harmony, dominance and control. Ph.D. dissertation, Rutgers. Becker, Michael, Andrew Nevins & Jonathan Levine (2012). Asymmetries in generalizing to and from initial syllables. Language 88(2), 231-268. Beckman, Jill (1998). Positional faithfulness. Ph.D. dissertation, Rutgers. Benua, Laura (1997). Transderivational Identity: Phonological Relations between Words. Ph.D. dissertation, UMass, Amherst. Bybee, Joan (2004). Restrictions on morphemes in affixes: a cross-linguistic test of a popular hypothesis. Ms., University of New Mexico. Fougeron, Cécile & Patricia Keating (1997). Articulatory strengthening at edges of prosodic words. JASA 101(6), 3728-3740. Itô, Junko & Armin Mester (1999). Realignment. In R. Kager, H. van der Hulst & W. Zonneveld (eds.) The Prosody-Morphology Interface, 188-217. Cambridge University Press. Keating, Patricia, Taehong Cho, Cécile Fougeron & Chai-Shune Hsu (1997). Domain-initial articulatory strengthening in four languages. Phonetic interpretation: Papers in laboratory phonology VI, 143-161. Kiparsky, Paul (1982). Lexical Phonology and Morphology. In I.-S. Yang (ed.) Linguistics in the morning calm, 3-91. Seoul: Hanshin. Taft, Marcus (1994) Interactive-activation as a framework for understanding morphological processing. Language and Cognitive Processes 9(3), 271-294. Taft, Marcus & Kenneth I Forester (1975). Lexical storage and the retrieval of prefixed words. Journal of Verbal Learning and Verbal Behavior 14, 630-647.

[A domain-general bias cannot explain the suffixing preference: Experimental evidence from English and Kîîtharaka](#)

Alexander Martin, Jennifer Culbertson
Université de Paris, LLF, CNRS CLE, The University of Edinburgh
alxndr.martin@gmail.com, jennifer.culbertson@ed.ac.uk

It has long been noted that suffixes are more common across the world's languages than prefixes (and deed other affix types). Dryer (2013), for example, notes that over half of the world's languages have more suffixes than prefixes, while only 17% have more prefixes than suffixes. One account of this asymmetry claims that the perceptual salience of the beginnings of words (and in turn their importance for word recog nition) makes suffixing a better affixation strategy than prefixing, as it allows quicker retrieval of the base word (i.a., Hawkins & Cutler, 1988). A stronger claim has been made, namely that the perceptual salience of word beginnings reflects the domain-general salience of beginnings of sequences of all kinds (Hupp et al., 2009). Hupp et al. provide evidence in favour of this account in the form of a similarity judgement task. In their study, participants judged that "suffixed" sequences were more similar to base sequences than "prefixed" ones were. This preference held whether the sequences were composed of syllables (e.g., ba-ta-be), pictures of shapes (e.g., ○ ◨ ↘), or arpeggiated musical chords (e.g., *do-mi-sol*), indicating that similarity judgements were not restricted to linguistic stimuli. Their participants, though, were speakers of English, a predominantly suffixing language; it is therefore unsurprising that they preferred suffixing-like linguistic sequences, and indeed a lifetime of experience with suffixing could plausibly affect preferences in other perceptual domains. Here, we replicate Hupp et al.'s experiments with English speakers and provide a much stronger test of their hypothesis by testing the similarity judgements of speakers of a predominantly prefixing language: Kîîtharaka, a Bantu language spoken in Eastern Kenya.

Experiment. Our experiment design was based on Hupp et al.'s tasks using sequences of syllables and shapes. We tested 51 self-reported native English speakers on Mechanical Turk and 72 self-reported native Kîîtharaka speakers in villages surrounding Marimanti in Kenya. For English speakers, the experiment was run online; Kîîtharaka speakers were tested by local experimenters using tablet computers. On critical test trials, participants were presented a sequence of two syllables or shapes (of the form *X-Y*) followed by two response options, one pre-changed (of the form *Z-X-Y*, akin to prefixing) and one post-changed (of the form *X-Y-Z*, akin to suffixing). Participants were asked to choose which of the changed sequences was most similar to the base *X-Y* sequence. If the perceptual salience of sequence beginnings determines how similar two sequences are to each other, participants should prefer post-changed sequences across the board.

±**Results.** The data were analyzed with mixed effects logistic regression models. On critical test trials, English speakers chose post-changed [suffixed] sequences significantly above chance level for both sequences of syllables and shapes (syllables: mean = 78%, sd = 29%; shapes: mean = 76%, sd = 29%; $\beta = 2.05 \pm 0.37$, $z = 5.30$, $p < 0.001$), replicating the findings of Hupp et al.[1] Kîîtharaka speakers, on the other hand, chose post-changed sequences significantly *below* chance level, showing a preference for the pre- changed [prefixed] sequences instead (syllables: mean = 37%, sd = 18%; shapes: mean = 37%, sd = 25%; $\beta = 0.69 \pm 0.15$, $z = 4.70$, $p < 0.001$).

Discussion. The results show that, contrary to previous claims, a domain-general bias does not determine preferences in a similarity judgement task. Rather, preferences track affixation patterns in the native language of participants: speakers of English (a mostly suffixing language) found post-changed [suffixed] sequences to be more similar to a base sequence, while speakers of Kîîtharaka (a mostly prefixing language) found *pre*-changed [prefixed] sequences to be more similar to a base sequence. Our results are a challenge to the claim that domain-general salience of sequence beginnings favours suffixing over prefixing in the world's languages. Accounts of the typological asymmetry which rely on linguistic constraints like prosody (e.g.,

Himmelmann, 2014) and their effects on grammaticalisation present more plausible alternative explanations.

[1]   We observed no significant difference between conditions in either population (both $\chi^2(1) <$ 1), so data for syllables and shapes conditions were combined for the statistical analyses.

References: Dryer, M. S. (2013). Prefixing vs. suffixing in inflectional morphology. In M. S. Dryer & M. Haspelmath (Eds.). Max Planck Institute for Evolutionary Anthropology. Hawkins, J. A., & Cutler, A. (1988). Psycholinguistic factors in morphological asymmetry. In J. A. Hawkins (Ed.), *Explaining language universals* (pp. 280–317). Blackwell. Himmelmann, N. P. (2014). Asymmetries in the prosodic phrasing of function words: Another look at the suffixing preference. *Language*, *90*(4), 927–960. Hupp, J. M., Sloutsky, V. M., & Culicover, P. W. (2009). Evidence for a domain-general mechanism under-lying the suffixation preference in language. *Language and Cognitive Processes*, *24*(6), 876–909.

The suffixation preference: Native language and information load in artificial language learning

Xinyi Wang, Itamar Kastner
University of Edinburgh
X.Wang-188@sms.ed.ac.uk, itamar@itamarkast.net

Background. The suffixation preference shows the asymmetry of prefixes and suffixes in natural languages – languages across the world have more suffixes than prefixes. Moreover, the suffix preference was also found in people's word acquisition and development as evidence showed both adults and children found it easier to master suffixes in both natural languages acquisition and Artificial Language Learning (ALL) paradigms. Yet the exact source of this preference is unclear.

One hypothesis suggests that a domain-general cognitive mechanism accounts for people's suffix preference because people have a suffixation-type preference for all sequences, not just suffixed words (Hupp et al, 2009). A related hypothesis suggests that prefixes add information load before the stem is seen or heard, and since the stem carries most of the information in the word, suffixes are a more efficient form of affixation (Pycha, 2015; Blazej and Cohen-Goldberg 2015).

However, most previous studies exclusively investigated the performance of native speakers of English (a predominantly suffixing language); additionally, studies that investigate people's suffix preference when prefixes contained different information load were also rare. Therefore, whether people's suffix preference is a domain general preference or come from the impacts of the affixation patterns in their native languages is not clear.

Results. In Experiment 1, it was found that participants' preferences in affixation patterns were consistent with the affixation patterns of their native languages: while English native speakers significantly showed the suffix preference in the experiment, Mandarin native speakers did not have the preference to extend either suffixed test words or prefixed test words. Experiment 2 suggests that participants' preference for affixes is modulated by both information load and native language. There was no significant difference between conditions for English speaking participants, while Mandarin speaking participants' preference for prefixes significantly increased when extra information load in prefixes was lowered in word perception.

Methods. This study used an ALL approach conducting two label extension experiments with 24 native speakers of English and 22 speakers of Mandarin (a morphologically impoverished language) to see the influence of native languages on the suffixation preference. Additionally, the stimuli in two experiments were designed to have different information loads in order to test its influence on people's affixation preferences. In Experiment 1, 11 English speaking participants and 11 Mandarin speaking participants were shown a target shape (e.g.,■) which was labelled by an unaffixed word (e.g., *pefi*) as the target word. Then they were shown an affixed word as the test word (e.g., *pefizo*) and were asked to extend the test word to one of the two test shapes (e.g., ■ or ◆). For example, "This is *pefi*, here are two items, which one is *pefizo*?" (on a suffixed trial). The experiment consisted of 8 prefixed trials and 8 suffixed trials and this experiment was designed to see people's preference to suffixes or prefixes by examining whether participants were more likely to extend test words to the identical shape of the target shape in suffixed trials or prefixed trials. The procedures in Experiment 2, in which 13 English participants and 11 Mandarin participants took part, were the same as those in Experiment 1 except that Experiment 2 used affixed words (e.g., *pefizo*) as target words and unaffixed words (e.g., *pefi*) as test words. This change intended to lower the information load contained in target words of prefixed trials because participants had been shown all the syllables before the forced-choice section to counteract the different information load contained in prefix and suffix syllables.

Conclusion. The study found that affixation patterns of native languages influence the suffix preference of both English participants and Mandarin participants in the reported

experiments (see also Martin and Culbertson, to appear): while English native speakers showed the suffix preference, Mandarin native speakers did not present any preference towards either suffixes or prefixes. Additionally, the study also suggests the influence of information load in prefixes on participants' preference for affixes: when information load contained in prefixes was lowered, Mandarin participants' preference for prefixes increased. As the suffix preference does not seem to be a general cognitive preference, an open issue lies in whether other language exposure except people's native language would also influence their suffix preference. Future work should investigate the affixation preference of bilingual or multilingual speakers who speak languages with different affixation patterns.

References: Blazej, L. J., & Cohen-Goldberg, A. M. 2015. "Can we hear morphological complexity before words are complex?" *Journal of Experimental Psychology: Human Perception and Performance* 41:50. Hupp, J. M., Sloutsky, V. M., and Culicover, P. W. 2009. "Evidence for a domain-general mechanism underlying the suffixation preference in language". *Language and Cognitive* Processes, 24:876-909. Martin, A., and Culbertson, J. To appear. "Revisiting the suffixing preference: Native language affixation patterns influence perception of sequences."Pycha, A. 2015. "Listeners perceive prefixes differently: Evidence from a noise-rating task. *Word Structure* 8:53-83.

Edge-biases in mutation

Daniel Gleim, Sören E. Tebay
Universität Leipzig
daniel.gleim@uni-leipzig.de, tebay@uni-leipzig.de

We claim that mutation is similar to segmental morphology in that both prefer the right edge of a word. The suffixing preference for segmental exponents has been known at least since Greenberg (1957). More recent work has tested this against larger crosslinguistic samples (Dryer 2013) and offered a variety of theorectical explanations (e.g. Cutler et al. 1986, Bybee et al. 1990) rooted in processing, diachronic and semantic background. All typological studies agree that there are more languages that primarily use segmental suffixes to express inflectional categories in morphology than there are prefixing languages. Note however, that these works have focused on concatenative morphology. Non-concatenative morphology, like mutation and tonal morphology, has remained outside the scope of the existing research.

This talk draws its data from the MAMPF database (Gleim et al. 2019) that collects data from a large set of languages that are both genealogically and areally diverse. Mutation is defined here in a broad sense as a change in segmental features, segmental length, tone or stress between two morphologically related forms that cannot be explained by the general phonology of a language and – alone or in combination with a segmental string – serves as a morphological exponent. The database thus includes well-known cases of non-concatenative morphology, such as German (Germanic, Central Europe) Umlaut. Here, the morphological category plural is expressed by fronting the rightmost full vowel, e.g /fatɐ/ 'father' vs. /fɛtɐ/ 'fathers'. It also features patterns from less well documented languages and languages that have only very recently been described, such as Lele (Asutronesian, Papua New Guinea) first person possessors. These are expressed by a suffix -o, as well as changing the rightmost vowel into an /o/ if it was an /a/ in the singular, e.g. /tam/ 'father' vs. /tomo/ 'my father' (Boettger 2015, 44). Note that in Lele there is no general phonological process that raises an /a/ to /o/ before a syllable with another /o/.

The MAMPF database includes information on the locus of mutation. We distinguish between initial, leftmost, final and rightmost targets for the purpose of this study. The difference between initial/final and leftmost/rightmost is the following: An initial/final segment is only affected if it occurs at the edge of a word. If not, the mutation does not apply. Rightmost/Leftmost targets are affected by the mutation as long as there is no other target closer to the designated word edge. If there is, only the target closest to the edge is affected. We show that rightmost/final targets are much more frequent in the languages of the world than leftmost/initial targets, with some important exceptions. This mirrors the suffixing preference found in segmental concatenative morphology. Crucially, this parallelism also extends to the distribution across different geographical areas and morphological categories. Papuanesia, for example, is an exception to the strong suffixation trend amongst the worlds languages (Dryer 2013) and also shows a more balanced picture with regard to mutation, in that a sizable amount of mutation is oriented towards the left edge.

We conclude that mutation is a special case of morphology. This assumption is pursued for example by the research programme known as Generalized Nonlinear Affixation (Bermudez-Otero 2012). The main assumption is that mutation comes about from affixation of phonologically deficient material. Mutation is thus initiated in the morphology and should therefore obey the same generalization for edge biases. We take the data from our study to be a strong argument in favor of such an approach and against approaches like Lexical Constraint Indexation and Cophology Theory, that rest on the assumption that non- concatenative morphology is just a special application of regular phonological processes. These would predict that mutation patterns with regular phonological processes and not with concateniative affixation. In contrast, our data clearly show that the crosslinguistic distribution of mutation parallels the suffixation preference found in regular concatenative morphology. Additional reasons for this parallelism might be found in diachronic paths to non- concatenative morphology.

References: Bermúdez-Otero, Ricardo. 2012. "The architecture of grammar and the division of labour in exponence." In Jochen Trommer, ed. The morphology and phonology of exponence Oxford: Oxford University Press. 8-83. Boettger, Juliane. 2015. Topics in the grammar of Lele: a language of Manus Island, Papua New Guinea. Diss. James Cook University. Bybee, Joan L., William Pagliuca, and Revere D. Perkins. 1990. "On the Asymmetries in the Affixation of Grammatical Material." In William Croft, Keith Denning, and Suzanne Kemmer, eds. Studies in Typology and Diachrony: Papers Presented to Joseph H. Greenberg on his 75th Birthday. Amsterdam: John Benjamins. 1-42. Cutler, Anne, John A. Hawkins, and Gary Gilligan. 1986. "The Suffixing Preference: a Processing Explanation." Linguistics 23. 723-758. Dryer, Matthew S. 2013. "Prefixing vs. Suffixing in Inflectional Morphology." In Matthew S. Dryer, and Martin Haspelmath, eds. The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. Gleim, Daniel, Katie McCann, Sören E. Tebay, Jochen Trommer and Eva Zimmermann. 2019. "Database on Morphological Affixation with Mutation /Phonological Features (MAMPF)." University of Leipzig. Greenberg, Joseph H. 1957. "Order of Affixing: a Study in General Linguistics." In Joseph H. Greenberg, ed. Essays in Linguistics. Chicago: University of Chicago Press. 86-94. McPherson, Laura. 2017. "Multiple feature affixation in Seenku plural formation." Morphology 27.2: 217-252.

Tonal asymmetry for tense-aspect at verbal phrase edges

Ronald P. Schaefer, Francis O. Egbokhare
Southern Illinois University Edwardsville, University of Ibadan
rschaef@siue.edu, foegbokhare@yahoo.com

Hyman (2020) highlights prosodic patterns for nominal and verbal forms in selected Bantu languages. He finds verb roots with fewer prosodic contrasts than noun roots, and noun phrases with more prosodic variation than verb phrases. For this paper we examine tonal prosody in Emai, an under described Edoid language, which though Benue Congo is not Bantu. Our focus is verbal left and right edge and their tonal expression of tense and aspect (TA), respectively. In Edoid studies, it is widely accepted that verbs are lexically toneless, whereas nouns show low, high and high downstep (Elugbe 1989). Similarly, proclitic subject pronouns are toneless, receiving tone from tense marking.

In Emai, grammatical aspect is suffixal to the verb. The basic contrast is between perfective (PFV) and imperfective (IPFV). PFV is signaled by high tone *-í*, whose tonal value spreads onto the preceding verb. Its segmental and tonal value has three realizations comparable to Bantu metatony (Hyman and Lionnet 2014). If followed by a verb argument, *-í* deletes; if followed by a non-argument, *-í* is retained; and if clause final, *-i* tone shifts from high to low. IPFV, with no segmental realization, exhibits a floating low tone (-`L); it too spreads onto a preceding verb.

In contrast, underlying tense precedes the verb or any auxiliary/preverb. Tense morphemes follow a subject pronoun in a proclitic complex. Within the complex, tone is primarily contrastive. There are three tense types. Past (PST) conditions are near (NPST) and remote (RPST), as are future (FUT: NFUT, RFUT). Present has values progressive (PROG) and habitual (HAB). For PST, pronoun *o* '3sg,' NPST ʼH, intransitive verb *muzan* 'halt,' and PFV *-í* surface as *ò-múzán-ì* 'He has halted.' When a transitive verb occurs, tone at verb left edge is variable. It is extraprosodic for RPST, where floating low tone -`L docks on verb initial syllable resulting in downstep high (↓*dú*). Replacive tone affects verb left edge for both FUTs (*dùmé*), where verb initial H preceded and followed by H becomes low: H → L ∕ FUTH#[___H. Verb left edge is not adjusted for NPST (*dúmé*).

/ó	ò	HAB	dume-`L	émà/	>	[ó	ò	dùmè	émà]	'she pounds yam'
/ò	ó	PROG	dume-`L	émà/	>	[ò	ó	dùmè	émà]	'she is pounding yam'
/ó	`L	RPST	dume-í	émà/	>	[ó		↓dúmé	émà]	'she pounded yam'
/ò	ʼH	NPST	dume-í	émà/	>	[ò		dúmé	émà]	'she has pounded yam'
/ó	ló	RFUT	dume-í	émà/	>	[ó	ló	dùmé	émà]	'she will pound yam'
/ò	ló	NFUT	dume-í	émà/	>	[ò	ló	dùmé	émà]	'she is about to pound yam'

When an auxiliary/preverb is present, e.g. *gbo* 'also,' tense again conditions extraprosodic and replacive tone at phrase left edge (↓*gbó*, *gbò*). In addition, extraprosodicity affects aux/preverb in NPST, i.e. *gbo* surfaces as contour HL tone (*gbô*). Consequently verbal phrase left edge is consistently non-high with right edge PFV aspect but high with IPFV.

/ó	ò HAB	gbo	dume-`L	émà/	>	[ó	ò	gbó	dùmè	émà]

/ò	ó PROG	gbo	dume-ˋᴸ	émà/	>	[ò	ó	gbó	dùmè	émà]
/ó	ˋᴸ RPST	gbo	dume-í	émà/	>	[ó		↓gbó	dúmé	émà]
/ò	ʼᴴ NPST	gbo	dume-í	émà/	>	[ò		gbô	dúmé	émà]
/ó	ló RFUT	gbo	dume-í	émà/	>	[ó	ló	gbò	dúmé	émà]
/ò	ló NFUT	gbo	dume-í	émà/	>	[ò	ló	gbò	dúmé	émà]

Drawing on Hyman (2014) and Hyman et al. (2020), we conclude with discussion of the demarcative function of Emai grammatical tone. Morphophonological processes for TA that impact verbal left and right edge evidence prosodic asymmetry. At LE processes are bounded, perseverative, and dissimilatory with respect to a right adjacent syllable of verb or aux/preverb, while at RE they are unbounded (spread up to verb onset), anticipatory (verb tone assimilating to -*í*) and dissolutive (-*í* dissolving before verb argument). Moreover, NPST's extraprosodic HL contour conditioned by aux/preverb but not verb reveals complexity in the co-variation between tense tone and one of its semantic values (Plank 1998). This suggests that prosodic expression of NPST (cf. perfect with present relevance; factative) may deserve greater scrutiny, especially across West Africa.

References: Elugbe, Ben. 1973. A comparative Edo phonology. Ibadan: University of Ibadan. Ph.D. dissertation. Elugbe, Ben. 1989. Comparative Edoid: Phonology and Lexicon. Port Harcourt: University of Port Harcourt Press. Franich, Kathryn H. 2014. Contour tones and prosodic structure in Medumba. BLS 40: 102-124.Hyman, Larry M. 2008. Directional Asymmetries in the Morphology and Phonology of words, with special reference to Bantu. Linguistics 46, 2: 309-350. Hyman, Larry M. 2014. Toward a canonical typology of prosodic systems. In Esther Herrera Zendejas (ed.), Tono, Acento y Estructuras Métricas en Lenguas Mexicanas. 13-38. México: El Colegio de México. Hyman, Larry M. 2020. Prosodic asymmetries in nominal vs. verbal phrases in Bantu. Invited talk. 6th International Conference on Phonetics and Phonology, December 15, 2019, Japan. Hyman, Larry M. & Florian Lionnet. 2012. Metatony in Abo (Bankon), A42. In Michael R. Marlo et al. (eds.), Proceedings of 42nd Annual Conference of African Linguistics, University of Maryland, College Park, 1-14. Cascadilla Proceedings Project. Hyman, Larry M. and Francis Katamba. 2010. Tone, syntax, and prosodic domains in Luganda. In Laura J. Downing et al. (eds.), ZAS Papers in Linguistics 53:69-98. Hyman, Larry M., Hannah Sande, Florian Lionnet, Nicholas Rolle & Emily Clem. 2020. Prosodic systems : Niger-Congo and Adjacent Areas. In Carlos Gussenhoven & Aoju Chen (eds.), Oxford Handbook of Language Prosody. Plank, Frans. 1998. The co-variation of phonology with morphology and syntax: A hopeful history. Linguistic Typology 2: 195-230.

Edge-asymmetries in affix order

Marie-Luise Popp
Leipzig University
marie_luise.popp@uni-leipzig.de

When examining the order of affixes, most theoretical approaches focus on discussing the relative order of affixes claiming that these are driven by syntactic (Baker 1985, among others) or semantic factors (Rice 2000, among others). The side of affixation, that is whether an affix will be attached as a prefix or a suffix is taken to be orthogonal to the relative order of affixes. However, this assumption is seriously questioned by two influential contributions by Julien (2002) and Trommer (2003).

First, a crosslinguistic investigation of the order of tense and aspect by Julien (2002) reaches the conclusion that aspect is closer to the stem than tense In addition, Julien (2002) reveals a typological gap in mixed systems where one category is a prefix while the other is a suffix. In these systems, only the order Tense-V-Aspect is attested. Second, Trommer (2003) examines the position of person and number markers on the verbs. Trommer (2003) shows that there is a strong tendency that person markers tend to be prefixes while number markers tend to be suffixes. Crucially, neither the generalization by Julien (2002) nor the generalization by Trommer (2003) can be explained by taking the side of affixation into account.

Moreover, there are numerous prefixing languages that have been argued to exhibit so-called *templatic morphology,* that is the relative order between affixes is rigid and does not seem to follow syntactic or semantic factors, e.g. Murrinh-Patha (Nordlinger 2010), Adyghe (see e.g. Ershova 2019 or Arkadiev 2020) and Oneida (Diaz et al 2019), among many other examples.

In this talk, I will examine edge-asymmetries in affix ordering patters of 24 languages pursuing the following questions:

– Do affix ordering patterns in prefixing languages differ from suffixing languages?

– How widespread are *gaps* in the relative order of grammatical categories in the sense of Julien (2002) and Trommer (2003)?

– How can asymmetries be modelled in linguistic theory (e.g. via one-directed morphlogical movement?)

References: Arkadiev, Peter. 2020. Non-canonical inverse in Circassian languages. *STUF – Language Typology and Universals* 73(1). 81–111. Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16(3). 373–415. Diaz, Thomas S., Jean-Pierre Koenig & Karin Michelson. 2019. Oneida prepronominal prefixes in Information-based Morphology. *Morphology* 29. 431-473. Ershova, Ksenia. 2019. Syntactic ergativity in West Circassian. University of Chicago dissertation Julien, Marit. 2002. *Syntactic heads and word formation*. Oxford: Oxford University Press Nordlinger, Rachel. 2010. Verbal morphology in Murrinh-Patha: Evidence for templates. *Morphology* 20(2). 321–341. Rice, Keren. 2000. *Morpheme order and semantic scope: Word formation in the Athapaskan verb.* Cambridge Studies in Linguistics. Cambridge: Cambridge University Press. Trommer, Jochen. 2003. The interaction of morphology and syntax in affix order. In *Yearbook of morphology 2002*, 283–324. Springer.

Arbeitsgruppe 12 (Kurz-AG)

Eye-tracking and language production
Martina Penke, Judith Schlenter & Elyesa Seidel

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Sentence planning and production in two Australian free word order languages

Gabriela Garrido Rodriguez, Sasha Wilmoth, Rachel Nordlinger, Evan Kidd

ARC CoEDL, University of Melbourne, Australia, ARC CoEDL, University of Melbourne, Australia, , ARC CoEDL, University of Melbourne, Australia, ARC CoEDL, The Australian National University, Australia; Max Planck Institute for Psycholinguistics, The Netherlands
gabriela.garridorodriguez@unimelb.edu.au, s.wilmoth@student.unimelb.edu.au, racheln@unimelb.edu.au, evan.kidd@mpi.nl

In this paper we present the results from two eye-tracking studies that investigated sentence planning and production in Murrinhpatha (non-Pama-Nyungan, Southern Daly) and Pitjantjatjara (Pama-Nyungan, Western Desert language), two unrelated Australian Indigenous languages. While both languages have been described as having flexible word order, they differ significantly on several relevant typological dimensions: Murrinhpatha is polysynthetic and head-marking, containing only vestigial dependent marking via the optional use of ergative marking in some contexts (Walsh 1976; Nordlinger 2010; Mansfield 2019). In contrast, Pitjantjatjara is ergative and dependent-marking, with no verbal agreement morphology (Bowe 1990). We ask: (i) what influences the production of different word orders, and (ii) how does speaking a free word order language influence sentence planning?

Native speakers of both languages (Murrinhpatha, N=43; Pitjantjatjara, N=49) completed a picture description task while their eye-movements were recorded. Our method closely followed Norcliffe, Konopka, Brown, and Levinson (2015). There were 48 target pictures that depicted two-participant events (e.g., a crocodile biting a man) interspersed amongst 93 filler pictures (e.g., intransitive events). The resulting picture descriptions were transcribed and coded for word order, and participants' eye movements were analyzed using multilevel logistic regression (Baayen, Davidson, and Bates 2008; Barr 2008; Jaeger 2008).

The results show that participants from both languages produced all possible orderings of S, O and V in the experimental corpus. As in past studies (Norcliffe et al. 2015; Sauppe et al. 2013; Ferreira and Yoshita 2003; Christianson and Ferreira 2005), differences in word order were sensitive to the different configurations of Agent and Patient humanness. Specifically, the humanness of patients plays an important role in A-initial sentences. In contrast, human agents were more likely to condition P-initial and V-initial sentences, but in interaction with P humanness. Our analyses of the eye-movement data suggest that sentence planning in these languages is best described as a weakly hierarchical process (Griffin and Bock 2000; Konopka and Meyer 2014), with no evidence to suggest that bottom-up perceptual cues drive word order selection (cf. Gleitman et al. 2007). Notably, the results suggest that speaking a free word order language results in a rather different pattern of sentence formulation than in languages with fixed word orders: speakers' gaze was more evenly distributed across the two characters, providing evidence of very early relational encoding during event apprehension that differed across A-initial and P-initial word orders. This suggests that Murrinhpatha and Pitjantjatjara speakers lay down a very early conceptual representation of the event, which guides their subsequent linguistic encoding and production (see Figure 1). This pattern of early relational encoding is consistent across the two languages, despite their typological differences.

Our results suggest that sentence planning is significantly affected by typological properties such as free word order and support the growing body of research revealing significant cross-linguistic differences in sentence production that are linked to grammatical properties of languages.
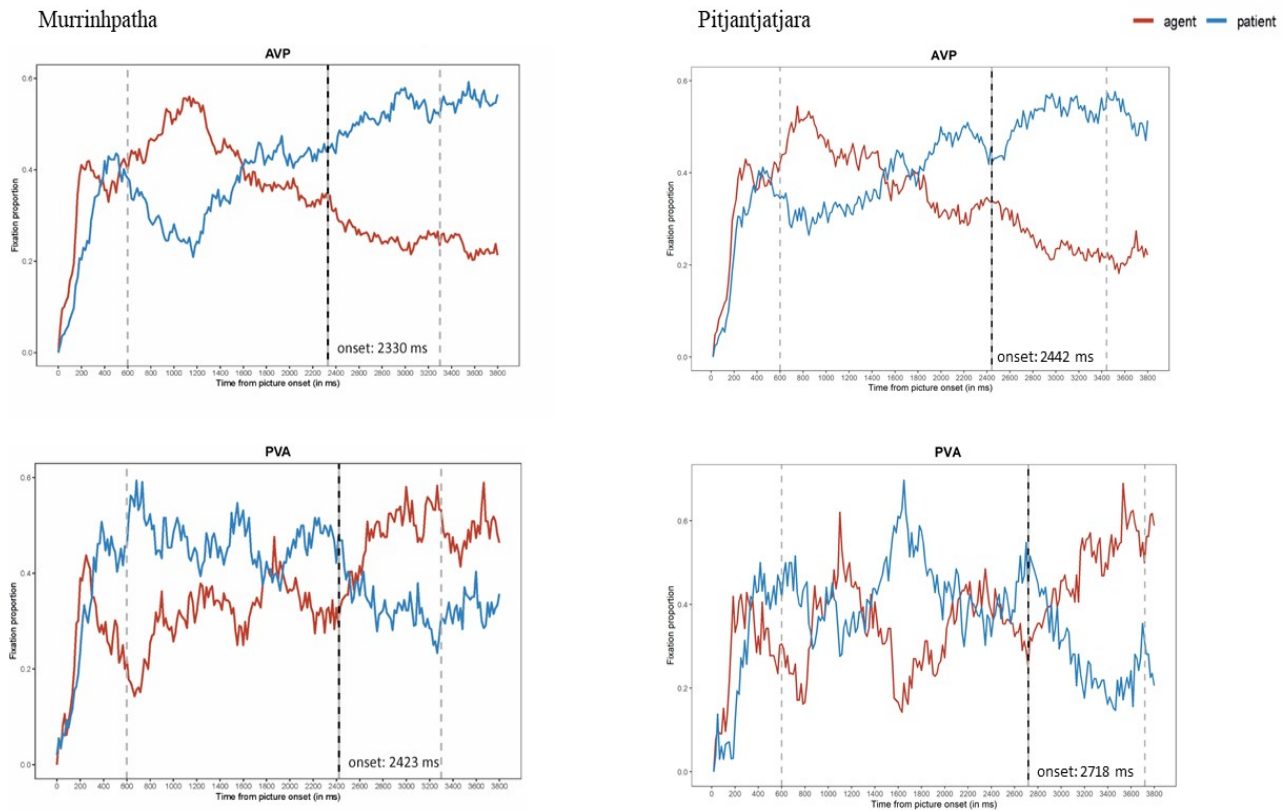
Figure 1. Time course graphs showing proportion of agent- and patient- directed fixations in AVP and PVA sentences in Murrinhpatha and Pitjantjatjara

References: Baayen et al. *J Mem Lang*, 2008, 59, 390-412; Barr, D.J. *J Mem Lang*, 2008, 59, 457-474; Bowe, H.J., *Routledge*. 1990; Christianson et al. Cognition, 2005, 98, 105-135; Ferreira et al. *J Psycholinguist Res*, 2003, 32, 669-692; Gleitman et al. *J Mem Lang*, 2007, 57, 544-569; Griffin et al. *Psychol Sci*, 2000, 11, 274-279; Jaeger, T. *J Mem Lang*, 2008, 59, 434-446; Konopka et al. *Cognitive Psychol*, 2014, 73, 1-40; Norcliffe et al. *Lang Cogn Neurosci*, 2015, 30, 1187-1208; Nordlinger, R. *Morphology*, 2010, 20, 321-341; Mansfield, J., *de Gruyter Mouton*, 2019; Sauppe et al. In *CogSci*, 2013; Walsh, M. *AIAS*. 1976)

Linear vs. structural incrementality in the face of sentence production in context

Xiaogang Wu, Johannes Gerwien
Guangdong University, Heidelberg University
Xiaogang.eu@hotmail.com, jo.gerwien@uni-heidelberg.de

In three picture description experiments with German native speakers, we manipulated the discourse status of the agent and patient in transitive scenes by adding a mini lead-in discourse before Ss were to produce the target responses (Prat-Sala & Branigan 2000), as well as the sequence of participant's visual attention directed to the event referents by means of an attention capture technique (Gleitman et al. 2007). Our aim was to determine (1) whether the discourse status of the referents as a factor for syntactic choices outranks thematic role information and visual salience; (2) how contextual information is mapped from the conceptual representation to linear positions of a sentence; and (3) whether sentence production in context is a linear or structural incremental formulation process.

Ss first saw an introduction picture that showed two animate referents which appeared again as the participants of a transitive scene in a subsequent target picture. While being exposed to the introduction picture, Ss heard a mini discourse in which the referents' names were introduced by a first sentence ("In this picture, you see a king and a soldier."), and a specific response about the target picture was evoked by a second sentence. In Exp 1, the second sentence was „Please tell me what happens in the next picture". In Exp 2 and 3, the format of the sentence was „In the next picture you will see the [Referent 1] again, please tell me what happens to [Referent 2]". In half of the trials, [Referent 1] was the agent in the target picture and [Referent 2] was the patient, in the other half of the trials this was reversed. This manipulation allowed us to determine whether the agent or the patient in the target pictures received the role of the discourse topic. After the introduction phase, Ss saw a fixation cross (500ms) and a blank screen (200ms) and then a visual cue (80ms). In Exp 1, the cue appeared at the position of the following agent or patient. In Exp 2, the cue appeared at the position of the following patient. In Exp 3, the cue appeared at the position of the following agent. The cue was immediately followed by the target picture. Filler items followed the general format of the critical items, but target pictures showed non-transitive scenes. Ss were instructed to respond as quickly as possible. Eye movements were recorded.

In Exp 1, the analysis of first fixation locations indicated that, without the manipulation of discourse status, Ss' visual attention was drawn to the cued entity in more than 75% of all trials. However, regardless of the first fixation location, Ss directed their second saccade (the first 'voluntary' saccade) in more than 90% of all cases to the agent. Almost 100% of the produced sentences mentioned the agent as the first referent. Thus, the manipulation of visual attention did not predict syntactic choices. With the introduction of topicality in Exp. 2 and 3, we found a modulating effect of discourse status on visual cueing. Ss' visual attention was drawn to the cued entity in more than 79% of all trials when this entity carried topic status, whereas the proportion dropped to 37% when the cued entity was not topical. With respect to syntactic choices, we found that topical referents, regardless of their semantic role, always appeared in the sentences' initial position. Moreover, we found that producing an active sentence (with agent as topic) was significantly faster than producing a passive sentence (with patient as topic). We conclude that (1) information structure specifications (ISS) influence sentence production from early on by directing processing attention to topical entities, (2) ISS outrank thematic role information, as well as the visual salience of referents, and (3) ISS are mapped onto the thematic grid before grammatical encoding (function assignment/linearization) begins. We expand the notion of structural incrementality to the level of information structure, showing that

information structure, together with the semantic structure influences syntactic encoding. These results indicate that German sentence production in context is mainly a structural incremental formulation process.

Implicit perceptual priming in context: When the prominent patient meets the eye

Yvonne Portele
Goethe University Frankfurt
y.portele@gmail.com

Implicit visual cues have been found to influence speakers' structural choices in English (e.g., Gleitman et al. 2007). Presenting a cue at the position of the subsequent patient promotes passive picture descriptions compared to agent cueing. In more flexible languages, patient cueing does not show reliable influences on speakers' structural choices (e.g., Hwang & Kaiser 2015 for Korean). Myachykov et al. (2011, p. 103) proposed that "*[…] speakers universally attempt to employ the grammatical-role assignment mechanism in order to represent the perceptually salient referent in the structural plan of the sentence.*" The absence of significant influences might therefore be attributed to the low availability of passives rather than typological differences of languages. **Experiment.** To investigate the question whether German speakers employ grammatical-role assignment to represent perceptually salient referents, perceptual priming was combined with a manipulation of *derived accessibility*. *Topic questions* (e.g., What happens to X?), have been found to influence syntactic choices cross- linguistically. Depending on the choices available, speakers employ grammatical-role assignment and/or word order linearization to promote the topic. Prior work in German has shown that topic patients are promoted by producing passives, which is why a topic manipulation was included to investigate the *availability* of passives. **Participants**. Forty-four students (native speakers of German) of the Goethe University Frankfurt participated. **Materials**. Twenty-four experimental sets consisting of two context sentences, a question, and a target picture were created. The first sentence (1) introduced the two male characters participating in the respective target picture (4) of the trial. In the second sentence (2), only the subsequent target patient was rementioned (additionally making him more *prominent* compared to the agent). The question following the context was either a general (3a) or a patient (3b) question, with the patient question serving as topicalization strategy.

1) Jetzt geht es um einen siegreichen Boxer und einen Trainer in einer Sporthalle.
   Now, there is a victorious boxer and a coach in a gym.

2) Der Boxer hat bereits seine Kampfkleidung an—Boxhandschuhe sowie Schuhe.
   The boxer already wears his sportswear—boxing cloves and shoes.

3) a. General question: Was passiert? (*What happens?*) or
   b. Patient question: Was passiert mit ihm? (*What happens to him?*)

4) Transitive black & white drawing (balanced for the referent positioned left), e.g., showing a coach measuring a boxer (see Fig 1).

Procedure. After reading the context and question aloud, a crosshair at the lower bottom of the screen had to be fixated for at least 150 ms. Afterwards, the visual cue (black dot with a diameter of 0.66 cm) appeared at the center of the interest area for the subsequent agent or patient for 60 ms, immediately followed by the target picture (see Fig 1). **Results**. A generalized linear mixed model showed a main effect of *Question* (i.e., topic status). Participants were significantly more likely to produce passives following patient compared to general questions (Fig 2). There was no effect of *Cue* (agent vs. patient cue) on participants' structural choices. Eye tracking data (Fig 3) showed the cueing manipulation itself was effective in attracting participants' fixations. **Discussion.** Whereas derived accessibility (here: topic status) seems to have universal influences on speakers grammatical encoding, perceptual accessibility in terms of implicit visual cueing does not seem to reliably influence speakers when producing syntactic structures. Missing effects due to implicit priming in German have now been reported in two further studies conducted by other research groups. The puzzling cross-linguistic findings are discussed in context of the question why flexible languages might not make use of options provided by

grammar and within the recent research context of word- vs. structure driven development during grammatical encoding.
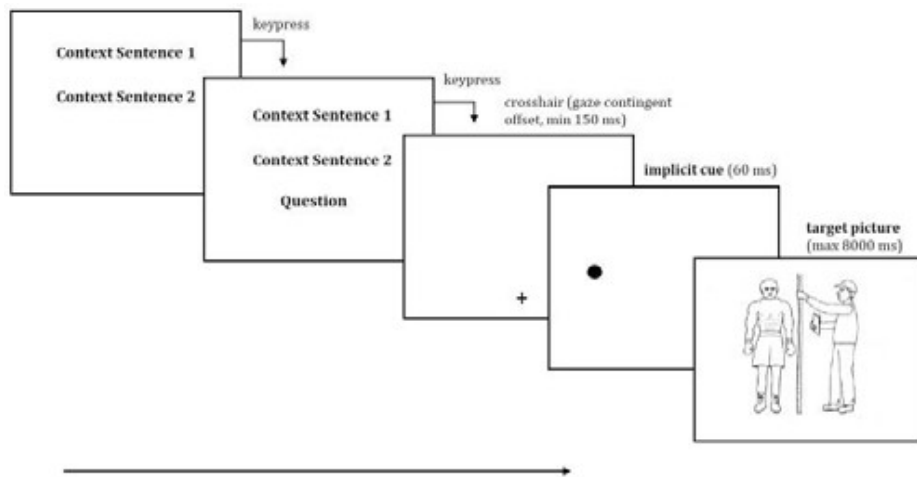

Figure 1: Procedure for a patient cued trial of the Experiment.
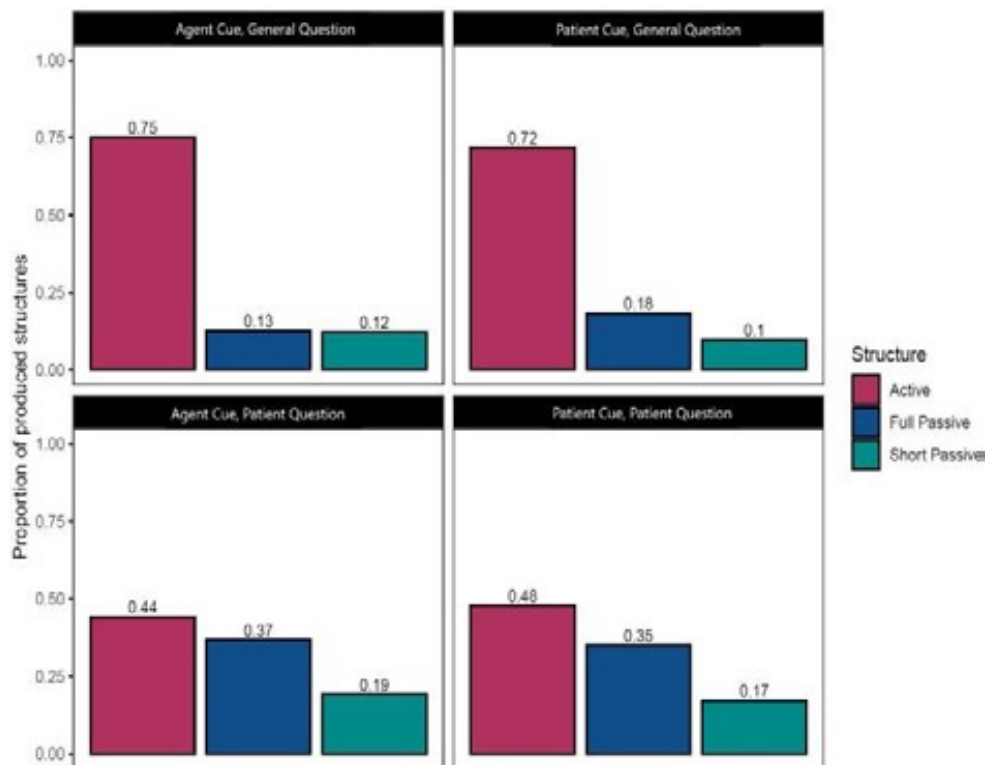

Figure 2: Mean proportions of produced target descriptions (structures) in the different conditions.
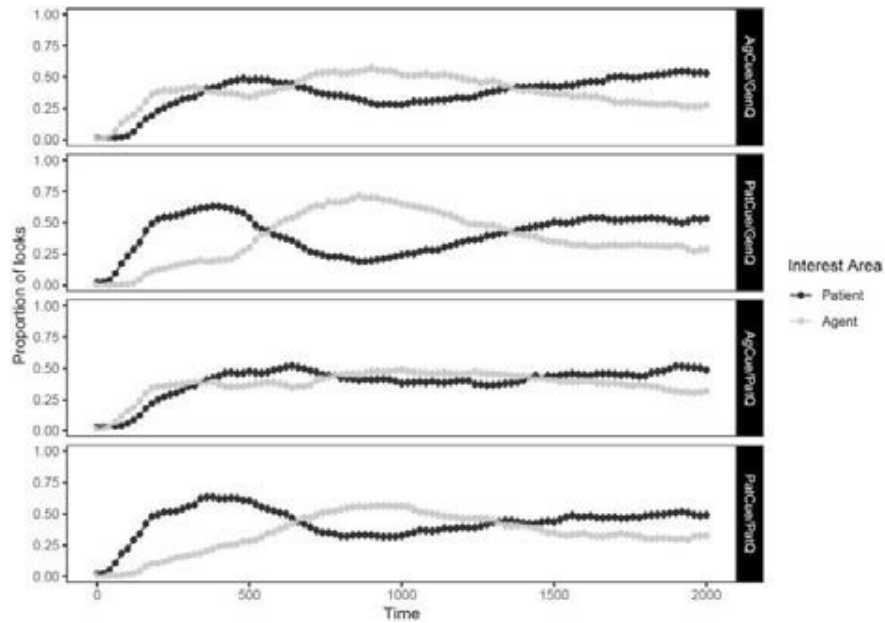
Figure 3: Proportion of looks to patient/agent interest area relative to time course (ms). 0 = picture onset.

References: Gleitman, Lila R., David January, Rebecca Nappa, and John C. Trueswell. 2007. On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language* 57(4): 544-569. Hwang, Heeju, and Elsi Kaiser. 2015. Accessibility effects on production vary cross-linguistically: Evidence from English and Korean. *Journal of Memory and Language* 84: 190-204. Myachykov, Andriy, Dominic Thompson, Christoph Scheepers, and Simon Garrod. 2011. Visual attention and structural choice in sentence production across languages. *Language and Linguistics Compass* 5(2): 95-107.

Case marking shapes the time-course of sentence planning: Crosslinguistic evidence from Hindi, Yélî Dnye, Japanese, Basque and Swiss German

Sebastian Sauppe, Elisabeth Norcliffe, Kamal K. Choudhary, Agnieszka E. Konopka, Aitor M. Egurtzegi, Nathalie Giroud, Shikha Bhattamishra, Mahima Gulati, Gabriela Garrido Rodriguez, Damian E. Blasi, Ina Bornkessel-Schlesewsky, Itziar Laka, Martin Meyer, Stephen C. Levinson, Balthasar Bickel
University of Zurich, University of York, Indian Institute of Technology Ropar, University of Aberdeen, University of Zurich, University of Zurich, Indian Institute of Technology Ropar, Indian Institute of Technology Ropar, University of Melbourne, Harvard University, University of South Australia, University of the Basque Country, University of Zurich, Max Planck Institute for Psycholinguistics, University of Zurich
sebastian.sauppe@uzh.ch, elisabeth.norcliffe@york.ac.uk, kamal@iitrpr.ac.in, agnieszka.konopka@abdn.ac.uk, aitor.martinezegurcegui@uzh.ch, nathalie.giroud@uzh.ch, shikha.mishra@iitrpr.ac.in, mahima.gulati@iitrpr.ac.in, gabriela.garrido@mpi.nl, damianblasi@gmail.com, Ina.Bornkessel-Schlesewsky@unisa.edu.au, itziar.laka@ehu.eus, martin.meyer@uzh.ch, Stephen.Levinson@mpi.nl, balthasar.bickel@uzh.ch

Speakers adapt the time course of sentence planning, e.g., in response to extralinguistic factors such as time pressure (Ferreira & Swets, 2002) or visual salience of referents (Gleitman et al., 2007; Myachykov et al., 2011). It remains largely unknown, however, how planning is adapted to different grammars and whether adaptations to grammatical affordances are the same across languages – mainly due to a lack of comparative cross-linguistic production research (Jaeger & Norcliffe, 2009; Norcliffe, Harris, et al., 2015). We present three separate eye-tracked picture description studies exploring how early planning is shaped by the signaling of agent-verb dependencies across a wide range of languages. Some languages assign the so-called ergative case to the subjects of agentive verbs, distinguishing them from most other arguments. Other languages assign the same nominative (or no) case to both transitive agents and most other arguments (Bickel et al., 2015). We hypothesized that planning ergative sentences requires more extensive and earlier relational and structural encoding (Griffin & Bock, 2000; Konopka, 2019) because verb agentivity is signaled on the sentence-initial noun through case, while nominative case does not reveal transitivity early on. In our studies, participants described drawings of transitive events while their gaze was tracked (Griffin & Bock, 2000). We contrasted the production of transitive ergative and nominative sentences in languages with different case marking properties: Study 1 compared canonical SOV sentences in Yélî Dnye, an ergative language from Papua New Guinea (with ergative case marking for all transitive agents), and Japanese sentences with nominative agents (*N* = 59). Study 2 compared subordinate SOV sentences in Basque (ergative marking for all agents) and Swiss German (nominative, *N* = 66). Study 3 employed a within-language comparison in Hindi by capitalizing on the split nature of the case marking system in this language, in which transitive agents in perfective aspect sentences receive ergative case, but agents in imperfective aspect sentences carry nominative case (*N* = 50).

Mixed-effects growth curve analyses of fixation likelihoods on the single-trial level (Cho et al., 2018; Jaeger, 2008; Mirman et al., 2008; Sauppe, 2017) revealed that speakers of all tested ergative languages gazed less towards agent referents in the to-be-described pictures during the first 800 ms of each trial when planning ergative sentences as compared to nominatives. This means that they distributed their visual attention more over relational information of the depicted events during the earliest phases of planning (Konopka, 2019; Sauppe, 2017). This pattern indicates an early prioritization of planning relational information in order to determine verb transitivity and thus to know whether to assign ergative case. By contrast, when planning nominatives, speakers prioritized encoding of the sentence-initial agent, which did not signal information about the verb. These findings demonstrate that the time-course of sentence planning is systematically shaped by grammatical affordances across diverse languages, and

therefore strongly support hierarchically incremental accounts of planning that require speakers to commit to structures spanning entire sentences (initial agent – final verb) at the outset of sentence formulation (Griffin & Bock, 2000; Norcliffe, Konopka, et al., 2015; Sauppe et al., 2013).

References: Bickel, B., Witzlack-Makarevich, A., Choudhary, K. K., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2015). The neurophysiology of language processing shapes the evolution of grammar: evidence from case marking. *PloS One*, 10(8), e0132819. https://doi.org/10.1371/journal.pone.0132819. Cho, S.-J., Brown-Schmidt, S., & Lee, W.-Y. (2018). Autoregressive Generalized Linear Mixed Effect Models with Crossed Random Effects: An Application to Intensive Binary Time Series Eye-Tracking Data. *Psychometrika*. https://doi.org/10.1007/s11336-018- 9604-2. Ferreira, F., & Swets, B. (2002). How Incremental Is Language Production? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums. *Journal of Memory and Language*, 46(1), 57–84. https://doi.org/10.1006/jmla.2001.2797. Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, 57(4), 544–569. https://doi.org/10.1016/j.jml.2007.01.007. Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11(4), 274–279. https://doi.org/10.1111/1467-9280.00255. Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, 59(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007. Jaeger, T. F., & Norcliffe, E. J. (2009). The Cross-linguistic study of sentence production. *Language and Linguistics Compass*, 3(4), 866–887. https://doi.org/10.1111/j.1749- 818X.2009.00147.x**.** Konopka, A. E. (2019). Encoding actions and verbs: Tracking the time-course of relational encoding during message and sentence formulation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 45(8), 1486–1510. https://doi.org/10.1037/xlm0000650. Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, 59(4), 475–494. https://doi.org/10.1016/j.jml.2007.11.006. Myachykov, A., Thompson, D., Garrod, S., & Scheepers, C. (2011). Referential and visual cues to structural choice in visually situated sentence production. *Frontiers in Psychology*, 2, 396. https://doi.org/10.3389/fpsyg.2011.00396. Norcliffe, E., Harris, A. C., & Jaeger, T. F. (2015). Cross-linguistic psycholinguistics and its critical role in theory development: early beginnings and recent advances. *Language, Cognition and Neuroscience*, 30(9), 1009–1032. https://doi.org/10.1080/23273798.2015.1080373. Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. C. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*, 30(9), 1187–1208. https://doi.org/10.1080/23273798.2015.1006238. Sauppe, S. (2017). Word Order and Voice Influence the Timing of Verb Planning in German Sentence Production. *Frontiers in Psychology*, 8, 1648. https://doi.org/10.3389/fpsyg.2017.01648. Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin, R. D., & Levinson, S. C. (2013). Dependencies first: eye tracking evidence from sentence production in Tagalog. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35.

Attention and syntactic choice: Evidence from Russian and English

Mikhail Pokhoday, Yury Shtyrov, Andriy Myachykov
Higher School of Economics Moscow, Aarhus University and Higher School of Economics Moscow, Northumbria University
mikhail.pokhoday@gmail.com, andriy.myachykov@northumbria.ac.uk, yury@cfin.au.dk

In a fully developed sentence production system, perceptual input provides initial information about the event to be described. The speaker's attention foregrounds relevant/important information for the conceptual analysis, and subsequent language production mechanisms collaborate to select between simultaneously available syntactic alternatives. Existing evidence suggests that the system responsible for assigning the grammatical roles is sensitive to the distribution of the speaker's attention within the described scene (Myachykov, Pokhoday and Tomlin, 2018, *for a recent review*). For example, a speaker of English is more likely to select a passive-voice frame when her attention is directed to the patient of the described event and she is more likely to use an active-voice frame when the agent is in her attentional focus (e.g., Gleitman, et al., 2007; Myachykov, et al., 2012) indicating a regular interplay between attention and syntactic choice. At the same time, these and other similar studies exclusively use variants of the visual cueing paradigm (Posner, 1980). As a result, the reported link between attention and syntactic choice cannot be generalized beyond the visual modality. A more ecologically valid proposal needs to consider sensory integration and a multi-modal nature of attention (Driver and Spence, 2004).

Here, we report the results of a series of experiments that compared effects of perceptual priming on syntactic choice in Russian and English. English and Russian native speakers described transitive events while their attention was directed to the agent or the patient by means of (1) a visual (lateral cue preceding event presentation), (2) an auditory (lateral beep presented via headphones), or (3) a motor (lateral key press prior to event presentation) cue. Hence, two factors were manipulated: (1) the Cued Referent (Agent/Patient) and (2) the Cue Type (Visual/Auditory/Motor). The likelihood of producing a passive-voice sentence was the dependent variable. First, we replicated previous findings by registering a main effect of Cued Referent (more passive-voice sentences in Patient-Cue condition). Second, there was a main effect of Cue Type (more passive-voice sentences with visual and motor cues compared to the auditory cue). Third, there was no interaction between the two factors suggesting that only one attentional modality at a time can impact syntactic choice. Also, Russian speakers in comparison to their English language counterparts relied more on linear ordering alterations in structure, rather that active-passive voice alterations. This is in line with previous literature (for example Myachykov and Tomlin, 2008) and the fact that Russian is a flexible word order language which allows various "unconventional" word orders (like SOV, OSV etc.). Overall, our findings replicate previous findings using visual cueing paradigm and generalize them to auditory and motor perceptual modalities.

References: Levelt, W. J. 1993. *Speaking: From intention to articulation (Vol. 1)*. MIT press. Myachykov, A., & Tomlin, R. S. (2008). Perceptual priming and structural choice in Russian sentence production. *Journal of Cognitive Science*, *6*(1), 31-48. Myachykov, A., Garrod, S., & Scheepers, C. 2012. Determinants of structural choice in visually situated sentence production. *Acta psychologica*, *141*(3), 304-315. Myachykov, A. and Tomlin, R. S. 2014. Attention and salience. In Dabrowska, E., & Divjak, D. (Eds.). 2015. *Handbook of Cognitive Linguistics (Vol. 39)*. Walter de Gruyter GmbH & Co KG. Posner, M.I.1980.Orienting of attention. *Quarterly Journal of Experimental Psychology,32*, 3-25

Extracting event structure at a glance: The role of case during scene apprehension for speaking

Arrate Isasi-Isasmendi, Sebastian Sauppe*, Caroline Andrews*, Monique Flecken, Moritz Daum, Itziar Laka, Martin Meyer, Balthasar Bickel
University of Zurich, University of Zurich, University of Zurich, University of Amsterdam, University of Zurich, University of the Basque Country, University of Zurich, University of Zurich (* - contributed equally)
arrate.isasi-isasmendilandaluze@uzh.ch, sebastian.sauppe@uzh.ch, caroline.andrews@uzh.ch, monique.flecken@mpi.nl, daum@psychologie.uzh.ch, itziar.laka@ehu.eus, martin.meyer@uzh.ch, balthasar.bickel@uzh.ch

When describing pictures, sentence production begins with the apprehension of the depicted event (Griffin and Bock, 2000) where speakers extract the event's "gist", including event roles (the "who does what to whom"), rapidly—in as little as 100–300 ms (Dobel et al., 2007; Hafri et al., 2013). Event apprehension has been argued to be a prelinguistic process (Bock and Levelt, 1994; Griffin and Bock, 2000), i.e., that grammar plays no role (yet) in speakers' gist extraction and only later impacts the message and linguistic encoding. Here, we present two brief exposure experiments (Gerwien and Flecken, 2016) on Basque and Spanish exploring whether the earliest stages of encoding events for speaking is influenced by a language's grammar. Basque and Spanish differ in their case marking systems: Agentive subjects are marked by ergative case (*-k*) in Basque, while patients (subjects of unaccusative intransitive verbs and transitive objects) receive absolute case (shown by pseudo-Basque in 1–2).

(1)   *Transitive:* Lisa-k_{ERG} Alex-$\oslash_{ABS}$ greeted → *Lisa greeted Alex*

(2)   *Intransitive:* Lisa-$\oslash_{ABS}$ arrived → *Lisa arrived*

In Spanish, by contrast, all subjects carry the same unmarked nominative case regardless of their thematic role, while objects are marked accusative. Prior picture description studies in our lab showed that for ergative-aligned sentences, speakers inspect event- relational information more intensively in early planning because they need to select the verb transitivity type early in order to select the right case marker (Sauppe et al., *forthcoming*). We hypothesized that the need to commit to argument structure early in ergative sentences affects not only linguistic encoding processes but can already shape early event apprehension.

       In our experiments, participants saw photographs of events with four different actors (e.g. a man watering a plant or a woman dragging a bag) for 300 ms in a randomized corner of the screen. As planning and executing saccades takes up to 200 ms (Pierce et al., 2019), this left ~100 ms to take up visual information foveally after the gaze shifted from a central fixation cross into the picture. This brief exposure paradigm provides access to event apprehension via fixation patterns and attention allocation because further planning steps can't be executed in this time frame (Gerwien and Flecken, 2016). After exposure, participants either typed a sentence description or performed a probe recognition task (not reported here). In Experiment 1 (online, w/out eyetracking) native speakers of Basque (*N*=90) and Spanish (*N*=88) typed out their descriptions of 58 images; in Experiment 2 (w/ eyetracking) descriptions were given orally. We analysed the accuracy with which participants identified agents and patients in questions or descriptions in Exp. 1 (e.g., *Paul* vs *someone*, *an orange* vs *something*) with Bayesian hierarchical logistic regression (Bürkner, 2017) to assess whether speakers focused on individual event roles or on the overall event and the relations therein. Patients were described less accurately than agents overall across both languages (log odds: mean $\hat{\beta}$ = -0.77, *P(β< 0)* = 1). Patients were, however, described more accurately by Basque speakers (log odds: mean $\hat{\beta}$ = 0.06, *P(β > 0)* = 0.93), suggesting that they directed more attention to them than Spanish speakers. In Exp. 2 we tracked the location of participants' fixations on the stimuli (*N*=64, SMI RED250 mobile, sampling at 250 Hz). We found that Basque speakers fixated more often on agents than Spanish speakers ($\hat{\beta}$ = 0.08, *P(β >0)* = 0.95), while Spanish speakers fixated more often on patients than Basque

speakers ($\hat{\beta}$ = -0.09, *P(β >0)* = 0.98).

Our results suggest that the specific grammatical features of different languages (here: the case marking system) shape not only relational, structural and linguistic encoding processes (Norcliffe et al., 2015; Norcliffe and Konopka, 2015; Sauppe et al., 2013) but also affect the earliest stage of sentence planning, event apprehension.

References: Bock, Kathryn, and Willem J.M. Levelt. 1994. "Language production: Grammatical encoding." In Morton A. Gernsbacher, eds. *Handbook of psycholinguistics*. 945–984. Bürkner, Paul C. 2017. "brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software, Articles*, 80(1): 1–28. Dobel, Christian, Heidi Gumnior, Jens Bölte, and Pienie Zwitserlood. 2007. "Describing scenes hardly seen." *Acta Psychologica*, 125(2): 129–143. Gerwien, Johannes, and Monique Flecken. 2016. "First things first? Top-down influences on event apprehension." In Anna Papafragou, Daniel Grodner, Daniel Mirman, and John C. Trueswell, eds. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*. Cognitive Science Society. 2633–2638. Griffin, Zenzi M., and Kathryin Bock. 2000. "What the eyes say about speaking." *Psychological Science*, 11(4): 274–279. Hafri, Alon, Anna Papafragou, and John C. Trueswell. 2013. "Getting the gist of events: recognition of two-participant actions from brief displays." *Journal of Experimental Psychology. General*, 142(3): 880–905. Norcliffe, Elisabeth, and Agnieszka E. Konopka. 2015. "Vision and Language in Cross- Linguistic Research on Sentence Production." In Ramesh K. Mishra, Narayanan Srinivasan, and Falk Huettig, eds. *Attention and Vision in Language Processing*. Springer India. 77–96. Norcliffe, Elisabeth, Agnieszka E. Konopka, Penelope Brown, and Stephen C. Levinson. 2015. "Word order affects the time course of sentence formulation in Tzeltal." *Language, Cognition and Neuroscience*, 30(9): 1187–1208. Pierce, Jordan E., Brett A. Clementz, and Jennifer E. McDowell. 2019. "Saccades: Fundamentals and Neural Mechanisms." In Christoph Klein and Ulrich Ettinger, eds. *Eye Movement Research: An Introduction to its Scientific Foundations and Applications* . Springer International Publishing. 11–71. Sauppe, Sebastian, Kamal K. Choudhary, Nathalie Giroud, Damian E. Blasi, Elisabeth Norcliffe, Shikha Bhattamishra, Mahima Gulati, Aitor M. Egurtzegi, Ina Bornkessel- Schlesewsky, Martin Meyer, and Balthasar Bickel. *forthcoming*. "Neural signatures of syntactic variation in speech planning". *PLOS Biology*. Sauppe, Sebastian, Elisabeth Norcliffe, Agnieszka E. Konopka, Robert D. Van Valin, and Stephen C. Levinson. 2013. "Dependencies first: eye tracking evidence from sentence production in Tagalog." *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35.

Using eye-tracking to gauge the effect of phonological dependencies on planning

Emiel van den Hoven, F.-Xavier Alario, Audrey Bürki
University of Potsdam, CNRS & Aix-Marseille Université, University of Potsdam
emiel.vandenhoven@uni-potsdam.de, francois-xavier.alario@univ-amu.fr, buerki@uni-potsdam.de

We investigated how the presence or absence of phonological dependencies affects the scope of phonological advance planning – how much of the utterance speakers phonologically encode before they start articulating. The form of some pre-nominal adjectives in French depends on the phonological and morphosyntactic context. For instance, the adjective *nouveau* ("new") takes the form *nouveau* [nuvo] when preceding a (singular) masculine noun with a consonant onset, but *nouvel(le)* [nuvɛl] when preceding a masculine noun with a vowel onset or a feminine noun. Such dependencies put constraints on the scope of phonological planning: Speakers must minimally have retrieved the onset of the noun before they can start to articulate the end of the adjective in a phrase like *le nouveau camion* ("the new truck") or *le nouvel avion* ("the new airplane"). Our aims were to find out whether planning scope varies as a function of utterance-specific constraints (variable vs. invariant adjectives; masculine vs. feminine noun gender), whether it varies between languages (French, in which such phonological dependencies exist, vs. German, in which they do not), or whether planning scope is similar across utterances and languages and set to satisfy constraints when present.

Following previous research, we measured both speech onset times and gaze duration, and took the difference between those two measures, the *eye-speech lag* as our dependent variable indexing the scope of advance planning, together with speech onsets. To clarify, if participants start to speak at an early point relative to the time spent looking at the picture, the scope of planning is small. On the other hand, if speech onset is late relative to gaze duration, the scope of planning is large. We gauged the scope of planning by means of an auditorily presented distracter word that was either phonologically related or unrelated to the onset of the noun: Facilitation from a related distractor relative to an unrelated one should only occur if the noun falls within the scope of planning.

First, we compared the production of utterances with adjectives that display phonological alternations to utterances with adjectives that do not vary. French speakers used Det + Adj+ N phrases to describe pictures of objects modified to elicit a variable or an invariant adjective. Speech onsets were longer for utterances with variable adjectives than for those with invariant adjectives. However, there was a similar but stronger effect of adjective variability on gaze durations, which meant that variable adjectives were associated with shorter eye-speech lags. We found moderate evidence that the effect of the distracter on speech onsets was stronger among masculine nouns, compared to feminine nouns.

Next, we investigated whether speakers of a language in which phonological dependencies exist in Det + Adj + N phrases use a larger planning scope than speakers of a language without such constraints. To this aim, we compared the French data to data from native German speakers tested in German, which does not have adjectives that depend on phonological context. We only included French trials with feminine nouns and invariant adjectives, and similar German words. We observed some (weak) evidence that French speakers used a larger scope of planning than German speakers.

In summary, we found some evidence that French speakers employ a larger scope of phonological planning when the head of the Det + Adj + N phrase is masculine rather than feminine. They also take longer before starting to articulate and look at the object longer when the utterance contains an adjective the form of which potentially depends on the phonological context. This pattern, along with the exploratory finding that adjectives are lengthened in contexts where they are potentially variable, suggests that part of phonological encoding may be done after the onset of articulation.

Arbeitsgruppe 13

Model and evidence in quantitative comparative linguistics
Gerhard Jäger & Johann-Mattis List

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Data in quantitative comparative linguistics

Johann-Mattis List
Department of Linguistic and Cultural Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig

With increasing amounts of digital cross-linguistic data being produced and shared, it becomes more and more important to make sure that data are comparable across sources. In order to achieve this goal, not only standard formats for data exchange are needed, but also meta-data-sets that may serve as a reference to indicate what kind of data a given datasets provides. With the establishment of the cross-linguistic data formats (CLDF) initiative (https://cldf.clld.org, Forkel et al. 2018), a first attempt towards the integration of cross-linguistic datasets across sources has been made. CLDF provides testable standards for data exchange and recommends to link individual datasets to reference catalogs which are maintained by the team of scholars who actively contribute to CLDF. These catalogs offer scholars the possibility to indicate with which language varieties they work (Glottolog, https://glottolog.org, Hammarström et al. 2020), which concepts they use in their questionnaires in lexical studies (Concepticon, https://concepticon.clld.org, List et al. 2020), or which speech sounds they document in their lexical data or their phoneme inventories (CLTS, https://clts.clld.org, List et al. 2019). As of now, CLDF has proven extremely useful in the aggregation of larger datasets, as witnessed by the Database of Cross-Linguistic Colexifications (CLICS, https://clics.clld.org, Rzymski et al. 2020), which documents colexifications observed for more than 2000 language varieties aggregated from 30 lexical datasets, the maintenance of well-known typological databases such as the World Atlas of Language Structures Online (https://wals.info, Dryer et al. 2013), or for the retro-standardization of datasets, as exemplified by the recent publication of the Tableaux Phonétiques des Patois Suisses Romands Online (https://tppsr.clld.org, Geisler et al. 2020). In the talk, I will present the most recent developments of the Cross-Linguistic Data Formats Initiative and point to future chances and challenges.

References: Dryer, Matthew S. and Haspelmath, Martin (eds.) 2013. The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info, Accessed on 2020-12-11.) Forkel, R., J.-M. List, S. Greenhill, C. Rzymski, S. Bank, M. Cysouw, H. Hammarström, M. Haspelmath, G. Kaiping, and R. Gray (2018): Cross-Linguistic Data Formats, advancing data sharing and re-use in comparative linguistics. *Scientific Data* 5.180205. 1-10. Geisler, Hans, Robert Forkel, and Johann-Mattis List. 2020. The Tableaux Phonétiques des Patois Suisses Romains Online. Jena: Max Planck Institute for the Science of Human History. (Available online at https://tppsr.clld.org, Accessed on 2020-12-11.) Hammarström, Harald & Forkel, Robert & Haspelmath, Martin & Bank, Sebastian. 2020. Glottolog 4.3. Jena: Max Planck Institute for the Science of Human History. https://doi.org/10.5281/zenodo.4061162 (Available online at http://glottolog.org, Accessed on 2020-12-11.) List, Johann-Mattis, Cormac Anderson, Tiago Tresoldi, Christoph Rzymski, Simon J. Greenhill, and Robert Forkel. 2019. Cross-Linguistic Transcription Systems. Jena: Max Planck Institute for the Science of Human History. (Available online at https://clts.clld.org, Accessed on 2020-12-11.) List, J., C. Rzymski, S. Greenhill, N. Schweikhard, K. Pianykh, A. Tjuka, M. Wu, and R. Forkel (2020): Concepticon. A resource for the linking of concept lists (Version 2.3.0). Version 2.3.0. Jena: Max Planck Institute for the Science of Human History. (Available online at https://concepticon.clld.org/, Accessed on 2020-12- 11.) Rzymski, C., T. Tresoldi, S. Greenhill, M. Wu, N. Schweikhard, M. Koptjevskaja-Tamm, V. Gast, T. Bodt, A. Hantgan, G. Kaiping, S. Chang, Y. Lai, N. Morozova, H. Arjava, N. Hübler, E. Koile, S. Pepper, M. Proos, B. Epps, I. Blanco, C. Hundt, S. Monakhov, K. Pianykh, S. Ramesh, R. Gray, R. Forkel, and J.-M. List. 2020. The Database of Cross-Linguistic Colexifications, reproducible analysis of cross- linguistic polysemies. *Scientific Data* 7.13. 1-12.

Models in quantitative comparative linguistics

Gerhard Jäger
Universität Tübingen, Seminar für Sprachwissenschaft
gerhard.jaeger@uni-tuebingen.de

The emergence of data science has inspired a surge in interest in the application of quantitative and computational methods in comparative linguistics in the broad sense. By this we mean any kind of research studying features of several natural languages in parallel.
High profile results touch upon three major topics:

- the study of deep history, both regarding reconstruction of past language stages and language change processes and of population history in general,
- statistical investigations of typological questions regarding, e.g., the (non-)universality of feature correlations,
- probing for – possibly causal – connections between linguistic properties and extra-linguistic variables such as language community size, climate, or diet.

These results are often met with a healthy skepticism within the linguistic community. It is tempting to discount the criticisms levelled against quantitative comparative linguistics – such as the insistence by practitioners of classical historical linguistics that historical linguistics must be based on the identification of sound laws – as inevitable side effects of a paradigm shift. However, computationally and statistically minded comparativists do not agree among themselves regarding the standards of model validation and model comparison.

In the talk I will review the state of the art of the field regarding model comparison and model validation. Furthermore, I will discuss possibly relevant techniques from neighboring disciplines such as cross-validation and posterior predictive simulations and sketch ways how they can be applied to computational historical linguistics.

Methods and models in historical comparative research on signed languages

Justin M. Power, Danny Law, David Quinto-Pozos
University of Texas at Austin, University of Texas at Austin, University of Texas at Austin
justin.power@utexas.edu, dannylaw@austin.utexas.edu, davidqp@austin.utexas.edu

While there has been a surge in interest over the past two decades in quantitative and computational approaches to questions in historical linguistics, this research program has focused on human languages in only one of the two main modalities, that is, spoken languages in the aural-oral modality. Signed languages in the visual-gestural modality, in contrast, have been largely absent from theoretical discussions concerning the evolution of language and from methodological discussions aimed at developing infrastructures for data sharing and data accessibility in historical comparative research. Perhaps in consequence, despite the ubiquity of computational and quantitative phylogenetic approaches in historical linguistics, there have been relatively few attempts to apply these approaches to the study of sign languages and their histories (Yu et al 2018, Power et al 2020).

In this presentation, we highlight two of the main obstacles to quantitative and computational approaches in historical comparative research on sign languages. These obstacles include the lack of consensus among researchers on a sign language transcription system or on alternative, computer-readable representations of signs. We provide a brief overview of recent approaches to the representation of sign languages in computational research (Hall et al 2017, Yu et al 2018, Power et al 2020, Börstell et al 2020). The lack of consensus in data representation schemes has direct consequences for the accessibility of historical comparative data, for the reproducibility of comparative studies, and for data sharing among historical linguists.

A second main obstacle to quantitative approaches in sign language historical research relates to models of diachronic change in signed languages. While most quantitative, computational approaches to historical research on spoken languages depend on the prior application of the comparative method for identifying cognate vocabulary (Gray & Atkinson 2003, Gray et al 2009, Sagart et al 2019), sign scholars have yet to successfully apply the comparative method to identify recurring correspondences across putatively cognate signs (Power et al 2019, Power 2020); and it remains unclear whether the regularity principle—a foundational assumption of the comparative method (Rankin 2008, Hale 2015)—holds for signed languages. How should sign scholars approach the comparison and validation of models of sign change while lacking a gold standard arrived at by independent methods?

In the final part of this presentation, we introduce the Sign Change project, a new research initiative aimed at exploring the theoretical and methodological foundations of sign language historical linguistics, with a focus on thirteen sign languages in two putative families, the French and BANZSL families. The Sign Change project has three main aims. First, the project studies the question of whether change in signed languages can be regular by applying the traditional comparative method in historical linguistics to basic vocabulary signs transcribed using HamNoSys (Hanke 2004), a computer-readable transcription system. Second, the project seeks to increase data sharing and accessibility by making transcribed data freely available to other researchers. Third, the project builds on initial attempts to develop a quantitative model of sign change for estimating evolutionary distances across putatively cognate signs (Power 2020). This presentation reports on these aspects of the project's work to address the two main obstacles outlined above.

References: Börstell, C., Crasborn, O., & Whynot, L. (2020). Measuring lexical similarity across sign languages in Global Signbank. In *Proceedings of the 9th Workshop on the Representation and Processing of Sign Languages* (LREC 2020), 21-26. Gray, R. D., & Atkinson, Q. D. (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature*, 426(6965), 435–439. Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science*, 323(5913), 479–483. Hale, M. (2015). The comparative method: Theoretical

issues. In *Routledge handbook of historical linguistics*, Bowern, C., & Evans, B. (Eds.), 146–160. Routledge. Hall, K. C., Mackie, S., Fry, M., & Tkachman, O. (2017). SLPAnnotator: Tools for implementing sign language phonetic annotation. In *INTERSPEECH,* 2083-2087. Hanke, T. (2004). HamNoSys - Representing sign language data in language resources and language processing contexts. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation* (LREC 2004), 1–6. Power, J. M., Quinto-Pozos, D., & Law D. (2019). Can the comparative method be used for signed language historical analyses? Conference presentation, 13th Conference on Theoretical Issues in Sign Language Research. Power, J. M., Grimm, G. W., & List, J. M. (2020). Evolutionary dynamics in the dispersal of sign languages. *Royal Society Open Science*, 7(1), 1–15. Power, J. M. (2020). The origins of Russian-Tajik Sign Language: Investigating the historical sources and transmission of a signed language in Tajikistan [Doctoral dissertation, University of Texas at Austin]. UT Electronic Theses and Dissertations. Rankin, R. L. (2008). The comparative method. *The handbook of historical linguistics*, Joseph, B., & Janda, R. (Eds.), 183–212. Malden, MA: John Wiley & Sons, Incorporated. Sagart, L., Jacques, G., Lai, Y., Ryder, R. J., Thouzeau, V., Greenhill, S. J., & List, J. M. (2019). Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proceedings of the National Academy of Sciences*, 116(21), 10317–10322. Yu, S., Geraci, C., & Abner, N. (2018). Sign languages and the online world of online dictionaries & lexicostatistics. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (LREC 2018), 4235–4240.

Language contact in the evolution of linguistic features

Harald Hammarström

As large databases of linguistic features are emerging, there has been an increased interest in explicit models of their dynamics over time. Typically one (or, several, codependently) typological feature is modelled as a Markov process on a genealogical tree (Dunn et al., 2011, Maslova, 2000, Pagel, 1994). While elegant, these approaches do not take language contact into account, yet language contact is known to ubiquitous and may have important ramifications (Koptjevskaja-Tamm, 2010). Fortunately, some quantitative work in this direction has been done, see e.g., Daumé (2009), McMahon (2010), Muysken et al. (2015) and in particular, the model of Murawaki and Yamauchi (2018) comes close to providing an adequate solution.

In the present work we modify the generative model of Murawaki and Yamauchi(2018) to be better capture known types of language contact (Muysken, 2010). In particular,

(i) we replace the distance-threshold of Murawaki and Yamauchi (2018, 17) with a parameter-free neighbourhood-relation since language in prehistory would have borrowed from concrete neighbours rather than an abstract vicinity, and

(ii) incorporate asymmetry in contact using relative population sizes as a proxy. To maximize applicability, the model will be illustrated on a very dense dataset of basic constituent order with data for over 5 400 languages.

References: Daumé, III, H. (2009). Non-parametric bayesian areal linguistics. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, NAACL '09, pages 593–601, Morristown, NJ, USA. Association for Computational Linguistics. Dunn, M., Greenhill, S. J., Levinson, S. C., and Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473:79–82. Koptjevskaja-Tamm, M. (2010). Linguistic typology and language contact. In Song, **J.** J., editor, *Oxford Handbook of Linguistic Typology*, pages 568–590. Oxford: Oxford University Press. Maslova, E. (2000). A dynamic approach to the verification of distributional universals. *Linguistic Typology*, 4(3):307–333. McMahon, A. (2010). Computational models and language contact. In Hickey, R., editor, *The Handbook of Language Contact*, pages 128–147. Oxford: Wiley-Blackwell. Murawaki, Y. and Yamauchi, K. (2018). A statistical model for the joint inference of vertical stability and horizontal diffusibility of typological features. *Journal of Language Evolution*, 3(1):13–25. Muysken, P. (2010). Scenarios for language contact. In Hickey, R., editor, *The Handbook of Language Contact*, pages 265–281. Oxford: Wiley-Blackwell. Muysken, P., Hammarström, H., Birchall, J., van Gijn, R., Krasnoukhova, O., and Müller, N. (2015). Linguistic areas, bottom up or top down? the case of the guaporé-mamoré region. In Comrie, B. and Golluscio, L., editors, *Language Contact and Documentation*, pages 205–238. Berlin: DeGruyter Mouton. Pagel, M. (1994). Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 255(1342):37–45.

Partial cognate comparison and pre-settlement history of the Dogon ethnolinguistic group

Abbie Hantgan-Sonko
CNRS-LLACAN
ahantgan@gmail.com

The Dogon languages and peoples who inhabit the rocky Bandiagara Escarpment of central-eastern Mali have long been the subject of research across scientific disciplines, but interdisciplinary investigations have only just begun to emerge. Archeological evidence attests to Dogon inhabitation of the cliff range for the last 700 years (Mayor & Huysecom 2016), yet preliminary linguistic data suggests a time-depth of at least 2000 years (Hantgan 2019; Moran & Prokić 2013; Heath 2015; Prokhorov, Heath & Moran 2012). Furthermore, if the classification of the Dogon language group constituting its own branch of Niger-Congo (Hammarström et al. 2020) is correct, its divergence from the higher order must have occurred at some time in the distance past. The geographic seclusion of the Dogon people with respect to other ethnolinguistic groups of the area has been attributed as the cause for their linguistic, and genetic (Babiker et al.), cohesion, to the exclusion of neighboring languages and peoples (Nunn & Puga 2012).

Dogon represents a continuum of at least 21 separate languages and upwards of 60 dialects (Moran & Prokić 2013). However, Dogon languages may not be as disparate as has been claimed. Despite Dogon languages such as Bondo So having incontestable evidence of a noun class system with at least six discernable classes (three of which are shown in Table 1), Niger-Congo specialists have remarked on the group's lack of noun classes as evidence for its outlier status (Bendor-Samuel, Olsen & White 1989; Dimmendaal 2008; Dimmendaal 2011).

BIRD/S	níí / níí=mbò	WHITE BIRD/S	níí pílɛ̀ / níí pílɛ̀=mbò
BEARD/S	bɛ̀ɛ́-ŋ / bɛ̀ɛ́=jè	WHITE BEARD/S	bɛ̀ɛ́ pílɛ̀-ŋ / bɛ̀ɛ́ pílɛ̀=jè
NOSE/S	kìndʒ-áà / kìndʒ-ɛ́ɛ̀	WHITE NOSE/S	kìndʒ-áà píl-áà / kìndʒ-ɛ́ɛ̀ píl-ɛ́ɛ̀

Table 1: Noun class suffixes in Bondu So

Heath (2015) splits the Dogon languages along a geographically and genealogical east-west line. He describes eastern Dogon languages, such Ben Tey and Bankan Tey, as simply having a binary animate-inanimate distinction in singular nouns (plural is unmarked), whereas western Dogon languages Bunoge and Dogul Dom have been noted for their noun class vestiges. Compare example lexemes in Table 2.

	Bankan Tey	Ben Tey	Bunoge	Dogul Dom	Yanda Dom
BIRD	nìí-m	nììɲ̃í-m	níí-bè	lègèè-gú	àɲà-n
BEARD	bɛ̃́-ɲ̃	bɛ̃́-ɲ̃	ʔɔ̀ndɔ́ kùlɛ̀	bɛ́ɛ̀	bìjà kùlà
NOSE	cíĩ̀̃	cíĩ̀̃	kìnà	kìndʒɔ́ɔ́	tʃìnzà

Table 2: Comparative Dogon singular noun stems

Yanda Dom is an exception in that its speakers currently reside on the eastern half of the Escarpment whereas the language fits in with those of the western side. Heath (2017: 94) describes nouns in the language as having an unmarked singular, yet note that many of the language's inanimate nouns and body parts (BIRD, EAR, FLY (INSECT), TOOTH), end in [n]. The lexeme for BIRD is an example in which roots and suffixes across the languages may have separate, yet cognate, forms.

Whereas cognates for the lexeme BEARD end in a nasalized [j], those for BIRD are suffixed with [-m]. At least for Ben Tey, Heath (2005: 80) considers the final [j̃] to be a frozen diminutive, derived from the noun for CHILD in other Dogon languages, and thus does not parse it as a suffix. Rather, based on lexical comparative data (Heath et al. 2015; Hantgan & List 2018), it seems that this suffix appears primarily with body parts (see also KIDNEY), but also adjectives (see RED), and other inanimates (TREE, ASH). Furthermore, a non- nasalized suffix [j] is found with many other inanimates including some numbers (TWO, SIX, EIGHT). One hypothesis is that the nasalization is residual evidence of prior noun class marking; a proposed evolution for the lexeme ASH is

given in Table 3.

Ampari	Bondu So	Bankan Tey	Jamsay	Yanda Dom	Bunoge	Dogul Dom
dɔ̀ɔ̀jé-gè	dɔ̀ dɛ̀ɛ́-ŋ	dɔ̀jɔ̌-j̃	dɔ̀jɔ́	dɔ̀jɛ̀	díwɔ̀	dɔ́ɔ̀

Table 3: Comparative Dogon singular noun stems for ASH

Thus, in order to accurately estimate the time depth of the Dogon languages, it is crucial not only to examine roots, but also suffixes. According to Wu et al. (2020), the sole computational method of comparing partial cognates is the algorithm proposed by List, Lopez & Bapteste (2016). This study proposes to use *computer-assisted methods* as provided by List et al. (https://digling.org/calc/) within the *computational historical linguistics* framework as defined by Jäger (2019: 155) using data formatted for this purpose https://digling.org/links/bangime.html.

References: Babiker, Hiba, Jeffrey Heath, Floyd Reed, Stephan Schiffels and Russell D. Gray. "Striking Genetic Diversity among Populations of West Africa Uncovers the Mystery of a Language Isolate." http://dx.doi.org/10.2139/ssrn.3631471 (5 August, 2020). Bendor-Samuel, John, Elizabeth Olsen and Anne White. 1989. "Dogon." In John Bendor- Samuel, eds., *The Niger-Congo languages*. Lanham MD/New York/London: University Press of America. 169–177. Dimmendaal, Gerrit J. 2008. "Language Ecology and Linguistic Diversity on the African Continent." *Language and Linguistics Compass* 2(5). 840–858. https://doi.org/0.1111/j.1749-818X.2008.00085.x. Dimmendaal, Gerrit J. 2011. *Historical Linguistics and the Comparative Study of African Languages*. Amsterdam and Philadelphia: John Benjamins Publishing Company. Güldemann, Tom. 2018. "Historical linguistics and genealogical language classification in Africa." In Tom Güldemann, ed., *The Languages and Linguistics of Africa*. Berlin, Boston: De Gruyter. 58–444. Hammarström, Harald, Robert Forkel, Martin Haspelmath and Sebastian Bank. 2020. *Glottolog*. Jena: Max Planck Institute for the Science of Human History. http://glottolog.org (20 August, 2020). Hantgan, Abbie. 2019. "Linguistic Support for an Early Dogon Diffusion." Poster presented at the Peopling History of Africa conference, Geneva, Switzerland. Hantgan, Abbie and Johann-Mattis List. 2018. *Dogon and Friends*. http://tsv.lingpy.org/?remote_dbase=dogon.sqlite3&file=dogon&css=menu:show|textfields:hide|database:hide|&preview=10&basics=DOCULECT|CONCEPT|IPA|COGID|ALIGNMENT&pinyin=CHINESE&sampa=IPA|TOKENS&highlight=TOKENS|ALIGNME NT. Heath, Jeffrey. 2015. "Dogon languages of eastern Mali." Paper presented at the World Congress of African Languages, Kyoto. Heath, Jeffrey, Laura McPherson, Kirill Prokhorov and Steven Moran. 2015. *Dogon Comparative Wordlist*. http://cdstar.shh.mpg.de/bitstreams/EAEA0-C97A-A1D2-2E76- 0/a.xls. (22 August, 2020). Jäger, Gerhard. 2019. "Computational historical linguistics." *Theoretical Linguistics*. Berlin, Boston: De Gruyter Mouton 45(3–4). 151–182. https://doi.org/10.1515/tl-2019-0011. List, Johann-Mattis, Philippe Lopez and Eric Bapteste. 2016. "Using sequence similarity networks to identify partial cognates in multilingual wordlists." In *Association of Computational Linguistics*, Berlin: Association of Computational Linguistics. 599–605. http://anthology.aclweb.org/P16-2097. Mayor, Anne and Eric Huysecom. 2016. ""Toloy", "Tellem", "Dogon" : une réevaluation de l'histoire du peuplement en Pays dogon (Mali)." In Marina Lafay, Françoise Le Guennec-Coppens, and Elisée Coulibaly, eds. *Regards scientifiques sur l'Afrique depuis les indépendances*. Paris: Karthala. 333–350. Moran, Steve and Jelena Prokić. 2013. "Investigating the Relatedness of the Endangered Dogon Languages." *Literary and Linguistic Computing* 28(4). 676–691. Nunn, Nathan and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *The Review of Economics and Statistics* 94(1). 20–36. Prokhorov, Kirill, Jeffrey Heath and Steven Moran. 2012. *Dogon classification*. Paper presented at the Niger-Congo Congress. INALCO/LLACAN, Paris. Wu, Mei-Shin, Nathanael E. Schweikhard, Timotheus A. Bodt, Nathan Hill W. and Johann- Mattis List. 2020. "Computer-Assisted Language Comparison: State of the Art." *Journal of Open Humanities Data* 6(1). 2. https://doi.org/10.5334/johd.12.

Why we need more studies of methods, not data, in computational historical linguistics

Philipp Rönchen, Tilo Wiklund

Department of Linguistics and Philology, Uppsala University, Uppsala, Sweden, Department of Mathematics, Uppsala University, Uppsala, Sweden

philipp.ronchen@lingfil.uu.se, tilo@wiklund.pm

After decades of research in computational historical linguistics, there is still no consensus on how phylogenetic methods should be employed and how much trust into particular models is warranted. The dating of Indo-European by computational methods, though long- attempted, has still not been conclusively resolved (Gray and Atkinson 2003, Bouckaert et al. 2012, Chang et al. 2015). There is evidence emerging that many phylogenetic analyses are sensitive to the use of ancestry constraints (Chang et al. 2015), methods of age calibration (Maurits et al. 2020), the inclusion of poorly attested languages (Chang et al. 2015, Rama 2018), the prior distribution of tree topologies used (Rama 2018, see also Ritchie and Ho 2019) as well as the treatment of loanwords (Kelly and Nicholls 2017).

There have been made several suggestions on how inferences in computational historical linguistics can be made more robust: One strain of thought is that it is the subjectivity of the input data, that is the variability of cognate coding and age calibrations, that leads to a lack of robustness. To counter this problem, Jäger (2019a) and Rama and Wichmann (2020) suggest to automatise more steps in the computational workflow. Another idea is that model fit has to be more seriously evaluated (Jäger 2019a, 2019b). The problem with both of these approaches is that they rely heavily on an extended use of data: The automatisation of tasks previously done by humans needs data as input as well as for evaluation. The tests of model performance as suggested by Jäger also most often measure the performance of a model on a given data set.

That a model agrees with a given data set is of course a necessary condition for it to be useful. However, there is a danger of *over-exploiting* a data set, that is to use one and the same data set to generate hypotheses, fit model parameters and evaluate model performance. Compared to other fields, the amount of data accessible in historical linguistics is small. In theory, any sufficiently high-dimensional model can be tweaked to fit any data set desired, a phenomenon known as overfitting. Any use of a data set to make modelling decisions increases the number of "meta"- or "hyper"-parameters of the analysis and therefore the risk of overfitting.

While there is no easy way around the problem that the amount of data limits the complexity of questions one can ask of the data set, we believe that progress can be made at least in the following areas:

- The better-understood the real-world processes are that a model tries to emulate, the higher is its explanatory power. Most phylogenetic models used in linguistics have been designed in biology for the purpose of DNA evolution, and are not well-suited for linguistic data. But not much effort has been spent on developing models which more closely model linguistic evolutionary processes.

- When the behaviour of a model is only partly understood, simulation studies often go a long way to show whether it behaves as expected. These can often be carried out without any utilisation of the data set.

- Phylogenetic models are often more complex and high-dimensional than they need to be with regard to the questions they try to answer. Simpler models are less sensitive to the types of overfitting outlined above.

We present some examples to illustrate how these ideas may be implemented.

References: Bouckaert, Remco, Philippe Lemey, Michael Dunn, Simon J. Greenhill, Alexander V. Alekseyenko, Alexei J. Drummond, Russell D. Gray, Marc A. Suchard, and Quentin D. Atkinson. 2012. "Mapping the origins and expansion of the Indo-European language family." Science 337: 957–960. Chang, Will, Chundra Cathcart, David Hall, and Andrew Garrett. 2015. "Ancestry-constrained phylogenetic analysis

supports the Indo-European steppe hypothesis." *Language* 91: 194-244. Gray, Russell D. and Quentin D. Atkinson. 2003. "Language-tree divergence times support the Anatolian theory of Indo-European origin." *Nature* 426: 435-439. Jäger, Gerhard. 2019a. "Computational historical linguistics." *Theoretical Linguistics* 45: 151-182. Jäger, Gerhard. 2019b. "Model evaluation in computational historical linguistics." *Theoretical Linguistics* 45: 299-307. Kelly, Luke J. and Geoff K. Nicholls. 2017. "Lateral transfer in stochastic Dollo models." *The Annals of Applied Statistics* 11: 1146-1168. Maurits, Luke, Mervi de Heer, Terhi Honkola, Michael Dunn, and Outi Vesakoski. 2020. "Best practices in justifying calibrations for dating language families." *Journal of Language Evolution* 5: 17-38. Rama, Taraka. 2018. "Three tree priors and five datasets: A study of Indo-European phylogenetics." *Language Dynamics and Change* 8: 182-218. Rama, Taraka and Søren Wichmann. 2020. "A test of Generalized Bayesian dating: A new linguistic dating method." *PLoS ONE* 15(8): e0236522. Ritchie, Andrew M. and Simon Y. W. Ho. 2019. "Influence of the tree prior and sampling scale on Bayesian phylogenetic estimates of the origin times of language families." *Journal of Language Evolution* 4: 108-123.

Theoretical (in)compatibilities of the comparative method and cladistics

Erik Elgh
Department of Linguistics and Philology, Uppsala University
erik.elgh@lingfil.uu.se

Since the break of the millennium cladistic analyses have been frequently employed in historical linguistics (e.g. Bouckaert et al., 2012, Gray & Atkinson, 2003, Gray & Jordan, 2000). These often use cognate coded word lists as their data for inferring phylogenetic relationships (see Nicholls & Gray, 2008, Greenhill & Gray, 2009, Dunn, 2014). The cognate assessments and thus homology hypotheses are based on the comparative method of historical linguistics (for the comparative method see Campbell, 2013, 107-158). This method assumes certain sound changes that occurred between the time of a reconstructed proto-language and its descendants. Both the methods of reconstructing a proto-language and the utilization of this language for character coding leads, in effect, to certain assumptions of tree topology already before the phylogenetic analysis. On the other hand, character coding of biological data, whether morphological characters or DNA, are not based on a reconstructed proto-organism (see e.g. Baum & Smith, 2013, 195-202 and 399-403 for character coding in biology). Thus, the theoretical framework of compiling the data matrices differs between the fields, even though the analyses they are used in are similar.

The presentation aims to detail the underlying assumptions of both the comparative method and biological cladistics when used to investigate linguistic relationships. Further, these assumptions will be compared to determine whether these two methods are compatible and thus whether they are appropriate to use in conjunction. If not, another method is needed. Hammarström (pers. comm.) suggests integrating the processes of cognate detection and reconstruction of the tree topology into the same algorithm. Efforts of this nature have been taken in biology with the software StatAlign which integrates Bayesian analysis of sequence alignment and phylogeny (Novák et al., 2008). These and other alternative methods will be discussed.

References: Baum, D. A., Smith, S. D. (2013). Tree Thinking. An Introduction to Phylogenetic Biology. Roberts and Company Publishers. Greenwood Village, Colorado. Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S. J., Alekseyenko, A. V., Drummond, A. J., Gray, R. D., Suchard, M. A., Atkinson, Q. D. (2012). Mapping the Origins and Expansion of the Indo-European Language Family. *Science*, 337(6097):957-960. Campbell, L. (2013). Historical Linguistics. An Introduction. (3rd edition). Edinburgh University Press. Edinburgh. Dunn, M. (2014). Language Phylogenies. *In: Bowern, C., Evans, B. (eds.). The Routledge Handbook of Historical Linguistics. 190-211. Routledge. New York.* Gray, R. D., Atkinson, Q. D., (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature*, 426(6965):435-439. Gray, R. D., Jordan F. M. (2000). Language trees support the express-train sequence of Austronesian expansion. *Nature*, 405:1052-1055. Greenhill, S. J., Gray, R. D. (2009). Austronesian Language Phylogenies: Myths and Misconceptions about Bayesian Computational Methods. *In: Adelaar, A., Pawley, A. (eds.). Austronesian Historical Linguistics and Culture History: a Festschrift for Robert Blust. 375-397. Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University. Canberra.* Nicholls, G. K., Gray, R. D. (2008). Dated Ancestral Trees from Binary Trait Data and their Application to the Diversification of Languages. *Journal of the Royal Statistical Society. Series B*, 70(3):545-566. Novák, A., Miklós, I., Lyngsø, R., Hein, J. (2008). StatAlign: An Extendable Software Package for Joint Bayesian Estimation of Alignments and Evolutionary Trees. *Bioinformatics*, 24(20):2403-2404.

Lexedata: Tying existing software to CLDF Wordlists

Gereon Kaiping, Natalia Chousou-Polydouri
Geographic Information Science Center, Universität Zürich, CH and Center for the Interdisciplinary Study of Language Evolution, University of Zürich, CH, Institut für vergleichende Sprachwissenschaft, Universität Zürich, CH and Center for the Interdisciplinary Study of Language Evolution, University of Zürich, CH
gereon.kaiping@geo.uzh.ch, natalia.chousou-polydouri@uzh.ch

The CLDF standard (Forkel 2017) defines an exchange format for comparative linguistic data in terms of comma-separated data files tied together by a metadata file that uses standardized terms from a dedicated ontology (Forkel et al. 2017). This makes it easy to exchange data and use it in the framework of cross-linguistic linked databases (Forkel & Bank 2016). It is, however, not the natural format in which linguists tend to collect, edit, and process data, in particular word list data.

Most comparative wordlists are extracted manually from documents of language documentation, such as grammars, dictionaries, or the field linguist's notes in physical or digital form (eg. PDFs, Shoebox/Toolbox/Flex corpora). Often, the first point of collection is in spreadsheet software (see eg. Kaiping & Klamer 2018) such as Microsoft Excel or Google Sheets, in a plethora of formats, which tend to be variants of either long wordlists in different sheets with columns corresponding to different properties of forms (and thus somewhat similar to the FormsTable of CLDF), or even matrices with multiple forms per cell, indexed by language and concept in row and column headers.

Other data formats frequently used in the context of working with word lists are the tab-separated value format used by LingPy (List et al. 2018) and Edictor (List 2017), tabular comparative cognate data used for quickly comparing and visualizing cognate sets, the SQL databases used as a backend for CLLD web applications, and coding sequences for use in phylogenetic analyses.

Using CLDF as the core of a dataset – but making the data easily available in these formats for the associated programs – will therefore be useful. The use extends to both the people working with the data, who can keep to their familiar tools, and to potential later consumers of the data, who will find it in a standardized and reusable format.

To achieve this, we present Lexedata, a collection of open-source software tools written in Python for translating lexical data between these different formats, and for enriching CLDF datasets.

We show the benefits of this software suite using as example two lexical datasets, of Arawak and Maweti-Guaraní languages respectively. The data sets were available in different Excel formats following different stages and philosophies of collection. Using tools from our Lexedata suite, we converted both datasets into CLDF, added unified sound segmentation checked with CLTS (List et al. 2019), used (on the Arawak dataset, which did not have cognate sets yet) the automatic cognate coder LexStat (List 2012) implemented in LingPy, made the results available for collaborative post-processing in a version-controlled, multi-user fashion (using Excel, Edictor and Google Sheets). A collaborative team of linguists with no computational background is using the suite in there workflow to improve and study the Arawak dataset, thus testing the usability and robustness of the software. The resulting cognate codes are, again using Lexedata, converted to NEXUS for a preliminary phylogenetic analysis in MrBayes, using two different coding schemes (Chousou-Polydouri et al. 2016).

The collaboration with linguists help us prioritize removing the biggest hurdles where currently, data carpentry still requires programming skills. Using loops of export and import, Lexedata can be used to generate changes to the CLDF dataset, and can be integrated in a version-controlled workflow so that the current version of the CLDF is always reflected in a derivative format, and changes made to that derivative file are automatically fed back into CLDF on commit-time.

As such, Lexedata provides a useful toolbox to vastly decrease the threshold of editing

lexical data in a data management context that upholds the FAIR principles (Wilkinson et al. 2016).

References: Chousou-Polydouri, Natalia & Birchall, Joshua & Meira, Sérgio & O'Hagan, Zachary & Michael, Lev. 2016. A Test of Coding Procedures for Lexical Data with Tupí-Guaraní and Chapacuran Languages. *Proceedings of the Leiden Workshop on Capturing Phylogenetic Algorithms for Linguistics*. Tübingen: Philosophische Fakultät. (https://bibliographie.uni-tuebingen.de/xmlui/bitstream/handle/10900/68644/Chousou-Polydouri_Birchall_Meira_O'Hagan_Michael.pdf?sequence=1&isAllowed=y) (Accessed August 9, 2020.) Forkel, Robert. 2017. CLDF: Cross-linguistic Data Formats. *Zenodo*. (doi:https://doi.org/10.5281/zenodo.835502) (https://zenodo.org/record/835502) Forkel, Robert & Bank, Sebastian. 2016. clld - a toolkit for cross-linguistic databases. *Zenodo*. (doi:10.5281/zenodo.55099) (http://clld.org/) Forkel, Robert & List, Johann-Mattis & Cysouw, Michael & Rzymski, Christoph & Greenhill, Simon J. 2017. The CLDF Ontology. (Resource Description Framework.) (https://cldf.clld.org/v1.0/terms.rdf) (Accessed August 27, 2020.) Kaiping, Gereon A. & Klamer, Marian. 2018. LexiRumah: An online lexical database of the Lesser Sunda Islands. *PLOS ONE* 13(10). e0205250. (doi:10.1371/journal.pone.0205250) List, Johann-Mattis. 2012. LexStat: Automatic detection of cognates in multilingual wordlists. *Proceedings of the EACL 2012 Joint Workshop of LINGVIS & UNCLH* (EACL 2012), 117–125. Stroudsburg, PA, USA: Association for Computational Linguistics. (http://dl.acm.org/citation.cfm?id=2388655.2388671) (Accessed August 12, 2015.) List, Johann-Mattis. 2017. A Web-Based Interactive Tool for Creating, Inspecting, Editing, and Publishing Etymological Datasets. *Proceedings of the Software Demonstrations of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, 9–12. Valencia, Spain: Association for Computational Linguistics. (http://aclweb.org/anthology/E17-3003) (Accessed April 25, 2018.) List, Johann-Mattis & Anderson, Cormac & Tresoldi, Tiago & Rzymski, Christoph & Greenhill, Simon J & Forkel, Robert. 2019. cldf/clts: Cross-Linguistic Transcription Systems. Zenodo. (doi:10.5281/zenodo.2633838) (https://zenodo.org/record/2633838#.Xzf0YM-xW00) (Accessed August 15, 2020.) List, Johann-Mattis & Moran, Steven & Forkel, Robert & Greenhill, Simon J & Tresoldi, Tiago & Frank Nagel & Christoph Rzymski et al. 2018. *LingPy. A Python Library for Quantitative Tasks in Historical Linguistics.* Zenodo. (doi:10.5281/zenodo.1438429) (https://zenodo.org/record/1438429#.W-xUQBBRfc8) (Accessed November 14, 2018.) Wilkinson, Mark D. & Dumontier, Michel & Aalbersberg, IJsbrand Jan & Appleton, Gabrielle & Axton, Myles & Baak, Arie & Blomberg, Niklas et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*. Nature Publishing Group 3(1). 160018. (doi:10.1038/sdata.2016.18)

Controlling for geographical, areal, and family biases in typology

Matías Guzmán Naranjo, Laura Becker
Université de Paris, Universität Freiburg
mguzmann89@gmail.com, laura.becker@linguistik.uni-freiburg.de

Introduction Controlling for family biases has mostly focused on ways of choosing a sample of languages in a way that most families are equally represented. Different methods have been proposed in e.g. Bickel (2008), Dahl (2008), Dryer (1989, 1991, 1992, 2011), Greenberg (1966), Jaeger et al. (2011), Maslova (2008), Perkins (1989), and Rijkhoff and Bakker (1998). Controlling for areal effects often relies on similar techniques: Choosing a sample of languages which are assumed to have as little contact with each other as possible. One important point noted in Rijkhoff, Bakker, et al. (1993, pp. 174–175) and Dryer (2018) is that distance between languages cannot be accounted for in absolute terms, but that distances depend on the specific geographical and ecological properties of a given area. Two languages spoken 100km apart in a region of low linguistic density may be in contact, whereas language contact across 100km would be very unlikely in regions of high linguistic density. Finally, most studies try to balance the number of languages selected from each macro area in some way (cf. Jaeger et al. 2011). These methods all face the same issue: The researcher can only include a portion of her data. We present an alternative approach that, can control for the types of biases mentioned without the need to exclude data.

Materials and methods For illustration, we focus on the relation between VO and OV word orders and the degree of prefixation vs. suffixation in a language. It has been argued that while VO orders can occur with both prefixes and suffixes, OV orders show a strong preference against prefixation (e.g. Bybee, Pagliuca, and Perkins 1990; Hawkins and Gilligan 1988; Siewierska and Bakker 1996; Song 2012). We use the datasets from WALS chapters 26 and 83 (Dryer 2013a,b). We fitted a Bayesian ordinal models with the degree of suffixation as the dependent variable (7 levels: strongly suffixing to strongly prefixing) with verb-object order as the predictor (3 levels: OV, no dominant order, and VO). All models were fitted with Stan (Carpenter et al. 2017) using the brms package (Bürkner 2018) in R. To control for geographic biases we included latitude and longitude information directly into our model. We do this by using a two-dimensional gaussian process for each macro-area. To control for areal biases we found the neighbors of each language, neighbors being defined in the following way: Given a language L1, we define a neighbor as any language which is at most 1.5 times as far from L1 as the nearest neighbor of L1 (the factor of 1.5 was chosen for providing the best results). Thus, if the distance between L1 and its nearest neighbor is 10 km, all languages within a 15 km radius will be considered neighbors of L1. However, if its nearest neighbor is 100 km away, we treat all languages within a 150 km radius as neighbors of L1. We then calculated the mean degree of suffixation across the neighbors and added it as a one-dimensional gaussian process to the model. Finally, to control for family biases, we included a phylogenetic term (Housworth, Martins, and Lynch 2004) in our regression model. Unlike simple group-level effects, phylogenetic regression can take into account a complete phylogenetic tree. Therefore, our model does not only control for language family or genus, but for all (known) relations between the languages in our sample. The phylogenetic term ensures that, e.g. Spanish, French, and Farsi are modelled as related, with a much closer genetic relation between Spanish and French than between those two languages and Farsi.

Results We compared a model including the three controls described above to a model using no controls and a model using simple group-level effects for family and macro area. We carried out the comparison using LOO cross-validation (Vehtari, Gelman, and Gabry 2016) on the full dataset and two oversampled datasets (I: duplicating 100 Indo- European languages; II: duplicating all South American languages). In all sampling experiments, our model (accuracy: 0.59) performed much better than the model with group- level effects (accuracy: 0.47) and the model without controls (accuracy: 0.2). With regards to the linguistic question, our model confirmed a very mild effect of verb-object order on the degree of suffixation of the language,

with most of the variance being accounted for by family and areal effects. In contrast, the group-level effect model and the model without controls strongly overestimated the effect of verb-object order on affixation preferences.

Revisiting typological universals with Grambank

Annemarie Verkerk, Haynie Haynie, Russell Gray, Simon Greenhill, Olena Shcherbakova, Hedvig Skirgård

Universität des Saarlandes, University of Colorado, MPI SSH Jena, MPI SSH Jena, MPI SSH Jena, MPI SSH Jena

annemarie.verkerk@gmail.com, Hannah.Haynie@colorado.edu, gray@shh.mpg.de, greenhill@shh.mpg.de, shcherbakova@shh.mpg.de, hedvig_skirgard@eva.mpg.de

Linguistic typology has a rich history of proposing universals; the Konstanz Universals Archive (Plank and Filimonova 2006) contains 2,029 entries (297 of them absolute). Most, if not all, have at some point been demonstrated as false given a certain language. As Bickel (2018) proposes, we should focus less on absoluteness, exceptions and rara, but instead recognize that all universals might be probabilistic. Typology, since Greenberg, has evolved as a discipline; we have found more nuanced and sophisticated ways of testing the power of a certain universal and developed larger cross-linguistic datasets for investigating these associations. One of the most important advances in the testing of rules like universals involves controlling for language history and areality. Here we contribute in two ways to this advance. First, we introduce Grambank, a database of 195 morphosyntactic features currently coded for over 2,000 languages, which enables us to investigate typological patterns within and across families. Secondly, we implement an explicitly evolutionary analysis to investigate the universality of these typological associations.

In recent years, several of Greenberg's word order universals have been claimed to be lineage-specific (Dunn et al. 2011) or alternatively to be divided between lineage-specific patterns and true universals (Jäger 2018a). Other universals have not yet been investigated using quantitative methods that control for language history. In this paper we investigate a large set of universals (100+) with a new large morphosyntactic dataset using phylogenetic comparative methods and a global tree. The typological data are taken from Grambank; the universals are gathered from Greenberg (1963) and Plank and Filimonova's (2006) massive collection. We select universals that dictate the presence or absence of two specific features (such as gender and number; or verby and nouny adjectives) and investigate whether these features are likely to co-evolve or not. Rather than investigating individual families (as done by Dunn et al. 2011 and Jäger 2018a), we make use of global language trees (Jäger 2018b, other global trees are in preparation). This allows us to incorporate small families and isolates, as well as quantitatively appreciate the fact that many separate families that are in a certain area probably share a past, even if we cannot currently confirm with the comparative method that they are indeed of one family.

Results indicate that some features change together while others do not; indeed universality is a matter of degree. Some of Greenberg's well-known universals, such as no. 43, "If a language has gender categories in the noun, it has gender categories in the pronoun", are close to the absolute universal end of the spectrum. Others, however, are not universal (for example, no. 27, "If a language is exclusively suffixing, it is postpositional; if it is exclusively prefixing, it is prepositional"). In some cases, these negative findings are due to a mismatch between the terms used in the original formulation of the universal and the Grambank questionnaire. In other cases, our findings support earlier falsifications of universals, for instance Greenberg's universals regarding Object-Verb and Adjective-Noun order (no. 5 and no. 17), which have been shown to be wrong by Dryer (1988). We also find new evidence against the universality of certain associations; for example, "If a language is tensed, it will have nouny adjectives" (Stassen 1997, Wetzer 1996). These may have held up in their original samples, but are not found to be universal in the current analyses. Viewing universality as a matter of degree has implications for the debate regarding culture and cognition: we put forward a more sophisticated view where both cultural evolution and cognitive factors play a part in investigating correlations between typological features.

References: Bickel, Balthasar. (2018). Estimating diachronic biases of typological traits in unknown phylogenies. Paper presented at the *12th Conference of the Association for Linguistic Typology*, Canberra. Dunn, Michael, Greenhill, Simon J., Levinson, Stephen C. & Gray, Russell D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473, 79-82. Dryer, Matthew S. (1988). "Object-Verb Order and Adjective-Noun Order: Dispelling a Myth." *Lingua*, 74, 185–217. Greenberg, Joseph H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of grammar*. Cambridge: The MIT Press. Jäger, Gerhard. (2018a). A Bayesian test of the lineage-specificity of word order correlations. Paper presented at the *12th International Conference on Language Evolution.* (Evolang XII), Torun. Jäger, Gerhard. (2018b). Global-scale phylogenetics linguistic inference from lexical resources. *Scientific Data*, 5, 180189. doi:doi.org/10.1038/sdata.2018.189 Plank, Frans & Filimonova, Elena. (2006). The Universals Archive: A brief introduction for prospective users. *STUF - Language Typology and Universals*, 53(1), 109-123. Skirgård et al. (forthcoming). *Grambank*. Jena: Max Planck Institute for the Science of Human History. Stassen, Leon. (1997). *Intransitive predication*. Oxford: Oxford University Press. Wetzer, Harrie. (1996). *The typology of adjectival predication*. Berlin: Mouton de Gruyter.

Towards richer multi-source machine-readable etymologies

Johannes Dellert
Seminar für Sprachwissenschaft, University of Tübingen
jdellert@sfs.uni-tuebingen.de

Digital etymological databases have so far appeared in two broad variants. The first type closely mirrors the contents of a single source for the purpose of searchability (e.g. UraloNet, Bakró-Nagy ea., 2013). The other, more machine-readable type integrates knowledge from multiple sources into a formal model, but has so far been limited to extending independently compiled lexicostatistical databases by flat annotations of some core aspects such as cognacy judgments (e.g. LexiRumah, Kaiping & Klamer 2018) and loanword annotations (e.g. WOLD, Haspelmath & Tadmor, 2009). Annotation standards for further aspects such as historical morphology have recently started to be developed (Schweikhard & List 2020), but central issues like reconstructed phonology and semantics remain unstandardized.

As a contribution to the ongoing development, we present our machine-readable text formats for representing complex etymologies. Unlike annotation formats for lexical databases, our primary format is source-driven, i.e. it aims to stay close to the internal organization of major etymological dictionaries, and will by default keep all the information from a single source in one file. A form is minimally identified by a language code and the representation of the form in the source, but can be linked to canonical representations, semantic glosses and complex category annotation. Each line anchors a list of such forms to a pivot form (very often a reconstructed form in a common proto-language). The default semantics for the listed forms (inheritance or cognacy) is determined by an underlying phylogenetic tree which can be specified separately for each source. A range of explicit etymological relation symbols can be freely configured into lists, branching structures (to represent word formation events), bracketed tree structures (e.g. to specify groups of full cognates in a collection of partial cognates), chains (e.g. of borrowings or derivational processes) and disjunctions (e.g. to model multiple equally valid theories). We systematically support underspecification, such as merely stating partial cognacy without making explicit which parts are cognate, or not fully specifying a language, e.g. in the case of a borrowing event which can only be narrowed down to a group of donor languages. Finally, there is support for epistemic modifiers to relation symbols, allowing to model the source author's stance towards an etymology (including rejection). The format has already been used in the NorthEuraLex project to model more than 15,000 etymologies from four language families. In addition to a parser and data model, our Python toolset provides methods for exporting full paths as well as simpler tabular outputs such as cognacy and loanword annotations.

Our next step is the integration of etymologies across sources modeled in this format. The first major challenge is the automatic identification of etyma across sources. Since reconstructions will often differ, we rely on anchoring via matchable forms, either by language-specific symbol equivalence definitions, or – more reliably – via explicit links to a common resource, such as a standardized orthography, or IDs in a lexical database. The second challenge is the treatment of contradictory information from different sources, which we approach by a user-definable preference ranking of sources, and an override procedure with subsequent propagation of edits (such as negated cognacies) in order to maintain a consistent model. Analyzing the first results of this procedure, we find that achieving both clean and high-coverage data will always involve some amount of original research, as many pieces of information necessary for a richer model are not represented explicitly enough, relying on the competent expert reader instead. Reconstructed forms and their semantics will often be heavily underspecified, and morphology is often discussed only very implicitly. Many etymologies (e.g. for smaller Germanic and Romance languages) are not explicitly documented anywhere. Work on high-coverage databases could proceed much more quickly if data modelers could enrich the information given in the sources by obvious morphological analyses and further cognacy judgments, but these decisions would cease to be fully traceable to a published source written

by an expert in the respective language family. Given our goal of maximizing acceptability of the resulting databases to expert communities, it will be necessary to find a good solution for this problem.

References: Bakró-Nagy, Marianne, Nikolett Mus, Beatrix Oszkó, Mária Sipos, Dávid Takács, and Zsuzsa Várnai. 2013. "Uráli etimológiák a világhálón". In: *Obi-ugor és szamojéd kutatások, magyar őstörténet. Hajdú Péter és Schmidt Éva emlékkonferencia 2012, Pécs.* Haspelmath, Martin and Uri Tadmor, eds. 2009. *WOLD*. Max Planck Institute for Evolutionary Anthropology, Leipzig. Available at https://wold.clld.org/. Kaiping, Gereon A. and Marian Klamer. 2018. "LexiRumah: An online lexical database of the Lesser Sunda Islands." *PLOS ONE* 13(10): e0205250. Schweikhard, Nathanael E. and Johann-Mattis List. 2020. "Developing an annotation framework for word formation processes in comparative linguistics." *SKASE Journal of Theoretical Linguistics* 17.1. 2-26.

Correlating borrowing events across concepts to derive a data- driven source of evidence for loanword etymologies

Verena Blaschke, Johannes Dellert
Seminar für Sprachwissenschaft, University of Tübingen, Seminar für Sprachwissenschaft, University of Tübingen
verena.blaschke@student.uni-tuebingen.de, jdellert@sfs.uni-tuebingen.de

Computational methods for approximating various aspects of the reasoning of a historical linguist have great potential as components of a future generation of systems for more rapid machine-aided theory development (List 2019). One of the main challenges for such methods is that some of the heuristics and reasoning patterns commonly used in historical linguistics are difficult to formalize completely. Etymological arguments frequently appeal more to the shared experience of experts than to a fully developed theoretical framework. Computationally emulating this process will require experience in the shape of data with annotations that represent the heuristics and preferences employed within human expert communities.

Our first application of this general paradigm focuses on informal evidence used for establishing loanword etymologies. Classical arguments for assigning a loanword etymology to a word rely on deviations from the sound laws which would have applied if the word had been inherited, or borrowed at a different point in time. For instance, it is clear that the German word *Person* is a borrowing and not strictly cognate with Latin *persona*, because otherwise the initial *p* would have had to undergo a sound shift to *f*. Such a criterion would be rather straightforward to formalize based on a formal description of the expected sound laws. However, this criterion is only helpful if some known sound law would have applied to a part of the phonetic material of the word in question. In many cases, we are not in this comfortable position, and the etymological discussion will be based on more elusive evidence.

In some cases, historical, geographical or archaeological knowledge will help to make the decision, but the most systematically exploitable type of evidence builds on the tendency for loanwords to appear in batches. For instance, if some language has already been established as a donor language for some words, it becomes more likely as a candidate donor for other words as well, even if the evidence from the individual words alone would not warrant such a conclusion. Even more crucially, arguments often rely on the observation that words from the same semantic field tend to get borrowed together. This applies to obvious cases like numbers and month names as well as to less obviously connected sets of concepts such as tools belonging to a certain craft (Tadmor 2009, Carling et al. 2009).

A helpful automated method for inferring possible loanword relations will have to emulate at least some of these types of informal reasoning. As a first step in this direction, we develop data-driven measures of how much evidence establishing one borrowing event provides for assuming others. We also explore in how far such a correlation structure of borrowing events can be extracted from the limited amounts of existing cross-linguistic loanword data.

Given a set of parallel wordlists annotated with loanword status and semantic concept information, we extract how often each concept was loaned and by which pairs of donor and target languages. To quantify the non-independence of borrowing events for each pair of concepts, we average the normalized pointwise mutual information across 1,000 bootstrap samples. In order to additionally retrieve some directional signal that can be interpreted as an approximation to implicational universals of borrowing, the same procedure is applied to the conditional probabilities of concept pairs given one of the concepts.

We execute our methods on WOLD (Haspelmath and Tadmor 2009), and find that even from this limited sample of 41 languages, it is possible to extract quite a few of the expected within-domain correlations (such as the ones between numbers or between kinship terms), which validated our approach. In addition, we also receive some more surprising cross-domain correlations (such as between NARROW and HOLE and between KNEEL and DEFEAT, but also between BEESWAX and KIDNEY) which require further investigation.

References: Carling, Gerd, Sandra Cronhamn, Robert Farren, Elnur Aliyev, and Johan Frid. 2019. "The causality of borrowing: Lexical loans in Eurasian languages." *PloS one* 14(10): e0223588. Haspelmath, Martin and Uri Tadmor, eds. 2009. *World Loanword Database.* Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at https://wold.clld.org/. List, Johann-Mattis. 2019. "Automated methods for the investigation of language contact, with a focus on lexical borrowing." *Language and Linguistics Compass* 13(10): e12355. Tadmor, Uri. 2009. "Loanwords in the world's languages: Findings and results." In Martin Haspelmath, and Uri Tadmor, eds. *Loanwords in the world's languages: A comparative handbook.* Berlin: De Gruyter Mouton. 55-75.

Modelling linguistic data in space using autologistic regression

Miri Mertner
University of Tübingen
miri.mertner@uni-tuebingen.de

The role of the environment in determining linguistic features and patterns of linguistic diversity has been studied extensively and continues to fascinate linguists and anthropologists. It is somewhat accepted that environmental factors have some effect on the global distribution of language diversity, although the reasons for this – and the exact aspects of the environment which lead to the observed patterns – are still being debated (e.g. Bentz et al. 2018; Hua et al. 2019; Nettle 1996). The statistical methodology clearly matters: where Hua et al. (2019) found that climate is the only determinant of linguistic diversity, others suggest that there is a variety of environmental features, such as altitude and distance to water, which may shape patterns of linguistic diversity (Bentz et al. 2018).

It is more controversial to state that any feature of the environment directly influences language structure. Nonetheless, it remains a pervasive idea. Everett (2013) argued for a causal link between the presence of ejectives and high altitude, which he states may have arisen due to the lower air pressure at high altitude and a the need to conserve air. Researchers rightly view such findings with some skepticism, as large datasets can reveal correlations which often turn out to be spurious (Roberts & Winters 2012). Hammarström (2013) suggested that Everett's (2013) correlation between high altitude and ejectives is indeed spurious and pointed out a number of problems with the study, including the way in which languages were classified as belonging to a high or low altitude area. He also argued that even if the correlation turned out to be statistically significant, this could be explained more plausibly by language contact and common ancestry than the physical effects of altitude.

The present study will re-examine the possible correlation between ejectives and altitude using an autologistic regression model (Besag 1972; Wolters 2017). The choice of model is motivated by its ability to deal with spatially autocorrelated data, meaning that neighbouring languages' effect on each other will be controlled for, and an estimate of this effect will be provided. Thus, if languages with ejectives show a high level of spatial clustering, this will show up as a large neighbourhood effect. This could indicate that areal diffusion is a possible explanation for the correlation. However, providing a causal explanation for the correlation (or lack thereof) will not be the main focus of this study, as it has a predominantly methodological focus.

This study will provide ample opportunity for model evaluation and possibly comparison between different model variants. The model will be coded in Julia and the data will come from the PHOIBLE database (Moran & McCloy 2019). If possible, this will be compared to the original dataset used by Everett (2013). This presentation will conclude with a discussion of the potential future applications of autologistic regression as a tool in linguistic typology.

References: Bentz, Christian, Dan Dediu, Annemarie Verkerk, and Gerhard Jäger. 2018. 'The Evolution of Language Families Is Shaped by the Environment beyond Neutral Drift'. *Nature Human Behaviour* 2 (11): 816–21. https://doi.org/10.1038/s41562-018-0457-6. Besag, J. E. 1972. 'Nearest-Neighbour Systems and the Auto-Logistic Model for Binary Data'. *Journal of the Royal Statistical Society: Series B (Methodological)* 34 (1): 75–83. https://doi.org/10.1111/j.2517-6161.1972.tb00889.x. Everett, Caleb. 2013. 'Evidence for Direct Geographic Influences on Linguistic Sounds: The Case of Ejectives'. Edited by Mark Aronoff. *PLoS ONE* 8 (6): e65275. https://doi.org/10.1371/journal.pone.0065275. Hammarström, Harald. 2013. 'C. Everett's Ejectives/Altitude Correlation Is Not Significant'. Billet. *Diversity Linguistics Comment* (blog). 2013. https://dlc.hypotheses.org/491. Hua, Xia, Simon J. Greenhill, Marcel Cardillo, Hilde Schneemann, and Lindell Bromham. 2019. 'The Ecological Drivers of Variation in Global Language Diversity'. *Nature Communications* 10 (1): 2047. https://doi.org/10.1038/s41467-019-09842-2. Moran, Steven, and Daniel McCloy. 2019. 'PHOIBLE 2.0'. 2019. http://phoible.org. Nettle, Daniel. 1996. 'Language Diversity in West Africa: An Ecological Approach'. *Journal of Anthropological Archaeology* 15 (4): 403–38. https://doi.org/10.1006/jaar.1996.0015. Roberts, Seán, and James Winters. 2012. 'Social Structure and Language Structure: The New Nomothetic Approach'.

Psychology of Language and Communication 16 (2): 89–112. https://doi.org/10.2478/v10057-012-0008-6.
Wolters, Mark A. 2017. 'Better Autologistic Regression'. *Frontiers in Applied Mathematics and Statistics* 3 (November): 24. https://doi.org/10.3389/fams.2017.00024.

Loanpy: A framework for computer-aided borrowing detection

Viktor Martinović
Universtiy of Vienna
viktor.martinovic@hotmail.com

Computational historical linguistics has seen significant advances in the recent past. One of those is in the field of cognate detection. Cognates are any two words that share a common history. While cognates are being detected with relatively high accuracy within the core vocabulary, there is, to the best of my knowledge, still no viable framework for detecting them in the periphery. In other words, searching algorithms for the inherited lexicon of genetically related languages are efficient, while those for loanwords are not.

One of the main problems of applying state-of-the-art algorithms to find cognates in the periphery is semantic change. The stable core is less prone to semantic change, thus allowing the use of so-called concept lists, whereas the instable periphery is not only more prone to semantic change, but also cross-linguistically more diverse, making concept lists almost impossible to use.

I will therefore propose a new framework that tackles this problem by using the Cartesian product of two wordlists, instead of matching each item by conceptual identity. Regardless of semantics, the algorithm will first evaluate for each word-pair whether its elements can be phonetically related to each other. Secondly, the semantic similarity of phonetically relatable words will be calculated with the help of word vectors. This way I hope to find potential candidates for yet-undetected loanwords between two genetically unrelated languages. I have used Hungarian and Gothic as a case study.

The effect of priors on tree topologies

Johannes Wahle
Seminar für Sprachwissenschaft Universtität Tübingen
johannes.wahle@uni-tuebingen.de

A typical software for inferring phylogenetic trees offers a wide array of different statistical models. The researcher's task is to find the correct model for the problem at hand. The software used in this kind of research was especially designed for computational biology in order to solve biological research questions using statistical models. (Ronquist & Huelsenbeck, 2003; Suchard, et al., 2018) Linguistic studies using such software tools focus for example on the issue of dating language families (Rama, 2018; Chang, Cathcart, Hall, & Garrett, 2015; Gray & Atkinson, 2003), or rates of lexical change (Greenhill, et al., 2017). Other scholars took up the challenge to test crucial modeling aspects. (Rama, 2018; Yanovich, 2018)

This study follows up on the model checking aspect. To aid in this process ten different combinations of priors are tested on their effect on the inferred topology. They are tested on the five datasets already used by (Rama, List, Wahle, & Jäger, 2018).

The results show that the different priors have almost no influence on the topological accuracy of the inferred trees. Using the visualization toolchain of the *sprspace* tool (Whidden & Matsen, 2015) and the *Cytoscape* software environment (Shannon, et al., 2003), the space of tree topologies can be analyzed. Every model is able to find the topologies with a high similarity to the gold standard tree as can be seen from the generalized quartet distances. This talk shows two things, first the choice of a particular model influences the reconstructed topology only marginally and second how additional tools such as *sprspace* and *Cytoscape* can be used to inspect and analyze the posterior distribution of tree topologies for phylogenetic linguistics.

References: Chang, W., Cathcart, C., Hall, D., & Garrett, A. (2015). Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language, 91*, 194–244. Gray, R. D., & Atkinson, Q. D. (11 2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature, 426*, 435-439. Greenhill, S. J., Wu, C.-H., Hua, X., Dunn, M., Levinson, S. C., & Gray, R. D. (2017). Evolutionary dynamics of language systems. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1700388114. Rama, T. (2018). Three tree priors and five datasets. *Language Dynamics and Change, 8*, 182–218. doi:10.1163/22105832-00802005. Rama, T., List, J.-M., Wahle, J., & Jäger, G. (2018). Are automatic methods for cognate detection good enough for phylogenetic reconstruction in historical linguistics? *Proceedings of the North American Chapter of the Association of Computational Linguistics*, (S. 393-400). Ronquist, F., & Huelsenbeck, J. P. (8 2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics (Oxford, England), 19*(12), 1572–1574. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . Ideker, T. (11 2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research, 13*, 2498–2504. doi:10.1101/gr.1239303. Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (1 2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution, 4*. doi:10.1093/ve/vey016. Whidden, C., & Matsen, F. A. (1 2015). Quantifying MCMC Exploration of Phylogenetic Tree Space. *Systematic Biology, 64*, 472–491. doi:10.1093/sysbio/syv006. Yanovich, I. (6 2018). The effect of dictionary omissions on phylogenies computationally inferred from lexical data. *Language Dynamics and Change, 8*,78–107.doi:10.1163/22105832-008010

Arbeitsgruppe 14

Grammatische Modellierung als Grundlage für sprachdidaktische Vermittlung

Anja Müller, Katharina Turgay

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Grammatiktheoretische Überlegungen zum Schulunterricht

Sandra Döring
Universität Leipzig

Im Grammatikunterricht soll Wissen über das System der deutschen Sprache erworben werden. Dabei geht es zum einen um Kenntnisse über sprachliche Strukturen und Regeln oder Muster, aber zum anderen auch darum, eine Außenperspektive auf einen so alltäglichen Gegenstand wie Sprache einnehmen und sich metasprachlich darüber verständigen zu können.

Ein Ziel des schulischen Grammatikunterrichts darin, Sprachbewusstheit zu fördern, indem sprachliche Strukturen entdeckt bzw. offengelegt werden. Dazu gehört auch explizites Wissen über grammatische Strukturen im Sinne einer Begriffsbildung zu erwerben. Allerdings ist festzustellen, dass gerade der Grammatikunterricht eine gewisse Scheu vor Strukturen und Systematisierungen aufweist. Nach den gescheiterten Versuchen, transformationsgrammatische Forschungsergebnisse (direkt) in die Schulen zu bringen, stellt sich die Frage, warum es auch danach scheinbar kein formales Modell geschafft hat, in den schulischen Grammatikunterricht vorzudringen. Sind also grammatiktheoretische Forschungsergebnisse gar nicht so wichtig für die schulische Sprachreflexion?

Daraus resultiert die Frage, welche Rolle grammatiktheoretische Forschungsergebnisse im Schulunterricht spielen können und sollen. In diesem Zusammenhang ist auch die Rolle formaler Modelle zu diskutieren: sollen diese im schulischen Unterricht Berücksichtigung finden (und wenn ja, in welcher Form), oder sollen sie (nur) als Hintergrundwissen für Lehrkräfte präsent sein, oder sollten sie im Lehramtsstudium gar keine Rolle spielen?

Im Vortrag werden auch Diskussionen des DFG-Netzwerks Grammatik für die Schule (GrafüS, 2016-2020) berücksichtigt.

Sind Haupt- und Nebensatz noch zu retten? Ein Plädoyer für eine widerspruchsfreie Satzanalyse

Matthias Granzow-Emden
Universität Potsdam
mgranzow@uni-potsdam.de

Die Unterscheidung von Haupt- und Nebensatz gehört zu den althergebrachten Inhalten des Grammatikunterrichts – es fehlt in keinem der verbreiteten Lehrwerke und wird auch im neuen „Verzeichnis grundlegender grammatischer Fachausdrücke" genannt. Wenn von Haupt- und Nebensatz die Rede ist, bedeutet dies in Schulbüchern, in Grammatiken und in den Köpfen der Nutzerinnen und Nutzer ganz Unterschiedliches. Der Beitrag zeigt an Beispielen aus Lehrbüchern für den Deutschunterricht, wie in der schulgrammatischen Tradition Form und Funktion wild vermischt werden. Eine Untersuchung aus Seminaren mit Germanistikstudierenden weist darauf hin, wie dies zu unterschiedlichsten Vorstellungen grammatischer Sachverhalte führt. Auch in der sprachwissenschaftlichen Tradition gibt es keine Klarheit – als Beispiel dient die von satzhierarchischen Aspekten bestimmte Darstellung in der Dudengrammatik. All dies führt zu einer Grammatik, auf die kein Mensch Lust hat oder haben kann.

Zukunftsweisender erscheinen Modelle wie die Feldgliederung (Felderstruktur) und Verbvalenz, die ebenfalls im neuen „Verzeichnis grundlegender grammatischer Fachausdrücke" genannt werden. Die lineare Darstellung bietet ausgehend vom Verb die Möglichkeit, in der Zeilengliederung zunächst die Sätze in ihrer formalen Gestalt wahrzunehmen. Dass sie über- oder untergeordnet sein können, zeigt sich durch die Einbettungen in den einzelnen Zeilen. Das verständige Eintragen eröffnet dabei strukturelle Einsichten und die Chance zu einem tieferen Text- und Sprachverständnis.

 References: Dudenredaktion (Hrsg.). 2016. *DUDEN. Die Grammatik. Unentbehrlich für richtiges Deutsch.* 9., vollständig überarbeitete und aktualisierte Auflage. Berlin: Dudenverlag. Granzow-Emden, Matthias. 2019. *Deutsche Grammatik verstehen und unterrichten. Eine Einführung.* Unter Mitarbeit von Johannes Luber. 3., überarbeitete und erweiterte Auf- lage. Tübingen: Narr Francke Attempto. Granzow-Emden, Matthias. 2020a. "Sind Haupt- und Nebensatz noch zu retten? Ein Plädo- yer für eine widerspruchsfreie Satzanalyse." *Der Deutschunterricht* 2/2020: 14-24. Granzow-Emden, Matthias. 2020b. "Das rote und das blaue Feld. Spielerisches Entdecken des Verbzweitsatzes." *Praxis Deutsch* 282/2020: 24-28 Leibniz-Institut für deutsche Sprache/KMK. 2019. *Laut, Buchstabe, Wort und Satz. Verzeich- nis grundlegender grammatischer Fachausdrücke*. Von der Kultusministerkonferenz zustimmend zur Kenntnis genommen am 7. November 2019: https://grammis.ids- mannheim.de/vggf

Das Stellungsfeldermodell im Sprachunterricht: Wie urteilen Deutschlehrer*innen über ein grammatisches Modell?

Elvira Topalović, Benjamin Uhl
Universität Paderborn, Universität Paderborn
elvira.topalovic@uni-paderborn.de, benjamin.jakob.uhl@uni-paderborn.de

Das Stellungsfeldermodell ist im sprachdidaktischen Forschungsdiskurs mittlerweile prominent vertreten:

Es wird in zahlreichen sprachdidaktischen Einführungswerken vorgestellt (vgl. z.B. Bredel 2013; Granzow-Emden 2019), in Monographien und Sammelbänden zum Thema gemacht (vgl. z.B. Metzger 2017; Wöllstein 2015; Peyer 2011) und aktuell in einem gleichsam transdisziplinären Zugang aus sprachdidaktischer, spracherwerbstheoretischer und linguistischer Sicht diskutiert (vgl. Uhl 2019; Müller/Schönfelder 2019; Gallmann 2019; Tophinke/Topalović/Rohlfing 2019).

Während das Stellungsfeldermodell in der DaF-Forschung bzw. DaF-Didaktik bereits seit Jahrzehnten etabliert ist, dürfte die Aufnahme des grammatischen Terminus „Felderstruktur" in das neue, von der Kultusministerkonferenz (KMK) im November 2019 zustimmend zur Kenntnis genommene „Verzeichnis grundlegender grammatischer Fachausdrücke" (vgl. unter: https://grammis.ids-mannheim.de/vggf) nicht nur seine Implementation in Lehrwerke beschleunigen – Hennig/Langlotz (2020, 23) sprechen von „erhellenden Ergänzungen zu den Kernbereichen der Schulgrammatik" –, sondern auch die empirische sprachdidaktische Forschung begünstigen. Denn inwieweit das Stellungsfeldermodell im schulischen Handlungsfeld bzw. im Fach Deutsch tatsächlich genutzt wird und falls ja, wie, ist bisher kaum empirisch erforscht.

Um erste Erkenntnisse unter anderem über die Bekanntheit des Stellungsfeldermodells (z.B. seiner theoretischen Grundlagen), seinen Einsatz (z.B. als Lernmittel oder Lerngegenstand) und vor allem sein angenommenes Potential (z.B. als Lesestrategie oder inklusionsdidaktische Ressource) zu gewinnen, haben wir die Fachkollegien Deutsch zweier weiterführender Schulen (n=26) befragt. Die Ergebnisse sollen auch Auskunft darüber geben, ob das Modell in besonderer Art und Weise an Bedingungen des grammatischen Lernens im Deutsch- bzw. Sprachunterricht angepasst wird bzw. werden würde.

References: Bredel, Ursula. 2013. *Sprachbetrachtung und Grammatikunterricht*. Paderborn: Schöningh. Gallmann, Peter. 2019. "Das topologische Modell". In Doris Tophinke, Elvira Topalović and Katharina Rohlfing: 344–352. Granzow-Emden, Matthias. 2019. *Deutsche Grammatik verstehen und unterrichten. Eine Einführung.* Tübingen: Narr Francke Attempto. Hennig, Mathilde and Miriam Langlotz. 2020. "Das ‚Verzeichnis grundlegender grammatischer Fachausdrücke' 2019. Anliegen, Konzeption, Perspektiven". In Sprachreport 36 (2020), Nr. 2: 20–31.Metzger, Stefan. 2017. *Grammatik unterrichten mit dem Feldermodell: Didaktische Grundlagen und Aufgaben für die Orientierungsstufe.* Hannover: Kallmeyer. Müller, Anja and Mandy Schönfelder. 2019. "Das topologische Modell aus spracherwerbstheoretischer Perspektive". In Doris Tophinke, Elvira Topalović and Katharina Rohlfing: 353–359. Peyer, Ann. 2011. *Sätze untersuchen. Lernorientierte Sprachreflexion und grammatisches Wissen.* Hannover: Kallmeyer. Tophinke, Doris, Elvira Topalović and Katharina Rohlfing, eds. 2019. *Sprachstrukturelle Modelle. Konvergenzen theoretischer und empirischer Forschung.* Mitteilungen des Deutschen Germanistenverbandes 4/2019. Göttingen: V&R unipress. DOI: 10.14220/mdge.2019.66. Tophinke, Doris, Elvira Topalović and Katharina Rohlfing. 2019. "Sprachstrukturelle Modelle. Linguistische, spracherwerbsorientierte und didaktische Zugänge". In Doris Tophinke, Elvira Topalović and Katharina Rohlfing: 313–319. Uhl, Benjamin. 2019. "Topologie und sprachliches Lernen – vier Argumente für das sprachdidaktische Potenzial des Stellungsfeldermodells". In Doris Tophinke, Elvira Topalović and Katharina Rohlfing: 360–368.Wöllstein, Angelika unter Mitarbeit von Saskia Schmadel, eds. 2015. *Das topologische Modell für die Schule.* Baltmannsweiler: Schneider Verlag Hohengehren.

Satzgliedmodelle in Schulbüchern: Eine qualitative Untersuchung der Klassen 5 und 6

Steffen Dyck
TU Braunschweig
s.dyck@tu-bs.de, s.dyck@mail.de

Der promotionsbasierte Vortrag fokussiert die Qualität schulgrammatischer Satzgliedmodelle in aktuellen Bildungsmedien verschiedener Schulformen. Hierbei richtet sich das Erkenntnisinteresse auf die logische Konsistenz, die Funktionalität der valenzgrammatischen, topologischen und semantischen Konzepte sowie das Definitionsmuster. Diese Aspekte stehen im Spannungsfeld zwischen den gleichermaßen wertvollen wie herausfordernden Charakteristika der Satzgliedlehre und den im Grammatikunterricht anzustrebenden Kompetenzen. So gilt die Satzgliedlehre neben der Wortartenlehre als eine der beiden Säulen der Schulgrammatik (Granzow-Emden 2019: 10). Ihr Konzept zeichnet sich durch ein komplexes Kategorieninventar sowie ein heterogenes formalgrammatisch, topologisch und semantisch geprägtes Kriterienspektrum aus (Gallmann & Sitta 1992: 143). Einerseits ermöglicht dies die Entwicklung optimaler Satzgliedmodelle zur Ausbildung von Kompetenzen, mit denen Schülerinnen und Schüler sowohl zur Analyse sprachlicher Daten als auch zur eigenständigen und gezielten Textüberarbeitung befähigt werden (vgl. Müller 2011: 144). Andererseits stellt der Umgang mit ebendiesen Charakteristika bei der Entwicklung von Satzgliedmodellen für Schulbuchverlage eine Herausforderung dar. Die vorgestellten Ergebnisse zur Qualität der logischen Konsistenz, der konzeptionellen Funktionalität und dem Definitionsmuster ermöglichen die Diskussion potenzieller Konsequenzen für den Grammatikunterricht sowie den Lernprozess und die Kompetenzen.

References: Gallmann, Peter & Horst Sitta (1992): Satzglieder in der wissenschaftlichen Diskussion und in Resultatsgrammatiken. *Zeitschrift für germanistische Linguistik (ZGL)* 20.2: 137- 181. Granzow-Emden, Matthias (2019) [2013]: *Deutsche Grammatik verstehen und unterrichten.* überarbeitete und erweiterte Aufl. Tübingen: Narr Francke Attempto Verlag GmbH + Co. KG. Müller, Christoph (2011): Warum fällt mir das nicht ein? Grammatikwissen als Impuls für selbstgesteuerte Schreibprozesse. *Osnabrücker Beiträge zur Sprachtheorie* 79: 141- 158.

Empirische Befunde zum Einsatz grammatischer Modelle im Deutschunterricht

Daniela Elsner
PH Vorarlberg
daniela.elsner@ph-vorarlberg.ac.at

Dass grammatische Modelle im schulischen Deutschunterricht Anwendung finden, steht außer Frage; schließlich kann auch die traditionelle partes orationis-Einteilung der lateinischen Schulgrammatik als Modell verstanden werden, also als verkürztes Abbild eines Originals, das einem bestimmten Zweck genügt (vgl. Stachowiak 1973). Ebenso hat der Strukturalismus über die Arbeiten von Glinz in Form verschiedener Proben (Umstellprobe, Ersetzungsprobe) seinen Weg ins Klassenzimmer gefunden. Vielmehr ist daher von Interesse zu erörtern, welche Modelle auf welche Art und Weise im schulischen Grammatikunterricht eingesetzt werden. Denn das vorrangige Ziel von in der Wissenschaft entwickelten Modellen ist nicht deren Anwendbarkeit im Schulunterricht, sondern eine konsistente Beschreibung und Erklärung sprachlicher Phänomene (vgl. Dürscheid 2010), woraus folgt, dass eine Modifizierung der Modelle unter didaktischen Aspekten erfolgen muss und dass es auch Modelle gibt, die sich nicht für einen Einsatz im Unterricht eignen.

Ich werde in meinem Vortrag zunächst erste Ergebnisse einer empirischen Untersuchung berichten, bei der auf Basis der Auswertung einer österreichischen Schulbuchreihe (Starke Seiten Deutsch) eine Befragung von Deutschlehrerinnen und -lehrern im österreichischen Bundesland Vorarlberg durchgeführt wurde. Ziel dieser Untersuchung ist in einem ersten Schritt die Beantwortung der Frage, welche Modelle im Deutschunterricht wie eingesetzt werden. In einem zweiten Schritt versuche ich herauszuarbeiten, welche Eigenschaften grammatische Modelle haben (müssen), um im schulischen Deutschunterricht genutzt werden zu können. Denn offensichtlich eignen sich nicht alle Modelle ohne weiteres, wie u.a. die 1950er/60er Jahre gezeigt haben, in denen die Inhaltbezogene Grammatik Weisgerbers einen Einfluss auf den Deutschunterricht hatte, oder die 1970er Jahre, in denen die Generative Grammatik Einzug in den Deutschunterricht fand.

References: Dürscheid, Christa. 2010. "Lateinische Schulgrammatik oder andere Modelle? Welche Grammatik eignet sich am besten zur Beschreibung des Deutschen?" In Mechthild Habermann, ed. Grammatik wozu? Vom Nutzen des Grammatikwissens in Alltag und Schule. Mannheim/Zürich: Dudenverlag. 47-65. Stachowiak, Herbert. 1973. Allgemeine Modelltheorie. Wien/New York: Springer.

Integrative Sprachbildung und sprachreflexive Vermittlung in der Grundschule am Beispiel des Projekts „wortreich"

Christina Noack, Anna Kurtz, Bastian Stöppler
Universität Osnabrück
cnoack@uos.de; kontakt@wortreich-sprachbildung.de

Der moderne Sprachunterricht ist von seiner curricularen Struktur her integrativ und kumulativ angelegt (BS MSA 2003). Vom Beginn der Grundschule bis zum Ende der Pflichtschulzeit – so die Zielsetzung – bilden Schüler lernbereichsübergreifend und spiralcurricular fachbezogene Kompetenzen aus, die teils aus der Fachtradition, teils aus aktuellen bildungspolitischen Bedarfslagen resultieren. In der Praxis stellt sich Sprachunterricht jedoch oftmals als ausgesprochen inkohärent dar: Gerade am Übergang von der Primar- zur Sekundarstufe fehlen einheitliche Begriffe und Methoden, Kompetenzen werden zu wenig vernetzt und Kategorien eher plakativ als funktional vermittelt, was einer Einsichtnahme in sprachliche Funktionen und Zusammenhänge kaum förderlich ist (vgl. u. a. Granzow-Emden [3]2019). Dabei ist forschungsseitig weitgehend unstrittig, dass ein reflexiv angelegter Unterricht einem auf Regelabruf und Reproduktion ausgerichteten vorzuziehen ist, zumal wenn Deutschunterricht, wie es in den Bildungsstandards intendiert ist, Einsichten in sprachliche Funktionen evozieren soll.

Im Vortrag wird am Beispiel der Kommasetzung – mit Fokus auf ihre syntaktische Fundierung (nach Primus 1993; Bredel 2008) – gezeigt, welche Wissensgrundlagen schülerseitig zu entwickeln sind, um sie in ihrer Komplexität zu verstehen und funktional anzuwenden. Um das Ziel einer umfassenden Kommakompetenz zu erreichen, so die Position des Beitrags, müssen bereits in den ersten Schuljahren Einsichten in den Aufbau von Sätzen vermittelt werden, die über die übliche anfängliche Satzgliedanalyse hinausgehen (Afflerbach 1997; Dauberschmidt 2016). Wie dies in der Praxis gelingen kann, wird am Beispiel des Transferprojekts „wortreich – Sprachbildung für alle" vorgestellt, in dem Grundschulkindern über einen niedrigschwelligen valenzorientierten Ansatz Zugänge zu einer nachhaltigen Reflexion über Sprachstrukturen eröffnet werden, die im Idealfall in eine automatisierte Kommasensitivität (Esslinger 2014) münden.

References: Afflerbach, Sabine. 1997. Zur Ontogenese der Kommasetzung vom 7. bis zum 17. Lebensjahr. Eine empirische Studie. Frankfurt/ Main u.a., Lang. Bredel, Ursula. 2008. Die Interpunktion des Deutschen. Ein kompositionelles System zur Online-Steuerung des Lesens. Tübingen, Niemeyer. [BS MSA 2003]: Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss. Hg. von der Kultusministerkonferenz(https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2003/2003_12_04-BS-Deutsch-MS.pdf) Dauberschmidt, Franziska. 2016. „Die Entdeckung des (syntaktisch fundierten) Kommasystems". In Birgit Mesch, and Christina Noack, eds. System, Norm und Gebrauch – drei Seiten einer Medaille? Performanz im Spannungsfeld zwischen System, Norm und Empirie. Baltmannsweiler, Schneider Hohengehren: 174-199. Esslinger, Gesine. 2014. Rezeptive Interpunktionskompetenz. Eine empirische Untersuchung zur Verarbeitung syntaktischer Interpunktionszeichen beim Lesen. Baltmannsweiler, Schneider Hohengehren. Granzow-Emden, Matthias. 2019. Deutsche Grammatik verstehen und unterrichten. 3., überarbeitete und erweiterte Auflage. Tübingen, Narr. Primus, Beatrice. 1993. "Sprachnorm und Sprachregularität: Das Komma im Deutschen." *Deutsche Sprache* 3: 244–263.

Modelle des Sprachwandels im Deutschunterricht: Sprachreflexive und fächerübergreifende Potenziale

Katharina Böhnert
RWTH Aachen University
k.boehnert@lfd.rwth-aachen.de

In den Bildungsstandards für den Erwerb der Allgemeinen Hochschulreife sind Kompetenzen wie „Phänomene des Sprachwandels […] theoriegestützt beschreiben" (KMK 2012: 26) fest verankert. In der Unterrichtspraxis beschränkt sich der Einsatz von Sprachgeschichte und Sprachwandel – sofern sie überhaupt eingesetzt werden (vgl. Böhnert 2017: 95) – jedoch auf den Einsatz vereinzelter Wortgeschichten, wie sie bspw. in den Lehrwerken für die gymnasiale Oberstufe zu finden sind. Im vorliegenden Vortrag soll demgegenüber ein anderes Konzept des Einsatzes von Sprachgeschichte und Sprachwandel in der Sekundarstufe II vorgestellt werden, das einerseits den o.g. Anforderungen in den Bildungsstandards Rechnung trägt und darüber hinaus auch fächerübergreifendes Potenzial bietet. So sollen Theorien des Sprachwandels schülernah didaktisch aufbereitet und anschließend gezeigt werden, wie sich diese Theorien – da sie oftmals auf generalisierte Grundmuster in der Onto- und Phylogenese rekurrieren – auch für fächerübergreifenden Unterricht eignen. Mögliche Sprachwandeltheorien sind hierbei

Kellers (1990) Theorie der Unsichtbaren Hand, mittels derer Sprachwandel als unbeabsichtiger Nebeneffekt unseres alltäglichen Kommunizierens beschrieben wird. Hier bietet sich der fächerübergreifende Blick auf die Gesellschafts-, insbesondere Wirtschaftswissenschaften, an, in denen Adam Smiths *Invisible Hand*-Theorie bereits breit rezipiert wird; Parallelen von Historio- und Ontogenese, wie u.a. Bredel (2006) sie für die Entwicklung bzw. den Erwerb der satzinternen Großschreibung aufzeigt. Theorien des graphematischen Wandels lassen sich hier lernbereichübergreifend im Kontext von Orthographiedidaktik behandeln oder auch fächerübergreifend, indem z.B. im Biologieunterricht auf Parallelen historio- mit onto- bzw. phylogenetischen Entwicklungsprozessen aufmerksam gemacht wird; dies gilt auch für Grammatikalisierungstheorien wie bspw. die Exaptationstheorie, die Simon (2003) auf die Entwicklung der Anredepronomen im Deutschen angewandt hat.

Im Vortrag soll anhand von konkreten Unterrichtsmodellen gezeigt werden, wie die genannten Theorien im Deutschunterricht eingesetzt werden können und hierbei einen ebenso curricular angemessenen wie lernertragreichen Sprachgeschichtsunterricht in der Oberstufe möglich machen, der Schülerinnen und Schüler zur Reflexion über Sprache – auch in ihrer diachronen Dimension – anregt.

References: Böhnert, Katharina 2017. Sprachwandel beobachten, untersuchen, reflektieren: Was Sprachgeschichte für den gymnasialen Deutschunterricht leisten kann. Frankfurt am Main: Peter Lang. Bredel, Ursula. 2006. "Die Herausbildung des syntaktischen Prinzips in der Historiogenese und der Ontogenese der Schrift." In Ursula Bredel, Hartmut Günther, eds. Orthographietheorie und Rechtschreibunterricht. Tübingen: Niemeyer. 139-163. Keller, Rudi.1990. Sprachwandel. Von der unsichtbaren Hand in die Sprache. Tübingen: utb. Kultusministerkonferenz. 2012. Bildungsstandards im Fach Deutsch für die Allgemeine Hochschulreife. http://www.kmk.org/bildung-schule/qualitaetssicherung-in schulen/bildungsstandards/dokumente.html [10.11.2020]. Simon, Horst 2003. "From pragmatics to grammar. Tracing the development of respect in the history of the German pronouns of address." In Irma Taavitsainen, I., Andreas H. Jucker, eds. Diachronic perspectives on address term systems.Amsterdam/Philadelphia: John Benjamins. 85-123.

Grammatische Modelle im Unterricht Deutsch als Fremdsprache: Indirekte Evidenzen aus Lernersprachenforschung und linguistischer Lehrwerkanalyse

Eva Breindl
Friedrich-Alexander-Universität Erlangen-Nürnberg.
eva.breindl@fau.de

In den jahrzehntelangen Kontroversen zum Stellenwert von Grammatik im DaF-Unterricht (vgl. etwa Diehl et al. 2000, Burwitz/Melzer 2012) hat sich als eine Art Minimalkonsens herausgeschält, dass explizite Vermittlung grammatischer Regularitäten nicht ganz obsolet, aber doch nur insoweit zu rechtfertigen ist, als sie ihre „dienende Funktion" für die Entwicklung umfassender kommunikativer Kompetenzen der Lernenden erfüllt. Diese Globalorientierung an der kommunikativen Kompetenz teilt der DaF-Unterricht mit der Muttersprachendidaktik (s. Holler/Steinbach 2015), wobei der Sprachreflexion in letzterer ein größerer Stellenwert zukommt als etwa in den Lernzielbeschreibungen des Gemeinsamen Europäischen Referenzrahmens. Somit müsste jede grammatische Theorie im DaF-Unterricht an ihrem Beitrag zum Erfolg des Spracherwerbs gemessen werden. Der Versuch einer empirischen Validierung steht hier aber nicht nur vor der generellen Herausforderung der Faktorenkomplexität im L2-Erwerbsprozess, sondern auch des methodischen Problems einer vergleichenden Untersuchung gesteuerten Spracherwerbs. Ein alternativer Weg besteht darin, lernersprachliche Da- ten mit dem potentiellen Input, den eine grammatische Theorie bieten kann, sowie mit dem konkreten Input, wie er sich in Lehrwerken manifestiert, abzugleichen. Die Leitfrage lautet dann: inwieweit ist die Theorie X bzw. ihre komplexitätsreduzierte didaktische Aufbereitung in Lehrmaterialien in Bezug auf ein bestimmtes Phänomen geeignet, typische Lernerfehler und Overuse-Underuse-Phänomene zu verhindern?

Mit Blick auf grammatische Theorien gilt für DaF nach wie vor die Aussage Roches (2008), dass „wegen der relativ einfachen Darstellungsmöglichkeiten […] außer schulgrammatischen Darstellungen […] die Valenzgrammatik eigentlich der einzige grammatische Ansatz ist, der systematisch in Lehrwerken zum Einsatz kommt." Traditionell wird die vermeintlich gute Passung, ja Überlegenheit dieser Theorie für den DaF-Unterricht mit ihrer Erklärungskraft für Wortstellungsregeln begründet (s. Fobbe 2010).

Im Vortrag soll an einem Lernerkorpus mit Hilfe einer Analyse von Wortstellungsfehlern herausgearbeitet werden, welche strukturellen und funktionalen topologischen Regeln typischerweise auf der Satz- und Textebene verletzt werden. Anschließend wird geprüft, ob und wie die zur Vermeidung solcher Fehler nötigen Informationen von der Valenzgrammatik, traditionellen Ansätzen und vom Feldermodell bereitgestellt werden. Die lernersprachliche Analyse wird durch eine exemplarische linguistische Lehrwerkanalyse (s. Ahrenholz/Grießhaber 2019) ergänzt. Im Mittelpunkt steht die Besetzung des Vorfelds als einer Gelenkstelle zwischen satzbezogenen „Grundwortstellungsregeln" und kontextbezogener Variation.

References: Ahrenholz, Bernt, und Wilhelm Grießhaber. 2019. „Texte in Schulbüchern und ihre Analyse." In: Ahrenholz, Bernt/Jeuk, Stefan/Lütke, Beate/Paetsch, Jennifer, und Heike Roll, eds. *Fachunterricht – Sprachbildung – Sprachkompetenzen*. Berlin: De Gruyter. 158-184. Burwitz-Melzer, Eva/Königs, Frank G, und Hans-Jürgen Krumm, eds. 2012. *Sprachenbewusstheit im Fremdsprachenunterricht*. Tübingen: Narr. Europarat. 2001. *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin: Langenscheidt. http://www.europaeischer-referenzrahmen.de/. Fobbe, Eilika. 2010. „Was von der Valenz übrig bleibt. Die Rolle der Valenzgrammatik in Lehrwerken des Deutschen als Fremdsprache." In: Fischer, Klaus, Fobbe, Eilika und Sefan Schierholz, eds. *Valenz und Deutsch als Fremdsprache*. Bern u.a.: Lang. 61-85. Holler, Anke, und Markus Steinbach. 2015. „Grammar in the classroom." In: Alexiadou, Artemis und Tibor Kiss, eds. *Syntax. An international handbook of contemporary research*. Berlin, New York: De Gruyter. 2095-2126. Roche, Jörg (2008): *Fremdsprachenerwerb, Fremdsprachendidaktik. 2., überarbeitete und erweiterte Auflage.* Tübingen: Narr.

Form follows function: Sprachdidaktisches Potenzial gebrauchsbasierter Ansätze von Sprache und Spracherwerb

Karin Madlener-Charpentier

Universität Basel/Zürcher Hochschule für Angewandte Wissenschaften

Gebrauchsbasierte Modelle von Sprache und Spracherwerb gehen davon aus, dass (1) Sprachwissen modellierbar ist als strukturiertes Netzwerk von Form-Bedeutung-Zuordnungen unterschiedlicher Abstraktheit und Komplexität (wobei ein Kontinuum lexikogrammatischer Ressourcen angenommen wird); (2) Spracherwerb primär auf impliziten und inzidentellen Prozessen der Rekonstruktion von Form-Bedeutung-Beziehungen und Kookkurrenzmustern aus dem Input beruht (N. Ellis & Cadierno 2009). Für (Erst-, Zweit- und Fremd-)Spracherwerb ist daher einerseits die Menge und Qualität des verfügbaren Inputs zentral, andererseits die Qualität der Inputverarbeitung durch Lernende (N. Ellis & Cadierno 2009: 117). Didaktische Maßnahmen zur Verbesserung der Inputverarbeitung im Sinne einer gebrauchsbasierten didaktischen Formfokussierung schliessen dabei nebst impliziten Optionen u.a. der Inputoptimierung (Madlener 2015) – durchaus auch ein Repertoire explizit bewusstmachender Verfahren der Sprach(-gebrauchs-)reflexion ein, die u.a. an das Konzept des *pushed output* anknüpfen können (Swain 2005); im Fokus stehen dabei das Konzept der Aufmerksamkeit (*attention*) bzw. des *noticing (the gap)* und die Frage, wie die Aufmerksamkeit der Lernenden auf schwierige Form-Bedeutung-Zuordnungen gelenkt werden kann, so dass auch für wenig saliente, intransparente oder redundante Konstruktionen Intake generiert und eine Restrukturierung der Lernersprache angestossen werden kann (vgl. N. Ellis 2008; R. Ellis 2016).

Dieser Beitrag diskutiert das Potenzial einer gebrauchsbasierten didaktischen Formfokussierung für verschiedene Vermittlungskontexte und Lerngegenstände. Folgende Fragen stehen dabei im Fokus: Bis zu welchem Grad bzw. in welchen Fällen ist explizites Wissen/Lernen/Lehren notwendig und zielführend? Unter welchen Bedingungen können explizite, be- wusstmachende Varianten einer didaktischen Formfokussierung für spezifische Zielgruppen erwerbsförderlich umgesetzt werden? Wie können lernerseitiges *noticing,* Hypothesentesten und metasprachliche Reflexion erreicht bzw. angeleitet werden? Welche Techniken und (ggf. kollaborativen) Aufgabenformate stehen dafür zur Verfügung? Und welche Befunde, Konstrukte und Vermittlungsoptionen aus der L2-Forschung können hier für die L1-Didaktik nutzbar gemacht werden? Der Beitrag schlägt dazu das Grundprinzip „form follows function" vor: Die Reflexion (der Verwendung) sprachlicher Mittel (→ Formen) muss in die Erfahrung und Reflexion sprachlich-kommunikativer Zwecke und Ziele (→ Funktionen) eingebettet sein (vgl. R. Ellis 2016). Drei Dimensionen des Grundprinzips werden skizziert: (1) von der Sprachhandlung zu den sprachlichen Mitteln; (2) von lexikalischen zu grammatischen Ressourcen; (3) von der Inputverarbeitung zum (*pushed*) Output.

References: Ellis, Nick C. (2008). Usage-based and form-focused language acquisition. The associative learning of constructions, learned attention, and the limited L2 endstate. In P. Robinson & N. C. Ellis (Hrsg.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York: Routledge, 372-405. Ellis, Nick C. & Teresa Cadierno (2009). Constructing a Second Language. Introduction to the Special Section. *Annual Review of Cognitive Linguistics* 7, 111-139. Ellis, Rod (2016). Focus on form: A critical review. *Language Teaching Research* 20 (3), 405-428. Madlener, Karin (2015). *Frequency Effects in Instructed Second Language Acquisition.* Berlin, Boston: De Gruyter Mouton. Swain, Merrill (2005). The output hypothesis: Theory and research. In E. Hinkel (Hrsg.), *Handbook of research in second language teaching and learning*. Erlbaum, 471-483.

Arbeitsgruppe 15

Contrastive corpus methodology for language modeling and analysis
Martin Klotz, Anke Lüdeling & Anna Shadrova

Ein Linksklick auf den Titel eines Vortrags oder den Namen eines Vortragenden im Programm führt – falls vorhanden – zum entsprechenden Abstract weiter unten. Umgekehrt führt ein Linksklick auf die Überschrift eines Abstracts zum entsprechenden Programmslot.

Left clicking on the title of a talk or the name of its speaker will lead you to the abstract of that talk further down below. Conversely, left clicking on the title of an abstract will lead you to its programme slot.

Linguistic modelling and analysis

Anna Shadrova, Martin Klotz, Anke Lüdeling
Humboldt-Universität zu Berlin
anna.shadrova@hu-berlin.de, martin.klotz@hu-berlin.de, anke.luedeling@hu-berlin.de

As we try to understand and empirically investigate language, a wide range of methods are at our disposal and many decisions are to be made. Not only do we have more and better (corpus) data and more sophisticated formal and quantitative models for analyzing it, we also have many more linguistic theories with various foci and goals, based on different (and sometimes incompatible) formal models. Ideally, each research question would require an adequate modelling, from which the operationalization of the analysis would derive, before we would even start looking at the data. Yet, often we find the same analytic method (be it a simple one or the most advanced method at hand) being used throughout a study without recurrence to the research question at hand. Which is not surprising as it may be very difficult to understand the (formal properties behind the) underlying linguistic models and the make-up of the data well enough to translate those into models of analysis.

In our introductory talk, which is meant to contextualize the workshop and generate discussions, we will focus on several examples, based on German learner and heritage speaker corpora. These are what we call mid-sized corpora, that is, corpora which are too small to do large-scale statistical analysis but small enough to be well-understood in design and annotation. All examples is concerned with the analysis of lexical and morphological composition and structure, and they require different models, methods, and interpretations. The investigation of the many different aspects of morphological productivity needs statistical analyses, the analyses of co-selection within the lexicon can be done adequately with graph- based methods, etc. We will discuss several aspects of finding the right method and model.

The group and the individual: Complementary dimensions of language development

Wander Lowie
University of Groningen
w.m.lowie@rug.nl

Most 'traditional' research into second language development tends to focus on groups of learners that are representative for larger populations based on statistics based on the Gaussian distribution. This allows us to make generalizations about second language development that describe general tendencies and relationships. Extensive corpora of language use allow us to validate these observations and to corroborate our generalizations. In contrast, studies inspired by Complex Dynamic Systems Theory (CDST) tend to focus on the development over time of using single or multiple case studies. This allows us to closely follow and describe the process of development of individual learners. The two approaches seem to cover contradictory perspectives on L2 development. Group studies allow for generalizations, but cannot say anything about the individuals in the group. Case studies show us the development of individual learners, but these observations cannot be generalized beyond the individual learner. In this presentation, I will discuss the contrastive dimensions of research into L2 development and consider solutions to solve the apparent incommensurability of the two approaches. I will make use of the construct of ergodicity as described in Lowie and Verspoor (2019), and will discuss a new (and still ongoing) study that focuses on learners' use of mobile devices to support language learning (Peng, Jager & Lowie, forthcoming), where the group meets the individual. Finally, I will plea for researchers' collaboration in creating and expanding longitudinal learner corpora to uncover the process of second language development.

References: Lowie, Wander M., and Marjolijn H. Verspoor. (2019). "Individual Differences and the Ergodicity Problem." *Language Learning* 69(1), 184-206.

A comparison of frequentist and Bayesian models of language variation: The problems of priors and sample size

Natalia Levshina
MPI for Psycholinguistics
natalia.levshina@mpi.nl

Bayesian inference is becoming increasingly popular in linguistic research (cf. Nicenboim & Vasishth 2016). In this talk I will compare frequentist (maximum likelihood) and Bayesian approaches to mixed-effect logistic regression, which is *de facto* the standard method for modelling linguistic variation. The main advantages of Bayesian inference are as follows:

- it provides the researcher with an opportunity to test the research hypothesis directly, instead of trying to reject the null hypothesis;
- one can use information from previous research as priors for subsequent models, which helps to overcome the recent crisis of reproducibility (Goodman et al. 2016). This also enables one to use smaller samples;
- it helps to solve such problems as overfitting, data separation and convergence issues, which often arise when one fits generalized mixed-effect models with complex structure.

The Bayesian approach to generalized mixed-effect models is illustrated by a multifactorial case study of help + (to-)infinitive in U.S. magazines, e.g. *These simple tips will help you (to) survive the Zombie apocalypse.* According to previous research (e.g. Rohdenburg 1996; Lohmann 2011; Levshina 2018), there are many factors that play a role in the choice for one or the other variant, such as the distance between *help* and the infinitive, the morphological form of *help*, the presence of *Helpee*, the *horror aequi* principle, formality and dialect. These factors are first tested on a dataset of 2,050 examples of the constructions from the Magazines section of the Corpus of Contemporary American English. The frequentist and Bayesian models yield very similar results. Importantly, Bayesian posteriors are not sensitive to priors. However, things change dramatically when we take a small dataset of 400 observations to test the principle of rhythmic alternation (Schlüter 2003), which requires meticulous and costly annotation of corpus data. The Bayesian model with informative priors converges, whereas the frequentist model is highly problematic due to data sparseness.

References: Goodman, Steven, Daniele Fanelli, and Johny Ioannidis. 2016. "What does research reproducibility mean?" *Science Translational Medicine* 8(341): 12. Levshina, Natalia. 2018. "Probabilistic grammar and constructional predictability: Bayesian generalized additive models of help + (to) infinitive in varieties of web-based English." *Glossa* 3(1): 1-22. Lohmann, Arne (2011). "Help vs. help to – a multifactorial, mixed-effects account of infinitive marker omission." *English Language and Linguistics* 15(3): 499-521. Nicenboim, Bruno, and Shravan Vasishth. 2016. "Statistical methods for linguistic research: Foundational Ideas - Part II." *Language and Linguistics Compass* 10: 591-613. Rohdenburg, Günter. 1996. "Cognitive complexity and increased grammatical explicitness in English." *Cognitive Linguistics* 7(2), 149-182. Schlüter, Julia. 2003. "Phonological determinants of grammatical variation in English: Chomsky's worst possible case." Günter Rohdenburg, and Britta Mondorf, eds. *Determinants of Grammatical Variation in English.* Berlin, and New York: Mouton de Gruyter. 69-118.

Corpora, inference, and models of register distributions

Felix Bildhauer, Elizabeth Pankratz, Roland Schäfer
IDS Mannheim, HU Berlin, HU Berlin
 felix.bildhauer@hu-berlin.de, pankrael@hu-berlin.de, mail@rolandschaefer.net

The workshop description relies on the concept of "in-depth theoretical modelling" in corpus analysis and design. It also draws a sharp distinction between lots of bad data ("shallow") and small amounts of good data ("deeper, manually obtained"). We agree with concerns about inaccurate automatic annotations, but we argue that such a strict dichotomy is not required and even potentially perilous. We take corpus-based register modelling as our example and describe a new method of modelling register distributions which solves related problems. A register is defined as a set consisting of pairs of signs (comprising words, constructions, rules, etc.) and situational parameters.

The core purpose of empirical methods in mature scientific inference is to find flaws in models of reality through some form of severe testing (Mayo 2018). With respect to registers, controlled experiments could be used probe flaws in a model of the cognitive representation of registers. A situation where the data *incorrectly fails to contradict* or *incorrectly contradicts* the model might arise from [i] an incorrect model (e.g., the assumption of discreteness in a probabilistic phenomenon), [ii] an inadequate experimental method, [iii] bad error control (e.g., confounding sociolectal variables), or [iv] flawed statistical analysis (statistical tests with too high or too low power). Unfortunately, it is often a mixture of the above. By relying on manual annotation of corpus data based on in-depth theoretical modelling, one always runs a significant risk of circularity: the corroboration of primary or auxiliary hypotheses of the specific models used in annotating the corpora, *even if the model is wrong*.

Turning to corpus-based register modelling, Multi-Dimensional Analysis (MDA; Biber 1988) is an established method. MDA assumes that the relevant registers (of English) and the associated linguistic signs are known (although it is unclear how they are known). In Biber's work, smaller corpora are manually annotated for those registers as well as known types of linguistic signs ("features"). A form of dimensionality reduction (factor analysis) is used to establish connections between signs and registers. Nothing in MDA is capable of probing for errors in the underlying model. The results of MDA will always be in accord with the model specifically *because* MDA relies on smaller, manually annotated corpora. Crucially, the substantive model is *tacitly* assumed to consist of known discrete text-wide situational parameter distributions as well as discrete mappings between those and the linguistic signs. Testing these two assumptions of discreteness, we posit, would be most relevant in light of current discussions about the non-discreteness of language (Divjak et al. 2016). Assessing the concrete register distributions of English should merely be a step in probing the adequacy of the substantive model.

In our work, we do not assume a given catalogue of registers for a fundamental reason: while the associated signs can be more or less exhaustively enumerated, there are as many situational parameters as there are types of situations. There is no way for linguists to enumerate these based on any criteria available today. We also assume *and test* a probabilistic model of register distributions where situational parameters probabilistically define situations and are probabilistically associated with signs— both signs and registers being probabilistically instantiated (and mixed) in corpus texts. Under this approach, Latent Dirichlet Allocation (LDA; Blei et al. 2003) is an ideal algorithm to uncover *potential registers* from the distributions of linguistic signs in very large corpora. We reduce the noisy output from LDA through exploratory manual annotation to a set of candidates for proper registerhood, creating a medium-sized corpus. However, these steps only solve the problem of the unknown concrete (probabilistically mapped) registers of German by inducing potential register distributions. They do not actually constitute a severe test of our model of probabilistic register distributions. Hence, we proceed to a true test of the model in the form of controlled experiments using data from the corpora created. Our talk describes all those steps in detail and relates them to the theme of this workshop by showing that none of them can be taken using any other sources of data.

Deviation of proportions as the basis for a keyness measure

Christof Schöch, Iuliia Dudar, Cora Rok, Keli Du
Universität Trier
schoech@uni-trier.de, NA, rok@uni-trier.de, duk@uni-trier.de

In the context of Corpus Linguistics, numerous statistical measures and instruments have been adopted to investigate and analyze large amounts of textual data, especially in a contrastive perspective (e.g. Rayson et al. 1997; Oakes and Farrow, 2007; Newman et al., 2008). Despite several important studies (e.g. Paquot & Bestgen 2009; Lijffijt et al. 2014), there is still a lack of in-depth understanding of their key characteristics and how these key characteristics impact the results. In our project "Zeta and company" we aim to enhance our understanding of statistical keyness measures that are used for comparative, quantitative analysis of two or more text collections. Based on literary texts, we are going to implement these measures in a Python framework and evaluate which measures perform best for different tasks and kinds of textual data.

The most widely used statistical keyness measures are based on word frequency (chi-squared, log likelihood etc.) and do not consider how the particular words are distributed within a corpus. This means that a word can appear to be important for the whole corpus, although it is just used very frequently in a small number of texts in this corpus. To deal with this challenge, several dispersion measures were suggested (Lyne, 1985). Stefan Gries (2008) gives a detailed overview of such measures and develops his own measure deviation of proportions (DP). DP compares the difference between observed and expected relative frequency of a word in the individual documents contained in a corpus in order to quantify how this word is dispersed. This measure seems to have several advantages compared to other dispersion measures. For example, it can handle different corpus parts, it is simple, and can distinguish between slight variations in distribution without being overly sensitive.

However, there is still a lack of empirical evidence supporting the use of DP. For this contribution, we are going to implement this measure of dispersion in our keyness framework (see Schöch et al. 2018; for a use of dispersion, though not of DP, for keyness analysis, see Egbert & Biber 2019). First, using a collection of 160 French novels from the 1980s belonging to four different subgenres (sentimental novels, crime fiction novels, science fiction novels and high-brow novels), we will examine how DP works with different numbers of texts, words and proportions of particular words in the corpus. For example, we aim to understand DP better by examining whether DP values change when the number of texts increases and whether DP values correlate with the relative word frequencies. One of the open questions about dispersion is whether it can be used to compare two collections of texts, especially when document length varies. Therefore, we will also investigate how useful DP is as a basis for keyword extraction in contrastive analysis.

References: Egbert, Jesse, and Douglas Biber. 2019. "Incorporating text dispersion into keyword analysis." *Corpora*, 14(1): 77-104. Gries, Stefan. 2008. "Dispersions and adjusted frequencies in corpora." *International Journal of Corpus Linguistics* 13(4): 403-437. Lijffijt, Jefrey, Terttu Nevalainen, Tanja Säily, Panagiotis Papapetrou, Kai Puolamäki, and Heikki Mannila. 2014. "Significance testing of word frequencies in corpora." *Digital Scholarship in the Humanities* 31(2): 374-97. Lyne, Anthony A. 1985. "Dispersion." *The Vocabulary of French Business Correspondence*. Paris and Geneva: Slatkine-Champion. 101-124. Newman, Matthew L., Carla J. Groom, Lori D. Handelman, and James W. Pennebaker. 2008. "Gender differences in language use: An analysis of 14,000 text samples". *Discourse Processes* 45: 211-36. Oakes, Michael P., and Malcolm Farrow. 2007. "Use of the chis-quared test to examine vocabulary differences in English-language corpora representing seven different countries." *Literary and Linguistic Computing* 22(1): 85-100.

Machine learning and syntactic theory: Focus on German and German varieties

Giuseppe Samo

Beijing Language And Culture University
samo@blcu.edu.cn

This paper discusses an empirical approach in investigating theory-driven linguistic proposal concerning syntactic functional architecture (Rizzi & Cinque 2016 *inter alia*) and computational principles (such as locality) with respect to degrees of confidence of machine learning algorithms fed with naturally occurring utterances from corpora transformed in vectorial representations.

Research Question: Can (supervised) machine learning environment fed with morpho-syntactic features detect the different varieties of German? We focus on Standard German and Swiss German. Similarly, we detect dimensions of variation according to the nature of the treebank in terms of *genres* (e.g. literature, social media, fiction, news, encyclopaedic entries, etc.).

Focus on German and its varieties: German and German varieties have interesting cases in syntactic literatures allowing restrictions (no more than one left-peripheral element, Roberts 2004) and freedom of movements (*scrambling*) of constituents in different areas of the structures. Both phenomena can be explained in terms of locality (Samo 2019a), which is quantitatively tested here.

Encoding syntactic features: To do so we encode presence/absence of specific lexical entries and the activation of syntactic functional projections syntactic elements as vectors of features to train a model for a classification task based on a Naïve Bayes / (supervised) artificial neural networks.

Mapping Syntactic properties and Machine Learning: We follow Samo (2019b)'s model of mapping cartographic projections into universal dependencies (Nivre 2015). By using machine learning techniques, we are able to capture the graduality of the variability between varieties and between *genres*. Not only would we find out which features are crucial in parameter setting, but we would also be able to set up an algorithm which can possibly automatically retrieve the varieties/genres in question and quantitatively account for the variation. The method we used here requires transforming naturally occurring sentences extracted from corpora into machine-readable vectors of featural representations (Merlo & Ouwayda 2018, Samo & Merlo 2019) in order to implement probabilistic classification techniques (in the spirit of Zimmerman 2014) and observe a measure of confidence. The advantage of using AI tools and the relevant statistics is the fact that they can easily spot the locus of microvariation in terms of morphosyntactic features, hopefully providing further insights on the theory and the linguistic proposal on linguistic variability.

Materials and Methods: Syntactically annotated corpora following Universal Dependencies (Nivre 2015) of German and German varieties (*treebanks* 5; *trees* 250,000 ca.). The sentences and combinations of features should be transformed into *n*-dimensions vectors: the mutual exclusivity of the features is not problematic for our analysis, because of the intrinsic nature of Naïve Bayes in considering every feature independent to each other. The resulting model is then run with the tool *Waikato Environment for Knowledge Analysis*, WEKA (Hall et al. 2009) to derive the accuracy of a model and quantifying the performance on the required linguistic tasks.

Ultimate goal and further improvements: These results aim to add a quantitative dimension to the qualitative descriptions provided in cartographic studies. A first improvement is certainly the increase of the data set. Secondly, a *human* control group can be built: the test set can be presented and classified (both naturally occurring examples in corpora and/or devised ex-novo sentences) by native speakers or experts (as "control groups" in the spirit of Gulordava et al. 2018).

References: Gulordava, Kristina, Piotr Bojanowski, Edouard Grave, Tal Linzen, Marco Baroni. 2018. Colorless Green Recurrent Networks Dream Hierarchically. *North American Chapter of the Association for Computational Linguistics: Human Language Technologies* 1: 1195-1205. Hall, Mark, Eibe Frank, Geoffrey Holmes, and Bernhard Pfahringer. 2009. "The Weka data mining software: An update." *SIGKDD Explorations Newsletter* 11(1): 10-18. Merlo, Paola, and Sarah Ouwayda. 2018. "Movement and structure effects on Universal 20 word order frequencies: A quantitative study." *Glossa* 3(1): 84. Nivre, Joakim. 2015. "Towards a universal grammar for natural language processing." In Alexander Gelbukh, ed. *Computational Linguistics and Intelligent Text Processing*. Heidelberg, New York, Dordrecht, and London: Springer International Publishing. 3- 16. Rizzi, Luigi,

and Guglielmo Cinque. 2016. "Functional categories and syntactic theory." *Annual Review of Linguistics* 2: 139-163. Roberts, Ian. 2004. "The C-system in Brythonic Celtic languages, V2, and the EPP." In Luigi Rizzi, ed. *The Structure of CP and IP.* Oxford and New York: Oxford University Press. 297-328. Samo, Giuseppe. 2019a. *A criterial approach to the cartography of V2.* Amsterdam and Philadelphia: John Benjamins. Samo, Giuseppe. 2019b. "Cartography and Locality in German: A quantitative study with Dependency structures." Rivista di Grammatica Generativa 5: 1-2. Samo, Giuseppe, and Merlo Paola. 2019. "Intervention effects in object relatives in English and Italian: A study in quantitative computational syntax." In Xinying Chen, and Ramon Ferrer-i-Cancho, eds. *Proceedings of the First Workshop on Quantitative Syntax*. 46–56. Zeman, Daniel, Joakim Nivre, Mitchell Abrams, et al. 2020. *Universal Dependencies 2.6.,* LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. Zimmermann, Richard. 2014*. Dating hitherto undated Old English texts based on text internal criteria.* Ms., University of Geneva.

Postersession der Computerlinguistik

Trafilatura: An open-source tool for web corpus construction

Adrien **Barbaresi**, Berlin-Brandenburgische Akademie der Wissenschaften
barbaresi@bbaw.de

Large "offline" web text collections are now standard among the research community in linguistics and natural language processing. The construction of such corpora notably involves "crawling, downloading, 'cleaning' and de-duplicating the data, then linguistically annotating it and loading it into a corpus query tool" (Kilgarriff 2007). Although text is ubiquitous on the Web, extracting information from web pages can prove to be difficult. Web documents come in different shapes and sizes mostly because of the wide variety of genres, platforms, and content management systems, and not least because of greatly diverse publication goals. Web crawling involves a significant number of design decisions and turning points in data processing, without which data and applications turn into a "Wild West" (Jo & Gebru 2020). Researchers face a lack of information regarding the content, whose adequacy, focus, and quality are the object of a post hoc evaluation (Baroni et al. 2009).

Comparably, web corpora usually lack metadata gathered with or obtained from documents. Between opportunistic and restrained data collection, a significant challenge lies in the ability to extract and pre-process data to meet scientific expectations with respect to corpus quality.

Trafilatura is a library and command-line tool used for corpus construction within the lexicographic information platform *dwds.de* (Geyken et al. 2017) which hosts and provides access to a series of metadata-enhanced web corpora (Barbaresi 2016). It seamlessly downloads, parses, and scrapes web page data. It handles the extraction of metadata, main body text and comments while preserving parts of the text formatting and page structure.

Link discovery in feeds and sitemaps is also included. The output is then converted to common formats (TXT, CSV, JSON, XML & XML-TEI). Distinguishing between a whole page and the page's essential parts helps to alleviate many quality problems by dealing with the noise caused by recurring elements (headers and footers, ads, links/blogroll, etc.), so that the software both facilitates text data collection and enhances corpus quality. As evaluations of extraction tools show significant domain-related disparities (Barbaresi & Lejeune 2020), the experiments at hand show that the tool performs better than known alternatives. It is freely available under an open-source license: https://github.com/adbar/trafilatura

References

Barbaresi, Adrien, 2016. "Efficient construction of metadata-enhanced web corpora". In Paul Cook, Stefan Evert, Roland Schäfer, and Egon Stemle, eds., *Proceedings of the 10th Web as Corpus Workshop*, Association for Computational Linguistics*, 7-16.

Barbaresi, Adrien, and Gaël Lejeune, 2020. "Out-of-the-Box and into the Ditch? Multilingual Evaluation of Generic Text Extraction Tools", *Proceedings of the 12th Web as Corpus Workshop (LREC 2020)*, ELRA, 5-13.

Baroni, Marco, Bernardini, Silvia, Ferraresi, Adriano, and Eros Zanchetta, 2009. "The WaCky Wide Web: a collection of very large linguistically processed web-crawled corpora". *Language Resources and Evaluation*, 43(3): 209-226.

Geyken, Alexander et al., 2017. "Die Korpusplattform des Digitalen Wörterbuchs der deutschen Sprache (DWDS)", *Zeitschrift für germanistische Linguistik*, 45(2): 327- 344.

Jo, Eun Seo, and Timnit Gebru, 2020. "Lessons from Archives: Strategies for Collecting Sociocultural

Data in Machine Learning", In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency,* 306-316.

Kilgarriff, Adam, 2007. "Googleology is bad science", *Computational Linguistics*, 33(1): 147- 151.

Advancing neural question generation for formal pragmatics

Speaker 1 (Kordula **De Kuthy**, Universität Tübingen), speaker 2 (Haemanth **Santhi Ponnusamy,** Universität Tübingen), speaker 3 (Madeeswaran **Kannan**, Universität Tübingen), speaker 4 (Detmar **Meurers**, Universität Tübingen),
{kdk,mkannan,hsp,dm}@sfs.uni-tuebingen.de

Question generation, creating questions for a given sentence or paragraph, is a challenging task with many applications, from question answering, via dialogue systems, to reading com- prehension tasks. The recent state-of-the-art approaches are generally based on neural net- works. The task of QG is typically formulated as a sequence-to-sequence (seq2seq) learning problem in which a sentence is mapped to a corresponding question (cf., e.g., Pan et al., 2019). In formal pragmatics, questions also play an prominent role in so-called Questions-under-Discussion (QuD, Roberts, 2012) approaches. Questions there make explicit the interface between the information structure of a sentence and the discourse structure that the sentence functions in. Under such a QuD perspective, for every sentence in a text, a question needs to be formulated – and indeed explicit guidelines have been defined to support reliable manual QuD annotation (Riester et al., 2018). De Kuthy et al. (2020) argue that such question generation should be automated for the analysis of large corpora, and they propose a seq2seq neural network approach to generate all potential questions for a given sentence. They show that the approach learned to (often) predict the correct question word for a given answer and generated questions that correctly reflect the word order properties of questions in German.

There are, however, clear challenges for such a seq2seq architecture that generates questions for any type of data set. One problem are rare or unknown words that have to be predicted. In most neural generation architectures, words are the basic tokens. Pretrained word embeddings are used to initialize the token embedding matrix with a fixed vocabulary. In any corpus material serving as input there are likely to be rare or unknown words that are not part of the fixed vocabulary and therefore cannot be predicted in the output, the generated question. This indeed is a major issue in De Kuthy et al.'s question generation approach. To overcome this problem, they implemented an ad-hoc post-processing step: Each generated question is checked for markers indicating the places where an OOV token appears. A heuristic then tries to identify that missing word in the source sentence and insert it in the output.

Here we propose to adopt a pointer-based neural architecture for QG. We show that such an architecture is more successful than the seq2seq based model, replacing the post- processing step used in De Kuthy et al. (2020) into a design feature of the neural architecture. Architecturally separating the copying from the generation component also readily supports the integration of linguistic information needed to determine the question phrase to be gene- rated. Furthermore, the pointer-based architecture is able to generalize the task of question generation in identifying the material that is identical between source sentence and question and that can simply be copied over. A quantitative evaluation using BLEU scores and an in- depth qualitative evaluation show that indeed the pointer-based model with additional linguistic features is the best system for the task of generating questions to advance discourse analysis.

References

De Kuthy, Kordula, Madeeswaran Kannan, Haemanth Santhi Ponnusamy & Detmar Meurers. 2020. *Towards automatically generating questions under discussion to link information and discourse structure.* In Proceedings of the 28th international conference on computational linguistics, Barcelona, Spain.

Pan, Liangming, Wenqiang Lei, Tat-Seng Chua & Min-Yen Kan. 2019. *Recent advances in neural question generation.* arXiv preprint arXiv:1905.08949 .

Riester, Arndt, Lisa Brunetti & Kordula De Kuthy. 2018. *Annotation guidelines for questions under discussion and information structure.* In E. Adamou, K. Haude & M. Vanhove (eds.), Information structure in lesser-described languages: Studies in prosody and syntax Studies

in Language Companion Series, John Benjamins.

Roberts, Craige. 2012. *Information structure in discourse: Towards an integrated formal theory of pragmatics.* Semantics and Pragmatics 5(6). 1–69. doi: 10.3765/sp.5.6.

Recognizing deliberate metaphors

Stefanie **Dipper**, Anna **Ehlert**, Doreen **Scholz**, Felix **Theodor**, Larissa **Weber** (Ruhr-Universität Bochum)
stefanie.dipper@rub.de, anna.ehlert@rub.de, doreen.scholz@rub.de, felix.theodor@rub.de, larissa.weber@rub.de

Metaphors are a widespread phenomenon that occurs frequently in various types of text. Metaphors involve a "mapping across two conceptual domains" (Steen, 2007): they refer to the properties of one concept in order to describe and clarify the properties of another concept. For example, in (1) the concept "skeleton" is used to refer to the function of a skeleton as a supporting structure of a body, which can be transferred to the supporting structure of buildings. In addition, the concept "skeleton" indicates that the associated body is no longer alive, otherwise the skeleton would not be visible at all and would not be able to rise up into the air. transferred to buildings this means that the buildings are destroyed.

> (1)   Skeletons of skyscrapers rose into the sky.

Many modern linguistic expressions go back to metaphors, which have now been conventionalized, however. In (2), for example, the concept "attack" is no longer associated with a warlike activity, but is directly understood as a form of argumentation.

> (2)   Lakoff attacked Glucksberg.

Especially metaphors of type (1), which we call "deliberate metaphors", pose a challenge for automatic processing, because certain expressions are not used literally or with a non- canonical meaning.

In our poster we want to present our work on the automatic recognition of deliberate metaphors. We present our annotation guidelines as well as results of a corpus of sermons currently being annotated according to these guidelines. Furthermore, we implement a recognizer based on the approach of Shutova et al. (2012), but adapting and extending it to German.

References

Shutova, Ekaterina, Simone Teufel, and Anna Korhonen. 2013. Statistical Metaphor Processing, *Computational Linguistics*, 39(2): 301–353.

Steen, Gerard J. 2007. Finding Metaphor in Discourse: Pragglejaz and Beyond. *Cultura, Lenguaje y Representación*, 5: 9–25.

Annotating and interpreting deliberate metaphors: An implementation of Steen's Five Step Method

Stefanie **Dipper**, Frederik **Elwert**, Tim **Karis** (Ruhr-Universität Bochum)
stefanie.dipper@rub.de, frederik.elwert@rub.de, tim.karis@rub.de

The main characteristic feature of metaphors is the mapping from one conceptual domain to another, with the goal of appropriately describing the concept of the target domain using the concept of the source domain. Metaphor fulfils a special role in religious language, where its capacity to express ideas about an abstract entity with reference to a well-known concrete entity works as a means to make statements about the transcendent. In (1), an extract from a religious text in Middle High German, the metaphor SALVATION IS HEALING is used to convey religious ideas: abstract theological notions such as original sin and salvation are mapped onto a more tangible domain by referring to the concepts of wounding and healing.

(1) so vnsir herre got alle die wnden virbindit die wir íe von adames svndon gefrvmeton 'Thus our Lord God binds up all the wounds we have suffered through Adam's sin.'

We distinguish between two steps in metaphor analysis: metaphor identification and metaphor interpretation. For the first steps, there are comprehensive guidelines (MIP, Pragglejaz Group 2007, and MIPV, Steen et al., 2010). For the second step, Steen (2007) proposed the 'Five Step Method'. In our poster, we present an implementation and extension of Steen's method that supports annotators in identifying and writing up explicitly stated propositions as well as implicit assumptions that are relevant and necessary to arrive at the metaphor's interpretation.

References
Pragglejaz Group. 2007. MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, 22(1): 1–39.
Steen, Gerard J. 2007. Finding Metaphor in Discourse: Pragglejaz and Beyond. *Cultura, Lenguaje y Representación*, 5: 9–25.
Steen, Gerard J., Aletta G. Dorst, J. Berenike Herrmann, Anna A. Kaal, Tina Krennmayr, and Trijntje Pasma. 2010. *A Method for Linguistic Metaphor Identification: From MIP to MIPVU*, volume 14 of Converging Evidence in Language and Communication Research. John Benjamins Publishing Company, Amsterdam/Philadelphia.

Machine learning approaches to analyzing German synthetic compounds

Carlotta J. **Hübener**, Universität Hamburg
carlotta.huebener@uni-hamburg.de

Synthetic compounding (e.g., *schönheitsliebend* 'beauty-loving') is a highly productive word-formation pattern in German (Neef 2015: 588), which can give insight into the interplay of word-formation and syntax. However, little is known about the internal argument structure of the words it yields. The poster presents an application of a machine learning model for analyzing the syntactic structure of synthetic compounds, focusing on noun-participle combinations.

There is a consensus in the literature that most synthetic compounds are based on accusative phrases (e.g., *ekelerregend* 'nauseating' ~ *Ekel*$_{ACC}$ *erregen* 'to arouse disgust'). It is unclear to what extent there is a correspondence to dative and genitive phrases as well (e.g., *zweckentsprechend* 'appropriate' ~ *Zweck*$_{DAT}$ *entsprechen* 'to correspond to the purpose'). Knowing this distribution is important, however, for instance when investigating word-formation restrictions and the interplay of grammar and the lexicon in general.

To examine the internal argument structure of synthetic compounds, the valencies of their base verbs have to be determined. For this purpose, automatic dependency parsing is advantageous: Large numbers of texts can be parsed within a short time and at low cost. The present study used a pretrained dependency parsing model from the Python library spaCy (Honnibal & Montani 2017) to identify and classify clause constituents. In a first step, a list of noun-participle combinations was extracted from the DWDS core corpus and the DIE ZEIT corpus. The corresponding noun-verb combinations were queried in the DWDS core corpus (e.g., *Kopf* 'head' and *schütteln* 'to shake' for *kopfschüttelnd* 'head-shaking'). Then, the dependency parser analyzed the syntactic dependency between noun and verb. For instance, the model identified 1,705 sentences with a syntactic dependency between the lexemes *Kopf* and *schütteln*, classifying this relation as "oa" (accusative object) in 98.3% of cases. Thus, *kopfschüttelnd* is obviously based on an accusative phrase.

With manually annotated data serving as a reference standard, the approach achieved a micro-average accuracy of 0.94 (average precision: 0.99, average recall: 0.89, average F_1 score: 0.94) for a sample of 404 noun-participle combinations. Restricted to well- attested verbal phrases in the corpus ($f > 10$), the accuracy increased to 0.97. Both the manually and the automatically annotated data confirm that most noun-participle combinations correspond to accusative phrases (99.5% or 94%, respectively). The results suggest that spaCy's dependency parser is an overall reliable tool offering promising possibilities for the further examination of synthetic compounds, for instance regarding the relationship between grammar and the lexicon.

References

Digitales Wörterbuch der deutschen Sprache. Berlin-Brandenburgische Akademie der Wissenschaften. DIE ZEIT corpus. https://www.dwds.de/d/korpora/zeit.

Digitales Wörterbuch der deutschen Sprache. Berlin-Brandenburgische Akademie der Wissenschaften. Core corpus. https://www.dwds.de/d/k-referenz#kern.

Honnibal, Matthew, and Montani, Ines. 2017. *spaCy 2. Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.*

Neef, Martin. 2015. "33. Synthetic compounds in German." In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen, and Franz Rainer, eds. *Word-formation. An international handbook of the languages of Europe.* Berlin: de Gruyter. 581-592.

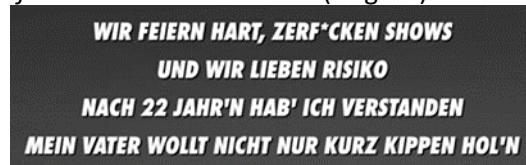# „Ich gehe kurz Zigaretten holen": Diskurs berechnen mit Word Embedding

Speaker 1 (Zakharia **Pourtskhvanidze**, Institut für Empirische Sprachwissenschaft, Goethe-Universität Frankfurt) pourtskhvanidze@em.uni-frankfurt.de

0. Grundannahme: (1) Die Diskurse lassen sich mithilfe von korpuslinguistischen Tools grundsätzlich berechnen. (Bubenhofer 2008); (2) Die Satzhypostasen sind Sprachgebrauchsmuster und indizieren die spezifischen korrespondierenden Diskurse.

1. Ontologie des Zigarettenholen-Diskurses im deutschsprachigen Gebrauch.

Heute will ich Ihnen erzählen, wie das damals wirklich war, als ich „mal eben Zigaretten holen ging" und erst 23 Jahre später an meinem Heimatort zurückkehrte… (Schottleitner).

"Ich geh mal Zigaretten holen"-Fälle. Hallo ihr! Wisst ihr, ob es wirklich so viele Männer gibt bzw. gegeben hat, die gesagt haben "ich gehe mal kurz Zigaretten holen" oder so etwas Ähnliches und dann einfach verschwunden sind? (Brigitte).

2. Problemstellung. (1) Überprüfung der Schlagwortfähigkeit der Satzhypostase „Ich-gehe- nur-mal-kurz-Zigaretten-holen", (2) Ermittlung von korrespondierenden Diskurse im korpuslinguistischen verfahren (Word Embedding).

3. „Zigarettenholen" Zwischen Word- und Phrase Embedding.

Es wurde ein Webkorpus aus ca. 1Mio. Token verwendet. Im 50-dimensionalen Tensor-Raum wurden die Vektorendaten analysiert und 14 miteinander zusammenhängende Knoten isoliert. Die verschiedenen phrastischen Versionen des Satzes wurden in einem Uni-Gram „Zigarettenholen" umgewandelt und der Vektor dieses kumulativen Uni-Grams wurde in der Analyse eingesetzt.

Durch das Netzwerk-Modell, dass aufgrund der Word Vectors Analyse entsteht, lässt sich eine allgemeine Diskurs-Klassifizierung vollziehen, in dem die folgenden Diskursschlag-wörter errechnet wurden:

1. 'warten'
2. 'geblufft'
3. 'weggehn'
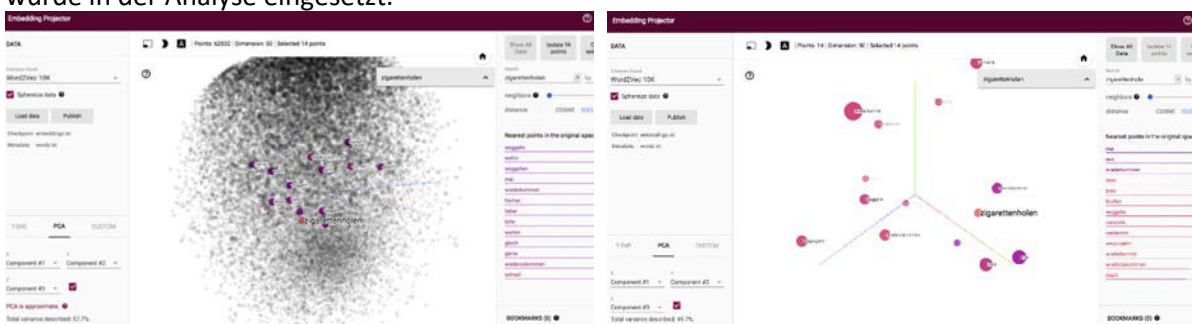
4. Ergebnis. Die Konstruktion „Ich-gehe-nur-kurz-Zigaretten-holen" hat sich mit der konzeptuellen Bedeutung „weggehen ohne sich zu verabschieden und/oder Gründe zu nennen" im Beziehungs- bzw. Liebes-Diskurs verselbstständigt. Die Konzepte wie {„weggehn", „bluff", „wieder(zu)kommen"} sind im Word-Embedding-Verfahren errechnet und stehen als korrespondierenden Diskurs-Schlagwörter.

References

Bubenhofer, Noah. 2008. Diskurse berechnen? Wege zu einer korpuslinguistischen Diskursanalyse. In: Spitzmüller/Warnke. Methoden der Diskurslinguistik: sprachwissenschaftliche Zugänge zur transtextuellen Ebene. Berlin / New York: de Gruyter, 407-434.

Schottleitner, Vera. „Männer u.s.w." Gedichte und Essays.

http://www.literaturlinie.de/Nicht_wahr.pdf. Eingesehen am 01.04.2020.

Annotating metonymic relations in a corpus-based resource for Italian verbs

Emma **Romani**, University of Pavia (Italy), Elisabetta **Ježek**, University of Pavia (Italy)
emma.romani01@universitadipavia.it, jezek@unipv.it

In this poster we address the results of a research thesis (Romani 2020) dedicated to the annotation of metonymies. Metonymy is the language phenomenon for which one referent is used to denote another referent associated with it. In our research, we investigated metonymy from a corpus-based perspective, through the analysis of corpus data and an annotation performed in T-PAS, a corpus-based resource for Italian verbs (Ježek et al.
2014). T-PAS consists in a repository of Typed Predicate Argument Structures (called *t-pas* or pattern, one for each meaning of each verb), i.e., verbal patterns with arguments signalled by semantic types, linked to manually annotated corpus instances.

The annotation of metonymies was performed starting from a list of 30 verbs contained in T-PAS. Our work was intended as an implementation of the resource; specifically, we annotated corpus instances of the verbs containing metonymies and created metonymic sub-patterns linked to them (Fig. 1). We followed a corpus-based methodology, which was also devised to distinguish metonymies from complex types (Ježek & Vieu 2014).

Fig. 1. Metonymic sub-pattern for t-pas 1 of the verb *bere* ('to drink') in T-PAS

We also conceived a theoretical framework to represent the metonymies found through the corpus analysis, by designing a map and by compiling a list of the metonymic *relations* occurring in the verbal patterns (in case of acceptance, the map and the list will be included in the poster). A *relation* is a brief description that illustrates how the *metonymic semantic type* is connected to the *target semantic type*; for example, [Container] (metonymic semantic type) 'contains' (the *relation*) [Beverage] (target semantic type). Both the map and the list depict the complexity and variety of the phenomenon, in terms of number of possible metonymic relations and of the semantic types interested.

In future perspectives, we intend to enrich the map and the list with new relations by extending the number of verbs investigated and to evaluate the annotation procedure. We are also interested in a crosslinguistic comparison of our results with those in the Croatian sister project of T-PAS (CROATPAS, Marini & Ježek 2020). The annotated corpus data, as well as the relations, will be useful for automatic detection of metonymies (Markert & Nissim 2009). To our knowledge, little work has been done on this for Italian language: it will be therefore intriguing to test our data in NLP tasks.

References

Ježek, Elisabetta, Bernardo Magnini, Anna Feltracco, Alessia Bianchini, and Octavian Popescu. 2014. "T-PAS: A resource of corpus-derived Types Predicate-Argument Structures for linguistic analysis and semantic processing". *Proceedings of LREC*. 890-895.

Ježek, Elisabetta, and Laure Vieu. 2014. "Distributional Analysis of Copredication: Towards Distinguishing Systematic Polysemy from Coercion". *Proceedings of the First Italian Conference on Computational Linguistics*. 219-223.

Marini, Costanza, and Elisabetta Ježek. 2020. "Annotating Croatian Semantic Type Coercions in CROATPAS". *Proceedings of the 16th Joint ACL-ISO Workshop ISA-16*, pp. 49–58.

Markert, Katia, and Malvina Nissim. 2009. "Data and models for metonymy resolution". *Lang Resources & Evaluation*. 43: 123–138.

Romani, Emma. 2020. *Searching for Metonymies in Natural Language Texts. A Corpus-based Study on a Resource for Italian Verbs*. BA Thesis, Pavia: University of Pavia.

Easy and reproducible WebAnno project management

Adam **Roussel**, Ruhr-Universität Bochum
roussel@linguistics.rub.de

In large annotation projects and in educational settings, you may need to create a large number of WebAnno (Eckart de Castilho et al., 2016) projects and/or user accounts at once, which can be tedious and time-consuming to do by hand. And, especially where it is important that the projects be configured in a particular way, it can also be error-prone to do everything with the graphical interface.

With this poster I'd like to introduce PyWebAnno, a Python script that helps you orchestrate collections of WebAnno projects and users. You can generate a large number of user accounts, notify these users of their login data by email, assign them automatically to WebAnno projects, and then, when the course has concluded, remove the generated projects and users – but only the right projects and users, leaving other projects on your WebAnno instance untouched. With PyWebAnno, you can also specify the documents that should belong in each project or have these assigned automatically. Finally, with the facility of uploading annotations to the generated projects, you can use the Curation function to compare the students' annotations with a gold-standard or automatically-generated annotations.

PyWebAnno is free and open source software available at:

https://git.noc.rub.de/ajroussel/pywebanno

References

Eckart de Castilho, R., Mújdricza-Maydt, É., Yimam, S.M., Hartmann, S., Gurevych, I., Frank, A. and Biemann, C. (2016): A Web-based Tool for the Integrated Annotation of Semantic and Syntactic Structures. In Proceedings of the LT4DH workshop at COLING 2016, Osaka, Japan

Differences between German and English text simplification

Regina **Stodden**, Heinrich-Heine-University Düsseldorf
regina.stodden@uni-duesseldorf.de

Text simplification is a natural language processing task which aims at automatically reducing the complexity of a given text. This research area is part of natural language generationand (monolingual) machine translation. Text simplification focus on generating a more easily understandable version of a complex text for language learners or low literacy. The simplified text should preserve the meaning of the complex text and should not contain grammatical errors (Alva-Manchego et al., 2020). So far, text simplification research mostly focuses on English (see Alva-Manchego et al. (2020) for an extensive list), and only a few studies exist for German (Klaper et al., 2013; Battisti et al., 2020; Mallinson et al., 2020).

The German text simplification research can nowadays benefit from an active community in easy-to-read German, including translation offices related to practices and research facilities related to theory. Two main versions exist of German easy-to-read languages, i.e., plain language (de: "Einfache Sprache") and easy language (de: "Leichte Sprache") (Maaß, 2020). Plain language seems more applicable to text simplification than easy language because the overall variant and its complexity are closer to everyday German. In a content analysis of recommendations on how to write German plain language and text simplification research papers, we found items that are more relevant in English than German and vice versa. These items specify the transformations during a simplification, e.g., substituting complex words or deleting superfluous information.

Both areas agree on deleting or replacing complex words and sentence splitting. In comparison to easy-to-read English, German plain language focuses more on compound splitting and compound segmentation. Furthermore, German plain language recommendations contain more frequent changes in the verb's voice, deletions of phrases and clauses, and explanations of complex words in a new sentence. In contrast, in text simplification research, sentence reordering is mentioned more often than in German plain language.

On the poster, we will explain more briefly text simplification and the differences between German plain language and German easy language. Furthermore, we will present text simplification transformations that are specific for German and English and give examples for the transformations in both languages.

References

Alva-Manchego, Fernando, Carolina Scarton, and Lucia Specia. 2020. "Data-driven sentence simplification: Survey and benchmark." *Computational Linguistics, 46(1)*:135–187.

Battisti, Alessia, Dominik Pfütze, Andreas Säuberli, Marek Kostrzewa, and Sarah Ebling. 2020. "A corpus for automatic readability assessment and text simplification of German." In *Proceedings of the 12th LREC*, pages 3302–3311, Marseille, France. ELRA.

Klaper, David, Sarah Ebling, and Martin Volk. 2013. "Building a German/simple German parallel corpus for automatic text simplification." In *Proceedings of the 2ndWorkshop on Predicting and Improving Text Readability for Target Reader Populations*, pages 11–19, Sofia, Bulgaria. ACL.

Maaß, Christiane. 2020. "Easy Language - Plain Language - Easy Language Plus. Balancing Comprehensibility and Acceptability." *Easy – Plain – Accessible*. Frank & Timme, Berlin.

Mallinson, Jonathan, Rico Sennrich, and Mirella Lapata. 2020. "Zero-shot cross-lingual sentence simplification." In *Proceedings of EMNLP 2020*, pages 5109–5126, Online. ACL.

Cross-Lingual Word Embeddings for extremely low-resource languages: Improving bilingual lexicon induction for Hiligaynon

Speaker 1 (Mary Ann **Tan**), Co-Author 1 (Dario **Stojanovski**, CIS LMU), Co-Author 2 (Alexander **Fraser**, CIS LMU)
anntanp@gmail.com, dario@cis.lmu.de, fraser@cis.lmu.de

Cross-Lingual Word Embeddings (CLWEs) have been experiencing a surge in popularity in the past couple of years due to the remarkable progress in machine learning techniques, the availability of large natural language processing (NLP) datasets and the exponential growth in computing power. CLWEs represent words from several languages in a shared embedding space; a more standard bilingual representation is called Bilingual Word Embeddings (BWEs). This research area has gained traction in the field of machine translation (MT) primarily because of its application to the task of Bilingual Lexicon Induction (BLI), which uses BWEs to learn word-pair translations with no or little supervision.

However, as with most research areas in NLP, progress is mostly limited to resource-rich Indo-European languages. Recent work on English (EN) and Hiligaynon (HIL), an extremely low-resource language and the 4th most spoken native language in the Philippines (10 million speakers), did not manage to produce BWEs of reasonable quality primarily due to a lack of a sizable monolingual corpus (Michel et al., 2020).

Mapping-based approaches to CLWEs have prevailed due to their simplicity, computational tractability and relaxed data requirements (Mikolov et al., 2013; Faruqi and Dyer, 2014; Dinu et al., 2015; Lazaridou et al., 2015; Xing et al., 2015, Artetxe et al., 2016). This approach requires only two (2) monolingual word embeddings (MWEs), pre-trained separately on large unannotated monolingual corpora, and a seed lexicon containing word pairs from the source and the target language. Its objective is to project the word embeddings of the source MWEs to the embedding space of the target MWEs by learning a transformation matrix using the seed lexicon as its bilingual supervision.

Previous studies on low-resource languages achieved zero or close to zero precision-at-1 (P@1) with EN-HIL (Michel et al., 2020), and a collection of other non-heterogeneous BWEs trained on 5M token corpora (Dyer, 2019). In this study, we showed that EN-HIL BWEs, trained on a target corpus containing just a little over 1M tokens, yielded a BLI performance of P@1 at 9.26%. This was achieved by adapting an iterative orthogonal mapping with generative adversarial approach (Conneau et al., 2018), by properly curating the seed lexicon and by employing resource-rich languages as pivots for transfer learning. The pivot languages used for our experiments were two (2) Philippine languages, Filipino and Cebuano, another Austronesian language, namely Bahasa Indonesia, and Spanish, a major source of foreign loan words in Hiligaynon (Kaufmann, 1934). Among the pivot languages used, Spanish performed best due to the high quality of its MWEs.

…

References

Artetxe, Mikel, Labaka, Gorka, and Agirre, Eneko. 2016. "Learning principled bilingual mappings of word embeddings while preserving monolingual invariance." In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 2289–2294, Austin, Texas. Association for Computational Linguistics.

Conneau, Alexis, Lample, Guillaume, Ranzato, Marc'Aurelio, Denoyer, Ludovic, and Jégou, Hervé. 2018. "Word translation without parallel data." In Proceedings of the 6th International Conference on Learning (ICLR 2018).

Dinu, Georgiana, Lazaridou, Angeliki, and Baroni, Marco. 2015. "Improving zero-shot learning by mitigating the hubness problem." In Proceedings of ICLR (Workshop Track).

Dyer, Andrew R. 2019. "Low supervision, low corpus size, low similarity! Challenges in cross-lingual alignment of word embeddings: An exploration of the limitations of cross-lingual word embedding alignment in truly low resource scenarios." Master's Thesis, Department of Linguistics and Philology, Uppsala University.

Faruqui, Manaal and Dyer, Chris. 2014. "Improving vector space word representations using multilingual correlation." In Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, pages 462–471, Gothenburg, Sweden. Association for Computational Linguistics.

Kaufmann, John. 1934. "Visayan-English Dictionary." CreateSpace Independent Publishing Platform.

Lazaridou, Angeliki, Dinu, Georgiana, and Baroni, Marco. 2015. "Hubness and pollution: Delving into cross-space mapping for zero-shot learning." Association for Computational Linguistics, volume 1 (Long Papers), pages 270–280.

Michel, Leah, Hangya, Victor, and Fraser, Alexander. 2020. "Exploring bilingual word embeddings for hiligaynon, a low-resource language." In Proceedings of the 12th Conference of Language Resources and Evaluation (LREC), pages 188-193. European Language Resources Association.

Mikolov, Tomas, Le, Quov. V., and Sutskever, Ilya. 2013. "Exploiting similarities among languages for machine translation." https://arxiv.org/abs/1309.4168

Xing, Chao, Wang, Dong, Liu, Chao, and Lin, Yiye. 2015. "Normalized word embedding and orthogonal transform for bilingual word translation." In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 1006–1011, Denver, Colorado. Association for Computational Linguistics.

A new system for rewriting linguistic annotations

Mark-Matthias **Zymla**, University of Konstanz
Mark-Matthias Zymla@uni-konstanz.de

We present a new system for expanding and rewriting linguistic annotations. As an example,we focus on the application of this system to syntactically annotated data. We showcase how the system can be used to add semantic annotations to some (syntactic) input and how it can be integrated into an annotation pipeline to produce semantic representations.

The present system was primarily inspired by the packed rewrite system (PRS) con- tained in the Xerox Linguistics Environment (XLE; Crouch et al. (2017)). The PRS has been successfully used to implement large-scale semantic parsing and even semantic reasoning (Bobrow et al., 2007). However, the system is not supported by XLE anymore and is not publicly available. We provide a new take on the PRS that aims to make the system available and useful to a broader audience. For this, inspiration is drawn from recent work in linguistic annotation. Ide and Bunt (2010) pointed out that linguistic annotations share a common core that can be expressed in terms of a graph-based format. The present system makes use of this by employing simple interfaces that translate syntactic annotations, which are acquired either via parsing with XLE or with Universal Dependencies, into such abstract graph representations.
These can then be modified by using, in principle, simple rewrite rules.

Rewrite rules consist of a query expression that serves to identify sub-graphs in a given annotation and an expansion graph that specifies the information that is added to the input provided that the query matches. Rules are specified in terms of a fact notation where a fact consists of a mother node, an attribute/relation, and a value/daughter node. Nodes are identified via variables, while attributes/relations are arbitrary strings without white spaces.

By engineering the output appropriately, it can be directly fed into further processing steps. We show this in terms of a syntax/semantics interface and a semantic interpretation component which produces semantic representations based on Glue semantics.

In summary, we present a system for expanding and rewriting linguistic annotations that can be applied to a wide array of linguistic resources given a simple translation interface,inspired by the ideas of Ide and Bunt (2010). Previous work on the PRS contained in XLE has shown that such a system can find a wide array of creative uses and opens up new possibilities for using formal computational methods in NLP. We concretely show this by presenting a syntax/seman- tics interface implemented with the system presented here. Since the system is implemented in a micro-service architecture, it can be easily integrated into linguistic annotation pipelines.

References

Bobrow, Daniel G., Bob Cheslow, Cleo Condoravdi, Lauri Karttunen, Tracy Holloway King, Rowan
Nairn, Valeria de Paiva, Charlotte Price, and Annie Zaenen. 2007. PARC's
Bridge
and Question Answering System. In *Proceedings of the GEAF 2007 Workshop*. 1–22.
Crouch, Richard, Mary Dalrymple, Ronald M. Kaplan, Tracy Holloway King, John T. Maxwell III, and Paula
Newman. 2017. *XLE Documentation*. Palo Alto Research Center.

Ide, Nancy and Harry Bunt. 2010. Anatomy of Annotation Schemes: Mapping to GrAF. In *Proceedings of the Fourth Linguistic Annotation Workshop*. 247–25.