The Fourth Annual

**FORMAL WAYS OF ANALYZING VARIATION**

29-30 June, 2017

Kings Manor, the University of York

**SPONSORED BY**

The University of York

LAGB
FWAV4 Programme

Venue: Huntingdon Room, King’s Manor, York

Thursday, 29 June 2017

08:00 – 08:45 Registration

08:45 – 09:00 Opening remarks

09:00 – 10:00 Plenary, Anton Karl Ingason (Chair: Susan Pintzuk)

*Spaghetti-shaped curves*

10:00 – 10:30 Coffee break

10:30 – 12:30 Session 1 (Chair: Betsy Sneller)

10:30 – 11:10 Luke Adamson

*Non-categorical Grammatical Conditioning of Past Tense Allomorphy*

11:10 – 11:50 Theresa Biberauer, Juliane Bockmuehl, Erika Herrmann, and Sheena Shah

*Imperative variation: the case of Afrikaans, Namibian German and Kroondal German*

11:50 – 12:30 Ingunn Hreinberg Indriðadóttir

*Heavy NP Shift in Icelandic and Faroese*

12:30 – 13:30 Lunch (provided on site)

13:30 – 14:50 Session 2 (Chair: Kajsa Djärv)

13:30 – 14:10 Theresa Biberauer

*Word-order variation and change in systems that maximize minimal means*

14:10 – 14:50 Henri Kauhanen, Caroline Heycock, and Joel C. Wallenberg

*Grammar Competition in Neutral Learning: A Reply to Han et al. (2016)*

14:50 – 15:20 Coffee break

15:20 – 17:20 Session 3 (Chair: Luke Adamson)

15:20 – 16:00 Betsy Sneller

*Rules with exceptions: Using the Tolerance Principle to diagnose allophones*

16:00 – 16:40 Jordan Kodner

*The Importance of Population Structure in Models of Language Change*

16:40 – 17:20 George Bailey

*Synchronic evidence for diachronic pathways of change: /g/-deletion and the life cycle of phonological processes*

18:00 – 20:00 Reception
Friday, 30 June 2017

08:30 – 09:00  Additional registration

09:00 – 10:00  Plenary, Nicole Holliday (Chair: Caitlin Light)
  Intonational Variation, Pragmatics, and Social Meaning: Critical Evidence from African American Language

10:00 – 10:30  Coffee break

10:30 – 12:30  Session 4 (Chair: Caroline Heycock)
  10:30 – 11:10 Lauren Ackerman and Joel C. Wallenberg
  Categorization of gender, modulated by experience, can constrain coreference
  11:10 – 11:50 Spencer Caplan and Kajsa Djarv
  Embedded V2 is anti-licensed by discourse familiarity: Quantitative and statistical evidence
  11:50 – 12:30 Göz Kaufmann
  Complex Verbal Heads in the Left Periphery of Pomeranian

12:30 – 13:30  Lunch (provided on site)

13:30 – 14:50  Session 5 (Chair: Joel Wallenberg)
  13:30 – 14:10 David Willis
  Modality and generalization of negative concord in northern Welsh dialects
  14:10 – 14:50 Katalin Gugan
  Variation and change in Hungarian negative sentences

14:50 – 15:20  Coffee break

15:20 – 17:20  Session 6 (Chair: David Willis)
  15:20 – 16:00 Bronwyn Bjorkman and Sali Tagliamonte
  What can explain it? The changing models of York English
  16:00 – 16:40 Amie Dejong
  Positive Anymore: Lexicalization and (De)grammaticalization
  16:40 – 17:20 Robert Truswell and Nikolas Gisborne
  Competing denotations in the diachrony of English wh

17:20 – 17:30  Closing remarks

19:30 – 21:30  Dinner at Red Chilli
  21-25 George Hudson St, YO1 6JL
  (0)1904 733668
Organizers

Caitlin Light  clight9@gmail.com
Susan Pintzuk  susan.pintzuk@york.ac.uk

Invited Speakers

Nicole Holliday  nicole.holliday@pomona.edu
Anton Karl Ingason  antoni@hi.is

List of Participants

Lauren Ackerman  lauren.ackerman@ncl.ac.uk
Luke Adamson  adamsonl@sas.upenn.edu
George Bailey  george.bailey@manchester.ac.uk
Theresa Biberauer  samtb23@gmail.com
Bronwyn Bjorkman  bronwyn.bjorkman@queensu.ca
Monja Burkard  monja.burkard@univie.ac.at
Spencer Caplan  spcaplan@sas.upenn.edu
Amie DeJong  acdejong@uw.edu
Kajsa Djarv  kdjarv@sas.upenn.edu
Katalin Gugán  gugan.kati@gmail.com
Caroline Heycock  caroline.heycock@gmail.com
Ingunn Indriðadóttir  ihreinberg@gmail.com
Göz Kaufmann  goez.kaufmann@germanistik.uni-freiburg.de
Jordan Kodner  jkodner@sas.upenn.edu
Betsy Sneller  betsysneller@gmail.com
Robert Truswell  rob.truswell@ed.ac.uk
Keith Tse  keith.tse@balliol.oxford.com
Joel Wallenberg  joel.wallenberg@gmail.com
David Willis  dwew2@cam.ac.uk
Abstracts
Spaghetti-shaped curves
Anton Karl Ingason

Linguistic change often involves a regular s-curve and it is common to assume that the s-shaped curve reflects some deep reality about the ways in which the use of linguistic variants spreads. As an example of the status of the mathematically well-behaved curve, Yang's Variational Model of language acquisition can be applied to calculate the expected usage trajectory of an ongoing change into the future. This method has, for example, been used to predict how Icelandic children will acquire and use an innovative new passive construction until the year 2050.

However, not all change is that regular and in this talk I will focus on aspects of individual lifespan change that highlight a certain type of case where the assumption of a neat curve is not valid. While individual lifespan change is sometimes regular, such regularity is really caused by systematic reactions of speakers to changes in the community. I will thus distinguish community-specific lifespan change from individual-specific lifespan change and show that when the factors that shape usage over time cannot be generalized over communities, usage patterns can change rapidly in reaction to idiosyncratic events in the life of a speaker.

I present a study of the use of Icelandic Stylistic Fronting in the parliament speeches of Steingrimur Sigfusson, former finance minister of Iceland. This study comes from joint work with Lilja Björk Stefánsdóttir. Stylistic Fronting is an optional syntactic movement process associated with formal style and we tracked its use in 8005 tokens collected from 23 continuous years of a political career. At first sight, the usage rate over time seems to fluctuate irregularly, but two crucial aspects of our methodology allow us to account for its development. First, the high time resolution of the study allows us to see when exactly important changes take place. Second, by carrying out a thorough interview with the research subject, we were able to enhance the quantitative findings with the qualitative depth needed to understand what was happening in any given year.

One conclusion of this work is that studies of individual-lifespan change demand a high time resolution as well as an emphasis on qualitative depth because individual-specific factors can overrule systematic reactions to community-specific factors. Another one is that while simple s-curves are to be expected when the crucial parts of reality generalize over communities, individuals are sometimes more complicated than communities, giving rise to spaghetti-shaped curves.
Intonational Variation, Pragmatics, and Social Meaning: Critical Evidence from African American Language
Nicole Holliday

Over the past 50 years, great strides have been made towards more comprehensive understandings of ethnolinguistic, pragmatic, and intonational variation, but research at the intersection of these subfields is still largely absent from modern sociolinguistics. This research aims to address that lacuna by examining the ways in which the meanings of intonational contours appear to differ between two varieties of American English, Mainstream U.S. English (MUSE) and African American Language (AAL). By challenging previous theories of the discourse-pragmatic meaning of various intonational contours, which have largely focused on white-speakers of MUSE to the exclusion of all others, this research expands our understanding of the nature of the intonational phonology of English. I will demonstrate that AAL may accomplish pragmatic functions related to declarative statements and polar questions in manners not predicted by earlier MUSE-based theories of the meaning of specific intonational contours. These results also illuminate the ways in which the assumptions of these earlier theories may be challenged by empirical data from AAL and other varieties, encouraging us to seek a more sociolinguistically-informed understanding of the variation possible in English intonation. The research also sheds further light on the ways in which interactions between speakers of MUSE and speakers of AAL may present challenges similar to those in other cross-cultural communication situations. Finally, I will highlight the social effects and consequences for speakers who do not participate in MUSE intonational patterns, especially those which can include difficulty interacting with hegemonic power systems such as education, criminal justice, and capitalistic markets. Via this analysis and discussion, I hope to highlight the ways in which an understanding of intonational variation is crucial for addressing issues of linguistic profiling, discrimination, and inequality.
Categorization of gender, modulated by experience, can constrain coreference
Lauren Ackerman (lauren.ackerman@ncl.ac.uk) and Joel Wallenberg, Newcastle University

English does not have grammatical gender, but it encodes something about conventional or natural gender in 3rd person singular pronouns, names, and gendered nouns (Baron, 1971; Barry & Harper, 1982, 1993, 2001, Kennison & Trofe, 2003, a.o.). Yet, natural gender is not biologically or socially binary, which is increasing in public awareness (Atkinson et al., 2017; Cheshire, 2002; Eckert, 2014; Gross, 2016; Gustafsson Sendén et al., 2015; Miltersen, 2016). This study investigates how individual differences modulate the linguistic processing of gender. We suggest that while the categorical representation of gender is difficult to alter, it interfaces actively with real-world knowledge of gender variation, and so affects sentence acceptability.

Linguistically encoded natural gender is used by psycholinguists to probe coreference dependency formation via the “gender mismatch effect” (GMME). A GMME, inferred from a slowdown in reading time, is detected when a ‘gendered’ name or noun is grammatically accessible for coreference with a pronoun of mismatched gender (Kazanina et al., 2007; Sturt, 2003; van Gompel and Liversedge, 2003, a.o.).

(1) # John, seemed to be upset with herself.

In (1), herself is linked to John by means of a grammatically constrained coreference relation, so readers struggle to link these words once the gender is determined to mismatch, herein indicated by ‘#’ (Chomsky, 1981; Foertsch and Gernsbacher, 1997; Sanford and Filik, 2007; van Gompel and Liversedge, 2003). However, natural gender is not grammatically constrained, demonstrated by specific pragmatic contexts, as in (2).

(2) ✓ At the Halloween party, the cowgirl, left his laasso prop in the kitchen.

As nonbinary genders become more widely acknowledged, understanding how natural or conventional gender is represented by the linguistic processing mechanism will be crucial for investigating processing phenomena that use GMMEs to detect grammatical phenomena.

Seventy-two native speakers of English were recruited on Amazon Mechanical Turk to rate the ‘naturalness’ of sentences on a 7-point scale, provide implicit and explicit measures of their gender identity/expression (BSRI; Bem, 1981a,b) and other demographic information. Of the demographic data collected, the strongest predictor of acceptability ratings for mismatched name+pronoun gender was age ($r=-0.33$, $p<0.01$), with younger participants rating mismatched pairs higher than older participants (Estimate=0.06, SE=0.01, LR.stat(1)=26.1,$p<0.001$). Neither age nor quantitative measures of gender correlated with the ratings of non-mismatched pairs ($p<0.1$).

We propose that the representation of natural gender in English pronouns does not contain formal syntactic features, but instead is categorized in a conceptual space based on real-world experience. The interaction between the representation of natural gender and the processing of coreference relations can vary based on the way the continuous dimension of gender is categorized. Our results suggest that older speakers and younger speakers have slightly different category means based on their differing experience of gender variation. Both still effectuate a linguistic binary due to their bimodal categorization of continuous gender, but the differences in overlap between “male” and “female” categories produce different acceptability judgments in a stochastic manner. Thus, older and younger speakers show the same broad patterns of grammaticality, but have a significant difference in their acceptance of non-conventional name-gender pairs.
References:


English past tense allomorphy is one of the best studied morphological phenomena, being widely discussed in psycholinguistics, in acquisition, and in morphology. One of the principal questions surrounding tense allomorphy concerns its conditioning factors: what kind of information matters for the choice between one form or another? The current study identifies two determinants of the past tense which have not received much attention, namely event semantics and Aktionsart. These two determinants are shown to impact the choice between doublet forms such as burned/burnt and weaved/wove. The analysis offered here integrates this choice into the grammar by appealing to operations applying variably in a postsyntactic module. The most common approach to doublet forms addresses the question of when the irregular form is unavailable, instead giving way to a regular form. Previous research has identified several factors that prohibit irregular past tense morphology (e.g. Pinker and Prince 1989). For example, true denominal verbs never allow irregular allomorphy (e.g. fled rather than flew). This prohibition is derivable in modern incarnations of Distributed Morphology. In Embick 2010’s formulation, for instance, the presence of both n and v heads ensures that the Tense morpheme is in a distinct Spell Out domain from a lexical root, making it impossible for Tense to be allomorphically sensitive. This gives rise to doublet forms even when the root would normally take irregular allomorphy, as in spat versus spitted the canvas up (applied spit to the canvas). While denominalization categorically bleeds irregular past tense allomorphy, other conditioning factors appear to have variable effects. I show in the current study that event semantics is relevant to the allomorphic calculus. For example, weaved/wove does not have uniform frequency distributions across different environments. The verb weave may be used either in its manner of motion sense (he weaved his way across the room) or its creation sense (he weaved a basket). While variation between weaved and wove is attested for both senses, frequencies in COCA suggest that the creation use greatly favors the irregular form (“weaved a” 6, “wove a” 57), whereas the manner of motion favors the irregular form to a lesser degree (“wove through” 69, “weaved through” 44; “wove [possessive pronoun] way” 56; weaved [possessive pronoun] way 43). Huang and Pinker (2010) also found from their experimental results that Aktionsart impacted speakers’ acceptability judgments for forms such as burned and burnt, the former being preferred to a greater degree in accomplishment environments where duration of the event is extended. Preliminary corpus work supports their findings that Aktionsart has an effect on tense allomorphy.

I propose that the frequency distributions can be accounted for by appealing to pruning operations (Embick 2010) with variable application. Embick 2010 and other work in Distributed Morphology assumes that linear adjacency is a necessary condition on two morphemes interacting for allomorphy, and that pruning of unpronounced material is necessary to bring together two morphemes that are not structurally adjacent to each other. In the case of weave, a pruning rule must prune a v head for the root to be adjacent to T. Failure to prune this head results in a configuration in which the root and T cannot interact for allomorphy. I adopt a flavors of v type of analysis (Folli and Harley 2005) to distinguish v(do) from v(create) (cf Levinson 2010), which distinguish the manner of motion use from the creation use. Crucially, the two heads are subject to different pruning operations that apply at different rates in the context of the root weave. The same analysis can extend to the Aktionsart cases, for which pruning operations can target an Inner Aspect head, which applies at different rates depending on its features.

Synchronic evidence for diachronic pathways of change: /ɡ/-deletion and the life cycle of phonological processes

GEORGE BAILEY - UNIVERSITY OF MANCHESTER

North western varieties of British English are unique in their lack of ng-coalescence, with surface forms that variably retain voiced velar nasal+stop clusters, e.g. sing /sɪŋɡ/ and swimming /swɪŋɡ/. In this paper, I claim that such variation stems from probabilistic application of a rule that deletes post-nasal /ɡ/ in syllable codas; using data from 24 sociolinguistic interviews conducted in the north west of England, modelled using mixed effects logistic regression, I show how quantitative analysis of its patterning indicates sensitivity to morphophonological structure in ways that are predicted by the life cycle of phonological processes (Bermúdez-Otero 2013). In doing so, I provide the first quantitative account of how synchronic variation in /ɡ/-deletion reflects the rule’s diachronic trajectory along the life cycle through centuries of linguistic change.

Assuming a modular feedforward architecture of grammar in which the phonological component is split into stem-, word-, and phrase-level strata, it is shown that the surface probability of /ɡ/-deletion is a function of the number of cycles in which the criteria for deletion is met, determined by the morphophonological environment in the following way:

(i) 1 chance to apply: word-internal, pre-vocalic in polymorphic words, e.g. singer
   • /ɡ/ syllabified as onset at word-level
(ii) 2 chances to apply: word-final, pre-vocalic, e.g. sing it
    • /ɡ/ resyllabified as onset at phrase-level
(iii) 3 chances to apply: word-final, pre-consonantal, e.g. sing verses
    • /ɡ/ invariably in coda position

Interestingly, /ɡ/-deletion is strongly inhibited in pre-pausal tokens in ways that are not predicted by a purely cyclic account. I provide suggestive evidence that this environment is involved in generational change within the community, with younger speakers reanalysing the pre-pausal category as one that favours /ɡ/-presence. The variability of pre-pausal environments with respect to their effect on /ɡ/-deletion mirrors that of similar lenition process, e.g. /s/-debuccalisation in South American varieties of Spanish (see Kaisse 1996) and /td/-deletion in varieties of English (see Santa Ana 1991 on Chicano English and Bayley 1994 on Tejano English).

I also provide evidence from lab speech, in which tokens of /ŋɡ/ are elicited before varying prosodic and syntactic boundary strengths à la Sproat & Fujimura 1993, which suggests that this pre-pausal retention of /ɡ/ seems to be a categorical effect of prosodic phrasal position rather than a gradient effect of rime duration by virtue of pre-boundary lengthening. Ongoing work aims to investigate the pre-pausal conversational tokens more closely to determine whether the same prosodic patterns are upheld in more naturalistic conditions.

This study lends quantitative support to the life cycle of phonological processes and adds to a growing body of knowledge regarding the behaviour of probabilistic lenition processes in pre-pausal environments.
References


This paper revisits the question of “basic” vs. “marked” word-order from an emergentist generativist approach to syntactic variation and change ([1] Biberauer 2011 et seq.). Our point of departure is Chomsky’s (2005) Three Factors model (1). Here UG is minimally specified, crucially lacking pre-defined parametric options like those traditionally assumed during the classic parametric era, e.g. the classic Head Parameter. Acquirers are assumed to tap into systematic departures from Saussurean arbitrariness to establish the identity of the formal features defining their target grammar. More specifically, we harness Chomsky’s (1995) semantic [S], phonological [P] and formal [F] features to distinguish “basic” vs “higher” degrees of arbitrariness in human language: (i) lexically stored, idiosyncratic conventionalized sound-meaning mappings involving [P]- and [S]-features, and (ii) grammatically regulated and thus more systematic conventionalized sound-meaning mappings requiring the postulation of [P]- and/or [S]- and [F]-features. Systematic departures in the PLD from one-to-one form-meaning mappings like those in (2) signal the presence of [F]s, i.e. grammaticalised [S] and/or [P]-features). Crucially, the postulation of [F]s is constrained by the general cognitive bias to Maximise Minimal Means (MMM), which, in the domain of language, underlies the acquisition biases in (3-4).

This has consequences for our understanding of word-order acquisition, variation and change. First, “basic” word order is acquired very early (Wexler 1996), and much work points to prosodic (i.e. exclusively [P]-type) properties, to which neonates are highly attuned, being key. Minimal means thus provide the stepping-stone into grammar. Since the linearization convention underlying “basic” word order instantiates a systematic idiosyncrasy, an [F] must be postulated to capture it; here: [I]ntial and [L]ast. Once present, MMM predicts that [I/L] will be the basis for further grammar structuring. Given (3-4), we expect ‘categorial harmony’ effects (Hawkins 1983), and, more generally, monotonicity effects within Extended Projections (e.g. all-[I/L] clausal or nominal projections, or all hierarchically contiguous heads within a given domain sharing [I/L]), with predictable opportunities for ‘disharmony’ (Biberauer 2017): monotonicity effects follow where properties are distributed across natural classes (4), with larger classes requiring the postulation of fewer [F]s (3); see (5-6). As “larger” (more macro) choices require fewer [F]s, we expect greater stability, which seems true: rigid head-finality seems very stable, while West-Germanic-style OV is less so (Biberauer & Roberts 2012). Change in the direction of “smaller” (more micro) choices is not ruled out, however; the expectation is that such change will reference [F]s already present in the system to account for other regularities. OV-loss/-restriction, for example, often seems to involve states where OV is limited to [negative]-, [topic]- and/or [focus]-marked objects or, in systems with prominent aspect-/tense-based structuring, to [aspect]-/[tense]-defined domains.

MMM further leads us to expect that the fixing of “basic” word order (e.g. a specific type of VO (V-in-situ, V-to-T, etc.) or OV (“rigid” OV, West Germanic-type OV, OVX-type OV)) will provide a reference point in relation to which other ordering patterns (e.g. (V)-to-T-to-C) may be interpreted. This is essentially “higher-level” duality of patterning (Hockett 1960), in terms of which meaningless “basic” word-order choices (OV vs VO), which are fixed early, pave the way for other meaningless obligatory-filling choices (e.g. V spellout position, Spec-TP, Spec-CP, etc.), and, crucially, also meaningful optional movements (e.g. T-to-C in English, the nature of the XP that raises to Spec-CP, etc.). Since (i) “basic” OV/VO (head-directionality) effectively “comes for free” via prosody in systems of the kind investigated by Mehler and colleagues, and (ii) higher-level duality of patterning permits a grammar to maximize the contribution of both the “virtually conceptually necessary” components of syntactic structure-building – Lexical Items and (External and Internal) Merge – it is difficult to sustain the idea of a grammar with truly “mixed”, i.e. “basic” order-eschewing OV/VO of the sort quite commonly assumed during the classic parametric era for systems like earlier English. Instead, we expect “mixed” OV/VO systems to exhibit [F]-regulated OV/VO patterning, with one of these being “basic”. Drawing on both historical and contemporary examples, I show that this seems to be correct. Looking beyond these, arguably not maximally “mixed” systems, I also briefly show for representative ‘free word-order’ systems that the MMM prediction that “basic” OV/VO/head-directionality should be fixed seems to be borne out. The potential for variation in the domain of “basic” and “marked” word order thus seems to be considerable, but nevertheless constrained.
(1) UG (=Factor 1) + PLD (=Factor 2) + Factor 3 → Steady-state Grammar/I-language

(2) a. **Doubling**: “extra” form (2/multiple forms and 1 meaning, 1 form with no meaning, etc.)
   – e.g. agreement, concord, and expletive phenomena.
   b. **Silence**: “missing” form (no form, but meaning) – e.g. null exponentence, null arguments, null complementisers, ellipsis, OCP effects (and thus also their failure).
   c. **Multifunctionality**: 1 form, multiple meanings – e.g. systematic homophony (cf. also Wiltschko 2014, Duffield 2013, 2014).

(3) **Feature Economy** (adapted from Roberts & Roussou 2003): Postulate as few features as possible to account for the regularities in the PLD.

(4) **Input Generalization** (adapted from Roberts 2007): maximise use of postulated features.

(5) FE- and IG-regulated formal-feature learning hierarchy: NO>ALL>SOME
   Is [F] present?
   - NO: YES: ALL heads?
   - YES: NO: Which subset of heads? (SOME)
   (postulate new [F])

(6) NO>ALL>SOME applied in the domain of word order
   Is head-final present?
   - NO: All head-initial
     - YES (=ALL): present on all heads?
     - YES: All head-final
       - NO: present on all [+/-V]heads? (SOME)
       - YES: Consistently head-final
down
   - ever more specific SOME
   options
**Imperative variation: the case of Afrikaans, Namibian German and Kroondal German**

Theresa Biberauer, Juliane Bockmuehl, Erika Herrmann, and Sheena Shah

That positive and negative imperatives frequently exhibit quite different syntax has often been noted (see Isaac 2015 for overview discussion). The acquisitional significance of the syntactic contrasts, and, more generally, of variation in this domain has never been systematically probed, however. Focusing primarily on three contact varieties spoken in southern Africa – Afrikaans, Namibian German (NG) and Kroondal German (KG) – we seek to demonstrate the central role that imperatives can play in establishing both core aspects of a language’s syntax and respects in which it permits variation.

The distinction between Afrikaans positive and negative imperatives is illustrated in (1). While (1a) mirrors the Dutch pattern, (1b) represents a contact-induced innovation, codified as standard in 1925. While the co-occurrence of moenie and nie signals Afrikaans’ Negative Concord status – leading, seemingly inevitably, to the non-standard postulation of generalised Negative Concord (Negative Spread), which must be explicitly “corrected” via prescriptive input (Biberauer & Zeijlstra 2012a,b) – “un-Germanic” initial moenie independently signals both Afrikaans’ V2 and OV properties: moenie transparently derives from moet (‘must’), and, like moet when it is the finite verb in a matrix structure, it blocks raising of the non-finite verb, which thus serves to signal the headness of VP. Significantly, moenie+VO\_DP structures are strongly ungrammatical in even the most English- (i.e. VO-) influenced varieties (e.g. Kaaps; McCormick 2002, Hendriks & Dyers 2016); cf. (2). By contrast, Moenie+V-PP\_adjunct/R-pronoun structures are possible alongside Moenie+PP\_adjunct/R-pronoun-V ones (3). Crucially, the contrast between negative-imperative options in (2) and (3) mirrors the difference between “leaking” and “non-leaking” options in Afrikaans: DP-objects cannot leak, while PP-adjuncts and R-pronouns can. More general inspection of Afrikaans “leaking” patterns reveals that negative imperative options very consistently match those in non-imperative contexts; in light of the further empirical evidence we consider here, we argue that this matching in fact follows from what imperative structures dictate.

Moenie also (i) cues the higher-than-Dutch default location of the Dutch-derived medial negation marker, with knock-on effects for the grammar of scrambling – a property which Afrikaans retains, but in modified form (4) – and (ii) reduces the contexts in which Dutch geen (‘no’) is employed: as (1b) and (4) show, nie readily precedes defines, which it would not do in Dutch, while the nie-component of moenie produces increased use of positive indefinites (geen > nie ‘n); compare (5) with their Dutch/German counterparts in (6). This, in turn, enhances the force of geen, which has “pragmaticalised” (Diewald 2011) as an emphatic sentential-negation substitute for clause-internal nie in denial contexts like (7); see also (5c).

Strikingly, contact with Afrikaans moenie has produced muss(t) nicht (‘must not’) as an imperative option in both NG and KG, with differing knock-on effects. Importantly, muss nicht has become the default negative imperative marker in NG, but not KG. (8-9) illustrate the negative-imperative options, available alongside standard-German V1- (9a) and V-final (9b) structures; the dominant patterns in KG are (9b,c). Research to date suggests that NG speakers accept (4a)-type Neg>Definite Object patterns as a neutral alternative to the prescriptively correct (4b)-type, which is neutral in Dutch and German (equivalent to (4a)), and also that nicht ein (‘not a’=nie ‘n) structures co-exist with kein (‘no’=geen) structures in neutral (6b)-type contexts. In KG, by contrast, kein appears not to be affected, while evidence of higher-than-usual nicht also seems more limited. Both varieties, like Afrikaans, also reflect overuse of the (deontic) must-modal in comparison to standard German (the same is true in South African English, which also has a distinctive You must-imperative).

What the discussion suggests, then, is (i) that Afrikaans’ innovated must-based negative imperative plays a key role in facilitating both stability (V2 and “basic” OV) and further innovation (the moenie knock-on effects) in that language; (ii) that it is a salient formal marker, which seems particularly amenable to contact-based propagation; and (iii) that its incorporation into a grammar may produce wide-ranging effects beyond the imperative domain, which, moreover, seem to “spread” in an at least partially ordered sequence, (10), which may also facilitate valuable insight into the process via which a grammar may reorganise itself in response to the incorporation of a novel structural option that initially becomes established in a high-frequency context. Looking beyond Afrikaans, our larger point is that syntax-rich, hearer-oriented structures like imperatives (positive and negative) and interrogatives deserve significantly closer attention in the study of what shapes I-language grammars than they have received to date.
Data

(1) a. Maak die deur oop!  
make the door open

b. Moenie die deur oopmaak nie!  
must.not the door open.make POL

‘Open the door!’

‘Don’t open the door!’

(2) * Moenie vergeet jou paspoort nie!  
must.not forget your passport POL ≠ ‘Don’t forget your passport!’

(3) a. Moenie [op straat] rondstaan op straat nie!  
must.not on street round.stand on street POL

‘Don’t stand around on the street!’

b. Moenie [daarvan] vergeet daarvan nie!  
must. not there.from forget there.from POL

‘Don’t forget about that!’ [ ] = prescriptively imposed order

(4) a. Ons het nie die boek gekoop nie. (neutral)  
us have not the book bought POL = ‘We haven’t bought the book.’

b. Ons het die boek nie gekoop nie. (marked)  
us have the book not bought POL = ‘The book we didn't buy.’ (contrast, etc.)

(5) a. Moenie ‘n das dra nie!  
must.not a tie wear POL

‘Don’t wear a tie!’

b. Hy het nie ‘n das gedra nie. (neutral)  
he has not a tie worn POL

‘He didn’t wear a tie.’

c. Hy het geen das gedra nie. (marked)  
he has no tie worn POL = ‘He did NOT wear a tie (denial)’

(6) a. Trage keine Krawatte!  
wear no tie

b. Er trug keine Krawatte. (neutral)  
he wore no tie

‘Don't wear a tie!’

‘He didn’t wear a tie.’

(7) Hy is G’N ondersteunend nie!  
he is no supportive POL

‘He is NOT supportive (contrary to what you’ve just said)’

(8) a. Musst nicht spät sein!  
must not late be

b. Net nicht spät sein!  
just not late be

[NG]

‘Don’t be late!’

‘Just don’t be late!’ (net = Afrikaans ‘just’)

(9) a. Vergiss deinen Reisepass nicht!  
forget your passport not

[standard German]

‘Don’t forget your passport!’

b. Nicht deinen Reisepass vergessen!  
not your passport forget

[standard German]

c. Du musst nicht deinen Reisepass vergessen!  
you must not your passport forget

[KG]

d. Musst nicht deinen Reisepass vergessen!  
must not your passport forget = the moenie-based structure

[KG]

(10) Muss(t) nicht > over-use of müssen beyond imperatives > earlier placement of nicht in clause-internal position { > reorganisation of the default order of nicht and object } > replacement of kein by nicht ein} { [ ] may not be ordered changes}
What can explain it? The changing models of York English

Bronwyn M. Bjorkman and Sali Tagliamonte

Across different varieties of English, the modal auxiliaries appear to be declining in use, being replaced by corresponding semi-modals. Recent studies have shown specifically that MUST is declining in favour of the semi-modal HAVE (GOT) TO (Tagliamonte and Smith, 2006; Tagliamonte and D’Arcy, 2007), and WILL similarly declining in favour of the semi-modal BE GOING TO (Berglund, 1997; Szmrecsanyi, 2003; Tagliamonte, 2002).

Against this backdrop, we might expect to find the same change with a third modal/semi-modal pair, CAN and BE ABLE TO, which stand in much the same relation as MUST/HAVE (GOT) TO and WILL/BE GOING TO. Instead, we find CAN charting a radically different path from the rest of the English modal system. Preliminary evidence from Google n-grams suggests that rather than declining, CAN has been steadily rising in frequency since at least the 1930s. The current study investigates the trajectory of CAN in more detail for one variety of English, using the York English Corpus (YEC: Tagliamonte, 1996-1998). We investigate the use of CAN (and its past counterpart COULD) not only in relation to the semi-modal BE ABLE TO, but also to the other modal auxiliaries expressing possibility, MAY and MIGHT.

The results confirm the initial observation that the use of CAN is increasing: across all clauses expressing modal possibility, the youngest speakers (<35) used CAN 53.5% of the time, but the oldest (>65) used it only 42.4% of the time. This profile raises at least two questions: first, does the change in the frequency of CAN represent grammaticalization into new domains of modal meaning, and second, can the split in frequency between CAN and the other modal auxiliaries be related to formal properties of the organization of the English modal system.

It is well established that modals tend to extend from purely deontic uses into epistemic domains (Traugott 1989, Hopper and Traugott 1993, Bybee et al. 1994, a.o.), and that dynamic modals such as CAN tend to develop into deontic modals. In the YEC, indeed, we find that CAN has fully displaced MAY as a modal expressing deontic possibility (=permission): of 271 tokens expressing permission, only 1 was expressed by MAY, while 98.9% were expressed by either CAN (169 tokens) or COULD (99 tokens).

Given the low rates of MAY by all ages of speakers, however, this is insufficient to explain the increase in CAN overall. More interesting is the extension of CAN into epistemic uses. It has often been noted in syntactic and semantic work on English modals that CAN differs from all other modal auxiliaries in resisting epistemic interpretations (except when negative or in some aspectual contexts). The YEC bears this out, with comparatively few pure epistemic uses of CAN overall (only 13 total), all of which are negative (samples in (1)). Despite the absence of pure epistemic uses, however, CAN does frequently exhibit quantificational uses (Brennan, 1997), illustrated in (2); while these are not epistemic interpretations per se, neither are they classic dynamic interpretations relating to physical, mental, or circumstantial ability.

This brings us to the second question: is the split between CAN and the rest of the modal auxiliaries due to independent properties of the modal system? Notably, CAN has long been identified as exceptional in a number of ways: beyond its lack of true epistemic interpretations, CAN is the only modal whose historical past tense form, COULD, is interpretively past—i.e. it can express was able to (by contrast, WOULD never means was going to).

Its exceptionality in these respects may explain why CAN has resisted replacement by BE ABLE TO: CAN not only lags behind the other modals in terms of grammaticalization (if we regard dynamic modals as the first step on the typical grammaticalization path), but it would also be exempt from any functional pressure to restore tense distinctions to the modal system, once the historically preterite modals became fully divorced from their present tense counterparts. It may be that these formal differences provide some of the underpinning to the divergence we now see between CAN and the rest of the modal system.
(1)  a. ...we thought “This can’t be it, this can’t be it.” (Speaker: 4; Age: 20)
   b. ...to be on the receiving end of that, you-know, it can’t be comfortable for them at all. (Speaker: 90; Age: 40)

(2)  a. the North-Sea [...] it can be so rough it sort-of goes along sideways (Speaker 30; Age: 20)
   b. some of the service can be a big dodgy sometimes (Speaker 90; Age: 40)
   c. They can be quite expensive, can’t they? (Speaker 64; Age: 46)

References


Embedded V2 is anti-licensed by discourse familiarity: Quantitative and statistical evidence
Spencer Caplan and Kajsa Djärv
University of Pennsylvania

Whereas all declarative main clauses in Mainland Scandinavian are obligatory verb second [V2] (1a), subject-initial embedded clauses are ‘optionally’ V2 (1b). We present the first large-scale corpus study addressing both the formal and the stylistic factors influencing the variable realization of embedded V2 [EV2].

**Grammatical Predictors:** A question of interest in the literature concerns the licensing of EV2. On the one hand, it has been proposed that EV2 is licensed (but not obligatory) in the complements of certain classes of matrix verbs (exemplified by *say, believe, deny, regret*, and *know*), defined in terms of their lexical semantics (Table 1). Such semantic selection accounts disagree, however, with respect to the ability of the factive/presuppositional verbs (*know, regret*) to license EV2. For Wiklund et al. (2009), *know* patterns with *say, believe*, in allowing EV2 (unlike *deny* and *regret*). For Haegeman (2010) and Kastner (2015) however, all factives disallow EV2. An alternative view takes EV2 to be pragmatically licensed, when the embedded proposition is used to update the conversational common ground (Julien, 2009; Jensen and Christensen, 2013). Regardless of the view adopted, most studies have assumed some a priori classification of the verbs, and have looked at a relatively small set of verbs to represent the different classes.

**External Factors:** Previous corpus studies also suggest that speech-style or genre plays a significant role in influencing the availability of EV2. Looking at spoken Danish, Jensen and Christensen (2013) find EV2 to occur at a rate of 45%. For Swedish, Heycock and Wallenberg (2013), observe a large discrepancy in the rates of EV2 across child-directed speech (26%) and blog-data (31%) as opposed to published novels (<1%). This leads us to expect that the rates of EV2 should be significantly higher in informal, than in formal, contexts.

**Current Study:** We extract a broad range of statistical information (code to subsequently be made available and open source) comparing across several Swedish corpora (Borin et al., 2012) representing different genres. From this investigate the formal and stylistic factors argued to influence the realization of EV2: Do some verb classes categorically disallow EV2, as predicated by semantic selection, or is the distribution of EV2 across verb classes (more) gradient, as expected on pragmatic selection? We also test the hypothesis that corpora representing more formal styles should show lower rates of EV2 across the board.

**Results:** Strikingly, our results reveal rates of EV2 at an order of magnitude lower than what previous studies have reported (Table 2). Although the rates of EV2 differ across genre, we observe overall rates of EV2 between approximately 1% and 5% in both the formal (academic texts) and informal contexts (a lifestyle website, and online forums). This is surprising, given previous findings. In terms of the formal factors argued to influence the availability of EV2, we observe contra Haegeman (2014) and Kastner (2015) that the *know*-class freely allows EV2 (Figure 2). Unlike previous accounts however, we find the *deny*-class to pattern similarly to the *believe*-class (Figure 3). Plotting the distribution of the rate of EV2 conditioned on individual matrix predicates shows a smooth declination (Figure 2). This is suggestive against a formal semantic account which predicts that a limited set of predicates with the requisite properties should have notably different rates of EV2. Both corpora with rates of EV2 under 3% (Table 2) show significant differences under a Wilcoxon signed-rank test for all pairwise comparisons. Interestingly, contrary to the hypothesis that speech-style effects the overall rates of EV2, we observe an interaction between genre and predicate type (Figure 3). This in fact provides a new form of evidence in favor of a pragmatic licensing account, according to which the use of a given predicate with a complement updating the common ground will vary depending on the type of discourse.
(1a) Han (gillar) inte (*gillar) kor. he likes not (likes) cows.

‘He doesn’t like cows.’

(1b) Han sa att han (gillar) inte (gillar) kor. he said that he (likes) not (likes) cows.

‘He said that he doesn’t like cows.’

<table>
<thead>
<tr>
<th>Examples</th>
<th>say, claim</th>
<th>believe, suppose</th>
<th>deny, accept</th>
<th>resent, regret</th>
<th>know, realize</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical semantic properties</td>
<td>Non-factive volunteered speech act</td>
<td>Non-factive doxastic state</td>
<td>Non-factive response speech act</td>
<td>Factive emotive state</td>
<td>(Semi-)factive doxastic state</td>
</tr>
</tbody>
</table>

Table 1: Verb classes identified to differ in their ability to take EV2-complements.

Figure 2: **Left:** Plot of probability of ev2 given matrix lemma by relative rank for the Flashback-Politik corpus. **Right:** EV2 under Factive vs. Non-factive verbs.

Figure 3: Distribution of EV2 across genre and predicate type.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Sentences</th>
<th>Proportion Non-ambiguous</th>
<th>p(ev2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Texts</td>
<td>60931</td>
<td>0.072</td>
<td>0.0261</td>
</tr>
<tr>
<td>Attasidor</td>
<td>8059</td>
<td>0.0768</td>
<td>0.0081</td>
</tr>
<tr>
<td>Familjeliv</td>
<td>458699</td>
<td>0.0801</td>
<td>0.0547</td>
</tr>
<tr>
<td>Flashback-Politik</td>
<td>2841872</td>
<td>0.0962</td>
<td>0.0449</td>
</tr>
</tbody>
</table>

Table 2: Rates of embedded v2 across corpora of varying styles. “Academic texts” are formal written texts in the humanities. Attasidor is a newspaper published in simple written prose. Familjeliv is a lifestyle website. Flashback-Politik is a political forum.

Positive *Anymore*: Lexicalization and (De)grammaticalization

Amie DeJong, University of Washington

“Positive” *anymore* (ex. 1, next page), which is ungrammatical for speakers for whom *anymore* is a negative polarity item (NPI), has caught the attention of linguists for close to a century (Malone, 1931). I propose an account of the original reanalysis of *anymore* from NPI to non-NPI in light of modern syntactic analyses of polarity sensitivity. I argue that the lexicalization of the NPI *any* and the lexical item *more* led to isolation from the paradigm of the indefinite *any* and its compound forms. This paved the way for the loss of polarity sensitivity, which involved the loss of uninterpretable polarity features - an instance of degrammaticalization, considered rare or impossible (by Roberts & Roussou, 2003; van Gelderen, 2011; a.o.).

Positive *anymore* has received attention in both sociolinguistic and formal frameworks, with both synchronic and diachronic proposals (Hindle & Sag, 1973; Labov, 1973; Murray, 1993, a.o.). Most, however, have concentrated on its spread through the United States, not the original reanalysis of NPI to non-NPI.

According to Merchant (2013), the NPI *any* and its positive polarity item counterpart *some* are identical D(eterminer) heads with uninterpretable polarity features before feature valuation because of the identity condition on ellipsis. Agree of these [uPol] features with a c-commanding ΣP with negative [iPol] is realized as *any*; with a positive ΣP, [uPol] is realized as *some*. The result may combine with an NP. This, I show, accounts for the DP *any more* (ex. 3) and the VP-internal reading of *anymore* (ex. 3) where *any more* takes immediate scope over the event (as in, “Sue won’t read additionally”), but not the VP-external *anymore*, a more time-related, “terminative” adverb (in the terms of Cinque, 1999), whose polar opposite is *still* (ex. 4).

This isolation from the paradigm of *any*- indefinites erased the cues for the more complex featural makeup of D[Indef; uPol:neg] plus NP or AdvP containing the lexical item *more* with semantic features like [additional]; I propose that it was reanalyzed as a single morpheme with the features Adv[continuation; uPol:neg]. Fusion of *anymore* into a single morpheme was key in making the change in polarity sensitivity possible: *anymore* could be further reanalyzed at some point as not having [uPol] features at all, again simplifying the featural makeup. Isolation from a paradigm resulting in opacity of a lexical item has been argued to be a factor in other instances of deggrammaticalization (Willis, 2007, a.o.). The process of deggrammaticalization of *anymore* was driven by the same universal tendencies that drive grammaticalization, according to theories like Roberts & Roussou, 2003: the language acquirer chooses the less complex of two structures that match a given input string.

The reanalysis of a VP adverb as a TP adverb recalls the grammaticalization of English auxiliaries from lexical verbs that Roberts & Roussou (2003) explain as a loss of movement of the verbs from V→T, resulting in a simpler structure. Because (English) adverbs are Externally Merged in their base position and do not undergo synchronic movement from VP to TP (Cinque, 1999), this reanalysis cannot involve loss of movement. VP→TP adverb reanalysis is not
uncommon, however (van Gelderen, 2011), and requires an explanation.

Examples

(1) Kids anymore are so spoiled.

(2) a. Matt doesn’t want any more, but Sue does [want some more].
   b. [CP [SP [xPol:neg] [TP Matt [t doesn’t [vP [VP want [DP [DIndef; uPol:neg] [NP more]]]]]]]]
   c. [CP [SP [xPol:pos] [TP Sue [t does [vP [VP want [DP [DIndef; uPol:pos] [NP more]]]]]]]]

(3) a. Sue won’t read any more (e.g., tonight), but Matt will [read some more].
   b. [CP [SP [xPol:neg] [TP Sue [t won’t [vP [VP read [DP [DIndef; uPol:neg] [AdvP more]]]]]]]]
   c. [CP [SP [xPol:pos] [TP Matt [t will [vP [VP read [DP [DIndef; uPol:pos] [AdvP more]]]]]]]]

(4) Matt doesn’t swim anymore (e.g., since he developed a phobia of water), but Sue does [*swim some more]/ [swim still].

References

Stable variation and change in Hungarian negative sentences

Katalin Gugán
Research Institute for Linguistics, HAS

The oldest written sources of Hungarian, as well as Modern Hungarian, display two patterns of negation in sentences containing a verbal modifier (VM): the VM may precede or follow the negative particle – verb unit. However, VM-NEG-V is only marginally present in Modern Hungarian, whereas it was predominant in Old Hungarian, and it can be assumed that this was inherited from Proto-Ugric (Klemm 1928-1942), whereas the other pattern is a Proto-Hungarian innovation.

According to É. Kiss (2014), VM-NEG-V pattern reflects the adjunction of the negative particle to V, whereas the other pattern is indicative of the movement of the negated verb. This movement is triggered by a Neg head in the left periphery, NegP being one of those functional projections that emerged in Proto-Hungarian. É. Kiss assumes the following phases of change. Initially, movement of the negated verb to NegP was optional, and it occurred mostly when the given sentence did not contain a negative pronoun. In time, negative pronouns went through phonological erosion, and they were no longer recognized as elements encoding negation, therefore, negator+verb movement became increasingly frequent. This, in turn, led to the reanalysis of the negative particle as base-generated in Neg, triggering verb movement.

The present paper adopts the structural explanation of the patterns offered by É. Kiss, but argues that the postulation of increasing frequency as a trigger of reanalysis is not supported by corpus data on the basis of two types of analysis. One of them focused on the two types of patterns in which the VM is a preverbal particle (see 1a and 1b). Both the analysed 14th-16th Old Hungarian sources and the investigation of the Corpus of Historical Hungarian Informal Texts (comprising 15th-18th century private letters and minutes of witch trials) showed that the rate of the marginal variant (NEG-V-VM) did not increase during the given centuries. Due to the nature of these sources, the possibilities to examine sociolinguistic and dialectal factors is limited, but it seems that these did not play a role in determining the distribution of the variants. As the investigation of selected preverbs in the Hungarian Historical Corpus showed, the change in the distribution started around 1820. However, prior to the start of this change, there seems to be no drift in the frequency of the innovative form (cf. Kroch 2005), even though it had been present in Hungarian for centuries.

The other analysis (in progress) focusses on analysing all negative sentences of selected individuals in CHHIT. The preliminary results show that grammatical features (main/subordinate sentence, type of the verb modifier) have a significant effect on the choice between the variants, for instance, the marginal order is predominant in sentences with nominal predicates (2). However, there are cases in which there seem to be no grammatical clue that would explain the choice of an individual, and one can only offer tentative explanations (e.g. that the NEG-V-VM order can be more emphatic, (1b, 3)) that are difficult to prove. Still, it seems reasonable to suggest that stable variation reflects a functionally driven optional movement, whereas the fairly radical change in the 19th century was brought about by a structural reanalysis. Due to this, the use of the formerly general pattern of
negation became restricted to given types of subordinating structures, or, when used in a different grammatical environment, it can encode emphatic negation in Modern Hungarian.

**Examples**

(1a) Szakmári uram még meg nem érkezet Debrecenbül.
    Szakmári lord.1SG yet PV NEG arrive.PST.3SG Debrecen-ELA
    ‘My lord Szakmári has not yet arrived from Debrecen’, K. Barkóczy, 1699

(1b) […] csak felprédálta, nem égette meg.
    only reave.PST.3SG>OBJ, NEG burn.PST.3SG>OBJ PV
    ‘He only reaved (the mill), he did not burn it.’ K. Barkóczy, 1704

(2) Az sátor, Szívem, nem lesz már szükséges.
    the tent heart.1SG NEG be.FUT.3SG already necessary
    ‘My dear, the tent won’t be necessary from now on.’ S. Károlyi, 1705

(3) lesz haszna a költségen felyül is a szőlőknek,
    be.FUT.3SG profit.3SG the expense.SUP above too the
    nem vész oda a pénzünk.
    NEG vanish.PRS.3SG PV the money.1PL
    ‘there will be indeed profit on the vineyards, our money will not be wasted.’ K. Barkóczy, 1705

**References**


**Corpora**

http://omagyarkorpusz.nytud.hu/en-intro.html (Old Hungarian Corpus)

http://tmk.ny wholeud.hu/ (Corpus of Historical Hungarian Informal Texts)

http://clara.nytud.hu/mtsz/run.cgi/first_form (Hungarian Historical Corpus)

---

1 DAT = dative, ELA = elative, FUT = future, NEG = negative particle, OBJ = objective conjugation, PL = plural, PV = preverbal particle, PST = past, SG = singular, SUP = superessive
Various studies on Heavy NP Shift (HNPS) in English converge on the conclusion that the frequency of HNPS is not only determined by the heaviness of the displaced object, measured in the number of words. The relative heaviness of the object vs. the word string between the main verb and the object plays a significant role as well (Hawkins 1994, Wasow 1997, Stallings & MacDonald 2011). For instance, the heavier a verb-modifying PP is, the less likely it is that a direct object of a fixed length is shifted to the right of the PP (cf. The radio listeners accepted without doubt the whole story about the defects in the new Mazda vs. The radio listeners accepted without doubt or any bit of concern the whole story about the defects in the new Mazda).

In this talk, we will discuss relative weight effects with HNPS in Icelandic and Faroese. A study of this kind has not been undertaken before. Our conclusions are based on a series of studies, including production tests, acceptability tests and reading tests. All the sentences tested contained an NP immediately preceding a PP and the relative length of NP vs. PP was controlled. In the production tests the beginning of the sentence including the main verb appeared in the middle of the screen and the NP and PP appeared above and below it in a randomized order. The speakers were asked to construct the sentences in the way they thought best and read them out loud from memory. The same sentences were then tested for acceptability and in a reading test. All produced sentences were recorded and diagnosed in Praat.

Our results suggest that the frequency of HNPS is significantly higher when the NP exceeds the PP in length by at least three words. HNPS rarely occurs when the NP and the PP are equally long and hardly ever when the PP exceeds the NP in length. Our results show that HNPS is accepted and used by all Icelandic speakers but with great variation in production frequency, whereas Faroese speakers accept it and use it to a much lesser extent. Our results show that both Direct Objects and Subjects can be shifted in both languages, contrary to previous claims about HNPS in Faroese (Barnes 1992, Vikner 1995, Thrainsson et al. 2012). Furthermore, our results showed clearly that NPs that cannot be considered long or heavy from a syntactic point of view, e.g. a single syllable noun, can undergo HNPS as long as the PP consists of a preposition and a personal pronoun. Neither prepositions nor personal pronouns are phonological words which means that the PP does not form a phonological phrase (Selkirk 2011). The syntactically “light” NP can therefore shift across the PP as it is phonologically heavier. This supports our theory that the weight of an NP is not merely defined by the amount of words or complexity of the NP itself, but by the relative weight of the word string next to it, and that phonological weight is just as important to HNPS as syntactic weight.
References


Complex Verbal Heads in the Left Periphery of Pomeranian

Göz Kaufmann, University of Freiburg / Germany

Translating the stimulus sentence ontem eu poderia ter vendido o anel (‘yesterday I could have sold the ring’), a substantial number of Brazilian speakers of Pomeranian produced one of the following two variants:

(1) jern yester hi kätt ik de ring verköff have

(2) jern yester hätt kütt ik der ring verköff hat have

While the verbal sequence in (1), probably an instance of verb projection raising, is common to West Germanic varieties, the translation in (2) has to be considered a rarum since it apparently features two verbal heads in C⁰. This rarum, which must not be confused with seemingly similar cases in Afrikaans (cf. DONALDSON 1993: 364), has already been discussed by POSTMA (2014: 639–642). POSTMA explains this phenomenon in Pomeranian from Espírito Santo by assuming two derivational mechanisms: first, Pomeranian verb clusters are supposed to be formed by head movement and second, Pomeranian is supposed to feature a weak T that “must move to C, incorporate into it, and form a stronger C+T complex” (POSTMA 2014: 642).

As our research is based on the actual translation of 56 stimulus sentences by 49 Pomeranian speakers from Rio Grande do Sul, we can describe the derivation of (2) in greater detail. Using statistical analyses, we are able to delineate the existing inter- and intra-speaker variation. In our opinion, the translation in (2) is indeed related to head movement. This can be seen in the fact that informants that produce (2) frequently produce the cluster sequence V3-V1-V2 in dependent clauses. An example for this is (3), the conditional clause of the stimulus sentence se ele realmente tivesse querido escrever essa carta, ele teria achado tempo (‘If he really had wanted to write this letter, he would have found the time’). The informants that are responsible for (1) hardly ever produce this sequence:

(3) wenn her auf ehrlich de Karte schriewe hätt wutt

If we derive the cluster in (3) by head movement and right adjunction of V2 wutt to V1 hätt in ᵠ, the fact that informants that produce (2) also produce (3) can be explained by assuming that complex heads in Pomeranian such as hätt wutt in (3) or hätt kütt in (2) can be moved as one morphosyntactic unit from ᵠ to C⁰. Interestingly, the non-existence of (2) in European varieties of German is precisely the reason why STERNEFELD (2009: 521) denies the existence of complex heads in the right periphery of German clauses. Obviously, assuming that the head position of CP can host two verbal elements is anything but a matter of course. Fortunately, another feature of (2), also absent from the data of the informants responsible for (1), may help to qualify this assumption. Many translations represented by (2) feature the syntactic doubling hat hewe (‘had have’; cf. KOENEMAN et al. 2011). If we assume that the non-finite participial features of kütt clash with the finiteness requirements of C⁰, it stands to reason that
they cannot move to this position. These stranded features may then have to be linked to a clause-final lexical entity and this entity could be a non-finite lexical copy of the highest verb hött, which itself is moved to C⁰. Hött in C⁰ may then not qualify as a “full-fledged” verbal head anymore; it may just serve as a carrier of finiteness features. This analysis is somewhat different from Postma’s (2014) weak T analysis, but enables us to explain several syntactic surface phenomena by means of one derivational mechanism, namely the formation of complex heads in clause-final verb clusters.

References
Grammar Competition in Neutral Learning: A Reply to Han et al. (2016)
Henri Kauhanen\textsuperscript{a}, Caroline Heycock\textsuperscript{b}, and Joel C. Wallenberg\textsuperscript{c}
Manchester University\textsuperscript{a}, University of Edinburgh\textsuperscript{b}, and Newcastle University\textsuperscript{c}

Han et al. (2016), building on Han et al. (2007), challenges the idea that grammar competition (Kroch 1989, inter alia) is a typical characteristic of language learning (Yang 2000, 2002). The (empirical) claim is that in a case of two plausible parameter settings and no positive evidence to learners, Korean speakers acquire one parameter setting or the other, essentially at random – but not both. Such a result would argue against the idea that grammar competition is a regularly occurring phenomenon in syntactic acquisition, as well as the idea that inherent derivational biases exist for particular parameter settings (Clark and Roberts 1993, and subsequent). We show, however, that Han et al’s (2007; 2016) own data are consistent with the presence of competing grammars in synchronic Korean. Additionally, we demonstrate that a competing grammars / variational learning model of language acquisition makes accurate predictions for the proportions of competing grammars in a neutral case (one in which neither grammar has an advantage). This, in turn, sheds light on the actuation problem in syntactic change.

The two Han et al. papers present experimental studies on the relative scope of quantifiers and negation in Korean, which they argue is diagnostic for whether V-to-T movement of finite verbs takes place. The two syntactic variants under investigation are not distinguishable except by virtue of possible interpretations in configurations that are plausibly very rare, and are arguably not be distinguishable by strictly positive evidence. Under these conditions, the authors claim that speakers converge on either a verb-raising Korean grammar or a non-verb-raising one, but do not acquire both variants. However, while the authors show that speakers are consistent in their linguistic behaviour across testing sessions, they do not establish that the overwhelming pattern for each individual is one of categorical use of a single system. Additionally, given the way participants’ scores are aggregated in Han et al. (2016) data, 4.8%–14.3% of adult participants plausibly had systems containing 25%–75% of each variant, and more may have had both variants with one at a frequency lower than 25%. In the Han et al. (2007) data, 20%–35% of the adult participants had robust mixed systems, as did 6.8% of the 2007 children.

We show that these results are predicted by the way a variational learning model with competing grammars (Yang 2002) reacts to a lack of evidence for parameter setting. We ran a batch of simulations of such a scenario, assuming a variational learner, a binary parameter, and the equivalent of 3 years for the learner to set the parameter. Using estimates from Shneidman et al. (2012) and Shneidman and Goldin-Meadow (2012) for the number of utterances overheard by or directed to children, cross-checked with independent estimates from Vosoughi and Roy (2012) and Vosoughi et al. (2010), we modeled a lower bound of 5 million utterances and an upper bound of 12 million utterances for the learning period, and assumed a learning rate parameter estimated from the application of the same learner to a non-neutral case of syntactic change in Heycock and Wallenberg (2013). We found that even under the assumptions of competing grammars and no inherent bias towards one parameter setting or the other, most learners are attracted to categorical states of 0% and 100% for each parameter setting in the case of no evidence, simply by stochastic drift. However, a residual number of learners exhibit intraspeaker variation (competing grammars), and the proportion of this residue is consistent with proportions reported in the two Han et al. studies (Figure 1). This variation constitutes an innovation in acquisition, and could be thought of as the beginning of a syntactic change.

Our study shows, contrary to Han et al. (2007, 2016), that a population of learners can come to a variety of different states of grammar competition in the absence of evidence. Furthermore, finite population simulations show that the proportions of learners in different states match those Han et al. observed experimentally. Finally, the variation in competing grammar states
that result from stochastic drift can be thought of as the actuation of a change, the emergence of new syntactic variants *de novo*.
Figure 1. Proportion of learners rejecting, accepting or ending up ambivalent with regard to V-to-T movement: data from Han et al. (2016, Fig. 7) and the variational learning model for a learning rate ($\alpha = 0.001$) independently estimated from Heycock and Wallenberg (2013) and both 5 and 12 million input tokens during the learning period.

References

The Importance of Population Structure in Models of Language Change
Jordan Kodner

We discuss the interaction between population size and structure with learning models on the path of language change. We present simulations of abstract neutral and S-curve change, the spread of the cot-caught merger in New England, and of the Northern Cities shift in the St. Louis Corridor to demonstrate these interactions. The simulations demonstrate how both population structure and learning must be considered for full understanding of both models and empirical data.

A recent paper by Kauhanen [3] seeks to address the impact that social networks have on neutral change with a series of agent-based simulations on centralized networks. He studies populations of size $n = 200$ and assumes categorical learners who settle on single grammars rather than multiple competing grammars. We contend that the baseline dynamics of these simulations are primarily a function of these assumptions rather than of neutral change itself.

We simulate change in a two-grammar network using a dynamical system model based on Niyogi & Berwick [6, etc.] augmented with an adjacency matrix to describe network structure. This model is efficient enough to compute with large $n$. If $n = \infty$ or competing grammars are allowed, each community homogenizes at a 50/50 grammar distribution (red curve). We simulate the change for 10 trials with categorical single-grammar speakers at $n = 200$, $n = 2000$, and $n = 20000$ (blue curves). Figure 1 shows the pattern of neutral change in one community of a two-community network where each community begins at 100% monolingual. Figure 2 shows the same simulation applied to a change with differential fitness. Such changes are expected to produces a logarithmic S-curve [4, 5, 6, 8, etc.]. At $n = 200$, the results are chaotic and unpredictable, and most simulations fix at 0 or 100% within 500 iterations. But as $n$ increases, the path of change begins to approximate the mathematically predicted curves. The categorical speaker assumption and population size dominate in determining the behavior of this change.

The spread of vowel mergers has been successfully modelled in terms of categorical individuals [7], and can be shown to produce an S-curve in infinite populations. As calculated for cot-caught in New England [2, 7], if $> 17\%$ of a learner’s input is merged, he or she will acquire the merged grammar. However, when simulated in a single-cluster network of any size, the population stabilizes on one grammar or the other in only a couple iterations. To correct this, we model the community as 40 loosely connected clusters of 18 individuals in a way more closely approximating actual population structures. In this scenario (Figure 3 left), each cluster acquires to the merger rapidly in fast succession (colored lines), but the community average (black curve) takes on the classic S-shape. Notably, the rate of change is dependent on both the grammatical advantage and the incidental of the network (Figure 3 center shows multiple averages), and network rewiring improves the S-curve shape (Figure 3 right).

The rise of the NCS in the St. Louis Corridor does not follow an S-curve [1]. The NCS rises and falls, first in communities on Route 66 (Figure 4 left red solid [1]) then in communities off Route 66 (Figure 4 blue solid). We reproduce this pattern in a network of community clusters to recreate the Corridor (Figure 4 center). Changing only a single parameter, the rate of movement from Chicago to On-Route communities, we reproduce a two-peak pattern (Figure 4 right).

The first results highlight the need for carefully planned simulations. Seemingly innocuous design choices can drastically change their outcomes. The second and third results demonstrate how populations and grammars interact to effect empirical rates and paths of change. Importantly, the slope of S-curved change cannot safely be attributed to grammar alone, since it is a function of both the learning algorithm and the network.
Figure 1: Neutral change, single-cluster community, $n = 200, n = 2000, n = 20000$.

Figure 2: S-curve change, single-cluster community, $n = 200, n = 2000, n = 20000$.

Figure 3: S-curve change, cot-caught merger, multi-cluster community.

Figure 4: Two-peak change in the St. Louis Corridor.

L. Friedman. The st. louis corridor: Mixing, competing, and retreating dialects. 2014.


Rules with exceptions: Using the Tolerance Principle to diagnose allophones

Betsy Sneller

Despite the vast phonetic variety in a language, these sounds can be reduced to a small inventory of contrastive units and abstract allophonic rules. While the difference between phonemes and allophones is robust and is taught to every first-year linguistics student, phonologists run into a notable problem in the boundaries, where some alternations are controversially analyzed as allophonic by some and phonemic by others. Perhaps the most widely cited case is that of the Philadelphia split in the TRAP vowel, which has been analyzed as both allophonic (e.g. Kiparsky 1995; Labov et al. 2016) and phonemic (e.g. Trager 1934; Labov 1989). Allophones are typically diagnosed using Predictability and Contrastiveness. In this paper, we argue that these diagnoses are insufficient and propose instead the application of Yang’s Tolerance Principle (2016) as an alternative, precise diagnostic of allophony.

First, we address Predictability. Many phonologists assume allophonic distinctions require full regularity, meaning that an alternation ceases to be predictable the moment a single lexical exception emerges. This suggestion results in the dissatisfying prediction that two speakers of the same language may have a different phonemic inventory based on whether they happened to acquire a single exceptional word or not. On the other end of the spectrum lies the phonemic LOT-THOUGHT distinction in English, which nevertheless shows some level of predictability by phonological environment which emerge as significant in a logistic regression, with pre-/l/ and post-/r/ tokens more likely to be a THOUGHT vowel.

The problem with LOT-THOUGHT predictability is easily resolved by appealing to Contrastiveness, since there are a number of minimal pairs between the two classes. However, as many phonologists have pointed out, some distinctions show only marginal contrastiveness (Ebeling 1960; Kiparsky 1995), appearing in near-minimal pairs or as systematically peripheral segments, leading to the proposal of an intermediate category of “quasi-phoneme” between an allophone and a phoneme, and the assertion that the boundary between these categories is “fuzzy” (Scobbie and Stuart-Smith 2006).

We hold that quasi-phonemes do not exist. Instead, we argue that diagnosing an alternation as allophonic or phonemic may be clearly defined using the Tolerance Principle (Yang 2016), which is a model of productive rules that both allows exceptions to the rule and defines the precise limit of exceptions. The Tolerance Principle states that a productive rule applying to $N$ lexical items may tolerate up to $N/\ln(N)$ exceptions to this rule. This has been shown to apply to a number of morphological and suprasegmental rules in a number of languages, including artificial language learning (Yang 2016; Schuler, Yang, and Newport 2015).

Using the Tolerance Principle as a diagnosis of productive allophonic rules has a number of implications for phonology and sound change. First, it extends Labov’s (1981) assertion that “words float on the surface of sound change” by allowing a precise number of lexical exceptions to a regular sound change. Second, allowing a list of exceptions to an allophonic rule opens up the possibility for distinct lexical items to join or leave the list of exceptions, as found in the diachronic instability of the lexical exceptions to the Philadelphia TRAP rule. Finally, this offers an elegant solution to regularity, defining productive rules in the same way across modules of the grammar.
Selected References


Competing denotations in the diachrony of English *wh*

Robert Truswell & Nikolas Gisborne, University of Edinburgh

For a learner acquiring the vocabulary of a language, the easy part is figuring out what the forms are. Figuring out the denotations of those forms is much harder. For instance, children know about the existence and approximate distribution of *the* long before they use it, or know what it means (Shipley et al. 1969). The challenge of figuring out the denotation of functional vocabulary is compounded by several factors: function words are frequently ambiguous (e.g. *that* is demonstrative and complementizer), and there is less of a mutual exclusivity pressure (e.g. complementizer *that* can alternate with *∅*). The division of labour between multiple function words in the compositional derivation of an utterance meaning may not be obvious, and the communicative penalty of a non-target-like choice of function word can be minor.

Where acquisition is hard and the consequences of mislearning minor, there is potential for change. This has been argued for syntax by Lightfoot (1979) and for semantics by Eckardt (2006), but little consideration has been paid to the consequences of changes in denotation for the inventory of syntactic structures available in a language. On a lexicalist conception of syntax and compositional semantics, syntactic structures are emergent results of the composition of lexical entries. The challenge faced by a learner in acquiring the denotations of function words directly implies variability in the constructions which can be derived using those function words.

In this talk, we illustrate the diachronic consequences of this challenge, by looking at the diachrony of *that*-relatives in English. For a brief period in Middle English, which we will call ‘peak *that*’, almost all relatives, whether headed or free, are formed by *that* (Fig.1). Partly, this is because of the disappearance of alternative relativizers (complementizer *pe* and inflected demonstratives) in early Middle English. However, during peak *that*, the rate of use of *wh*-forms in relatives also declined. During Old English, free *wh*-relatives had been common, but these largely disappeared during the peak *that* period, before re-emerging (along with headed *wh*-relatives) in later Middle English (Fig.2).

This decline and revival of *wh*-relatives can be understood in terms of a shift between competing denotations of *wh*-forms. Old English *wh*-forms had an indefinite denotation, as indicated by *wh*-indefinites like (1). Later Middle English *wh*-forms had a definite denotation, as indicated by nonrestrictive *wh*-relatives like (2). There is no direct evidence for a definite denotation in Old English (before the emergence of nonrestrictive *wh*-relatives), or an indefinite denotation in later Middle English (after the disappearance of *wh*-indefinites). Free *wh*-relatives are an ambiguous context, as they can be construed with either a definite or indefinite *wh*-phrase. Although the *wh*-form in (3) can be analysed as a modal definite (von Fintel 2000), it could also be an indefinite within the scope of a conditional operator, argued by Belyaev & Haug (2014) to be the diachronic source of *wh*-correlatives. Examples like (4), diachronically later, are unambiguously definite. This shift in denotation is possible, we argue, precisely because the *wh*-operator in (3) is in the scope of other propositional operators (here, a quantifier over situations), masking the precise compositional contribution of the *wh*-form.

Peak *that* falls between the disappearance of *wh*-indefinites and the emergence of nonrestrictive *wh*-relatives (Fig.3) — during the peak *that* period, *wh*-relatives were almost exclusively used in interrogatives, as they did not have a denotation conducive to their use as relativizers. As a result, there were few alternatives to *that* as a relativizer.

In conclusion, an understanding of changes in the inventory of English relative structures requires an understanding of the shifting model-theoretic denotations of the forms from which those structures are composed, such as the shift from indefinite to definite *wh*-forms. Such a shift is facilitated by scopal interactions with other operators which limit the truth-conditional effects of the different denotations.
Figure 1: Rate of use of *that* in relative clauses in YCOE/PPCME2/PPCEME

Figure 2: Free *wh*-relatives decline in frequency during the peak *that* period

Figure 3: Peak *that* falls between the disappearance of *wh*-indefinites and the rise of definite *wh*-relatives

(1) and gif hwa hyt bletsað, þonne ablinð seo dydrung.
   ‘and if who it blesses then ceases the illusion
   ‘and if anyone blesses it, then the illusion is dispelled.’ (coaelhom,+AHom_30:4.4082)

(2) Asa, kyng of Juda, . . . had sore feet, whech passioun oure bokys sey it was podegra
   ‘Asa, king of Judea, had sore feet, which suffering our books say was gout’
   (cmcapchr-m4,33.43)

(3) Swa hwylc eower swa næfð nane synne on him, awyrpe se ærest ænne
   So which you.GEN.PL so NEG.have no sin in him, cast.out.SBJ he first one stan on hy
   ‘He that is without sin among you, let him first cast a stone at her.’
   ‘If anyone is without sin among you, let him . . . ’ (coaelhom,+AHom_14:214.2117)

(4) Gemyne, [hwæt Sanctus Paulus cwæð]
   remember what Saint Paul said
   ‘Remember what Saint Paul said.’
   (cogregdC,GDPref_and_3_[C]:15.207.28.2739)

References


Modality and generalization of negative concord in northern Welsh dialects

David Willis, University of Cambridge

Some northern varieties of Welsh use *cau*, historically a reduced form of a verb found in literary Welsh as *nacáu* ‘refuse’, as a negative modal meaning roughly ‘won’t’, as in (1). In the nineteenth century, while the phonological reduction of *nacáu* to *cau* is well attested, it is not used in this negative-modal sense. This paper traces the development of this new modal over the last century using written sources and data from fieldwork for the ongoing *Syntactic Atlas of Welsh Dialects* (SAWD), to argue that the following stages need to be recognized:

(i) *cau* ceases to impose the requirement that its subject is volitional/agentive, and therefore goes from being a control verb ‘refuse’ to being a raising verb (loss of argument structure);
(ii) *cau* enters the negative system proper and begins to trigger negative concord; thus, the affirmative form of ‘be’, *mae*, in (1), is replaced by the negative-concord form, *(dy)di*, in (2).

These stages are reflected in ongoing dialect variation, different varieties experiencing change to differing extents: data from the SAWD fieldwork show that the pattern in (2) is grammatical for speakers under 50 in a contiguous area of central north Wales and is accepted sporadically by older speakers from this area, while elsewhere only (1) is accepted. Minor patterns, accepted by a subset of innovating speakers, suggest further stages:

(iii) *cau* is extended to clauses with an *n*-word, as in (3), where the shift in the polarity of *cau* plausibly triggers a shift from *pawb* ‘everyone’ to *neb* ‘no one’;
(iv) *cau* begins to trigger ‘strong’ negative concord i.e. both negative concord on the verb, as with *(dy)di* above, and presence of the sentential negator *ddim* ‘not’, as in (4).

Among SAWD speakers, the innovations are related in an implicational scale, with presence of an innovation implying the presence of all prior innovations in a speaker’s grammar:

- no neg. concord > weak neg. concord > concord with indefinites > strong neg. concord
  - (33 speakers) > (15 speakers) > (12 speakers) > (1 speaker)

This suggests that the various manifestations of negative concord form a single structured innovation. Formally, *cau* has ceased to impose the requirement that its subject is volitional / agentive, going from being a control verb, as in (5) in 19th-century dialects, to a raising verb (via loss of independent argument structure), as in (6), today. This change, attested in other cases of the grammaticalization of modals (cf. English *will*) (Kuteva 2001), eliminates structure, implying a Minimize Structure preference in acquisition (cf. van Gelderen 2004).

Furthermore, *cau* enters the negative system. In (5), it is a lexical verb, lacking polarity features. In the innovative analysis in (6), it bears a negative feature [+NEG]. Initially acquirers note that *cau* does not trigger negative concord, even though other [+NEG] items (e.g. the negator *ddim* ‘not’) do. They adapt their output to match the observed production of others in an ad hoc way. In (6), we expect the unvalued Pol feature on C to be valued [uPol: +NEG] (from *cau*). For a time, speakers override this, maintaining the conservative output in (1) despite reanalysis. Variation arises once this override is eliminated by some speakers, giving rise to the innovative pattern, (2). This approach conceptualizes innovation of negative concord as a form of extension in the wake of a 19th-century semantically based reanalysis.

Co-option of *cau* into the negative system proper is at first sight unexpected, but makes sense if the wider verbal system of the language is considered. *Cau* instantiates negation and mood, thereby amounting to a composite head. Welsh has other such heads paired for polarity, as with *gallu* ‘be able’ vs. *methu* ‘be unable’ (negation + mood) or *wedi* ‘perfect marker’ vs. *heb* ‘negative perfect marker’ (negation + aspect). The existence of complex heads of this type is likely to have favoured the same treatment for *cau*, suggesting that the tendency for acquirers
to make cross-category generalizations of this kind is a factor in language change.

(1) Mae ’r drws (yn) cau agor.
be.PRES.3SG the door (PROG) CAU open.INF
‘The door won’t open.’

(2) (Dy)di ’r drws (yn) cau agor.
NEG.be.PRES.3SG the door (PROG) CAU open.INF
‘The door won’t open.’

(3) Oddna neb cau dod efo fi!
was.3SG there no.one CAU come.INF with me
‘No one would come with me!’ (attested example)

(4) (Dy)di ’r car ddim cau cychwyn.
NEG.be.PRES.3SG the car NEG CAU start.INF
‘The car won’t start.’

References