The 21st Diachronic Generative Syntax Conference

Wednesday 5 June- Friday 7 June 2019
Memorial Union (MU) 085 (Union Stage)

Tuesday, June 4, 2019

8:30am – 9:30am Conference and Workshop Registration:
Lattie F. Coor Hall 4403, 976 S Forest Mall

9:00am – 5:00pm Pre-conference Workshop: Comparative Approaches to the Diachronic Morpho-Syntax of the Indigenous Languages of North and Central America.
Location: Lattie F. Coor Hall 4403 (See separate program.)

6:00pm – 8:00pm ***Workshop Closing Reception/Conference Opening Reception***
Location: TBA

Wednesday, June 5, 2019

8:00am – 9:00am Conference Registration: Memorial Union (MU) 085

8:45am – 9:00am Welcome: Johanna Wood, (Aarhus University)

9:00am – 10:15pm Chair: Johanna Wood, (Aarhus University)
Invited speaker: Will Oxford (University of Manitoba)
Probe respecification

10:15am – 10:30am ***Morning Break***

10:30am-12:15pm Session 1. Left Periphery (Romance), Chair: Mireille Tremblay (University of Montreal)
10:30am – 11:05am The rise of C-systems from Latin to Romance: a micro-diachronic analysis.
Valentina Colasanti (University of Cambridge)

11:05am – 11:40 am The diachrony of Old French verb-second: not just an inside job.
Espen Klævik-Pettersen (University of Oslo)

11:40am – 12:15pm Merged and moved - topics in Old French.
Christine Meklenborg Salvesen (University of Oslo)
12:15pm – 1:15pm  ***Lunch on your own***

1:15pm – 3:00pm  **Session 2. Left Periphery/Negation (Germanic)**, Chair: George Walkden, University of Konstanz

1:15pm – 1:50pm  Paratactic Negation in the History of German.
Elisabeth Witzenhausen (Ghent University)

1:50pm – 2:25pm  From '(nawther) ne...ne' to 'neither...nor'.
Ian Kirby (Harvard University)

2:25pm – 3:00pm  ‘Non-Coordinating und’ in the History of German.
Sophia Oppermann (University of Cologne)

3:00pm – 3:35pm  He then said...: (Understudied) deviations from V2 in Early Germanic.
Nicholas Catasso (Bergische Universität Wuppertal), Marco Coniglio (Georg-August-Universität Göttingen), Chiara De Bastiani (Università Ca' Foscari Venezia) and Eric Fuß (Ruhr-Universität Bochum)

3:35pm – 3:50pm  ***Afternoon Break***

3:50pm – 5:35pm  **Session 3. Number/Gender**, Chair: Judy Bernstein, William Paterson University

3:50pm – 4:25pm  The Evolution of Russian Paucal Numerals: From Dual to Paucal.
Tatyana Slobodchikoff (Troy University)

4:25pm – 5:00pm  The journey, not the endstate: finding innovation in the dynamics of L1A.
Naomi Lee and Ailís Cournane (New York University)

5:00pm – 5:35pm  Historical changes in Sub-word formation: The case of Arabic -a(t).
Myriam Dali and Éric Mathieu (University of Ottawa)

6:00pm – 8:00pm  ***Evening Reception***

    Appetizers: Cash Bar
    Graduate Hotel,
    225 E Apache Blvd, Tempe, AZ 85281
Thursday, June 6, 2019

8:30am – 9:00am  Conference Registration: Memorial Union (MU) 085

9:00am – 10:15am  Chair: William Kruger (Arizona State University)
Invited speaker: Bridget D. Samuels (University of Southern California)
Deep diachrony of human language & cognition

10:15am – 10:30am  ***Morning Break***

10:30am – 12:15pm  Session 4. Cycles of Change, Chair: Matthew Maddox
(University of Illinois at Urbana-Champaign)
10:30am – 11:05am  Formalizing the notions of re-bracketing and ad sensum agreement.
Timothy Gupton and Chad Howe (UGA Department of Romance Languages)
11:05am – 11:40am  A Plural to Singular Reanalysis Cycle.
Jonah Bates and Andrew McKenzie (University of Kansas)
11:40am – 12:15pm  Gothic Preverbs at the Juncture of Morphosyntactic Structure, Grammaticalisation, and
Cyclic Change.
Tamisha Tan (Harvard University, Nanyang Technological University)

12:15pm – 1:15pm  ***Lunch on your own***

1:15pm – 3:00pm  Session 5. Determiners and Pronouns, Chair: Paola Crisma, Università di Trieste
1:15pm – 1:50pm  Emphatic pronouns and the development of definite articles: Evidence for a layered DP
in early Romance.
Judy Bernstein (William Paterson University), Francisco Ordonez (Stony Brook University)
and Francesc Roca (Universitat de Girona)
1:50pm – 2:25pm  The Reanalysis of Schwa in Old French Determiners: Gender, Number and the Nominal
Spine.
Mireille Tremblay (University of Montreal) and Monique Dufresne (Queen's University)
2:25pm – 3:00pm  The genesis of third person pronouns: Insights from Late Latin and Old French.
Lieven Danckaert (CNRS), Liliane Haegeman (Ghent University) and Sophie Prévost
(CNRS)
3:00pm – 3:35pm  Gaps in the paradigm: Tracking the emergence of indefinite pronouns in Basque
Ricardo Etxepare (CNRS)
3:35pm – 3:50pm  ***Afternoon Break***

3:50pm – 5:35pm  **Session 6. Mood and Aspect**, Chair: Paul Kiparsky (Stanford University)

3:50pm – 4:25pm  Conditionals in Greek: operator movement & upward reanalysis from Attic to Modern. Katerina Chatzopoulou (Aristotle University of Thessaloniki)

4:25pm – 5:00pm  *Have*-doubling and the grammaticalization of *have* in West Germanic. Joanna Wall (Utrecht University and the Meertens Institute)

6:30pm – 9:30pm  ***Conference Dinner***

Gordon Biersch Brewery, 420 South Mill Ste 201, Tempe, AZ 85281

---

**Friday, June 7, 2019**

8:30am – 9:00am  **Conference Registration**: Memorial Union (MU) 085

9:00am – 10:15pm  Chair: Angela Schrader (Arizona State University)

**Invited Speaker**: Paul Kiparsky (Stanford University)

*Title TBA*

10:15am – 10:30am  ***Morning Break***

10:30am – 12:15pm  **Session 7. Grammaticalization**, Chair: Giuseppe Longobardi (University of York)

10:30am – 11:05am  Accusative Clitic/Null Object Variation in Spanish and the Object Agreement Cycle. Matthew Maddox (University of Illinois at Urbana-Champaign)

11:05 am – 11:40am  From clitic to affix: the loss of morphological parameters & grammaticalization in Western Iberian. Brian Gravely (University of Georgia)

11:40 am – 12:15pm  The Birth of an Epistemic Indefinite: *vaegy* in Transylvanian Hungarian. Tamás Halm and Ágnes Bende-Farkas (Research Institute for Linguistics (Hungarian Academy of Sciences))

12:15pm – 1:15pm  ***Lunch on your own***
1:15pm – 2:25pm  **Session 8. Agreement**, Chair: Elly van Gelderen (Arizona State University)

1:15pm – 1:50pm  Two Directions for Change: Case Studies in the Loss of Null Subjects.
Julianne Doner (University of Toronto)

1:50pm – 2:25pm  Greenberg’s Sixth Universal Revisited: The VSO/SVO Word Order Contrast in Early Egyptian.
Chris Reintges (Centre National de la Recherche Scientifique)

2:25pm – 2:45pm  ***Afternoon Break***

2:45pm – 4:30pm  **Session 9. General Change**, Chair: Éric Mathieu (University of Ottawa)

2:45pm – 3:20pm  Diachronic syntax is child’s play.
Paola Crisma (Università di Trieste), Cristina Guardiano (Università di Modena e Reggio Emilia) and Giuseppe Longobardi (University of York).

3:20pm – 3:55pm  Against mechanisms: towards a minimal theory of change.
George Walkden (University of Konstanz)

4:00pm – 4:30pm  **Business meeting**

We wish to thank the Department of English at Arizona State University, the Interdisciplinary Committee on Linguistics, and the Dean of Humanities for their support.
Much is known about the diachrony of agreement systems, particularly in regard to the agreement cycle (e.g. van Gelderen 2011) and changes in agreement strength (e.g. Roberts 1985). Both of these diachronic phenomena can loosely be said to involve the “rise and fall” of agreement. Not all agreement changes can be characterized in this way, however. Some polysynthetic languages, in which the agreement system carries the full weight of expressing grammatical relations, show a resistance to losing their agreement morphology. The Algonquian languages, for example, have maintained a system of rich subject and object agreement for well over 3000 years (Bloomfield 1946; Siebert 1967: 39; Pentland 1979: ii). However, even though the richness of the agreement system has remained stable, there have nevertheless been changes in the patterning of agreement. Less is known about the diachrony of such well-entrenched agreement systems. What kinds of changes take place? Does a generative analysis provide insight? Is anything at stake theoretically?

I approach these questions through the lens of the Algonquian language family of North America. The phonological and morphological history of this family has been reconstructed in great detail (Bloomfield 1925, 1946; Goddard 1967, 1969, 1974, 2007; Pentland 1979, 1999), giving us a rare glimpse of the grammatical history of a group of polysynthetic languages. I discuss two significant changes that have affected the patterning of Algonquian agreement. Both changes involve the well-known direct-inverse agreement pattern (Jacques & Antonov 2014). The first change is the emergence of the inverse in pre-Proto-Algonquian (Goddard 1974; Proulx 1982). The second change is a set of analogical extensions of the inverse that have taken place sporadically and to varying degrees across the family (Oxford 2014). These changes are descriptively complex, but a simple formal account is possible. I argue that all changes in the patterning of Algonquian inverse agreement reflect respecification of the probe on Infl—where “respecification” refers to a change in the uninterpretable features that the probe seeks to value. I further show that these respecifications follow pathways such as broadening, narrowing, and simplification, which are familiar from other kinds of language change.

**Agreement change 1: Emergence of the inverse.** Inverse agreement initially arose in pre-Proto-Algonquian as a consequence of a dramatic reanalysis that created a new paradigm of verb inflection. The reanalysis involved the transfer of possessed noun inflection to the verb (Goddard 1974; Proulx 1982), with possessor agreement reanalyzed as subject agreement and nominal gender/number marking reanalyzed as object agreement. The newly-created verb inflection could easily handle a 1→3 transitive form such as ‘we see them’, which shows the same inflection as a 1→3 possessed noun such as ‘our friends’. But a 3→1 transitive form such as ‘they see us’ would be problematic: it should show the same inflection as a 3→1 possessed noun, but no such possessed nouns exist (*‘their us’). The problem was solved by repurposing an existing passive construction: by adding a passive suffix to indicate the reversal of grammatical relations, a 3→1 transitive form can be expressed using the same agreement morphology as a 1→3 form. This “repurposed passive” is the construction now known as the inverse. I show that the emergence of the inverse construction can be understood simply as a consequence of respecifying the probe on Infl from [uϕ] to [uPers, uPart], with the more richly-specified probe establishing an agreement pattern that happens to mimic that of a passive. On a formal level, then, there was no “repurposing” of the passive, but rather just a change in the features of the probe.

**Agreement change 2: Analogical extensions of the inverse.** In the newly-created verb inflection, which is used only in main clauses, inverse agreement (i.e. the “repurposed passive”) appears in
all forms in which an SAP is acted upon by a non-SAP (e.g. 3.ANIM→1s ‘she sees me’, 3.INAN→1s ‘it sees me’, IMPERS→1s ‘people see me’). In conservative languages, the original verb inflection, which continues to be used in embedded clauses, does not show inverse agreement in any of these forms. Some languages, however, have begun to analogically extend the inverse pattern from the new verb inflection (called the “independent order”) to the original inflection (called the “conjunct order”) (Dahlstrom 1989; Oxford 2014; Despić & Hamilton 2017; Jacques & Antonov 2018). These extensions usually apply to some but not all of the forms in which a non-SAP acts on an SAP. As shown in Table 1 (Oxford 2014), the exact distribution of the extensions varies enormously, but not haphazardly: extensions of inverse marking display a strikingly regular “staircase” cline conditioned by the features of both the subject (animate, impersonal, inanimate) and the object (1st, 2nd, singular, plural).

<table>
<thead>
<tr>
<th>SUBJ→OBJ</th>
<th>INDEPENDENT ORDER</th>
<th>PA CONJUNCT ORDER</th>
<th>CENTRAL LANGUAGES</th>
<th>EASTERN LANGUAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.ANIM→1s</td>
<td>INV N N N N N N INV N N N N N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.ANIM→2s</td>
<td>INV N N N N N N INV N N N N INV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.ANIM→1p</td>
<td>INV N N N N INV N N INV INV INV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.ANIM→2p</td>
<td>INV N N N N INV INV INV INV INV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPERS→1s</td>
<td>INV N N N INV INV INV INV INV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPERS→2s</td>
<td>INV N N N INV INV INV INV INV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPERS→1p</td>
<td>INV N N N INV INV INV INV INV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPERS→2p</td>
<td>INV N N N INV INV INV INV INV</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Distribution of inverse marking in forms with SAP objects
(INV = form shows inverse marking; N = non-inverse form)

Building on my own recent work (Oxford 2017, forthc.), I will show that the entire set of changes represented in Table 1 can be understood as reflecting various respecifications of the probe on Infl. The respecifications follow several familiar pathways: narrowing the probe (e.g. [uφ] > [uPers]), broadening the probe (e.g. [uAddr] > [uPart]), simplifying the probe (e.g. [uPart.Pl] > [uPart]), and adding features to the probe (e.g. [uPers] > [uPers, uAddr]). Although all of the changes in Table 1 can be understood as having a functional/teleological motivation, the availability of a simple and intuitive formalization for the entire set of changes suggests that the changes nevertheless follow channels that are determined by the formal structure.

References


In recent years, much attention has been focused on neurodevelopmental changes that may have enabled the remarkable human capacity for language. For example, relative to other primates, the human brain has been claimed to show “neotenous” features: slow maturation, the retention of structurally immature neurons throughout life, and enhanced synaptic plasticity, which have been attributed to uniquely human changes to genes including *FOXP2*. This suite of characteristics has been claimed to increase our cortical complexity and thereby cognitive capacity by allowing a prolonged period of postnatal learning, such as is required to learn language, and enabling complex syntactic structures. However, recent fossil finds and new laboratory techniques have painted a new picture of our extinct relatives and the roles of genes implicated in language, suggesting that hypotheses regarding the role of neoteny and neuroplasticity in human cognitive evolution need refinement. In light of this, re-examine the deep diachrony of the human language faculty, combining evidence from paleogenomics, molecular biology, and developmental neuroscience.
**Summary:** Reanalysis of 1st or 2nd person marker from plural to singular reference is a multi-stage cycle involving two distinct changes: 1. dropping of the plural presupposition to become number-neutral and 2. pragmatic blocking of plural reference for the marker (resulting in default singular reference) through competition with a newly innovated competing presuppositional plural. These findings are consistent with general models of reanalysis and cyclical change in other functional domains.

**The Number Cycle:** Crosslinguistic comparison across several language families reveals a robust diachronic cycle for loss and renewal of plural/singular number contrast in person markers of the first and second person. Examples of this cycle can be found in Bantu, Dravidian, Indo-European, Semitic, and Mayan as in the examples below.

1. **Arabic prefix conjugation:** *ʔa-‘1SG’* → *ni-‘1PL’
   - New first plural innovated based on –u suffix seen in Semitic 2PL and 3PL

2. **Gondi (Dravidian):** *yaːm ‘1SG’* → *nαm ‘1PL(INCL)’
   - New first plural innovated based on plural –ot suffix

At first glance, the cycle appears simple. A plural person marker (αpl) is reanalyzed as singular (αsg) and a new plural person marker is innovated based on αsg plus some available pluralizing morphology (αsg+PL) resulting in a renewed number contrast between the original marker, singular αsg, and plural αsg+PL. The only reanalysis is αpl > αsg. This is a cycle in that the end result sets the stage for changes to begin again.

**Intermediate Stages:** However, some languages appear to have only completed portions of the cycle, providing evidence for more steps in the change. First, languages like Yao (Bantu-P21) below, which have not renewed the number contrast are not left with a singular marker but a general one, which can reference both singular and plural entities. In fact, it would be quite a stretch to propose a language that has a singular marker but no plural.

3. **Proto-Bantu *mu- ‘2PL’ > Yao (P21) mu- ‘2’**

   Another familiar example is English ye ‘2PL’ to Modern English you ‘2’. The existence of systems like these implies two things, that the first stage of the full cycle is not a reanalysis from PL > SG, but from plural to NUMBER-less reference, and that the second shift from general > SG is dependent on the language optionally innovating a new plural. That makes at least three stages: PL > general, innovation of new plural, and general > SG. Is it possible that the singular reading itself is somehow dependent on a contrast with a plural?

Even in languages that do renew the number contrast, the general marker is often only pragmatically constrained to singular readings. The force of that pragmatic constraint extends along a spectrum that is language-dependent. On the weaker side, many Mayan languages, including Tzeltal, allow the bare 1st person form to have either a general or singular interpretation, depending on discourse context. On the stronger side are American English dialects that have innovated a new second person plural *ya’ll* where you is heavily restricted to singular reading. These facts divide the ‘third’ stage, general > SG, into two steps; one in which the singular meaning...
is on a spectrum of pragmatic force, and another where it is fully confined to singular interpretation.

**Stages of the Plural to Singular cycle:**
1. PL marker > general (NUMBER-less) marker
2. Innovation of new plural marker
3. Spectrum of degrees of pragmatic forcing of general marker to SG readings
4. Full pragmatic force: general marker completely confined to SG interpretation

**Pragmatic Singular:** If singular meaning depends on having a contrast with a plural in person marking systems, is it possible that true singular person marking reference may not exist at all? Outside the realm of person marking contrasts, recent work on the semantics of number have come to this exact conclusion, claiming that “competition between singular and plural forms drives their interpretation in a process that intertwines semantics and pragmatics” (Farkas & de Swart 2010:66). Specifically, the singular feature is ‘semantically weak’, contributing no semantic content, while the plural is ‘strong’, having true presuppositionally plural syntactic structure. This results in a strong/weak Horn scale <PL,SG> where the singular is weak and the plural is strong. For person markers this means that ‘singular’ markers reference arguments with no number feature while plural morphemes are used when plural content is present. The more general ‘singular’ is blocked from having reference both singular and plural arguments then by an effect of Maximize Presupposition, namely that since a plural structure has more presuppositional content, it should be used with plural entities. This creates a pragmatic effect where the more general marker is pragmatically blocked from plural reference through competition with a contrasting presuppositional plural. The strength of this blocking is language- and context-dependent.

**Plural to General as Presupposition Dropping:** Applying this to the stages of our cycle above, we conclude that the first stage involves the dropping of a plural presupposition, while the last two involve pragmatic blocking. This story has significant support in literature surrounding linguistic reanalysis. To begin with, dropping unbacked presuppositions is claimed to be one of the primary mechanisms of reanalysis (Eckardt 2011, Schwenter & Waltereit 2010). Additionally, evidence exists that the strategic use of plural presupposition for sociolinguistic reasons (leading to them being unbacked and subsequently dropped) is a common factor in the plural to singular changes for which we have historical context, a process referred to as plurification (Heine & Song 2011). For example, the ‘royal we’ seeks to associate the speaker with an unseen group excluding the hearer. The hearer is then faced with accommodation of the plural presupposition or reanalysis of ‘we’ as number-neutral allowing singular reference. Finally, the directionality of the change, from strong (plural) to weak (general) on a Horn scale is consistent with other proposed diachronic cycles such as the Jespersen cycle in negation and the progressive to imperfective cycle in aspect (see discussion Deo 2015).

**SELECTED REFERENCES**
Emphatic pronouns and the development of definite articles:
Evidence for a layered DP in early Romance

Judy B. Bernstein, William Paterson University
Francisco Ordóñez, Stony Brook University
Francesc Roca, Universitat de Girona

The definite articles in modern Romance languages like French and Spanish developed from the Latin distal demonstrative *ille* and display *l-* forms. However, definite articles also developed from the Latin emphatic pronoun *ipse* in other Romance varieties such as Sardinian and Balearic Catalan (Aebischer 1948, Vincent 1997, Ledgeway 2012). The articles that descended from *ipse* display *s-* forms and are illustrated for Sardinian in (1a) and Balearic Catalan in (1b) (Sardinian forms from Jones 2003, his (3)).

(1) a. *su babbu de Maria* (Sardinian)
   the father of Maria
b. *ses taules* (Balearic Catalan)
   the.pl tables

In this paper we study the historical evolution of forms descended from *ipse* and also how they competed with *ille*. In Classical Latin *ipse* was used as an adjunct emphatic pronoun (Hertzenberg 2011, her (1)):

(2) Ancillae tuae credidi [...] tu mihi non credis *ipsi.* (Latin; Cicero, *Orat.* I)
   ‘I believed your handmaid and you won’t believe me (myself)?’

This marker of emphasis started to be used as a demonstrative in Latin (Aebisher 1948) after the loss of first person proximal *hic* and the expansion of second person medial *iste* (Aebisher 1948, Vincent 1997). As the deictic feature weakens, *ipse* starts to appear as a determiner and eventually yields the modern usage of *s*- as an article in Sardinian. The evolution of *ille* and *ipse* seems to correspond to the schema below, illustrating a process of grammaticalization in which a phrasal XP (emphatic pronoun and demonstrative) evolves into a head X (article) (see Giusti 2001, Iovino 2015, Renzi 1992, van Gelderen 2011).

(3) a. *ipse* (emphatic pronoun) \rightarrow *ipse* (demonstrative) \rightarrow *ipse* (article)
   b. ------------------ *ille* (demonstrative) \rightarrow *ille* (article)

From the representation of the evolution above one might think that the modern *s*- definite article is no different than the *l*- definite article apart from its form, which would simply boil down to a diachronic quirk of etymology and a synchronic quirk of phonology. However, relevant examples show that even in medieval Latin *ille* and *ipse* did not have the same use (Vincent 1997, Ledgeway 2012). Trager (1932) observed these differences in Latin in *Peregrinatio Silviae ad loca sanctos* (see also Ledgeway 2012: 92):

(4) ergo quarta pervenimus in *summitatem illam* montis Dei sancti Syna, ubi data est lex
    thus fourth we arrived in summit that mount God saint Sinai, where given is law
... tunte Deo persubissetem in *ipsa* *summitate* (IV-V)
   was given God reach we in *that-very* summit

In the first example *illam* refers to a unique referent; in the second example *ipsa* is discourse oriented and anaphoric to the first one. Thus these demonstrative-like elements have different functions and semantic features.

The process of grammaticalization schematized in (3) oversimplifies the historical development of what will become *ipse*-derived articles. In fact, we find evidence that Latin *ipse* and *ille* developed in different stages in accordance with the semantic properties of the noun.

In Old Sardinian and Old Catalan, *ipse*-derived articles tend to be absent with unique nouns in (5) *(iudice 'judge'; clesia 'church')* as well as those following a preposition in (6) *(ab senyor 'with master'; a mestre 'to teacher'; Old Sardinian examples from Cennamo et al. 2018):

(5) a. Et ego pregai a *iudice* et *assos curadores* (Old Sardinian; CSMB; XII-XIII)
   and I prayed.1sg to.DOM judge and to.DOM-the administrators

'And I begged /asked the judge and the administrators'
b. et levait clesia Arzoco
and took.3sg church Arzoco
'And the church took Arzoco'

(6) a. avant no gós · affermar o procurar de metre alcun misatge ab senyor (Old Catalan; Manual de consells de la Ciutat de Valencia; XIV)
before not dared to confirm or procure to put some message with master
'bfore s/he dared to confirm or procure giving some message to the master'
b. Axi mateix, vos tretemet una letra qui · s’adreça a mestre Jaconi (Epistolari Ferran I; XIII)
this way, we send a letter that is directed to teacher Jaconi

In contrast, ipse-derived articles appear with individuated nouns, as in (7) and (8) (Ledgeway 2012):

(7) ...assu frate Gitimel to.DOM-the brother Gitimel
'and his brother Gitimel'
(8) ...e · gat com hac lepada s'escudella (Old Catalan; Matrimoni desavingut, XIV)
and the cat as had licked the porridge

This historical development calls for a syntactic explanation. In particular, what distinguished the early nouns appearing with ipse from the ones that didn't? We argue that the semantic nature of ipse (emphatic), as well as the nature of the nouns it appears with, is tied to a higher position in the DP structure, what we label DP1. This contrasts with the semantic nature of ille (anaphoric demonstrative). Its absence of emphasis naturally ties it to abstract nouns and uniques. We propose that this corresponds to a position lower in the DP structure, which we label DP2. We schematize the multilayered DP as follows:

(9) [DP1 ip(se) · · · · · · [DP2 il(le) [NP]]]

Sardinian generalized the ipse article in DP1 to the contexts without an article, namely those corresponding to DP2. Standard Catalan is interesting in that it eventually lost the ipse-derived articles and adopted the ille forms for reasons that remain unclear but might relate to pressures from nearby Spanish and the literary register.

Interestingly, Balearic Catalan implements both articles contemporaneously. It reserves ipse-derived articles for individuated DPs and l- articles for non-individuated DPs, as illustrated by the interpretations associated with the noun muntanya in (10a) vs. (10b):

(10) a. Sa · muntanya que veus · és molt alta. (Balearic Catalan)
the.F mountain that see.2PSG is very tall
'The mountain that you see is very tall.'
b. Anem a · la · muntanya.
go.1PPL to the.F mountain
'Let's go to the mountain.'

Furthermore, in some restricted cases, we find both articles spelled out with a single noun:

(11) es l’amo, es l’avi · (s- + l-)
'the mister', 'the grandfather'

The structure in (9) extends beyond ipse and its descendants. Picard has developed a system similar to the one in Balearic Catalan, in which the higher DP1 is occupied by the descendant of Latin emphatic ecce. Lower down, DP2 is reserved for the ille-derived forms. As in Balearic Catalan, DP1 and DP2 are associated with different interpretations, shown in (12a), and can even co-occur, shown in (12b):

(12) a. che monde vs. le monde (Boulogne Picard; Haigneré 1901)
'the earth' vs. 'the entirety of the world'
b. ch’l’esprit, ch’l’infant · (ch- + l-)
'the spirit', 'the child'

Finally, our structure in (9) above makes sense of the absence of varieties in which emphatic demonstratives are employed exclusively with non-individuated DPs. In other words, the descendants of emphatic demonstratives are always initially merged in DP1.
Nicholas Catasso, Bergische Universität Wuppertal  
Marco Coniglio, Georg-August-Universität Göttingen  
Chiara De Bastiani, Georg-August-Universität Göttingen  
Eric Fuß, Ruhr-Universität Bochum

He then said…: (Understudied) deviations from V2 in Early Germanic

The structure of the left periphery in the Old Germanic root clause has been a much debated topic in the generative literature of the last three decades (cf. e.g. Lenerz 1984, Fuß 2003, Axel(-Tober) 2007, Speyer 2008, Walkden 2014, Petrova 2012, Hinterhölzl 2017, Speyer & Weiβ 2018 for German, Petrova & Speyer 2011, van Kemenade 2012, van Kemenade & Westergaard 2012, van Kemenade & Los 2018, Kroch & Taylor 1997, 2000 for English). Old High German (OHG), Old English (OE) and Old Saxon (OS) present variation in the position of the finite verb, displaying V2/V1, as well as V3 or verb-late(r) word orders.

In this paper, we discuss a non-correlative V3-pattern in OHG, OE and OS which has so far not been considered independently, namely one in which the prefIELD hosts an XP in clause-initial position followed by the originally adverbial element OHG do / OE pal/nonne / OS tho (lit. ‘then’, ‘there’), which typically immediately precedes the finite verb, as in (1).

This construction, which has generally been considered to instantiate the rarely attested XP > adverb > Vfin pattern (cf. e.g. Axel 2007, Walkden 2014), can be taken to be native in West Germanic languages, since – for example – it is attested independently of the Latin source in OHG (1a) and in non-translational texts as in OE (1b).

OHG do has been the object of a number of studies in configurations in which it occurs either clause-initially in V2 clauses or middle-field-internally in V1 clauses, its structural position being due to discourse- and information-structural principles (Betten 1987, Simmler 1998, Donhauser & Petrova 2009). Thus, the use in (1a) has been hastily considered as a deviation from the more frequent do-V2 and V1-do patterns (but see Ruhrs 1897 for an exception). For OE, it has been established that the originally temporal adverbial þa triggers categorical V2 when it occurs at the beginning of a main clause, whereas it serves to separate the topic from the focus domain when it is found in clause-internal position of subordinate clauses (cf. Los & van Kemenade 2006, van Kemenade 2009); moreover, it probably covers the same function when it is realized in the same position of matrix clauses (cf. Trips & Fuß 2009). The structure in (1), however, has neither been studied as an independent V3-construct nor from a comparative Old Germanic perspective in previous research.

In a first pilot study, the OHG translation of Tatian’s gospel harmony (830, East Franconian) and the OE text of the Blickling Homilies (10th c., West Saxon) have been primarily considered. The data were extracted from the Referenzkorpus Altddeutsch (Donhauser et al. 2018) and the YCOE corpus (Taylor et al. 2003), respectively. A closer look at the contexts in which this pattern occurs in the two languages reveals the following scenario: in the OHG text, in which this structure typically occurs in dialogic/interactional frames, the clause-initial element is a pronoun in 30 out of 31 occurrences, and the preverbal do seems to mark turn-taking, the subject of the corresponding clause almost systematically differing from that of the previous one, as in (2).

In OE, however, the referents preceding the discourse marker are shifting topics in half of the cases, while they are better analyzed as Continuity Topics in the remaining cases (cf. (3)).

To account for this variation, we will explore the possibility that:

(i) OE root-clause þa performs a discourse function similar to that found in subordinate structures, in which this element serves to mark the boundary between topic and focus domains. This is corroborated by the fact that the finite verb is often located in a lower structural position in the attested data (cf. (4));
(ii) OHG *do* is fronted to the prefield of the clause together with a shifting topic, giving rise to an apparent V3-configuration. The OHG pattern is attested – with very similar functions – in the Old Saxon text *Heliand* (first half of the 9th c.). More generally, we will show that these patterns are possibly Germanic in nature, as will be shown based on data from the Gothic Bible (4th c., cf. Klein 1994).

The adverbial elements included in the construction are etymologically related to the demonstrative paradigm and are used to mark a sequence of actions or events in the narration (cf. van Kemenade & Los 2006, Donhauser & Petrova 2009). However, it will be contended that the elements *do, þaþonne* and *tho* are formally distinct from their referential-deictic counterparts. We propose that OE *þa* is first-merged in the head of ΣP, a projection located in the higher middle field whose specifier is reached by shifting and continuing topics by movement (cf. van Kemenade 2009).

Moreover, we will argue that both the position of the verb and the further development of the construction under scrutiny are subject to different constraints in OHG, OS and OE.

**Examples**

(1) a. *sie theo antalengitun imo. neín*  
   they *do* answered *him.*DAT no
   ‘They said to him: ‘No’’

   **Respondentur ei: non (T. 337, 10-11)**  
   (coblick.HomS_8_[BIHom_2]:15.17.186)

(2) context:

   *tho quad In ther heilant, uuâr uuâr quidu ih Iû er thanne abraham uuari er bim ih. tho namun sie steina thaz sie vvurphin In Inan*
   ‘Jesus said to them: ‘Verily, verily, I say to you, before Abraham was, I am’. Then they took up stones to cast at him.’

   **clause: her theo barg sih …**
   he (= Jesus) *do hid* REFL.
   ‘Jesus hid himself’ *(T. 220, 11)*

(3) context:

   & *þa æt nehstan he let his lichoman on rode mid næglum gefæstian, & deap *he gebrowode for us, forjon þe he wolde us þæt ece lif forgifan.*
   ‘Then later he let his body be nailed on the cross, and he suffered death for us, because he wanted to give us the eternal life.’

   **clause: & he þa onsende his þone wuldorfaestan gast to helle grunde**
   and he *pa sent* his DET glorious spirit to hell abys
   ‘and he then sent his glorious spirit to the abyss of Hell.’

   *(coblick.HomS_26_[BIHom_7]:85.30.1056)*

(4) *Baloham ðonne fulgeorne feran wolde*

   **Baloham ðonne gladly go.INF wanted**
   ‘Ballam then very much wanted to go over’ *(cocura,CP:36.255.22.1674)*

**Selected References**

Conditionals in Greek: operator movement & upward reanalysis from Attic to Modern

Katerina Chatzopoulou, Aristotle University of Thessaloniki, National State Foundation

This presentation examines the expression of conditionals in Greek based on quantitative data from three major stages of vernacular Greek (Attic Greek, Koine and Late Medieval Greek) and qualitative evidence from Homeric Greek to Standard Modern. The history of Greek conditionals provides individual support on the analysis of Bhatt & Pancheva (2002, 2006, cf. also Haegeman 2003, 2006) that conditional clauses involve leftward movement of a world operator, in particular an operator located in IrrealisP (Cinque 1999), as argued in Danckaert & Haegemann (2012). In the case of Greek this operator is the morpheme ἄν /an/, already analyzed as a modal universal quantifier for Attic in Beck et al. (2012) with conditional formation requiring leftward operator movement. The Attic Greek ἄν /an/ was used both (i) as a complementizer introducing conditional antecedents with subjunctive and (ii) as a modal particle used to form potential indicative or potential optative in the apodosis (cf. Goodwin 1890, Smyth 1956, Lightfoot 1975, also with infinitives or participles). In its use as a particle introducing conditional antecedents the Attic Greek ἄν /an/ was in complementary distribution with the particle εἰ /e:/, with the main generalization being that ἄν /an/ selected for subjunctive mood, while εἰ /e:/ selected for indicative or optative (table 1). By Koine Greek the potential function of ἄν /an/ is much less attested in apodosis of conditionals, following also the restructuring of the mood system after major phonological changes and the isomorphic emergence of the Nonveridicality syntactic projection headed by ἰνα/ίνα/ (Chatzopoulou 2012, 2019). In Koine Greek ἄν /an/ began to generalize across different types of conditionals, since the phonological contrast between indicative and subjunctive was lost (Joseph 1990, Horrocks 2010), while also optative turns obsolete in these functions. Consecutio modorum (sequence of moods/mood assimilation) and consecutio temporum (sequence of tenses) are further lost (Moulton 1908, Meillet 1975). By Late Medieval Greek εἰ /e:/ > ἰ/ (after iatricism) is rare, often fossilized and in free variation with ἄν /an/ (ex. 1-2), while ἄν /an/ is statistically dominant, covering all the semantic functions of previous stages. The situation is stabilized in Modern Greek, where εἰ /i/ is no longer attested and all conditional types are formed through ἄν /an/ and various tense/aspect morphological combinations, among which are ‘fake’ past (Iatridou 2000), the perfective nonpast, the future morpheme θα /θα/ also as modal/epistemic marker (Giannakidou 2009, 2014) and other finite forms (table 2).

This presentation analyzes this change as upward reanalysis of the particle ἄν /an/, a leftward movement of the world operator involved in conditional formation, as an effect of structure economy in the sense of van Gelderen (2001, 2004) and Roberts & Roussou (2003). This process appears to have started already in Attic Greek given that εἰ /e:/ represents 34% of all conditional particles and in spite of its raise in Koine, by Late Medieval it is reduced to 10% in the examined vernacular texts (table 3). In Standard Modern Greek εἰ /i/ is no longer attested. This change in conditional formation is further examined along with changes in the syntactic status of negators, from phrases in Attic Greek to heads in Late Medieval and Modern Greek and general changes in the left periphery of the clause in conditionals. It is plausible to consider that this change in syntactic status of negators facilitated ἄν /an/—which remained phrasal from Attic to Modern Greek—to become generalized in all conditional environments. Crucially, in Modern Greek the particle ἄν /an/ never cooccurs with the particle va /να/, which along with the ban of NEG2 μή /μ/ from conditionals, means that they compete for the same syntactic position, one linked to the Irrealis Phrase or Nonveridicality Phrase.

Table 1. Distribution of conditional particles ἄν /an/ & εἰ /e:/ in Attic Greek (5th-4th century BC)

<table>
<thead>
<tr>
<th>protasis</th>
<th>Apodosis</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ἄν /an/</td>
<td>subjunctive</td>
<td>Future indicative</td>
</tr>
<tr>
<td></td>
<td>subjunctive</td>
<td>Nonpast indicative</td>
</tr>
<tr>
<td>εἰ /e:/</td>
<td>indicative</td>
<td>Any mood</td>
</tr>
<tr>
<td></td>
<td>optative</td>
<td>Past tense indicative (+ἄν /an/)</td>
</tr>
<tr>
<td></td>
<td>optative</td>
<td>Optative + ἄν /an/, nonpast indicative, potential infinitive or participle</td>
</tr>
<tr>
<td></td>
<td>Past tense indicative</td>
<td>Past tense indicative + ἄν /an/</td>
</tr>
</tbody>
</table>
Table 2. The conditional particle αv/an/ in Modern Greek (today)

<table>
<thead>
<tr>
<th>protasis</th>
<th>apodosis</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfective nonpast</td>
<td>Imperfective nonpast/ Future θα</td>
<td>Future more/less vivid</td>
</tr>
<tr>
<td>Perfective/imperfective nonpast</td>
<td>Imperfective nonpast/ Future θα/NA</td>
<td>Iterative in nonpast</td>
</tr>
<tr>
<td>Any tense/aspect combination</td>
<td>Any tense/aspect combination</td>
<td>Realis</td>
</tr>
<tr>
<td>Imperfective past</td>
<td>Imperfective past</td>
<td>Iterative in the past</td>
</tr>
<tr>
<td>Imperfective past/nonpast</td>
<td>Perfective/imperfective past/nonpast</td>
<td>Generic/habitual</td>
</tr>
<tr>
<td>(Fake future θα) + fake past</td>
<td>Fake future θα + fake past</td>
<td>Counterfactual</td>
</tr>
</tbody>
</table>

Table 3. Conditional particles from Attic to Late Medieval Greek

<table>
<thead>
<tr>
<th>Greek period</th>
<th>Text</th>
<th>ε(δ)(ν)</th>
<th>ει</th>
<th>Total</th>
</tr>
</thead>
</table>
| Attic Greek (Lysias, Plato Republic, Aristotle De anima) 5th-4th c. BC | Εάν γιρέω δίκαια, τίποτες μη με ποιήσις;  
If search.IN.1SG rightfully nothing NEG2 me.ACC do.PNP.2SG | 1955    | 1027 | 2982  |
| New Testament, Epictetus Dissertations 1st c. AD | Εάν γιρέω δίκαια, τίποτες μη με ποιήσις;  
If search.IN.1SG rightfully nothing NEG2 me.ACC do.PNP.2SG | 1320    | 1199 | 2519  |
| Digenis Akritis, Chronicle of Moreos, Floris and Blancheflour 14th c. AD | “If I am searching rightfully, do not harm me in any way.” (Floris and Blancheflour 1390, Wagner 1870) | 267     | 32   | 311   |

(2) Ει δε γιρέω δίκαια, κόψε γυρεύω την κεφαλήν μου.  
If 2P search.IN.1SG wrongfully cut.IMP.2SG the.FEM.ACC head.ACC my.GEN | 1395    | 1027 | 2422  |

“If I am searching wrongfully, cut my head.” (Floris and Blancheflour 1391, Wagner 1870)

ABBREVIATIONS


REFERENCES

Introduction. Sentential complementation in Latin represents one of the biggest structural differences with respect to Romance. Although there is already extensive evidence in early Latin for the structure of the CP (Vincent 1999), the clausal domain in Romance exhibits a much more complex structure. In particular, old Indo-European strategies of marking clausal embedding (i.e. Accusativus cum Infinitivo, AcI henceforth) in Latin have been lost in favour of more complex systems of complementation in Romance (i.e. multiple finite C-systems). In this paper I introduce new micro-parametric evidence from Italo-Romance in order to give a more fine-grained understanding of the huge structural change in sentential complementation that took place during the development from Latin to Romance. The precise working of multiple C-systems in Italo-Romance proves crucial in outlining the micro-diachronic path zero > multiple > single from Latin to Romance, which I situate within an acquisition-based approach to diachronic change (Lightfoot 1999; i.a.).

Complementation in Latin. Latin underwent an earlier archaic stage predating the emergence of a fully-fledged CP structure (Ledgeway 2012). The AcI, which seems not to involve any overt embedded CP structure (Calboli 2009), can be considered the unmarked strategy since it is found in every context: it can be the complement of say-verbs (e.g. ‘to say, to believe, to assume’), regret-verbs (e.g. ‘to regret, to know, to hate, to like’), and want-verbs (e.g. ‘to want, to wish’). However, in Latin we find the AcI co-occurring alongside indicative complements introduced by quod selected by factive verbs, indicative complements introduced by quia selected by declarative verbs, and subjunctive complements introduced by ne/ut selected by volitive verbs.

Finite complementation in Romance. In typological terms, we can divide Romance languages into two big groups on the basis of synchronic and diachronic evidence on complementation. The first group contains languages which are characterised by multiple finite C-systems (i.e. some early Romance varieties, most of southern Italo-Romance, and Romanian). The examples in (1) show the triple C-system of the Italo-Romance variety early Salentino (Ledgeway 2005), in which the complementizer ca is selected by the declarative verb penza ‘s/he.thinks’ (1a), and the complementizer che is selected by the factive verb fa parere ‘it.makes to.seem’ (1b), introducing realis and factive indicative complements respectively. By contrast, the complementizer cu is selected by the volitive verb volce ‘wanted’ and introduces an irrealis subjunctive complement (1c).

(1) Early Salentino (Ledgeway 2005:368:369)

a. Penza ca illu non è solo al mondo.
   ‘He.thinks that he not is alone in the world.’

b. Fa parere che lo sole et la luna et le altre pianete et le stelle girano.
   ‘It.makes to see that the sun and the moon and the other planets and the stars revolve.’

c. Illu volce cu nuy sappessemu...
   ‘He wanted that we should know.’

The second group contains languages which have single C-systems (i.e. most of early and modern Romance varieties, and some southern Italo-Romance varieties). Most modern Romance languages and many modern southern Italo-Romance varieties have single C-systems. For example, the modern Cosentino variety has only one complementizer (i.e. ca) which is selected by both say-verbs (3a) and want-verbs (3b). However, the early Cosentino variety had a dual C-system: the complementizer ca selected only by both declarative verbs (2a) and the complementizer chi selected only by volitive verbs (2b). The contrast between (2) and (3) demonstrates that the single C-system of modern Cosentino is the result of the generalisation of the lexicalisation of ca to the detriment of the complementizer chi (Ledgeway and Lombardi 2014:40).
Early Cosentino

a. Un pienzu ca vi canuscia buonu.
   ‘I don’t think that he knows better.’

b. Vulìa chi m’ accompagnassa a ra casa.
   ‘I wanted that he would accompany me home.’

Modern Cosentino

a. A dittu ca sgarrati.
   ‘He has said that you are mistaken.’

b. Idda vo ca ci fazzu na picca ‘i spisa.
   ‘She wants that I do a bit of shopping.’

A micro-diachronic analysis. Data from Latin and early/modern Romance can be reinterpreted through the diachronic path zero \(\rightarrow\) multiple \(\rightarrow\) single. The Zero stage refers to the absence of finite complementizer systems. This correlates to the presence of the AcI in early/classical Latin, which is productive in every context: the AcI can be a complement of say-, regret-, and want-verbs. The Multiple stage illustrates the situation found in classical/late Latin, in which subordinate clauses are mostly introduced by an overt complementizer (i.e. quod/quia/ut). This is what we find in most multiple and dual C-systems in Italo-Romance. Similar to Latin, early and modern southern Italo-Romance varieties and Romanian possess triple and dual C-systems (but not AcI). The Single stage illustrates the situation found in the majority of modern standard Romance varieties (e.g. Italian, French, etc.) and some modern southern Italian dialects (e.g. Cosentino, Salentino, etc.). Both these language types have single C-systems. Specifically, the majority of Romance show today finite C-systems with only one complementizer. At the same time, many modern southern Italo-Romance varieties show generalisation of one complementizer to the detriment of another, hence single C-systems. This will be analysed in detail during the talk assuming Chomsky’s (2008) idea and reinterpretation of the relation between the functional heads C and T, coupled with Ouali’s (2008) idea that C may parametrically choose to keep features that could potentially be shared or donated to T. It seems, then, that the higher head C is able to determine the content of I in different parametric ways, thus explaining observed microparametric variation in the C-systems.

An acquisition-based approach to zero \(\rightarrow\) multiple \(\rightarrow\) single. I situate the zero \(\rightarrow\) multiple \(\rightarrow\) single diachronic path within an acquisition-based approach to diachronic change. Following Biberauer’s (2017) Maximise Minimal Means idea, the learner makes maximal use of minimal means. Specifically, I assume that the 3rd Factor of language design (viz. non-language-specific cognitive optimisation strategies; Chomsky 2005) represents a general cognitive economy principle which has the two linguistic manifestations, namely Feature Economy and Input Generalisation (Roberts & Roussou 2003). The interaction between the two creates the NONE>ALL>SOME learning path. With this in mind, the zero \(\rightarrow\) multiple \(\rightarrow\) single path reflects the ability of C to host NONE>ALL>SOME features diachronically. NONE is the default choice as the acquirer experiences the non-existence of [F] (i.e. no modality features on C in early/classical Latin). ALL represents a generalisation of [F] following the acquirer’s initial ‘ignorance’ (i.e. all modality features on C in Latin quod/quia constructions and multiple C-systems in Italo-Romance), and SOME is the situation in which [F] is present in some domains and absent in others (i.e. some features on C in single C-systems).

Selected references


Fundamental evidence for investigating the syntax of contemporary languages is provided by native speakers’ grammaticality judgements, in particular by the contrast between grammatical and ungrammatical structures: virtually all the syntactic analyses proposed in a generative framework crucially rely at some point on access to this type of contrasts. This is one of the main reasons why, in generative syntax, explanatory adequacy has been much more effectively pursued than descriptive adequacy, given that evidence from ungrammatical structures is unlikely to be accessed by a child and then to be part of a realistic model of language acquisition. However, the main goal of generative linguistics ought to be precisely that of explaining how knowledge of language arises in the mind of speakers on the basis of fragments of E-language.

Now, when investigating ancient or mediaeval languages, the underlying grammar must be reconstructed on the basis of what qualifies as a random sample of E-language (a collection of extant texts), which may be non-uniform (different dialects, genres, registers, etc.), variously corrupted (scribal errors, late copies, physical damage to the manuscripts, etc.), and generally with no explicit indication as to which structures are ungrammatical (apart from the accidental existence of some grammatical treatise). Though this situation may not be the ideal one for the linguist, it resembles some conditions of first language acquisition and may represent a good testing ground for models of the latter. In this presentation, we work out a practical system of parameter setting which enables one to decide parameter values on the basis of closed and limited corpora, with interesting consequences for the study of language history and language acquisition. This model will be tested and validated through OE, ME and Ancient Greek datasets.

The model is based on the following tenets:

1. UG (S_0) does not contain a list of parameters, but at most a set of formats for possible parametric choices (schemata: Longobardi 2005, 2018), as well as some invariant principles;
2. only positive evidence is used to set parameters: simple existential assertions of the form ‘The property X occurs in the E-language Y’;
3. also, no parameter is set on the basis of peripheral patterns/structures.

In the classical theory of parameters, each parameter may be associated with a cluster of manifestations, with different degrees of saliency (Rizzi 1977, Taraldsen 1978, Chomsky 1981). Our experiment supports this approach (pace Newmeyer 2005). The clustering structure of parameters predicts three types of linguistic facts, of which only the first is normally visible in samples of E-language:

a. the grammaticality of some core manifestations
b. the grammaticality of some peripheral manifestations
c. the ungrammaticality of some other structures

All this evidence can be used by a linguist working with a native speaker, but is presumably not available to the historical linguist, nor, in a realistic model of acquisition, to the language learner.

For the present purposes, we selected and described a dataset of 24 parameters, constrained by the schemata of point (1) above, that define some broad (co-)variation patterns in the subdomain of the syntax of nominal determiners. This was done by comparing 87 modern language varieties from at least 19 different families, using all the evidence available to linguists, including grammaticality judgements by speakers and rare patterns. This process lead to the formulation of simple existential assertions (positive evidence only) that define the manifestations of the parameters considered, on the basis of synchronic, diachronic and typological evidence.

Then, we made some preliminary hypotheses about which manifestations are more salient, so as to tentatively meet (3) (thus obtaining a core subset of manifestations, called ‘the restricted list’).

We applied this restricted list of manifestations to corpora from 6 historical varieties: Beowulf, OE prose, ME (all supported by annotated corpora) and Homeric Greek, Classical Attic, New Testament Koiné (not provided with annotated corpora).

As an example, take our Parameter 14, which distinguishes languages that grammaticalize definiteness from those which do not. The table in (4) shows which manifestations are relevant to
set the parameter for the varieties considered, compared to their modern descendants, English (E) and Greek (Grk). This parameter has 6 manifestations, 3 of which (un-shaded boxes) can set the parameter on the basis of the inspection of relatively limited portions of text; the other 3 (shaded boxes) may be occasionally encountered. On the basis of our experience with speakers of modern languages, the empirical evidence linked to these manifestations is hard to elicit: thus, their rare occurrence in the texts is expected. Though it is improbable that the speakers actually need/use them to set the parameters, they are predicted by the more salient manifestations and represent in our experiment the classical *poverty-of-stimulus* situation (Chomsky 1980):

(4) Parameter 14: grammaticalization of definiteness

<table>
<thead>
<tr>
<th>a. Det on entities inferred from the previous context, though not mentioned</th>
<th>Beow</th>
<th>OE</th>
<th>ME</th>
<th>E</th>
<th>Hom</th>
<th>CIG</th>
<th>NTG</th>
<th>Grk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y &gt; +</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>b. Det on <em>situative Unika</em> Y &gt; +</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>c. Det on singular kind names Y &gt; +</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Det on mass/pl kind names Y &gt; +</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Bare nouns bounded Y &lt; -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Bare nouns definite Y &lt; -</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(5a) & he geseah twa scipu standeande wið þæne mere. *Da fisceras* codon & wohson heora nett.

(ID cowsgosp, Lk [WSCp]: 5.2.3895)

b. καὶ ἐδὲν πλοία δῷο ἑσπῆτα παρὰ τὴν λίμνην, *oi δὲ άληζες αὐτῶν ἀποβάντες ἔπλωσαν τὰ δίκτυα.* 'and he saw two boats standing by the lake. (but) The fishermen had gone and washed their nets.'

(6a) Aledon ða *leofne peoden* beaga bryttan on bearm scipes (ID cobeowul, 4.34.32)

‘they laid down *the beloved king*, the giver of rings, in the center of the ship’

b. στήσαμένη μέγαν ἱστόν ἐνὶ μεγάροις ὑφαίνει [...] ἔνθα καὶ ἡματίῳ μὲν ϕαινόκεκεν μέγαν ἱστόν

She set up a great web in the palace, [...] then day by day she weaved at *the great web* (β, 94-104)

Given the assumption that negative evidence is unusable/unavailable, only the answers YES are relevant to set the parameter in (4). Owing to the formulation of the manifestations, it is apparent that even a small portion of text is sufficient to set the parameter.

Since the other 23 parameters in our dataset have similar properties, the restricted list of their manifestations (along with general hypotheses about default states for certain parameters) provides a decision procedure that sets the 24 parameters about determiners solely on the basis of relatively short portions of texts, proving 1) to attain a very good approximation to descriptive adequacy for each of them individually, 2) to minimize the differences between contiguous stages of the same E-language, 3) to make strong and correct typological predictions on wider sets of languages.

This tool, which can be further extended to other domains of syntax beyond the 24 parameters considered here, offers a new perspective for the study of closed-*corpora* languages, because it opens up the possibility of mechanically inferring the I-language that generated each single specimen of E-language (=text). This has a series of practical consequences: (i) it makes it feasible to follow the parametric development of a language in a stepwise fashion as a series of minimal and constrained resets; (ii) it enables one to isolate synchronic variation at each single stage of development (in terms of dialect, genre, etc.); (iii) it makes it possible to compare languages that are only scantily attested (e.g. Gothic) to languages characterized by abundant textual sources; (iv) owing to the moderate amount of text necessary to set the parameters, it becomes possible to pursue syntactic investigations on languages for which annotated *corpora* are not yet available.

As a theoretical consequence, our experiment shows that the heuristic tools used to investigate historic languages produce the best approximation to a plausible model of language acquisition, one that successfully sets the parameters of a given I-language from datasets with the limitations usually attributed to primary *corpora*, therefore as a more plausible device to attain *explanatory* adequacy than most synchronic approaches just focusing on *descriptive* adequacy.
Historical changes in Sub-word formation: The case of Arabic -a(t)

Myriam Dali and Eric Mathieu
University of Ottawa

**Aim** In the theoretical context of Distributed Morphology, the aim of this paper is to provide evidence for the idea that historical changes are not limited to changes in the status of linguistic terminals from M-word to Subword or vice versa (Roberts and Roussou, 1999, 2003; Van Gelderen, 2011), but that historical changes can occur *within* an M-word (see Diertani 2011). Although the possibility that a derivational morpheme can change into an inflectional one was mentioned in Kuryłowicz’s 1965 classical definition of grammaticalization, it has been seldom studied (but see Diertani 2011, Koutsoukos and Ralli 2012): “Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status, e.g. from a derivational formant to an inflectional one” (Kuryłowicz 1965:52). In this paper, we focus on the development of gender and number in Arabic and, in particular, on the evolution of the suffix -a(t), originally a derivational morpheme, which we claim first underwent internal changes and then external changes to become an inflectional morpheme.

**Puzzle** The feminine marker -a(t) in Arabic is used, not only as a gender marker as in (1), but also to produce nouns from adjectives (2), collectives from singulants (3) and singulatives from collectives (4). That such nouns are all marked feminine is somehow strange. For example, why should ‘butchers’ become “feminine” when viewed as a group? Diachrony provides us with clues for this state of affairs: there is evidence from reconstruction studies (Hasselbach, 2014a,b) that -a(t) was originally a derivational morpheme and that it was only secondarily associated with feminine gender. On this view, the original function of the suffix was to mark derivatives of adjectives, specifically abstract nouns (2) and other usages of -a(t) derived from this basic function.

**Proposal** We propose to formalize this series of historical changes using Harbour’s (2011, 2014) theory of number and Kramer’s (2009, 2015) theory of gender. We hypothesize that the grammaticalization of -a(t) involved internal changes of a derivational morpheme and external changes involving mutation of a derivational morpheme into an inflectional morpheme. First, we assume, following Harbour (2011, 2014) – and many others – that classificatory features occupy their own projection, namely n (= Class) and that n takes a root as a complement. Furthermore, n labels the root as a noun and makes it visible to the computational system. Second, we assume that n structures the root as a join semilattice (Harbour, 2011,
2014; Zabbal, 2002; Martí, 2018). Third, we assume gender features are associated with \( n \) and not roots (Kramer, 2009, 2015). We follow reconstruction studies in claiming that \(-a(t)\) was originally a nominalizer and propose that the initial construction had the structure in (5).

One of the common sources of morphosyntactic change is a misunderstanding by language learners of which structural position an exponent is associated with (Diertani, 2011). We propose this is what happened with \(-a(t)\). A derivational morpheme in \( n \) was reanalysed as an inflectional morpheme in Num, as shown in (6) (for evidence that singulative markers are inflectional rather than derivational, see Mathieu 2012). Note that the singulative \(-a(t)\) is \([+\text{atomic}], [+\text{minimal}], [-\text{additive}]\) and that its input is a collective noun under \( n \) with the feature \([+\text{group}]\). This is a change of function involving a Subword. We propose further that, via a novel system of polarity (Hasselbach, 2014b), Arabic \(-a(t)\) came to denote groups under Num, as shown in (7). Note that, while the input to the singulative is a collective noun, the input to group formation is a count noun. This case involves a Subword changing its features to realize a different function in the grammar.

While Proto-Semitic had no gender (*bint ‘daughter’ developed from *bin now meaning ‘son’ but originally probably meaning ‘child’ or ‘youth’ and some nouns in current TA are still not marked for gender morphologically, e.g. \( b\text{u} \) ‘father’, \( o\text{mm} \) ‘mother’), \(-a(t)\) grammaticalized as a feminine gender marker over time. Since there is evidence that feminine gender was first marked on adjectives and was marked only secondarily on nouns (Hasselbach, 2014b), we propose that \(-a(t)\) was reanalysed in Num as a gender marker via the addition of a feature \([u +\text{Fem}]\). This feature spread to D where it could enter in an agreeing relationship with other elements (this was probably due anaphoric pressures), but this was made possible, we claim, because Arabic had acquired a polarity system where \(-a(t)\) could act on a given input with various results. Via a system of contrast, a neutral/masculine noun under \( n \) was made feminine via agreement under Num. The last step involved the spreading of \([u +\text{Fem}]\) on \( n \), as in (7). This involves a Subword changing its function from inflectional to derivational.

(1) ameer ameer-a(t)  (3) jazzar jazzar-a(t)  ‘prince, princess’ prince.MASC.SG prince-FEM.SG  butcher.MASC.SG butcher-pl

(2) sa\text{a}eed sa\text{a}ed-a(t)  (4) bee\text{d} bee\text{d}-a(t)  ‘happy, happiness’ happy  ‘happy, happiness’ egg.COLL egg-SING

(5) nP  (6) NumP  (7) NumP

\[
\begin{array}{c}
\text{n} \\
\downarrow \\
\text{aP} \\
\downarrow \\
\text{-a(t)} \\
\end{array}
\quad
\begin{array}{c}
\text{Num} \\
\downarrow \\
\text{aP} \\
\downarrow \\
\text{-ah(t)} \\
\end{array}
\quad
\begin{array}{c}
\text{n} \\
\downarrow \\
\text{aP} \\
\downarrow \\
\text{-ah(t)} \\
\end{array}
\]

([+atomic], [+minimal], [-additive])

\([+\text{group}]\)
References


The genesis of third person pronouns: Insights from Late Latin and Old French

Lieven Danckaert (CNRS/Université de Lille), Liliane Haegeman (Ghent University) & Sophie Prévost (CNRS, ENS & Université Sorbonne Nouvelle Paris 3/PSL University & USPC)

1. Introduction

As is well known, in the history of French the null subject property was gradually lost, on a par with ‘rich’ subject-verb agreement (for recent discussion see e.g. Prévost 2018; Simonenko et al. to appear). This talk focuses on one aspect of this development, viz. the rise of 3rd person pronouns derived from the Latin demonstrative ILLE. We inventorize and explain a series of puzzling – and hitherto unnoticed – contrasts between Late Latin and Old French (OFr), paying special attention to the syntactic and pragmatic properties of ILLE (‘that, he’) and IPSE (‘he himself’), the main Late Latin candidates for grammaticalization as genuine 3rd person pronouns.

2. ‘Clausetype’ and ‘Person’ effects in Late Latin and OFr

Two corpus studies comparing the rate of overt pronominal subjects in Latin and OFr reveal that the two languages behave very differently with respect to two parameters, viz. ‘[± root clause]’ and ‘[± 3rd person subject]’. The basic generalizations are the following: (i) in Late Latin, overt subject pronouns are generally more common in main than in embedded clauses. In OFr, we find exactly the opposite. (ii) In Late Latin, overt subject pronouns are slightly more frequent for 3rd person than for 1st/2nd person. In OFr on the other hand, 1st and 2nd person pronouns are generally more common. Crucially, closer inspection reveals complex interactions between the factors ‘Clausetype’ and ‘Person’.

2.1 To study the diachrony of subject realization in (late) Latin we analyzed 11 Merovingian Latin texts from ca. 520-830 AD, which were composed in the geographical area roughly corresponding to present-day France. Only finite clauses with a pronominal (null or overt; personal or demonstrative) subject were taken into account, excluding clauses appearing as a second, third etc. conjunct. The results of a logistic regression modelling the likelihood of subject pronouns to be overt are plotted in Graph 1. Surprisingly, later texts show overall lower rates of overt subjects than earlier ones. Furthermore, overt subjects are dispreferred (i) in embedded clauses (cf. the thinner lines in Graph 1) and (ii) with 1st and 2nd person subjects (cf. the dotted lines) especially after 600 AD. However, this last tendency is absent in embedded clauses, witness the general scarcity of embedded overt 3rd person subject pronouns.

2.2 For OFr, we analysed 11 texts dating from ca. 1000-1250 AD: in this period, the rate of null and overt subjects changes very quickly. As shown in Graph 2, in strong contrast with (Late) Latin, OFr overt subjects are generally preferred in embedded clauses. This environment initially favours overt 3rd person pronouns, a result which in light of the Late Latin situation is highly unexpected. However, after 1200, person distinctions no longer play a role in embedded clauses. In main clauses, overt 3rd person pronouns are always less frequent than their 1st/2nd person counter-parts (again in contrast with Late Latin from around 800), and this tendency becomes stronger over time.

3. Late Latin third person pronouns

The key to understanding the changing role of the factors ‘Clausetype’ and ‘Person’ resides in the distribution of various Late Latin demonstratives (HIC, ILLE, IS and ISTE) and identity terms (IPSE). Table 1 shows that in Late Latin, ILLE is generally the most
frequent 3rd person pronominal reference device. The distribution of *ILLE* is however heavily skewed towards main clauses, the most frequent embedded 3rd person pronoun being *IPSE*.

4. Towards an analysis The data sketched above can be interpreted as follows. First, Late Latin *ILLE*’s preference for main clauses suggests that the pronoun typically functions as a syntactic topic, a type of constituent whose distribution is known to be restricted in embedded clauses (Haegeman 2012). This observation can be taken to mean that *ILLE* took over some of the functions of its erstwhile competitors *HIC* and *IS*, which in Danckaert (2012) were shown to readily function as left-peripheral topics in Classical Latin main clauses. Second, to explain why in (Early) OFr null subjects are (and remain) least frequent in main clauses with a 3rd person subject, one also has to take into account the diachronic fate of *ILLE*’s most tenacious Late Latin competitor, namely *IPSE* (which in its bare form does not survive in Gallo-Romance). We suggest that OFr *IL* came into being as the result of a process of grammar competition (Kroch 1989) between *ILLE* and *IPSE*. In Late Latin, these elements were originally not used with the same pragmatic connotations: in subject function, Late Latin bare *ILLE* usually marks topic continuity (as in (1)). On the other hand, in the Merovingian corpus, subject *IPSE* is either used with its original Classical Latin force (roughly meaning ‘he himself’), or it indicates that anaphoric reference is made to a familiar discourse participant other than the current discourse topic (cf. (2)).

1. [DP *Stilla*, [CP quem comitem superius nominauil], apparuit, *ut* in circuitu eius, magna nigrido esset; *et illa* [...] inter tenebras relucbat. (= Greg. Hist. 6.14)  
   ‘A star, which I earlier called a comet, appeared, in such a way that everything around it was cloaked in thick blackness. And this star shone a bright light through the darkness.’

2. *Eo tempore multae ecclesiae a Chlodouecho, exercitu depredatae sunt. Eratque *ipse*, tunct fanaticus et paganus. (= Liber Historiae Francorum 10)  
   ‘At that time many churches were looted by Clovis’ army. He was a fanatic, and a pagan.’

We propose that to resolve the clash between the interpretive properties of *ILLE* and *IPSE*, the newly created OFr lexical item *IL* ended up retaining the intersection of the features of *ILLE* and *IPSE*. As a result, *IL* became a *bona fide*, unmarked personal pronoun, devoid of *ILLE*’s [+topic continuity] feature and of *IPSE*’s [+topic shift] feature. This correctly predicts that *IL* should in principle be readily available in both main and embedded clauses, at rates similar to those of 1st and 2nd person pronouns. In embedded clauses, this is indeed what we see in OFr: by ca. 1250, person distinctions virtually no longer play a role in predicting rates of overt subject pronouns. We can assume that any remaining OFr embedded null subjects are disappearing remnants of the Latin, agreement-based *pro*-drop system. Still to be explained are (i) the robust main/embedded asymmetry (cf. Graph 2), and (ii) the role of person distinctions in main clauses. As to (i), with Sigurðsson (2011: 296, 298) we assume that some languages allow for a type of ‘topic drop’ whereby null subjects are licensed by a left-peripheral head via non-local Agree, i.e. without movement to the left periphery. For the case of OFr, a topic drop account along these lines correctly accounts for the well-known observation that null subjects in the language typically occur in genuine V2 (i.e. ‘XP-Vfin-pro’) clauses. The main/embedded asymmetry can be reduced to the independent fact that in many V2 languages (among which OFr), syntactic dependencies can more easily be established across a finite verb in C than across a complementizer, yielding V2 (XP-Vfin-*XP*) but not COMP2 (*XP-COMP-*XP*) effects. Related to (ii), adopting a variant of Sigurðsson’s (2011: 290) ‘Relative Specificity Constraint’ we also understand why OFr main clause null subjects are more likely to be 3rd persons: given that 3rd person subjects are ranked lower on the Accessibility Hierarchy than 1st and 2nd persons and thus qualify as ‘less specific’ in Sigurðsson’s terms, the features of left-peripheral Topic heads licensing 3rd person null subjects are less likely to count as interveners for additional phrasal movement to SpecForceP than those of heads licensing 1st and 2nd person null subjects.

<table>
<thead>
<tr>
<th></th>
<th>All clauses</th>
<th>Main</th>
<th>Embedded</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>HIC</em></td>
<td>81</td>
<td>49</td>
<td>32</td>
</tr>
<tr>
<td><em>ILLE</em></td>
<td>309</td>
<td>266</td>
<td>43</td>
</tr>
<tr>
<td><em>IPSE</em></td>
<td>111</td>
<td>57</td>
<td>54</td>
</tr>
<tr>
<td><em>IS</em></td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td><em>ISTE</em></td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 1: Rates of HIC, ILLLE, IPSE, IS and ISTE in Late Latin.

\[
\text{Table 1: Rates of HIC, ILLLE, IPSE, IS and ISTE in Late Latin.}
\]
Two Directions for Change: Case Studies in the Loss of Null Subjects

Julianne Doner ● University of Toronto

It has long been assumed that null subjects are lost as a result of the loss of subject-verb agreement (e.g., in the history of French, or in the development of Brazilian Portuguese). However, Bilgin (2017) points out that the reverse direction of change is also logically possible. I will demonstrate that both directions for change are attested, and hypothesize that the direction of change has consequences for the resulting null subject properties of a language. Furthermore, I show that, even in languages where the impoverishment of agreement precedes the loss of null subjects, there is a delay of several centuries, indicating that the impoverishment of the agreement paradigm alone is insufficient to trigger a change in the licensing of null subjects.

According to Roberts (2014) and Bilgin (2017), two directions of change are possible:

1. Null Subject Chicken-Egg Problem (Bilgin 2017: 86)
   a. Hypothesis 1: Lose Person-Agreement Markers First
      Person-agreement markers (typically verbal endings in Romance, Turkish, and many other languages) are lost, which triggers the usage of overt pronouns.
   b. Hypothesis 2: Increase Overt Pronoun Usage First
      Overt pronouns become more common, thereby rendering distinct person-agreement markers redundant, which may cause these markers to fall into disuse, and eventually disappear.

There is evidence that both logically possible timelines of change are possible. The development of modern Swedish is an example of a language where person agreement was lost first, preceding the loss of null subjects. In Old Swedish, verbs agreed in person and number with their subjects and null subjects were acceptable, as shown in (2). Additionally, the verb underwent raising to T, as illustrated by its position preceding negation in (3). It could thus be classified as a consistent null subject language (NSL), in the sense of Barbosa (2011).

(2) Oc rängde ower iordhina fyretighi dagha och fyretight nätter
    And rained over the earth forty days and forty nights [Old Swedish; Falk 1993]

(3) … hvi kristne män räddos ei pino
    … why Christian men feared not pain [Old Swedish (1385); Falk 1993]

Person agreement in Swedish was lost in the 15th century, followed by the loss of number agreement in the 17th century (Falk 1993). However, it was not until the 17th century, when number agreement was lost, that the use of null subjects began to decrease, as shown in (4).

(4) | Year of Birth | 1530-1570 | 1571-1610 | 1611-1650 | 1651-1690 | 1691-1730 |
   | Null Subjects | 62% | 62% | 28% | 22% | 17% |

Impersonal constructions without overt subjects in Swedish (Platzack 1985)

In Swedish, then, the verbal paradigm was impoverished well before the change in NSL type. An intermediate grammar with an impoverished agreement system, verb raising, and null subjects was maintained as a stable system for several generations, from the 15th to the 17th centuries.
Thus, although Swedish follows the order of change outlined above in Hypothesis 1, there was a delay of multiple centuries between the loss of agreement and the loss of null subjects, indicating that although the loss of agreement was perhaps a factor in the change, it was not in and of itself the trigger for change. A similar pattern also occurs in the history of French (Roberts 1993).

On the other hand, Borges and Pires (2017) demonstrate that the increase in the frequency of overt subjects and the loss of free inversion both preceded the impoverishment of the verbal paradigm in Brazilian Portuguese, based on a corpus study of letters collected from the Goiás district of Brazil. They show that the proportion of overt subjects increased from 22.48% to 64.00% between the 18th and 19th centuries. However, in the 19th century data, there was only marginal impoverishment in the agreement. In the same period, the Goiás dialect also developed other properties of a partial NSL, such as the loss of free inversion constructions and the development of a generic null subject (cf. Holmberg 2005). This indicates that the modern impoverished paradigm developed following the shift from consistent to partial NSL.

Both French (Roberts 1993) and Swedish (Falk 1993) followed the path of language change outlined in Hypothesis 1 above, where agreement became impoverished first, followed by the subsequent loss of null subjects. Both French and Swedish have developed expletives and can be classified as non-NSLs.

In contrast, both Brazilian Portuguese (Borges and Pires 2017) and Dominican Spanish (Toribio 2000) have followed the path of change outlined in Hypothesis 2, where overt subjects increase in usage first, followed by the impoverishment of agreement. In Brazilian Portuguese, null subjects are restricted to certain persons and a null subject can be used in generic contexts. It can thus be classified as a partial NSL. Dominican Spanish is much harder to classify, as there appear to be multiple grammars currently in competition (Toribio 2000). However, it is clearly no longer a consistent NSL, as impersonal and generic pronouns are being innovated and a pseudo-cleft construction has evolved in order to replace free inversion (Toribio 2000).

Based on these four languages, I propose the following hypothesis. If the trigger for change is the loss of agreement (i.e., Hypothesis 1), the result is a non-NSL; however, if the trigger for change is an increase in the usage of overt pronouns (i.e., Hypothesis 2), the result is a partial NSL. This hypothesis also explains why partial NSLs with fully inflected verbal paradigms, such as Finnish, are possible. This hypothesis also raises several questions. In particular, why does the order of change from Hypothesis 1 cause expletives to be innovated, but not the order of change in Hypothesis 2? How can there be a delay of several generations between the loss of agreement and the loss of null subjects? What, ultimately, triggers the loss of null subjects, if not the loss of agreement?

Gaps in the paradigm: tracking the emergence of indefinite pronouns in Basque

Ricardo Etxepare (CNRS, IKER UM5478)

Basque has a rich system of quantificational expressions based on so-called «indeterminate pronouns» (Kuroda, 1968): indefinite bases formally identical to wh-words. Among the wh-word based expressions are existential pronouns equivalent to the English «someone» series. Those pronouns are formed by combining the wh-word and a suffix, formally identical to the prefixal Complementizer \textit{bait-}, used in relativization: (1) 	extit{zer-bait} « something »

Existential indefinite pronouns in Basque are akin to epistemic indefinites of the Spanish \textit{alguien} or the German \textit{irgendein} type (Kratzer and Shimoyama, 2002; Alonso-Ovalle & Menendez-Benito 2013; Chierchia, 2013). Those indefinites are incompatible with a known reference for the speaker and admit a free-choice interpretation under certain modals. Unlike other quantificational expressions constructed on the basis of a logical operator and the wh-pronoun, existential pronouns in Basque present a high degree of dialectal morphosyntactic variation. Here’s an illustrative sample (kase=morphological case):

(2) a. wh.bait.kase \quad \text{cf. Nor-bait-i} \quad \text{(Someone.dat)}
   b. wh.kase.bait \quad \text{cf \ Nor-i-bait} \quad \text{(Someone.dat)}
   c. wh.kase.bait.kase \quad \text{cf. Nor-i-bait-i} \quad \text{(Someone.dat)}
   d. wh.bait.Aux \quad \text{cf. Nor-bait-a} \quad \text{(Someone.abs)}
   e. wh.kase.focus.bait \quad \text{cf. Nor-i-ere-bait} \quad \text{(Someone.dat)}

There is however no attestation of forms that one would have expected, given (2), as (3a,b):

(3) a.*wh.kase.bait.Aux \quad \text{cf. *nor-i-bait-a}
   b.*wh.focus.bait.kase \quad \text{cf. *nor-ere-bait-i}

I show that the range of morphological variation can be properly understood under the hypothesis that existential pronouns come from correlative protases, via clausal integration as free relatives. Free relatives are a common source in the emergence of indefinite pronouns (HASPelmuth, 1997). Correlative constructions exist as a productive form in a few areas of the Basque country, and they were general before. (1) is thus to be compared to (4):

(4) \text{Nor (ere) bait-a}
   \text{Who.abs even Comp—is «Who (ever) it is»}

The source form of the indefinite pronouns is transparent in the Biscayan dialectal variants which have kept a fossilized auxiliary \textit{da « is »} following the complementizer (Azkue, 1925), as in (2d). The existing variation depends on: (i) the position of nominal inflection (case and number) vis-à-vis the complementizer \textit{–bait}; (ii) the optional presence of a scalar particle \textit{ere} « even/too » in between the inflected wh-word and \textit{-bait}; and (iii) the presence/absence of a fossilized copula \textit{da « is »} in the pronoun. I show how a careful examination of both the dialectal and the historical record sheds light on some of the finer details of the process leading from correlative protases to existential pronouns, including morphological gaps. The starting point is the following representation of correlative protases in Basque (Rebuschi, 2009): (5) \[[\text{IP \_CP Wh-word, \_FoeR, even \_Finiteness t, bait \_IP pro INFL t1]}], \ [[\text{IP…pro}…]]

« Who\_0 (ever) it is, …pro\_i …»

The correlative protasis, as is the case in correlative constructions generally, is outside the domain in which case and agreement relations are valued. For simplifying purposes, we represent this as a case of adjunction to IP. The process involves the integration of the correlative protasis in case and agreement positions and its progressive reduction to a nominal base. Note in this regard that Basque correlative constructions present the following relevant structural property in relation to the hypothesized evolution: the matching element in the correlative construction cannot have a nominal restriction, it can only be a pronoun (\textit{pro} in (5). Overt pronouns in Basque, on the other hand, are quite rare, as the language is pro-drop in all its arguments. The potential for ambiguity in the relation between the adjoined correlative protasis and the case/agreement position is thus considerable. The ambiguity is favored by the
loss of the final copula in the protasis (8a) in most of the Basque dialectal domain, a well established case of phonological reduction in Basque historical finite forms, not limited to correlative protases, as noted by Michelena (1985). The elimination of the copula makes morphological case, which I take to be a realizational case established at Spell Out (see Rezac et al., 2014 for ergative, and Etxepare and Oyharçabal, 2013 for dative), unvaluable at the PF interface inside the adjunct. The absence of a correlative internal case valuator triggers case attraction (8b) from matrix Tense morphology (see among others, Alexiadou and Varlokosta, 2007; Vogel, 2001). This is only possible if the clause is reanalyzed as a free relative and occupies an argument slot in the predicate. The relation between the absence of the copula and the dependent status of case is well supported by the residual cases maintaining a copula: none of them presents variants in which case inflection other than unmarked absolutive case shows up in between the wh-word and the complementizer bait, one of the prominent gaps in the existing variation (3a). Case attraction introduces a fundamental ambiguity in the status of the correlative protasis, whose clausal status is only evidenced by the optional presence of the focal particle ere « even » preceding the complementizer. As noted by Rebuschi (2003), by the XVIIIth century focal particles stop being realized close to the wh-word in correlative protases. An illustrative case of the change is provided by the translation of the following (same) line of the new testament by Haraneder, a labourdin writer of the middle of the XVIIIth century, and Duvoisin, another labourdin writer of the middle of the XIXth century:

(6)  a. [Zer ere neurri] egin baidiokezue bertzei,... (Haraneder, 1740)
  what even measure done bait.you.may.have others.DAT
  “Whatever measure you have applied to others…”
  b. [Zer neurri] ere neurtu bai{"\u00e9}ukezue... (Duvoisin, 1865)
  what measure.INSTR even measured bait.you.may.have
  “By whatever measure you have measured others,…”

Leaving aside the slightly different choice of predicate, irrelevant to the point, (6a,b) are representative of the change arising in the placement of focus sensitive particles like ere « even » in Basque. Modification of the wh-word is now done via association with focus (a.o. Bayer, 1996): (7)  a. [Wh-wordFOCUS even NP] -> b. [Wh-wordFOCUS NP] even

With the scalar particle outside the protasis, and dependent case in the wh-word, there is no cue anymore for the language learner to postulate a clausal status for the sequence wh-word-bait. There is furthermore no cue either for the Comp status of bait, which is reanalyzed as an existential quantifier Q (8d), following the general pattern of wh-based quantification in Basque. This step precedes the externalization of nominal inflection (case and number) (8e). (8a-f) provides a synthetic overview of the diachronic process:

(8)  a. WH-CASE-FOC-BAIT-COP -> COP=∅
  b. WH-CASE-FOC-BAIT -> CASE ATTRACTION/CLAUSAL INTEGRATION
  c. WH-CASEDEPENDENT-FOC-BAIT-> EXTERNALIZATION OF FOCUS PARTICLE
  d. WH-CASE_3-BAIT (FOC)-> REANALYSIS OF COMP AS Q
  e. WH-CASE.BAIT-{FOC} -> EXTERNALIZATION OF NUM/CASE
  f. WH-Q-CASE-FOC

That the externalization of the focus particle precedes the externalization of nominal morphology is shown by the intriguing gap in (3b), which corresponds to the status of the focal particle as a relevant cue for the clausal interpretation of –bait forms. The paper presents a view of grammaticalization that is closely tied to the learners’ role in hypothesis formation, and exploits the successive situations of input ambiguity that arise as a result of each independent morphophonological change. It provides a detailed historical road-map for the typological observation that free relatives are a possible source of indefinite pronouns, and shows that careful attention to morphological gaps can provide a basis for relative chronology in syntactic evolution.
From clitic to affix: nanoparameters & grammaticalization in Western Iberian

Brian Gravely, Jr.
University of Georgia

In this presentation, I examine the development of enclitic accusative pronouns into what appear to be agreement affixes in the grammaticalization of the presentative complementizers (PCs) velai and velaquí in Galician. I show that the transition from ‘clitic to affix’ in this construction, a well-attested change in diachronic literature (cf. van Gelderen 2011), was an emergent nanoparameter (cf. Biberauer & Roberts 2017, Roberts 2016) that led to subsequent stages of reanalysis and the grammaticalization of these PCs.

The modern-day PCs velai and velaquí derive from the phrasal sequence ve-lo aí/ve-lo aquí (‘You see it there’) (Ferreiro 1999). Originally, this consisted of a prosodically divided, bi-clausal construction in which the 2s person indicative ves (‘You see’) paired with a morphosyntactic variant of Galician accusative clitics lo, la, los, las (Uriagaereka 1996) formed the first clause, followed by aí (‘there’) or aquí (‘here’) and a verb of any class (1); furthermore, the enclitic accusative pronoun always agreed with the subject of the verb in the second clause in gender and number (2) (Freixeiro 2006). In Galician, enclisis is the unmarked parameter in finite clauses (3) with proclisis occurring in a number of specific contexts, such as wh- questions, negation, focus fronting, and the complementizer que (4) (Raposo & Uriagereka 2005). Following the generative analysis laid out by Uriagereka (1995) and Raposo & Uriagereka (2005), clitics in Galician move to an ‘active’ functional projection (FP) in the lower left periphery in order to check morphological φ-features on the head of F. These authors attribute the enclisis/proclisis alternation to a morphological parameter that was widespread in Medieval Romance (cf. Benincà 2004, 2006) but currently only survives in Western Iberian variants, marking the difference between Western Iberian and other Romance varieties which show finite proclasis across the board (5). In Western Iberian, this morphological parameter “provides a PF representation for a syntactically active F” (R&U 2005:643), represented as f, which requires a suitable host to its left in order for the derivation to converge. An inactive F, as shown in the Romance varieties of (5), holds a fixed clitic position in finite clauses, showing behavior characteristic of affixes rather than of true clitics (Franco 1991 et seq.).

I claim that the presentative interpretation of this phrase began before its present-day grammaticalized form had arisen, a reading easily interpreted in (6). As in the case of inflected infinitivals, I relate the movement of ver to f by way of morphological φ-features, resulting in affix-like enclisis. This is shown with finite enclisis in aforementioned contexts of proclisis in which the former accusative clitic continues to agree with the subject of the second verb, as in (7). Equally important to this analysis is recognizing that this stage of the development of Galician PCs should not be confused with the present-day form for the following reasons. I take modern-day Galician PCs to be complementizers base generated in the head of ForceP, an idea borne out due to the fact that they must be clause initial (8) and are not found in any other position typical of locative adverbs (9) (cf. the French counterparts voilâ and voici (cf. Morin 1985) (10)). There exist further distinctions in their identity as pure presentative adverbs and not quasi-verbs due to their lack of argument structure (i.e., a categorial [V+] feature), shown by the fact that they cannot stand alone or host clitics (11). Instances that show this
construction being preceded by the adverb *xa* (‘already’) (12) or a topic phrase *a proba* (‘the test’) (13) provide evidence that at the stage of development in question the verbal phrase is below ForceP, a position that I consider to be f. I affirm that morphological φ-features of *ver* are what provoke forced enclisis at this intermediate stage of the development of Galician PCs. These features, as well as all verbal properties, are ultimately lost when reanalysis of the modern-day form is engendered, at which point it adopts a semantic [+presentative] feature (cf. van Gelderen 2015).

Examples
(1) Ve-lo, aí vai o home ‘You see him, there goes the man’
(2) Ve-las, aquí veñen as nenas ‘You see them, here come the girls’
   *Ve-lo, aí anda a muller de Xabier ‘You see him, there goes Xabier’s wife’
(3) Colliche-lo? (*O colliche?) ‘Did you take it?’
(4) Foi o caso que a collera de antemán (*collera-a)
   ‘The thing was that she had taken it beforehand’
(5) [Spanish] Lo vemos todos los días (*Vémoslo)
   [Italian] Le vediamo ogni giorno (*Vediamo-le)
   [Catalan] El veiem tots els dies (*Veiem-lo)
   ‘We see him every day.’
(6) I as cardas de lá, vé-las eli’stán. ‘And the wool cards, there they are.’
(7) Torna Uxía, que ve-laá vai! ‘Uxía is coming back, (because) look at her go there!’
(8) Velaí están os teus cartos ‘There is your money.’
(9) De quereres, podemos ir aí/*velaí ‘If you want, we can go there.’
(10) Ne le voilà-t-il pas?/*Non velaí vai el? ‘Isn’t that him there?’
    Personne ne dit que voilà une situation facile/*Ninguén di que velaí algo fácil
    ‘No one said that this was an easy situation’
(11) Voilâ!/*Velaí! ‘There you are!’
    Le voilâ/*Velaí o! ‘There he is!’
(12) Meu meniño, come... e xa ve-lo aí vai correndo!
    ‘My son, eat... and there you see him go running!’
(13) Abofellas que n’hai tal i a proba ve-la aquí tes.
    ‘Certainly, there is no such thing, and the test here you have it.’

Selected references
Formalizing the notions of re-bracketing and ad sensum agreement

Timothy Gupton and Chad Howe
University of Georgia

In her work on the evolution of constructions like English *a lot of NP* and *a sort of NP*, Traugott (2008) maintains that these structures undergo a series of developments, beginning primarily as partitive constructions and then acquiring more general quantificational meanings. The focus of this analysis is to bring to bear a formal analysis (à la Roberts and Roussou 2003) of the development of binominal quantifiers in Spanish, focusing specifically on patterns of verbal agreement. Both diachronically and synchronically, these constructions in Spanish present a number of curious phenomena. First, the lead noun (N1) passes from a literal meaning (*un montón* 'a mountain/a heap' as in (1)) to acquiring a quantificational meaning (*muchos* 'many') in (2). Second, once N1 acquires a quantificational meaning, N2 ostensibly participates in number agreement processes (3). Following Verveckken (2012: 422), these quantifying nouns may refer to containers, configurations of masses, or collectives. Verveckken & Cornillie (2012) propose that, in cases of grammaticalization of binominal quantifiers in Spanish, the nucleus, which resides in the first noun (N1), is reanalyzed as a quantifier, and its nominal modifier (N2), is reanalyzed as the nucleus of the construction.

Given the apprehension among generative grammarians toward grammaticalization phenomena (see e.g. VanGelder 2011 for a brief review), formal accounts of this structure are in short supply. Not surprisingly, generative accounts of the processes responsible for reanalysis are also hard to come by (though Roberts & Roussou 1999, 2003 are notable exceptions). We build from Saab's (2010) account of noun ellipsis, targeting nP for deletion. An account of *ad sensum* data largely fall out from this proposal. His analysis captures the fact that n-ellipsis (4) does not target nominals in epithets (5) or possessive PP adjuncts (6). For Saab, quantificational nouns and PP adjuncts that escape ellipsis both target Spec, NumP, which lies to the left of nP, outside the ellipsis domain. While elegant, this analysis leaves N2 within a PP in Spec, nP, a state of affairs that leaves us with an opaque explanation for plural agreement processes like those in (3).

We seek to account for variable agreement and ostensible proximal agreement phenomena between subjects and verbs, which are unexpected assuming standard notions of locality of Agree (e.g. Chomsky 1998). We combine Corbett’s (1994) observation that syntax and semantics are both involved in agreement processes with Chierchia’s (1998) suggestion that grammatically singular nouns expressing masses/collectives may come from the lexicon pluralized, proposing a separation of grammatical and semantic Number, with NumP as the locus of semantic Number. Adopting the feature-valuation approach to agreement (Pesetsky & Torrego 2007), NumP values semantic number on the nP. Therefore, in *un montón de flores* 'a heap of flowers', [montón] internally merges in Num, valuing semantic number [s: sg] on NumP (7), followed by grammatical number [g: sg] feature-matching with the D [un]. In the quantificational reading of *un montón de flores* 'a ton of flowers', lexical features of [montón] are [Q=quantificational], with a series of aftereffects. Lack of the nominal features to merge in nP: the QP [montón de flores] merges in Spec, NumP, valuing [s: pl] number on NumP. At the merge of NumP and DP, what ensues is a labeling problem (Chomsky 2013) and an associated mismatch between grammatical and semantic number values, setting the stage for variable verbal agreement with finite verbs. We also show that a separation of grammatical and semantic number can explain diachronic processes and makes testable predictions for synchronic data.

Inspired by Adger (2006), we show that the only DP-NumP featural mismatches not ruled out by independent means are those involving plural semantic number [s: pl] and singular grammatical number [g: sg]. We propose that this scenario leads Spell-Out to “count-up” number features within the DP and calculate combinatorially variable outcomes.
Examples
(1) él y el vestido manchado de sangre, entre un montón\textsubscript{N1} de piedras\textsubscript{N2} también en sangrentadas (CdE, 17th century)
‘He and the dress stained with blood, between a heap of stones, also bloodstained’
(2) Estamos igual que hace - un montón\textsubscript{N1} de años\textsubscript{N2} (CdE, oral)
‘We have been the same for a long time’
(3) la mayoría\textsubscript{N1} de los niños\textsubscript{N2} eran negros (Verveckken & Cornillie 2012: 240, ex. 44, from COREC)
‘the majority of the children were black’
(4) el [burro] de Juan = ‘Juan’s (burro)’
(5) el [burro] de Juan ≠ el burro de Juan
‘Juan’s (burro)’ ≠ ‘that idiot Juan’
(6) el [nP [n’ [nburro] de Juan]] = ‘Juan’s (burro)’
(7) [DP un\textsubscript{D, g: sg} [NumP [Num [n montón\textsubscript{N, s: sg, g: sg}]+Num [uN, s: sg, g: sg]] … [pP [PP de flores]]]]
(8) [DPun\textsubscript{D, g: sg} [NumP [QP [montón de flores\textsubscript{Q, s:pl, g: sg}]+Num [uN, s:pl, g: sg]]]]

Corpora
COREC = Corpus de Referencia de la Lengua Española Contemporánea
CdE = Corpus del Español

References
The Birth of an Epistemic Indefinite: vaegy in Transylvanian Hungarian
Tamás Halm (halm.tamas@gmail.com) and Ágnes Bende-Farkas (agnesbf@gmail.com)
Research Institute for Linguistics, Hungarian Academy of Sciences

1 Introduction
The emergence and evolution of indefinites has steadily been receiving more attention in the typological and theoretical literature (Haspelmath 1997, Jäger 2010, Aguilar-Guevara et al 2010, Gianollo 2018). The emergence and reinforcement of the so-called epistemic meaning component (the cognitive agent's ignorance and/or indifference, plus a very weak existential commitment) has also been the focus of recent work (Aguilar-Guevara et al 2010, Jayez and Tovena 2011, Gianollo 2018, Kamp and Bende-Farkas 2019). In our talk, we discuss the birth of the epistemic indefinite vaegy ‘some or other’ in Transylvanian Hungarian. Vaegy is a recent addition to indefinite determiners, having developed over the course of the last two centuries: this recency makes it possible to map the various steps of grammaticalization precisely with the help of corpus data. The study of vaegy is cross-linguistically relevant as, unlike French quelque, Spanish cualquiera, Dutch wie dan ook, and Standard Hungarian valami; vaegy has not evolved from a free relative structure; accordingly, its epistemic meaning component did not originate from the free choice meaning component typical to free relatives. Rather, we will show that vaegy developed from a compound involving the disjunctive coordinator vagy ‘or’ and the numeral egy ‘one’ through several successive steps of reanalysis. We will also claim that the hypothesis of fossilization of pragmatically inferred information as the general mechanism of the grammaticalization of epistemic indefinites (Aguilar et al 2010) cannot be upheld in its strongest form: more complex grammaticalization pathways are also attested.

2 Data
For our study, we compiled a corpus containing 172 utterances of vaegy. Sources for the historical data include the Hungarian Historical Corpus and various digitized print works accessible in repositories such as the Hungarian Electronic Library (the oldest attestation is from 1860). Data reflecting current usage were collected from the Hungarian National Corpus and also from newspapers, blogs and discussion forums. Since vaegy is mainly used in colloquial registers; ethnographic interviews, private letters, diaries, blogs and discussion forums were prominent sources. As far as current usage is concerned, grammaticality judgements were also tested with native speakers.

Vaegy is attested in three different but related uses in Transylvanian Hungarian (a dialect group of Hungarian spoken in North-Western Romania):

- indefinite determiner (‘some’): \[ \text{vaegy}_{\text{some}} = \{x: f(x) = 1\} \cap \{x: g(x) = 1\} > 0 \]
  (Barwise and Cooper (1981); note that we use equivalent reformulations of the well-known formulae here and below in order to accentuate the deep similarities between the various uses of vaegy):

  (1) Férfiak nemigen vótak velünk, vaegy légyé üttötte bé magát.
  men not.really be:PAST:3PL.with.us some youth hit:PAST:3SG in self:ACC
  ‘We did not really have any grown men with us, some youths did show up.’

- epistemic indefinite determiner (some or other) \[ \{x: g(x) = 1\} \cap \{x: h(f(x)) = 1\} = 1 \]
  where h is a choice function (Reinhart 1997) such that vaegy ‘some or other’ is antispecific (Farkas 2002ab, Jayez and Tovena 2006, Aloni and Port 2010, Giannakidou and Quer 2013) and is characterized by weak existential commitment:

  (2) Ha létezik vaegy update le fogod tudni tiileni.
    if exist:3SG some.or.other update down will:2SG be.able.to:INF load:INF
    ‘If some update or other is available, you will be able to download it.’

- approximator of a numeral: \[ \text{vaegy}_{\text{approximately}} = \{x: f(x) = 1\} \cap \{x: g(x) = 1\} \approx n \]
  (3) Ma elmentek vaegy tízen bányamunkára alsórákos:ak.
    today leave:PAST:3PL approximately ten mine.work.onto Alsórákos:native:PL

1 Note that we used the formulation above in order to steer clear of some of the controversial topics in the analysis of approximators which, however, do not directly affect our analysis. For a recent discussion, see Khrizman and Rothstein (2015) and references.
‘Today, about 10 people from Alsórákos left for work in the mines.’

3 Analysis

The first step of grammaticalization involved the reanalysis of \textit{vagy} ‘or’ into \textit{vagy} ‘approximately’:

\begin{equation}
\text{vagy eggy } \frac{1}{2} \text{ kilo kenyer} \rightarrow \text{vagy eggy } \frac{1}{2} \text{ kilo kenyer}
\end{equation}

\begin{equation}
\text{appr. half kilo bread} \rightarrow \text{appr. half kilo bread}
\end{equation}

This was followed by the reanalysis of \textit{vagy egy} ‘appr. one’ sequences as ‘approximately’:

\begin{equation}
\text{vagy egy } \frac{1}{2} \text{ kilo kenyer} \rightarrow \text{vagy egy } \frac{1}{2} \text{ kilo kenyer}
\end{equation}

\begin{equation}
\text{appr. half kilo bread} \rightarrow \text{appr. half kilo bread}
\end{equation}

This was followed by phonological weakening (lenition): \textit{vaegy} \rightarrow \textit{vaegy} \rightarrow \textit{vaegy}. The final step was the reanalysis of \textit{vaegy} ‘approximately’ as an epistemic indefinite. One facilitator was that in Hungarian (as in many other languages), epistemic indefinites are used as approximators (cf. some five years ago). Also, Transylvanian Hungarian has been in strong contact with Romanian, and the Romanian epistemic indefinite \textit{vreun} (derived from ‘want+one’, cf. Farkas 2002b, Fălăuş 2014 a.o.) is very similar to \textit{vaegy} both phonologically and also, partially, in its morphology (both containing a remnant of the numeral ‘one’). This means that it was easy for language learners to recategorize \textit{vaegy} as an epistemic indefinite:

\begin{equation}
\text{vagy egy } \frac{1}{2} \text{ kilo kenyer} \rightarrow \text{vagy egy } \frac{1}{2} \text{ kilo kenyer}
\end{equation}

\begin{equation}
\text{appr. half kilo bread} \rightarrow \text{appr. half kilo bread}
\end{equation}

We will show based on corpus data that, similarly to \textit{vreun}, \textit{vaegy} ‘\textit{ei}’ is only licensed in environments that are compatible with weak existential commitment (in epistemic modals, the protasis of conditionals, imperatives, adversatives, desideratives, purposives, questions and habituals).\textsuperscript{2}

\textit{Vagy} as a standard indefinite determiner (‘some’) was the result of a separate grammaticalization pathway: \textit{vagy egy} ‘appr. one’ (|.| \approx 1) being reinterpreted as \textit{vagyegy} ‘some’ (|.| > 0). This reinterpretation is natural: while \textit{some} technically signals any positive cardinality, it is pragmatically taken to signal a small cardinality, which is close to ‘approximately one’. The reanalysis was also facilitated by the fact the numerically premodified nouns in Hungarian are in the singular.

\begin{equation}
\text{vagy egy alma } \rightarrow \text{vagyegy alma}
\end{equation}

\begin{equation}
\text{appr. one apple} \rightarrow \text{some apple}
\end{equation}

The steps of grammaticalization can be mapped as follows:

\begin{equation}
\text{vagy egy } \frac{1}{2} \text{ kilo kenyer} \rightarrow \text{vagy egy } \frac{1}{2} \text{ kilo kenyer}
\end{equation}

\begin{equation}
\text{appr. half kilo bread} \rightarrow \text{appr. half kilo bread}
\end{equation}

\begin{equation}
\text{vaegy ‘appr.’} \rightarrow \text{vaegy ‘epist. indef.’}
\end{equation}

\begin{equation}
\text{vreun, vaegy ‘ei’}
\end{equation}

Selected references


\textsuperscript{2} This differentiates \textit{vaegy} (‘\textit{ei}’) from Standard Hungarian \textit{valami} (‘\textit{ei}’), which does not have a weak existential supposition. The better-studied dependent-distributive reduplicated indefinites (Farkas 2002ab) such as \textit{egy-egy} (‘one-one’) differ from both \textit{vaegy} and \textit{valami} in that they are not antispecific, but rather, their value assignment is dependent on another operator in the environment.
This paper examines historical changes in English coordinating conjunction structures using corpus data from Old English through the 18th century (Taylor et al. 2003; Kroch and A. Taylor 2000; Kroch, Santorini, and Diertani 2004). The main focus is changes in *neither...nor*-type coordination, but it is also necessary to give an account of *either...or* due to its transparent morphological similarity. (1) provides the overlapping stages of *neither...nor*-type coordination.

(1) | STAGE 1  | (nawper) ne X ne Y | (neither) ne X ne Y | (850-1500) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>STAGE 2</td>
<td>{neither/nawper} X ne Y</td>
<td>(1350-1570)</td>
<td></td>
</tr>
<tr>
<td>STAGE 3</td>
<td>neither X neither Y</td>
<td>(1350-1640)</td>
<td></td>
</tr>
<tr>
<td>STAGE 4</td>
<td>neither X nor Y</td>
<td>(1500-1710)</td>
<td></td>
</tr>
</tbody>
</table>

I argue that the loss of *ne* in coordinate structures is a direct consequence of morphosyntactic changes brought about by the Jespersen cycle (Wallage 2017).

**Stage 1:** From 850-1500 the main negative disjunction is *ne...ne* (2-a). The structure could be reinforced with *nawper* (< *na* + *awper* ‘outher, one of two’) (2-b) or *neither* (2-c).

(2) a. and cwæð to hym dygollice: ne do ge na swa for þan ic na yfel on hym næbbe gemet, *ne* be hælinge *ne* be restedaga gewemminge
   ‘and he called to him secretly: do not do such, for I have found no evil in him, neither by healing, nor by profaning the sabbath’ (conicod,A,Nic,[A]:4.2.159), undated OE
b. ne reade him mon nauðer ne Moyses boc, ne Regum
   ‘he read neither Moses’ book, nor Kings’ (cobenuul,BenR:42.66.18.819), c950-1050
c. ...dat me of him ne skal neiðer ne spoken ne þench
   ‘That one shall neither speak nor think of him’ (cmvices,61.675) c1150-1250

I argue that conjunction *ne* is the same morpheme as the sentential negator *ne*. *Ne* is generated as the head of a NegP and has an interpretable [NEG] feature. *Neither* and *nawper* are focus adverbs with an uninterpretable [NEG] feature and they are generated in Spec,NegP (similar to *not*). This proposal is shown in (3-a), with a sketch of the syntax in (3-b), inspired by den Dikken’s (2006) analysis of PDE correlative conjunction.

(3) a. Features: {neither/nawper}[*FOC, UNEG* ... *ne*[iNEG] ... *ne*[iNEG]
   b. [FP neitheri,j [ConjP [ ... [NegP ti ne * ... ]]][Conj][Conj θ] [ ... [NegP tj ne * ... ]]]

In (3-b) it is shown that *neither/nawper* is generated in both disjuncts, but undergoes ATB movement. This is argued to explain the abrupt appearance of *neither...neither* during Stage 3. Finally, it should be stated that there is a lot of PF deletion, in ways familiar to all coordinate structures.

**Stage 2:** From 1350-1570, the first disjunct begins to be marked with *neither* or *nawper* alone (4-a)-(4-b).

(4) a. ...havyng nothyr mete *ne* drynke whthe
   ‘...having neither food nor drink with’ (cmgregor,200.1631), c1420-1500
b. For neyther by theyr prudence *ne* theyr sapynce...
   ‘for neither by their prudence nor their sapience’ (cmfitzja,B3V.158), c1420-1500

This deletion of left disjunct *ne* overlaps with the point in the Jespersen cycle that sentential negator *ne* was losing its interpretable [NEG] feature and *not* was acquiring interpretable [NEG].
(Wallage 2017). This cannot be a coincidence. I will argue that this emerging neither...ne pattern is a reanalysis from the structure in (3-a) to (5).

(5) Features: \{neither/naw\(\bar{p}\)\}[iFOC,iNEG] ... ne[iNEG]

As shown in (5), at this point neither/naw\(\bar{p}\) have gained an interpretable [NEG] feature, and ne has maintained its [iNEG]. Structurally, there is still ATB movement as in (3-b). The proposal in (5) makes an empirical prediction—if neither/naw\(\bar{p}\) has interpretable [NEG], we should see a drop in the frequency of (neither/naw\(\bar{p}\))...ne clauses with a negation in the preceding clause (or else the meaning would be \(-\neg(\neg\phi \land \neg\psi) \equiv (\phi \lor \psi)\)). Indeed we do: from 850-1350, 90% have a negation higher in the structure (N=618 with neg, 67 without). From 1350-1500, the frequency drops to 49.4% with a preceding negation (N=78 with neg, 80 without).

**Stage 3:** From 1350-1640, the right disjuncts begin to lose ne. Rather, both the left and right disjuncts are marked with neither (6-a), nor (< naw\(\bar{p}\)> (6-b), or naw\(\bar{p}\)). Crucially, aside from neither...nor, the commonest pattern is for the same word to appear with both disjuncts.

(6) a. ...thou...thei take not virginite, neither countynence, neither alle her goodis to pore men (cmpurvey, I, 56.2267), c1350-1420
b. and o\(\bar{p}\)ure \(\bar{p}\)er ben \(\bar{p}\)at han noug\(\bar{p}\)t of richesse, nor louen hit, nor \(\bar{p}\)ei sechen not to hauen hit (cmedver, 251.483), c1350-1420

I propose the following features for this stage:

(7) Features: \{neither/naw\(\bar{p}\)/nor\}[iNEG] ... \{neither/naw\(\bar{p}\)/nor\}[iNEG]

(7) shows a key development: neither/naw\(\bar{p}\)/nor have lost their [FOC] features, thus are no longer required to move out of their disjuncts to check their focus features.

**Stage 4:** From 1500-1710, there is a rapid increase in neither...nor. I argue that the morphosyntactic features at this stage are the same as Stage 3 (7) and that the changes in form from Stage 3 is caused by analogy with the then-developing either...or construction. Either has a complicated relationship with or—in OE it appeared in æg\(\bar{p}\)er...ge X ...ge Y ‘both X and Y’ (Gast 2013). In ME, it appeared in either X and Y, which competed with both X and Y. Either...or predominates from 1500 on:

(8) Frequency of either...or, either...and, both...and, neither...nor, and neither/naw\(\bar{p}\)/nor...neither/naw\(\bar{p}\)/nor in sentences with two or more conjunctions:

<table>
<thead>
<tr>
<th>DATE</th>
<th>eith...or %</th>
<th>N</th>
<th>eith...and %</th>
<th>N</th>
<th>bth...and %</th>
<th>N</th>
<th>neith...nor %</th>
<th>N</th>
<th>n/n/n/n/n %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1350-1420</td>
<td>.1</td>
<td>4</td>
<td>.8</td>
<td>53</td>
<td>.2</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>.5</td>
</tr>
<tr>
<td>1420-1500</td>
<td>.1</td>
<td>4</td>
<td>.04</td>
<td>2</td>
<td>2.1</td>
<td>61</td>
<td>.1</td>
<td>4</td>
<td>.4</td>
</tr>
<tr>
<td>1500-1570</td>
<td>2.4</td>
<td>147</td>
<td>.05</td>
<td>3</td>
<td>3.8</td>
<td>235</td>
<td>1.8</td>
<td>111</td>
<td>1.4</td>
</tr>
<tr>
<td>1570-1640</td>
<td>4</td>
<td>215</td>
<td>.03</td>
<td>2</td>
<td>4</td>
<td>237</td>
<td>1.7</td>
<td>99</td>
<td>.01</td>
</tr>
<tr>
<td>1640-1710</td>
<td>3.4</td>
<td>186</td>
<td>0</td>
<td>0</td>
<td>3.3</td>
<td>184</td>
<td>1.8</td>
<td>97</td>
<td>.3</td>
</tr>
</tbody>
</table>

**References**

Gast, V. (2013). “From æg\(\bar{p}\)æ\(\bar{p}\)ær to either”. In: Beyond ‘Any’ and ‘Ever’. Ed. by E. Csipak.
Kroch, A., B. Santorini, and A. Diertani (2004). PPCEME.
Taylor, A. et al. (2003). YCOE.
The diachrony of Old French verb-second: not just an inside job

Espen Klavik-Pettersen, University of Oslo. e.j.klavik-pettersen@ilos.uio.no

Synopsis I offer an account of the diachronic evolution of Old French verb-second based on the idea that the two components of verb-second, namely (i) V-to-C movement and (ii) the restricted prefield, have independent diachronies. First, I present data from a 4th century Latin text to illustrate that V-to-C movement was a Romance-internal process. Secondly, I argue that Germanic influence gave rise to a reanalysis of the prefield in pre-Old French. Finally, I provide a formal account of this new system.

Old French V2 Old French (OF) was a verb-second (V2) language. In the 13th century, this V2 system had two salient surface properties: the possibility of G-inversion (Poletto2002) with all kinds of predicates, and a general linear restriction to a single XP in the prefield that could have a variety of different information-structural readings, including various kinds of topic and focus. Both of these properties are illustrated in (1), where there is inversion and the initial direct object represents topical information:

(1) car [ce] avoit elle requis a Nostre Seignor (La Vie de Saint Eustace, ca.1230)
   'For this she had asked of Our Lord...'

There is considerable consensus that OF V2 was derived by V-to-C movement of the finite verb combined with XP-fronting to a left-peripheral specifier position (Roberts1993, Vance1997, Salvesen2013). This analysis is the same as the one normally assumed for Modern Germanic V2 (den Besten 1983 et seq.). Furthermore, due to these striking similarities with Germanic, it has repeatedly been claimed in traditional (Meillet 1931) and recent (Mathieu 2009) literature that OF V2 was inherited from the language of the Franks. This hypothesis is too simplistic, however, since G-inversion was readily available in many Old Romance varieties for which no Germanic influence can be claimed (Wolfe2018).

Internal diachrony: Late Latin I present detailed data from a full annotation of the late 4th century prose text Itinerarium Atheriae (IE). These data indicate that Late Latin had evolved into a verb-initial language with generalised inversion, as 350 of 643 (50.5%) main clauses with overt subjects feature inversion. Furthermore, inversion is equally available with transitive and unaccusative verbs:

(2) [iterata oratione] benedixit nos episcopus (Itinerarium Egeriae, late 4th century)
   repeated.ABL prayer.ABL  blessed us.ACC  bishop.NOM
   'After another prayer, the bishop blessed us.'

At the same time, there are no restrictions against multiple constituents in the prefield. This suggests that the left-periphery can be accessed by many constituents at the same time, as argued in recent cartographic work on Old Romance. I hypothesize that the grammar of IE is the Pan-Romance antecedent of the later Old Romance G-inversion systems in general (see tree on the left). The next challenge is to account for the restricted prefield in Old French.

External diachrony: Bilingualism in Merovingian Gaul The period from the 5th century until the 7th or 8th centuries was a time of bilingualism in Northern Gaul. The Old Francorian language was a V2 language (Walkden2015). This means that the two languages shared the property of V-to-C movement. However, in contrast to the Gallo-Romans, the Franks generally respected a constraint against multiple preverbal constituents.

I argue that significant overlap in the word of main clauses led to interference effects in the bilingual acquisition process. This scenario is consistent with the claim that transfer is triggered by overlap in the input structures (Hulk and...
Müller 2000). In the earlier stages of the contact, bilingualism in Gaul will predominantly have involved adults acquiring Latin or Franconian as a second language. It is well-documented in the literature on language change that imposition of native patterns is common in L2 acquisition (Lucas 2015). In the domain of V2, studies have shown that Swedes acquiring German and Germans acquiring Swedish impose their native use of the prefield, even at very high levels of proficiency (Bohnacker and Rosen 2008, Bohnacker2010). It is plausible that the Franks imposed a restricted use of the prefield, particularly since the constraint which existed in their language did not violate anything in the Latin language. As for the Gallo-Romans, although it is unlikely that they borrowed linear V2 directly, more moderate effects on the usage of the prefield in discourse (increased tendency for V2) may also have played a key role. Such gradual frequency changes may be considered a case of extension in the sense of Harris and Campbell (1995), as an existing pattern (linear V2) is extended at the expense of competitors (V3, V4). This had repercussions on the input to subsequent generations, a process which corresponds to incrementation in the model proposed by Labov (2007), whereby a change is brought about as a collective enterprise of several generations. Eventually, linear V2 was reanalysed as a grammatical constraint during acquisition. I argue that, as the evidence for distinct projections in the left-periphery reached a critical threshold in the PLD (co-occurrence of XPs being taken as the relevant evidence), the children syncretised or bundled (in the sense of Hsu 2017) the different IS-features topic, focus, contrast, etc. on a single left-peripheral head, to which an EPP-feature was also added. I have labelled this projection CP+ (see tree on the right).

Many features, one head The derivation of V2 involved a single projection in every case, regardless of whether the initial constituent was a topic or a focus, contrastive or not. I will show that this analysis fares significantly better than recent cartographic work on Old French postulating movement of the verb to a high left-peripheral position like Force° (Wolfe 2018). The latter analysis makes several incorrect predictions for Old French, particularly with respect to the syntax of embedded clauses. In a nutshell, the limited and highly predictable cases of linear V3, which all involve left-dislocation structures of some kind, are also freely available in embedded clauses, as shown in (3):

(3) sachiez que [tuit cil qui l' avront en memoire] (...) [il] avront hastif conseill

know that all those who him.CL will-have in memory they will-have speedy counsel 'Know that all those who remember him, they shall receive prompt guidance.' (La vie de Saint Eustace)

The left-dislocation projection, FrameP, must simply be situated below ForceP, since it is assumed that the complementiser occupies Force° in these cases (Wolfe2018, Salvesen and Walkden2017). This in turn is clear evidence that the verb does not move to Force°. Recomplementation facts make it clear that the Force-V2 analysis cannot be salvaged by postulating additional complementisers above ForceP.

Developing diachrony from learning patterns, not vice versa. First language acquisition is often taken to be the explanatory source of syntactic innovations (a line of thinking inaugurated over a century ago by the Neogrammarians). However, most influential generative models for syntactic change have back-engineered decisions made by child innovators, on the basis of comparing stable adult synchronic states at either end of the change (e.g., Roberts & Roussou 2003, van Gelderen 2011). Though this has been essential to link general economy and markedness principles from synchronic theory to historical patterns (e.g., Head Preference, Late Merge, Feature Economy), if the child learner theory of change is to be falsifiable, we must now focus explanation directly on the learning task, informed by over a half century of acquisition research.

We argue that investigations linking acquisition to diachrony should draw from what is independently known about the kinds of interim analyses acquisitionists ascribe to children, because that hypothesis space forms the diachronic linguist’s pool of potential innovations. Instead of supposing that children mitigate between adult-like grammars, we take innovations to occur constantly through development as interim analyses, only actuating into changes in the historical record if and when a child remains at an earlier stage in development (Cournane 2017).

Goals. This paper focuses on the morphosyntactic learning task for DP-internal syntax, asking: what language acquisition theory makes the right predictions for syntactic development and for syntactic change? What paths through learning form-feature mappings actually pass through interim analyses that match innovations? We emphasize that the architecture of the morphosyntactic theory and the learning algorithm you use will constrain the kinds of innovations that children may make. Crucially, though every framework requires children to map phonological strings to particular feature values (say [SG] or [PL]), some more inductive approaches leave more for the learner to discover from experience. Others (e.g. nanosyntax, cartography) start the child with more enriched representations: how subgroups of those values relate to each other, and/or where such subgroups go in a fixed sequence of syntactic projections.

A number of cross-linguistic cases, including the historical development of Somali’s crossed gender system that is discussed in this paper, suggest that those more representationally fixed frameworks are too rigid. We therefore propose a learner that initially posits an unstructured feature bundle at a single syntactic position, which is then unraveled inductively throughout learning. We show how a child learning the pre-change grammar using the algorithm we describe is likely to move through an interim analysis matching the innovation (Kramer 2015, Lecarme 2002).

Acquisition data supports structure unraveling. Children are known to pronounce fewer functional elements than their input, with omissions persisting (less and less frequently) for years; (1)-(2) demonstrate the phonological absence of determiners.

\begin{align*}
(1) & \quad \text{I got horn} \\
& \quad \text{(Adam, 2;03, Brown 1973)} \\
(2) & \quad \text{p’ese sur bouton} \\
& \quad \text{press on button (French, Max, 2;01, York)}
\end{align*}

But the phonological absence of [ð̆] and [la] does not mean that children are syntactically omitting functional material, which would require children to discover and add currently absent syntactic structure above the lexical noun, from the bottom up (e.g. Radford 1990). To the contrary, we find that children pronouncing apparently bare nominal Roots nonetheless produce them with the referential properties of DPs (Panne-mann 2007) and attentively process the morphosyntactic material of adultlike DPs (Lew-Williams & Fernald 2007, van Heugten & Shi 2009). This shows that the syntactic material required for anchoring nouns in the discourse really is present in children’s representations. So we can’t explain this production delay by saying that feature values like [+DEF] or [SG] aren’t represented anywhere; it’s that they haven’t been disentangled and relegated to distinct positions with distinct pronunciations, as not enough learning has yet occurred (Gleitman et al. 2005, Snyder 2011). Such unraveling in development creates upwards reanalyses in diachrony (Cournane 2016).

Unraveling an unstructured feature value bundle. How does that still-entangled interim analysis come to be from the child’s language endowment and experience with the world? What does it consist of? If we work within the realizational family of theories (e.g. DM), in which the morphological component reads off the abstract structure of narrow syntax, then a productive grammar will need a syntactic component that dictates where morphosyntactic feature values go, as well as a list of pronunciations to spell them out.

Let’s assume children start with what their experience (and cognition) provides them: feature values drawn from their situational contexts, and phonological strings to map to them. Where do children go
from this unsegmented pronunciation and its associated undifferentiated bundle of features?

Fig 1. Number & gender in nanosyntax (one-feature-one-head)

Two major morphosyntactic architectures currently on the market enrich that input in different ways, providing different ways forward for a learner. One option is nanosyntax, where each individual feature value has a fixed position (Fig. 1), and the morphology tries to insert pronounceable parts into adjacent parts of this ordered hierarchy (Starke 2009, Caha 2009). Nanosyntax is a reaction to more traditional assumptions that feature values arrive subgrouped into “prefabricated bundles” hosted in particular functional heads, as with the NUMBER attribute in the Num head (Fig. 2). The morphology tries to insert pronounceable parts in each functional position.

The fixed feature value subgroupings and functional sequences in these theories are meant to encode recurrent typological tendencies in the morphosyntactic architecture itself. An inductive learner would pick these patterns up from experience, subgrouping values into attributes and syntactically order them by reasoning over co-occurrence and complementary distribution patterns extracted directly from the learner’s input (Chan 2008). Importantly, an inductive approach leaves open the possibility that learners will seriously entertain grammars with syntactic positions that “mix” feature values across the putative GENDER and NUMBER attributes.

Somali gender “polarity”. Somali nouns can be pluralized via eight distinct strategies, some of which appear to “switch” the noun’s gender from its value in the singular; this is “signaled by the /k/g/ vs. /t/d/ gender alternation of the agreeing definite article” (Lecarme 2002:112). The masculine noun ‘thief’ in (3), when pluralized with a tone change, becomes feminine in (4) (Kramer 2015:157).

\[
(3) \text{thief-the.M}\quad(4) \text{thief.PL1.F}\quad(5) \text{thief-PL2.M}\quad(6) \text{thief-PL3.F}
\]

Kramer 2015 & Lecarme 2002 argue convincingly that modern Somali has a set of nominal functional heads that bear both PLURAL and gender information. The ability of these plural affixes to impose their own gender features is demonstrated by the alternate plural marking strategies for the very same ‘thief’ root in (5)-(6).

Unlike the nanosyntactic and cartographic approaches, an inductive learner will not snap the various PL affixes in (4)-(6) to a fixed syntactic position as soon as it discovers its PL feature value mapping. It will instead continue to consider a bundle consisting of any feature values available in context (Fig. 3). This learner’s interim state crucially resembles the crossed-gender-number modern Somali situation, as innovated from a Cushitic binary gender system (Fodor 1959). And it would exist alongside a number of competing preliminary hypotheses (see Valian 1996, Yang 2000), leading to avoidance of the pronunciation of that functional material (Snyder 2011), but not a lack of syntactic representation.

Conclusion. This proposal renders syntactic change theory as a testable extension of learning theory and as a new evaluation metric for morphosyntactic theory (in the spirit of Lightfoot 1979). While our case study focuses on the DP, finding innovations that match the interim analyses of children engaged in unraveling syntax is an approach that works across domains (e.g. modals from v ➔ T (Cournane 2016)).

Accusative Clitic/Null Object Variation in Spanish and the Object Agreement Cycle
Matthew L. Maddox
Southeast Community College-Lincoln

Intro.: Spanish has clitic-left dislocation (CLDT) (1) and accusative clitic doubling (ACD) (2).

(1) Los libros, los compré ayer,  (2) Juan la_i abrazó (a ella)/(a María)_i
the books them I-bought yesterday  Juan her DOM she DOM María
‘I bought the books yesterday.’  ‘Juan hugged her / María.’

CLDT developed prior to ACD historically. Standard Modern Spanish (ModS) disallows null referential objects (NROs); non-standard varieties like Rioplatense allow NROs. I argue these patterns result from the reanalysis of object clitics. I extend Holmberg’s (2005) D-in-Ṭ analysis of null referential subjects to NROs. My main claim is that a D-feature is present on v due to the Object Agreement Cycle (OAC), by which object clitics become agreement morphemes. In CLDT the clitic merges with a D-feature and moves to Spec,v where it is reanalyzed as a feature of v. After reanalysis there is D-in-v and the clitic does not occupy comple ment position so ACD and later NROs are licensed. Two predictions fall out of this analysis: i) a language with ACD will have developed CLDT first; ii) a language with NROs will have developed unrestricted ACD first. These predictions hold throughout Romance. As a consequence I propose a typology of null object languages similar to Holmberg’s (2005, 2010) typology of null subject languages.

Patterns/Analysis: The OAC has three stages per Gelderen (2011). At stage (a) the clitic is a full DP. At stage (b) the clitic merges as DP and moves as a D-head following Chomsky (1995). At stage (c) the clitic and its features are reanalyzed as v. Licensing of CLDT, ACD, and NROs varies diachronically. CLDT occurs in OldS as in (3). ACD (4) is very rare until the 16th century, and putative instances are clitic-right dislocation (Fontana 1993, Gabriel & Rinke 2010).

(3) [La tierra del Rey Alfonso]... la_i podemos quitar.  (4) no lo_i quieren a él_i
the land of-the king Alfonso it we-can leave not him
‘We can leave King Alfonso’s land.’  ‘…they do not love him.’

CLDT occurs in ModS (1); ACD is restricted to pronominal objects (2). This constraint does not hold in Rioplatense (Ormazabal & Romero 2013), where doubling is unrestricted as with María in (2) above. Only Rioplatense has CLDT with and without epithets as in (5) from Suñer (2006).

(5) [A mi mejor amiga]_i, la_i vi [a esa loca linda]_i el jueves.
DOM my best friend her I-saw DOM that crazy beautiful the Thursday
‘I saw my best friend, that crazy beautiful girl, on Thursday.’

OldS and ModS only allow non-referential null objects (6). Rioplatense allows NROs, as in (7).

(6)a. ¿Compraste pan_i / el libro?  (7)a. Queremos el postre_i. -- (Schwenter 2006)
you-bought bread the book  we-want the dessert
‘Did you buy bread / the book?’  ‘We want the dessert.’

b. Sí, compré pro_i / pro_o_i
yes I-bought some it
‘Yes, I bought some / it.’

CLDT, ACD, and NROs are tied to the stages of the OAC as follows. CLDT becomes possible when the clitic is a DP/D, stage (b). In OldS the clitic is a full DP that merges as complement and moves to Spec,v. In ModS the clitic merges as complement and moves as a D-head to v. Rioplatense is at stage (c); the clitic is the spellout of v and either pro or an epithet merges as complement. Unrestricted ACD occurs at stage (c), as in Rioplatense. Since the clitic is a v-head, the complement position is open for pro or a lexical DP; i.e., renewal. The D-feature in v licenses NROs. D is present in v due to reanalysis of the clitic at stage (c). As in Holmberg et al’s (2009) analysis of null subjects, for NROs D-in-v is valued by a topic which may be null; pro merges as complement. Rioplatense CLDT differs from ModS CLDT since they are at
different stages of the OAC. In ModS CLDT the clitic merges as DP complement and moves as D-head to v. The topic is base-generated. In Rioplatense CLDT pro is licensed as complement by D-in-v and it agrees with the base-generated topic. Rioplatense CLDT involves pro; ModS does not. CLDT occurs at stage (b) and (c) of the OAC; unrestricted ACD only becomes possible at stage (c), when the clitic is v and argument position is open for a lexical DP or pro.

**Consequences:** My analysis accounts for the distribution of CLDT, ACD, and NROs in Romance and makes two predictions: i) if a language has ACD it will have developed CLDT first; ii) if a language allows NROs and has clitics, it will allow unrestricted ACD since its clitics are agreement. Prediction (i) holds in Romance. OldS, Italian, and French have CLDT and lack ACD. Evidence that clitics are less advanced in French is they can be omitted in VP conjuncts (Kayne 1975). OldS (8) and Italian (9) also pattern like French (Luraghi 1997, Maddox 2018).

(8) lo mató y Ø despedazó… (9) l’ho baciat e Ø abbracciato.

‘It killed him and tore him apart…’

Rioplatense has unrestricted ACD and inherited CLDT from OldS. Rioplatense allows NROs, which is consistent with both predictions. European Portuguese (EP) has CLDT but lacks ACD (Barrie 2000), so its clitics should be at an early stage in the OAC. This is the case since EP clitics pattern like French in VP conjuncts (10) and, as in OldS, occur in interpolation, as in (11) from Luis & Kaiser (2016). EP’s constrained NROs are variables rather than pro (Raposo 1986).

(10) Apenas a minha mãe me ajudou e (me) incentivou. (11) Se me não engano… only the my mother me helped and me encouraged if me not mistake

‘Only my mother helped me and encouraged me.’

Brazilian Portuguese (BP) is more advanced than Rioplatense Spanish. BP has CLDT but a full pronoun tends to replace the clitic (p.c. J. Carvalho). ACD was present in Old Portuguese but lost in the 19th century (Castilho 2005), at which point NROs developed (Cyrino 1997). Clitics are being lost and NROs are much less restricted in BP than in EP. BP NROs are pro rather than variables (Galves 1989, Kato 1993). BP clitics were at stage (c) and the cycle is now being renewed by strong pronouns, so they are back at stage (a). NROs now occur frequently in BP.

**Final Remark:** An additional consequence of this analysis is that Holmberg’s (2005) proposed typology of null subject languages may apply to null objects, since both rely on the D-feature, but in distinct domains; i.e., T for subjects or v for objects. Consistent null subject languages have D-in-T which licenses null referential subjects. Partial null subject languages lack D-in-T and thus only have null nonreferential subjects. Null referential subjects can occur when bound by a DP subject. Extending this to null objects, ModS is a partial null object language, because its object agreement morphology (object clitics) is not fully grammaticalized, and only nonreferential null objects are licensed; i.e., it lacks D-in-v. Pashto is a consistent null object language that has “rich” object agreement morphology and NROs are common (Roberts 2010).


**Diachronic data:** (3)-Anon., El Cid, c. 1207; (4)-Juan de Arce de Otárola, Coloquios de Palatino y Pinciano, c. 1550; (8)- Pedro Mejía, Silva de varia lección, para. 264; c. 1540-1550.
Since the Old High German (OHG) period, the main coordinating conjunction in German is *und* (also spelled *inti/enti/unți/unde*). However, in some Middle High German (MHG) and Early New High German (ENHG) texts, *und* also appears in positions in which we would rather expect a relative or equative particle or even a (temporal) conjunction, cf. (1) (see Schröbler 1966; Ferraresi & Weiß 2011). In these constructions, the clause introduced by *und* features Verb-End (VE) order; besides the semantics, this strongly indicates that it is syntactically dependent on the matrix clause.

(1)

(a) *Vnd'ër div viñ fi an dirre verte waren.*
Under DET and they on this journey were
‘When/While they there on this journey…’(lit.: ‘Under that and they…’ (Spec. Ecclesiae 17va)

(b) *dì e benant hofstatmit alle vnd dar zve gehört*
DET before mentioned estate with all and there to belongs
‘the before mentioned estate, including everything which belongs to it’ (HZU CXLI)

(c) *zvain lechen div meinerhausvrowen[...] aigen sind gewesen vnd sev*
two fiefdoms REL.PRON. my wife’s property have been *und* they
von ir vater anerstorben waren
fromher father inherited were
‘two fiefdoms, which have been my wife’s property, as they were inherited from her father’ (HZU LXXII)

(d) *als dike vnd er daz brichet. als dike gibt er Sehzic phennig*
as often and he breaks that, as much gives he sixty pence
‘as often as he breaks that [law], he has to pay sixty pence’ (City Book Nürnberg 10rb,13-15)

(e) *Unt (do B) diž alexander uernam. niwhit er ne-beite*
*And* (when B) that.ACC Alexander,NOM heard not he NEG.PTV-waited
‘When Alexander heard of that, he didn’t wait any longer’ (Pfaffe Lambrecht: Alexander V 297)

Even though these ‘non-coordinating’ functions of *und* have been described by scholars quite early, corpus-based studies are lacking so far. In this talk, I discuss new empirical data obtained via search queries in the Old German Reference Corpus (ReA), the Reference Corpus of Middle High German (ReM), the Middle High German Conceptual Database (MHDBDB)\(^1\) and via a manual analysis of 302 Upper German (UG) legal documents of the Herzogenenburg monastery, dating from 1292 to 1450 AD (HZU). Data indicate that non-coordinating *und* originates in the UG dialect area during the 12th century, spreading to the north over time and reaching the peak of its usage during the 14th century.

In contrast to Schröbler (1966) and Ferraresi and Weiß (2011), I propose a slightly different syntactic classification of VE-clauses featuring non-coordinating *und*, distinguishing between *und* functioning as (i) relative particle in adverbial relative clauses, cf. (1)(a) and in simple relative clauses, cf. (1)(b), (ii) non-degree equative particle, cf. (1)(c), (iii) degree equative particle, cf. (1)(d), (iv) temporal conjunction, cf. (1)(e) and (v) other subordinating conjunctions. I argue that non-coordinating *und* did first appear in adverbial relative clauses (cf. (1)(a)); these constructions provide an ideal precondition for the equative and temporal function of *und* to develop.

Adopting the assumption that adverbial relative clauses are derived by moving an (adverbial) element from a position deeper down in the clause to its specifier (Geis 1970; Haegeman 2007; 1

---

\(^1\) Despite the name, the MHDBDB also contains a large number of annotated ENHG texts. Since the *Bonner Frühneuhochdeutschkorpus* (see Korpora.org, http://www.korpora.org/FnhdC/) is not fully annotated yet, the MHDBDB was used to obtain data from younger texts.
van Gelderen 2009, 2015), I assume a doubly-filled comp with the prepositional/ (pro-)nominal element being located in the SpecCP position and \textit{und} in the C-head, cf. (2)(a). While \textit{und} does not carry any temporal or modal semantics in adverbial relative clauses, these contexts constitute a precondition for \textit{und} to be reanalyzed as a temporal conjunction, making the presence of an adverbial element in the SpecCP-position obsolete, cf. (2)(b). The same holds true for adverbial clauses of manner which allow for a reanalysis of \textit{und} from a relative particle to a non-degree equative particle and from that point to a degree equative particle.

(2)  
(a) \([CP [\text{vnd}\text{r div} C]\_c [\text{vfl} [CP [\text{vnd}\text{diu} fi an dirre verte waren]])] 
(b) \([CP [\text{unt}] CP [\text{diz alexander verman...}]]\]

The reanalysis of \textit{und} to a relative particle may have been licensed by Subject Lacking in F-Structure (=SLF)-Constructions (see Höhle 1991; Büiring & Hartmann 1998; Reich 2009; Ferraresi & Weiß 2011), cf. (3)(a), which are already attested in the oldest OHG prose texts. Due to a subject gap in the second conjunct, these constructions provide a precondition for \textit{und} to be analyzed as a relative particle. Another possibility is to take structures which are ambiguous between paratactic constructions and non-coordinating \textit{und} like (3)(b) as a starting point for non-coordinating \textit{und} to develop. Note that (3)(b) is not an SLF construction, but either a simple paratactic structure (and therefore a regular CP-coordination) or an \textit{und}-relative clause with an extraposed object.

(3)  
(a) \textit{Quemans angila \textit{Enti ar scheidant} \_ dea \_ abilan \_ fona mitem \_ dem reht \_ uuisigom}  
\textit{Come.2 Pl. angels and separate.2 Pl. the. bad.Pl. from in between DET right-ones}  
\textit{‘Angels will come \textit{and/ who} separate the bad from the good’ (lit.: ‘There will come angels \textit{and} separate the bad from the good’)} (Monsee Fragments X, 21-23)  
(b) \textit{daz ist fone diu. \textit{unt ich sag iu. daz ein pruno da springit}}  
\textit{that is from that and I tell you that DET well there flows}  
\textit{,that is – and I tell you that – because a well flows there’/ ‘that is because of that, what I tell you: that a well flows there’} (Merigarto 201-203)

References:
MHDBDB = \textit{Mittelhochdeutsche Begriffsdatenbank/Middle High German Conceptual Database (MHDBDB)}.  
ReA= Donhauser, Karin/ Gippert, Jost/ Lühr, Rosemarie: \textit{Referenzkorpus Altdeutsch/Old German Reference Corpus. ddd-ad (Version 1.0)}. Berlin: Humboldt-Universität. http://hdl.handle.net/11022/0000-0003-37E5-D  
1. **Introduction.** Greenberg’s (1966:79) Universal 6 states: “All languages with dominant VSO order have SVO as an alternative or as the only alternative basic word order”. Furthermore, VSO languages are always prepositional (Pr) (Universal 3), which is harmonic with VS as well as VO orders. In complex NPs, basic VSO order correlates with possessum–possessor (NG) and noun–adjective (NA) order (Universal 17). VSO languages of this (first) type [I/Pr/NG/NA] display a broad crosslinguistic distribution. Greenberg mentions Hebrew, Aramaic, and Arabic (Semitic), Berber, and Ancient Egyptian as cases in point (Appendix II.1, p.108). I shall revisit the issue, focusing on the VSO/SVO alternation in Early Egyptian, the earliest stage of the Ancient Egyptian language (2650–1990 BCE).

2. **Major claims.** Early Egyptian meets the syntactic profile of the Sixth Universal insofar as it has VSO as the basic, and SVO as the alternative basic word order. However, several problems arise with respect to the notion of “basicness” the in Greenbergian word order correlations. In particular, VSO and SVO surface orders differ in the morphology of the finite verb and its semantic interpretation. Accordingly, VSO and SVO cannot be syntactic paraphrases of each other. There is another problem in regard to the VSO clausal pattern, which can be derived via different derivational routes.

3. **Word order variation that correlates with morphological variation.** The VSO pattern is typically used for the description of dynamic eventualities (events, actions, and accomplishments). A morphological correlate of this is that the initial verb can be inflected for the entire inventory of the language’s tense-aspect-mood (TAM) and passive voice markers. Consider in this regard (1), in which the main verb j–rx ‘learns (about)’ occurs in a perception verb (“to recognize”). in the perfective aspect form and assumes an eventive interpretation that comes close to that of a perception verb (“to recognize”). The marked SVO alternative, on the other hand, denotes states resulting from prior events (resultatives) or states irrespective of their origin (qualitatives). The Static verb form rx–w ‘knows’ is inflected for 3rd PERS. MASC. and denotes the possession of secret knowledge by the subject referent. Further, note that the presence of subject agreement on Static verb forms excludes the presence of independent tense–aspect morphology. Regardless of how the VSO–SVO contrast is analyzed in structural terms (in this regard, Kramer 2009 and Reintges 2016 offer two different analyses), what is relevant here is that the VSO/SVO alternation is correlated with variation in other domains of the grammar (TAM/voice marking vs. agreement; eventive vs. stative predication).

4. **Word order variation that does not correlate with morphological variation.** The basic VSO order does not line up with a single derivation. Rather, as shown by the contrast between (3) and (4), at least two types of VSO can be distinguished in terms of the exact placement of the subject and the verb. In (3), the subject Hemen is placed after the postverbal negation adverb w ‘not’. The verb jzp ‘will accept’ appears in its perfective–neutral aspect form, with the obtained future denotation being a contextual feature. In (4), the subject Thoth-nakht precedes both the negation w and the adverbial focus particle js. Once again, the perfective aspect form of the verb swr assumes a future meaning ‘will drink’. In cartographic work (e.g., Pollock 1989; Cinque 1999), clause-internal adverbs qualify as landmarks in the syntactic structure with fixed positions, while the verb, the subject and the object are syntactically more mobile. From this perspective, the resulting order VERB > NEGATION w > SUBJECT > DIRECT OBJECT would be indicative for a lower, VP-internal subject position. By contrast, the VERB > SUBJECT > NEGATION w > DIRECT OBJECT variant is indicative for a derived VP-external subject position. The two subject positions also correlate with different targets for finite verb movement. The tree structures in (5a–b) further illustrate this point.

5. **Concluding remarks.** Considering the semantic, morphological, and syntactic differences, the VSO and SVO word orders cannot be regarded as free syntactic variants, as Greenberg’s Sixth Universal would seem to predict. Rather, one must distinguish between an unmarked VSO structure, in which the canonical subject resides in a lower position, and a marked VSO structure, in which non-canonical focused, quantified, and pronominalized subjects have been raised out of the VP to the specifier of TP. In sum, the dominant VSO order displays more syntactic diversity than meets the eye.
(1) Basic VSO order with eventive interpretation
\[ j-rx \quad Pjpj \quad Pn \quad mwt=f \]
AUG-learn.PFV Pepi DEM.M.SG mother.F.SG=POSS.3.M.SG
“This (King) Pepi (here) learns about his mother” (Pyramid Text 910a/P)

(2) Alternative SVO order with stative interpretation
\[ D^3\text{hw}t(j)-nxt \quad pn \quad rx-w \quad m(n(j)) \quad wh\text{=}w \]
“This Thoth-nakht (here) knows the names of the fowlers.” (Coffin Text VI 22o/B1Bo)

(3) VERB > NEGATION w > SUBJECT > DIRECT OBJECT
\[ fzp \quad w \quad Hmn \quad jft=f \quad nb \]
accept.PFV NEG Hemen thing.F.SG=POSS.3.M.SG each.M.SG
“(The god) Hemen will not accept any of his property.” (Mo‘alla Inscription nr. 8, III.6)

(4) VERB > SUBJECT > NEGATION w > EMPHATIC PARTICLE js > DATIVE CLITIC > DIRECT OBJECT
\[ swr \quad D^3\text{hw}t(j)-nxt \quad w \quad js \quad n=sn \quad wzf \]
go.PFV Thoth-nakht NEG EMPH for=3.PL urine.F.SG
“This Thoth-nakht will surely not drink urine.” (Coffin Texts VII 115j/B4Bo)

(5a) Tense Phrase                      (5b) Topic Phrase
<table>
<thead>
<tr>
<th>Tense</th>
<th>Negation Phrase</th>
<th>VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negation w</td>
<td>Subject</td>
<td>Verb Object</td>
</tr>
</tbody>
</table>

Unmarked VSO structure with canonical ‘low’
subject NP within the VP constituent

Marked VSO structure with non-canonical ‘high’
subject NP within the Tense Phrase

(N.B. Arrows indicate movement; strikeout mark the original position of the moved constituent)

REFERENCES


Rebracketing (Gliederungsverschiebung), Minimize Structure, and the Early Merge Principle.

Reanalysis was firstly introduced as a systematic means to explain syntactic change by the Neogrammarians. It was presumably Holzmann (1875: 483) who proposed for the first time that the complementizer dass ‘that’ emerged in this way, i.e. that the pre-OHG demonstrative pronoun thaz that refered cataphorically to a following main clause was eventually reanalyzed as part (i.e. complementizer) of this second clause (which then was reanalyzed as subordinated). Reanalysis as originally conceived by the Neogrammarians and subsequently adopted by many linguists (Harris & Campbell 1995) consists of two separate, though often combined processes: rebracketing (boundary shift or loss) and relabeling (recategorization).

My talk is concerned exclusively with the shift of boundaries between sentences, that is, with cases where a lexical item originally belonging to the first sentence is reanalyzed as belonging to the second one. The main purpose of the talk is threefold: I will demonstrate, firstly, that rebracketing (i.e. the shift or loss of sentence boundaries) exists and occurs relatively frequently (what is sometimes questioned, cf. Whitman 2012); secondly, that it is not reducible to other types of syntactic change, but constitutes a separate type of its own; and, thirdly (and most importantly), that it is subject to a hitherto unnoted condition, namely that it is restricted to cases where the sentence whose boundary is shifted, is already subordinated. Therefore, the prototypical case for this kind of boundary shift – the development of the complementizer that/dass – can be excluded for principal reasons. A consequence of this condition on rebracketing is that the assumption that hypotactic structures developed from paratactic ones, which is widespread among functional linguists (cf. Hopper & Traugott 2003: ch. 7), loses considerable empirical evidence.

I will argue that sentence boundaries can be crossed or shifted in the process of rebracketing only if the item that is to be reanalyzed as part of the following sentence is a head that selects the respective sentence or a phrase that contains the respective sentence. Therefore, rebracketing can have the forms as given in (1a, b):

(1) a. [XP α [CP …]] > [CP α …]]
b. [XP α [YP (β) [CP …]]] > [CP α(β) …]]

The development of the temporal complementizers bis ‘until’ – out of the respective preposition (2a-c) – and seidem ‘since’ – out of the construction seit dem mal dass ‘since the time that’ (3a-d) – are concrete examples for (1a, b).

(2) a. [PP bis [CP daz …]] → [CP [C° bis] …]]
b. ich will bî dir bestân, biz daz ich nû geleben mac (R. von Ems, Barlaam, V. 5727)
   ‘I will stand by you as long as I may live’
c. daz her alle tage sîchte / sine uenie also ofte / biz ime der sweiz abe ran (Ägidius 295)
   ‘that he all days searched his genuflection so often until him the sweat run down’
   ‘that he fall down on his knees every day so often until the sweat run down on him’
I will present and discuss several further cases mainly (but not exclusively) from the history of German (e.g., *weil*, *als*, *seit*) to empirically corroborate my claim. In addition, I will try to specify the selectional relation holding between the head and the sentence and discuss why it seems to be a necessary precondition. This condition on rebracketing follows from the fact that otherwise, the respective clause would be a phase in the Minimalist sense (Chomsky 2005, Citko 2014) and no longer accessible for operations from outside (thus excluding rebracketing). Additionally, I will argue that rebracketing (as other kinds of syntactic changes, cf. van Gelderen 2004) is economy-driven in that it minimizes the structure and it is motivated by the Principle of Early Merge that applies whenever no further movement is necessary.

Russian paucal numerals *dva*, *tri*, četyre exhibit an interesting property regarding number agreement with a quantified noun (Zholobov 2002, Pereltsvaig 2010, Bailyn and Nevins 2008). The quantified noun appears in the so-called ‘paucal form’ marked morphologically by the suffix -a/ ja (1). Historically, the ‘paucal form’ corresponds to the masculine dual suffix –a/-ja (2-3). How and why did the dual number come to encode the paucal? This paper addresses the evolution of morphosyntactic category of number of Russian paucal quantified phrases. I argue that the loss of dual number in Russian resulted in the reinterpretation of the number features encoding the paucal numerals *dva*, *tri*, četyre. Following Harbour (2014), I suggest that [±minimal] feature of the dual does not get acquired anymore; instead, the [±additive] feature gets chosen to reflect paucities. The analysis presented here is supported by novel data obtained from the digital corpus of birch bark documents (www.gramoty.ru)

(1) dva/tri/ četyre god-a/rubl-ja/kon-ja
   two/three/four year-pauc/ruble-pauc/horse-pauc
   ‘two/three/four years/rubles/horses’

(2) vozmite dva roubl-ja
    take two ruble-du
    ‘Take two rubles.’ (1360-1380, Birch Bark #949)

(3) dva god-a
    two year-du
    ‘two years’ (1340-1360, Birch Bark #32)

Old Russian had a singular/dual/plural number system. In Old Russian, nouns modified by the numeral *dva* (‘two’) agree with the numeral in dual number. Dual agreement is also confirmed by the dual suffix -a of the modifying adjective (4). The numerals *tri* (‘three’) and četyre (‘four’) induce plural agreement with the nouns they modify (5-6).

(4) dv-ä kon-ja i storov-ä
    two-du horse-du and healthy-du
    ‘and two healthy horses’ (1120-1140, Birch Bark #842)

(5) tri god-y
    three year-pl
    ‘three years’ (1320-1340, Birch Bark #45)

(6) četyre rubl-i
    four rubl-pl
    ‘four rubles’ (1392-1427, Duxovnaja Gramota)

However, as a marked morphosyntactic category, the dual remained unstable. As agreement facts attest, the plural suffix –i shows that featurally the dual became syncretic with the plural (7).

(7) te dov-a celovek-a te pobegl-i
    those two-du man-du they run-pl
‘Those two men ran.’

Gradually, paucal numerals *tri* and *četyre* begin to trigger paucal agreement with the nouns they modify (8-9).

(8) po tri rubl-*pauc* proda
    for three ruble-*pauc* sell
    ‘Sell (them) for three rubles.’

(9) *četyre* kols-*pauc*
    four *pauc* sell
    ‘two rings’

Assuming (Harbour 2014) that in the syntax number is featurally represented by two features [±atomic] and [±minimal], a singular-dual-plural number systems will have the following representation (10).

(10)  Singular [+atomic –minimal]
      Dual [-atomic +minimal]
      Plural [-atomic –minimal]

Over time, Old Russian speakers began to favor the feature [±additive] to represent the paucal as they were not finding enough dual forms to keep the feature [±minimal] activated. The loss of the dual number in the pronominal and nominal systems in Old Russian was the trigger of the reanalysis of the numerals *dva*, *tri*, *četyre* as paucal numbers. In featural terms, as a result of the loss of the dual, the feature [±additive] replaced the [±minimal] feature. By the end of the 17th century, the Old Russian number system was characterized by the two features - [±atomic] and [±additive] (11).

(11)  Singular [+atomic –additive]
      Paucal [-atomic – additive]
      Plural [-atomic +additive]

The proposal predicts that in languages with dual number the numerals *tri* and *četyre* will not be treated as paucals. This prediction is borne out in Slovenian, Upper Sorbian, Lower Sorbian, and other languages. In Standard Contemporary Slovenian, the numeral *dva* (“two”) triggers dual agreement on the noun and verb whereas the numerals *tri* (“three”) and *štiri* (“four”) induce plural agreement (12-13).

(12)  *Dva* otrok-*a* hodit-*a* sê v sólo
      two child-*a* go-*a* still to school
      ‘Two children still go to school.’

(13)  Enota obsega: 3 stanovanjsk-*e* object-*e*
      unit comprises 3 residential-*pl* object-*pl*
      ‘The complex comprises 3 residential buildings.’

Selected References


Gothic Preverbs: Morphosyntactic Structure, Grammaticalisation, & Cyclic Change
Tamisha L. Tan (Harvard University, Nanyang Technological University)

1. Introduction
While the utility of conducting syntactic analyses on Gothic data has been questioned due to the lack of autochthonous texts, in this paper I argue that Gothic instantiates significant morphosyntactic constraints in its preverb compounds (PVC) that are generalisable to other Indo-European languages, and that these compounds provide evidence for a potentially cyclic process of diachronic grammaticalisation. i) I model the relative ordering of multiple preverbs in Gothic within DM, drawing from work on the feature geometry of local cases, while also capturing parallels in cognate preverb ordering in Sanskrit, Ancient Greek, and Old Irish. ii) I account for P-Copying in Gothic as the result of P-incorporation, morphological fusion, and multiple copy Spell-Out under the Copy Theory of Movement (CTM), which predicts the environments in which P-copying is licit as well as its interaction with multiple preverb compounds (MPCs), idiomatisation, and tmesis. iii) I argue that Gothic PVCs represent an intermediary stage of diachronic change, exemplifying grammaticalisation and its subsidiary components, including syntactic reanalysis via loss of movement steps, phonological reduction, and semantic bleaching. I also discuss the possibility of P-Incorporation and -Copying being cyclic in nature.

2. Preliminary Data
As in the other early Indo-European languages like Sanskrit, Ancient Greek, and Old Irish, Gothic possesses several indeclinable prefixes that attach to verbs:

<table>
<thead>
<tr>
<th>Pre-Verb</th>
<th>Meaning</th>
<th>E.g.</th>
<th>Constituents</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. af</td>
<td>‘away, from, out of’</td>
<td>af-niman ‘take away’</td>
<td>‘away’ + ‘take’</td>
</tr>
<tr>
<td>b. fra</td>
<td>‘ahead, forward’</td>
<td>fra-kunnan ‘forgive’</td>
<td>‘forward’ + ‘give’</td>
</tr>
<tr>
<td>c. miþ</td>
<td>‘with, among’</td>
<td>miþ-qiman ‘accompany’</td>
<td>‘with’ + ‘come’</td>
</tr>
</tbody>
</table>

These preverbs can append transparently adpositional meaning as in (1)a or idiomatic meaning as in (1)b. A verb may take more than one preverb, producing a MPC:

(2) a. inn-at-tiuhan ‘bring in to’ b. ana-in-sakan ‘add to, contribute’

These preverb compounds can also be reinforced with an additional copy of the preverb:

(3) jah afnimands ina af managein sundro
and take-away.PPL he.ACC away multitude.DAT many
‘And [he] took him aside from the multitude’ (Mark 7:33)

3. Preverb Ordering
Attested multiple preverb orderings comprise the following in (4):

(4) ana-in inn-at miþ-fra miþ-us
   du-at inn-uf miþ-in ut-us
   faur-bi miþ-ana miþ-inn

Considering the structure in (5), proposed by Radkevich (2010) for suffixing local cases, I argue that Gothic preverbs are ordered ASP-MOT-DST-PLC-V:

- *ana* ‘into, onto, upon’ > *ina* ‘in, on’
  ASP > PLC
- *du* ‘towards, against’ > *at* ‘at, by’
  ASP > PLC
- *inn* ‘into’ > *at* ‘at, by’; *uf* ‘under’
  MOT > PLC
- *faur* ‘in front of, before’ > *bi* ‘at, by’
  DST > PLC

Some diagnostics for assigning head-status include the obligatoriness of a goal argument (ASP), and selecting for verbs of inherent motion (MOT). Preverb *miþ* ‘with’ instantiates a MannerP external to Radkevich’s L (Place) and M (Path) projections and hence occurs peripherally. Furthermore, incorporating data from Vedic, Old Irish, and Ancient Greek (Papke, 2010), cognate preverbs display a similar relative ordering in attested MPCs:

1P-Copying also occurs in Latin and Ancient Greek. While (3) calques the original Greek text, other examples show P-Copying in the Gothic translation only, e.g. Gothic *bivojandem bi ina* for Greek *gongyzontos peri autou* ‘murmuring about him’ (John 7:32)

2Bracketed forms are unattested in multiple preverb constructions; null symbols indicate an absence of confirmed cognates
4. P-Copying

As suggested by Eyþórsson (1995), Gothic preverbs may undergo incorporation (Baker, 1988). Taking this as head-to-head adjunction within the CTM, we can account for sentences like (3) with the structure in (7). The tail copy of af is not deleted by Chain Reduction (Nunes, 1995) due to univerbation between the incorporated preverb and verb. I argue that lexicalisation renders the higher copy of af invisible to ordering contradictions under Kayne’s (1994) Linear Correspondence Axiom, enabling multiple copy Spell-Out. Evidence for lexicalisation comes from the complementary distribution of P-Copying and tmesis (i.e. separation by clitics), as a complex head requires contiguity of its elements. This predicts *[PreV-Clitic-(Clitic)-V PreV] to be illicit, which holds true in the Gothic Bible data. Diachronically, some of these lexicalised complex heads were then reanalysed as non-decomposable roots, directly inserted at VI. This reanalysis is closely tied to idiomatic interpretation, such that PVCs with tmesis lack idiomatic readings (since tmesis suggests a decomposable compound.) In MPCs, when a second preverb is incorporated, the morphology cannot ‘see inside’ the [PreV₁ + V] root. This correctly predicts the impossibility of P-Copying with the first-merged preverb *[PreV₂-PreV₁-V PreV₁], while PreV₂ can be copied as in (8). For the same reason, in (9), tmesis is possible between PreV₂ and PreV₁, but not between PreV₁ and the Verb:

(8) miþ-inn-galaiþ miþ Iesua in rohso with-into-go.3SG.PST with Jesus in temple

‘[He] went with Jesus into the temple’

(John 18:15)

(9) ga-uh-pan-miþ-sandidedim imma brobar GA-and-then-with-send he.DAT brother.ACC

‘And then [we] sent the brother with him’

(CorII 8:18)

This correctly predicts that PreV₂ cannot add idiomatic meaning to the MPC. We hence constrain the distribution of P-Copying and its interaction with tmesis, idiomatisation, and MPCs.

5. Grammaticalisation & Diachrony

PVCs thus instantiate a loss in both the independence of the preverb and the internal structure of the compound. I argue, following Booij & Van Kemenade (2003) and others, that this is a case of unidirectional grammaticalisation and its subsidiary processes. P-Incorporation (without Copying) provided the necessary surface ambiguity for syntactic reanalysis, where the preverb is lexicalised as part of a complex head rather than as the result of head movement, and subsequently an non-decomposable root. This change from internal to external Merge follows Roberts & Roussou’s (2003) account of diachronic change as the loss of movement steps, motivated by economy, and explains why P-Copying is no longer productive in Modern German with verbs. Furthermore, these preverbs have undergone phonological reduction (e.g. adverb faura vs. preverb faur-, Gothic /ga-/ vs. Mod. German /go-/ and semantic bleaching (cf. loss of locative meaning, rise of idiomatic meaning.) The question of the extent to which this change is cyclic remains open, but may parallel how prepositions came to reinforce directional meaning upon the weakening of case-suffixes.

6. References


3With possible exceptions for ein~in. Additionally, some analyses have proposed P-copying with R-pronouns.
1. **Introduction.** While in most Romance languages, gender and number appear as separate morphemes on D, as in Spanish *la, las*, and *los*, in French, D is marked either for gender, as in *la* [la] or *le* [le], or number as in *les* [le]. In terms of structure, this indicates that Gen and Num are distinct functional heads (F-heads) in most Romance languages (including proto-Romance), but that in French, they are realized on the same F-head (Déchaîne et al. 2018). When did Gen and Num cease to be distinct F-heads in French? What triggered the change? We provide evidence that restructuring of the nominal spine took place in 12th c. Old French (OF). The reduction of feminine -*a* in Gallo-Romance created an allomorph: feminine schwa -*er*. This feminine schwa was subsequently reanalyzed, either as an epenthetic vowel [e] (in *les*), or as a new morpheme marking singular number -*esG* (in *une*). This change is responsible for the fusion of Gen and Num in the nominal spine. Evidence comes from changes in the variable use of determiners in 12th c. OF.

2. **Variable D in Old French.** In OF, several factors determine whether N occurs with a determiner or as a bare N. Previous studies have shown that bare Ns are favoured by predicate Ns, object Ns, non-count Ns, indefinite Ns, plural Ns, and feminine Ns (Boucher, 2005; Buridant, 2000; Carlier, 2007, 2013; Carlier & Goyens, 1998; Déchaîne et al., 2018; Foulet, 1928/1974; Mathieu, 2009; Moignet, 1976). We show how these factors interacted to yield the Modern French system.

3. **Methodology.** Our study relies on two 12th century Anglo-Norman texts: *Le voyage de St-Brendan* (c.1106-21) and *Les voies du français* (Martineau, 2008). The data was extracted with *Corpus Search*. Only argument nominals in subject and object position were considered. All were coded for semantic class, definiteness, number, and gender. Translations from the target texts were used to determine whether a bare N should be construed as definite or indefinite. This yielded a corpus containing a total of 1650 occurrences (445 in *B*, and 1205 in *MdeF*.)

4. **Overall results.** *Goldvarb* analyses show that DEFINITENESS, NUMBER, and GRAMMATICAL FUNCTION remain constant factor groups across the two texts. The study also reveals striking differences. First, while there is a decrease of bare count Ns in *MdeF*, there is an unexpected increase of bare non-count Ns (Table 1).

<table>
<thead>
<tr>
<th></th>
<th>Count Ns</th>
<th>Non-count Ns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td><em>Brendan</em></td>
<td>71.3</td>
<td>355</td>
</tr>
<tr>
<td><em>Marie de France</em></td>
<td>85</td>
<td>1105</td>
</tr>
</tbody>
</table>

Second, for **count Ns**, GENDER emerges as significant only in *B*, where feminine Ns favor bare Ns. We find the exact opposite in **non-count Ns**: GENDER emerges as significant only in *MdeF*, where masculine Ns favor bare Ns.

<table>
<thead>
<tr>
<th></th>
<th>Count Nouns</th>
<th>Non-count Nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Masculine</td>
<td>Feminine</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td><em>Brendan</em></td>
<td>77.6</td>
<td>228</td>
</tr>
<tr>
<td><em>MdeF</em></td>
<td>84.5</td>
<td>644</td>
</tr>
</tbody>
</table>

Our results show that these changes can be traced back to 1- the neutralization of gender effect in *les*, and *une* in count Ns, and 2- the decrease of expletive masculine *le* used with non-count Ns.
5. Analysis. We attribute the changes in the variable use of determiners to the loss of feminine – \( \partial_{\text{FEM}} \), a reduced form of feminine \(-a_{\text{FEM}}\) found in unstressed word-final position and before \(-s/\). Between \(B\) and \(MdeF\), this feminine \(-a_{\text{FEM}}\) was reanalyzed either as an epenthetic vowel \([\partial]\), or as a new singular morpheme \(-\partial_{\text{SG}}\). These two changes had three consequences. First, the degrammaticalization of feminine \(-\partial_{\text{FEM}}\) as an epenthetic vowel \([\partial]\) led to syncrétism between masculine \(\text{les}\) and feminine \(\text{les}\), and the loss of the gender asymmetry in object position. Second, as a result of the reanalysis of feminine \(-\partial_{\text{FEM}}\) as a singular morpheme \(-\partial_{\text{SG}}\), \(un-e_{\text{SG}}\) was reanalyzed \(un-\partial_{\text{SG}}\), which accounts for the loss of the gender asymmetry in singular indefinites. Finally, the epenthetic vowel \([\partial]\) in \(\text{le}\), was reanalyzed as the singular morpheme yielding \(l-\partial_{\text{SG}}\). While this reanalysis had no effect on the distribution of determiners with count Ns, it triggered the loss of expletive \(\text{le}\) (non-count Ns being incompatible with the new singular morpheme \(-\partial_{\text{SG}}\)). This accounts for the gender asymmetry in \(MdeF\) non-count Ns.

The reanalysis of feminine -e\(_F\) as the singular marking morpheme -e\(_{SG}\) is supported by two changes in the variable use of determiners in 12\(^{th}\) c. OF: the sharp increase of the use of determiners \(\text{les}\) and \(\text{une}\) with feminine count Ns and an equally sharp decrease in the use of determiners with masculine non-count Ns. While the first change follows from the fact that the determiners \(\text{les}\) and \(\text{une}\) used with feminine count Ns are no longer marked for gender, the second change is a consequence of the fact that the determiner \(\text{le}\) is now marked for number. The change did not affect definite singular masculine count Ns, which are compatible with singular marking, but had a major impact on the use of determiners with non-count nouns, which are incompatible with the singular marking morpheme -e\(_{SG}\). This provoked the loss of expletive D with masculine Ns, and explains why masculine Ns, and only masculine Ns, disfavor the use of determiners in \(MdeF\).

6. Consequences The reduction of unaccentuated [-a] to [-\(\partial\)] during the 8\(^{th}\) century created an allomorph of -a\(_F\): -e\(_F\). This feminine schwa -e\(_F\) shared the same morphosyntactic feature as its non-reduced counterpart -a\(_F\). Consequently, this morphophonological change had no effect on the nominal spine, and Gen and Num still co-occurred as distinct F-heads.

(1) \[\text{KP} [\text{DP} I [\text{NP GenP om/\partialF/af} [\text{np}]]]] \] Gallo-Romance and early OF
This analysis is supported by the fact that in early Old French, the count determiners \(\text{les}\) and \(\text{une}\), with the reduced allomorph -e\(_F\), participated in the same type of gender asymmetries in their variable use as expletive D \(\text{la}\), with the non-reduced allomorph -a\(_F\).

Reanalysis of the nominal spine took place as a result of the reanalysis of feminine -e\(_F\) as a singular -e\(_{SG}\) during the 12\(^{th}\) century. With the loss of -e\(_F\), the reduced allomorph of feminine -a\(_F\), feminine marking on D is now restricted to singular \(\text{la}\). This means that the morphemes lexicalizing Num (singular/plural) and Gen (feminine) no longer co-occur. The complementary distribution of Num and Gen caused the reanalysis in (2), where they are no longer distinct F-heads.

(2) \[\text{KP} [\text{DP} I [\text{NP/GenP spl/ af} /\partialSG [\text{np}]]]] \] Old French

Against mechanisms: towards a minimal theory of change
George Walkden (University of Konstanz)

A widespread view in the literature on language change is that there exist a finite number of mechanisms of change to which attested instances of change can be assigned (e.g. Harris & Campbell 1995: reanalysis, extension, and borrowing; Croft 2000: hypo-, hyper-, crypt- and metanalysis; Blevins 2004: CHANGE, CHANCE, CHOICE). This view is also reflected in the debate on whether grammaticalization constitutes ‘a mechanism of change in its own right’ or is instead derivative of other mechanisms such as reanalysis (e.g. Campbell 2001), and plays a role in the more recent reanalysis-vs.-analogy debate, including in diachronic generative circles (e.g. Garrett 2012).

In this paper I argue against reifying ‘mechanisms’ as primitives of the theory of language change, on multiple grounds. First, the mechanisms in question, such as reanalysis and analogy, are commonly invoked in multiply ambiguous ways: as (dynamic) processes in general, as specific events, as causal factors, as the result of a process of change, and sometimes also as constant properties of human cognition or behaviour. At a minimum, we need to distinguish between mechanisms and motivations (Traugott 2011). Secondly, and more importantly, the ontological status of mechanisms is extremely suspect: where do they reside, and/or what are they properties of? Thirdly, and relatedly, in view of the complexity of human behaviour, human language acquisition, and population dynamics, there is no plausible reason to suspect that the changes (or motivations for change) in language can be categorized exhaustively into a finite, small number of discrete types, and so the purpose of such taxonomizing is unclear.

Lightfoot (2002: 127) holds that ‘there is no theory of change to be had independent of theories of grammar and acquisition’. In this paper I take the slightly weaker position that a theory of change should be entirely derivative of i) a theory of language in the individual (cognition, acquisition, and use) and ii) a theory of (human) populations, with at least the latter containing no principles or stipulations specific to language. From this it follows that ‘a mechanism of change that cannot be straightforwardly linked to the strategies of ordinary language use [or cognition, or acquisition–GW] is automatically suspect’ (De Smet 2009: 1731) and should be abandoned. Mechanisms, insofar as they have a role in our diachronic narratives at all, are epiphenomenal. If so, the reanalysis-vs.-analogy debate and the grammaticalization debate may both turn out to hinge on less contentful issues than previously thought.

In the talk I will contrast mechanism-based theories of change with mechanistic models of change (in the sense of Baker et al. 2018) and argue that the latter (e.g. Niyogi & Berwick 1997, Yang 2002, Blythe & Croft 2012, Kauhanen 2017, Kauhanen & Walkden 2018, Burnett & Bonami 2019) represent a more promising approach to understanding language change in the 21st century. This latter class of approaches is still internally extremely diverse. This is true both in terms of their theory of the individual and in terms of their theory of the population. As regards the theory of the individual, in some models (e.g. Niyogi & Berwick 1997, Kauhanen 2017) grammatical competence is conceived of as a matter of discrete states, while in others (e.g. Yang 2002, Blythe & Croft 2012, Kauhanen & Walkden 2018) a probabilistic component is introduced. Moreover, in some models (e.g. Niyogi & Berwick 1997, Yang 2002) the emphasis is on child language acquisition, while in others (e.g. Blythe & Croft 2012, Burnett & Bonami 2019) language use in its social context by adults is central; and even those models which assume a key role for child language acquisition often understand this process in dramatically different ways (e.g. Niyogi & Berwick 1997 vs. Yang 2002). At the population level, some models (e.g. Niyogi & Berwick 1997, Blythe & Croft 2012) assume a homogeneous well-mixing population, whereas others
(e.g. Kauhanen 2017) assume a more realistic social network structure. Thus, the major challenge for mechanistic modellers is to develop criteria by which the adequacy of models can be evaluated. This talk will present operationalizations of criteria of simplicity and empirical adequacy to this end.

Have-doubling and the grammaticalization of have in West Germanic

Joanna Wall (j.h.wall@uu.nl)

Utrecht University & Meertens Institute

1. Introduction. West Germanic languages feature a range of singular have-participial constructions, in which one form of have combines with a morphologically past participle. These include the English have-participials in (1): present perfects like (1a), eventives like (1b), and resultatives like (1c).

(1) a. I have locked up John. b. I have John locked up by Mary. c. I still have John locked up.

West Germanic languages also feature various have-doubling constructions, like those in (2).

(2) a. I have had John locked up (by Mary).
   b. Ik heb vandaag nog niet gerookt gehad.
      ‘I have today still not smoked.
   (English)

(2a) can be straightforwardly analysed as the combination of a present perfect (i.e. (1a)) and an eventive or resultative (i.e. (1b)/(1c)). However, so-called perfect doubling constructions like (2b) have proven more analytically puzzling. These constructions, attested in modern German (e.g. Rödel 2011) and south-eastern Dutch dialects (Koeneman et al. 2011), have strong parallels to present perfects. For example, like present perfects, perfect doubling constructions can embed unergative predicates, e.g. gerookt ‘smoked’ in (2b). This contrasts with eventives and resultatives (e.g. (1b), (1c), (2a)) which require an embedded DP object.

In this talk, I present new data on have-doubling constructions in historical varieties of Dutch (HVDs). This leads me to propose a new restructuring-based analysis for both doubling and singular have-participials, with implications for have’s grammaticalization as an auxiliary.

2. Have-doubling in HVDs. I present the first large-scale corpus study of have-doubling in HVDs. 512 have-doubling constructions were found in an approx. 83,000,000 word collection of texts from authors born between 1050 and 1649. These instances were analysed according to both distributional characteristics (diachronic frequency, genre, geography) and formal characteristics (syntax, semantics). Two important findings of these analyses are the following. Firstly, there is evidence for different types of have-doubling in HVDs: the resultative and eventive variants exemplified in (2a), but moreover the perfect doubling variant of (2b). Secondly, the perfect doubling variant is present in Hollandic HVDs. Crucially, it was Hollandic varieties which formed the basis for modern Standard Dutch where the construction is no longer attested. Thus, this corpus study reveals the attestation of perfect doubling in a historical variety whose closest descendant lacks it. From a comparative perspective, this provides fresh impetus to capture the variation in have-participials in West Germanic with the minimalist possible analytic means.

3. Analysis. My departure point is Brandner and Larsson’s (2014) proposal that perfect doubling constructions are a combination of two semantically distinct present perfects: the first of which functions as a true perfect requiring current relevance and the second of which functions as a temporal past, lacking current relevance. I show that the true perfect use of the present perfect is found in HVDs, Dutch dialects, modern Standard Dutch, English and German, whilst the temporal past use is found in all varieties but English. The net consequence of this is that perfect doubling constructions like (2b) can be derived from two singular have-participials just like the straightforward English have-doubling variants in (2a).

Next, I argue that part of the variation in West Germanic have-participials can be reduced to a parametrization of the clausal positions in which have can merge. More specifically, I adopt Wurmbrand (2001), according to which there are different types of restructuring (functional, semi-functional, lexical) configurations dependent on the clausal position the matrix verb merges into. I modify Wurmbrand’s clause structure for German by amalgamating her AuxP and TP. This results in (3) which I adopt for all West Germanic varieties.

(3) TP > ModP > vP/AspP > VP
I further assume that *have* is a transitive predicate which must probe for an active DP goal when it merges or moves into a case assignment position (i.e. vP/AspP). The various *have*-singular constructions are accounted for as follows. Firstly, resultatives like (1c) are argued to be lexical restructuring configurations in which *have* merges in VP and selects a bare VP complement. Secondly, eventives like (1b) are argued to be semi-functional restructuring configurations in which *have* merges in vP/AspP. These first two proposed configurations account for the inability of resultative and eventive *have*’s to embed unergative predicates: *have* moves or merges into vP/AspP, meaning that there must be an active DP goal for it to probe for. Thirdly, Wurmbrand already proposes that present perfect *have* can merge in two distinct functional projections (her ModP and AuxP) but without any resultant semantic distinction. Here, I argue that *have* merging in the lower projection (ModP) results in a true perfect, whilst *have* merging in the higher projection (here: TP) results in a past. This proposal is satisfying from a theory-internal perspective: Wurmbrand already posits that distinct merger positions for modals correspond to distinct interpretations (i.e. epistemic, deontic, dynamic), the current account extends this to present perfect *have*’s, implying a systematic one-to-one mapping between merger site and interpretation.

Whilst the analysis thus far correctly rules out perfect doubling constructions in English, it does predict that perfect doubling should be possible in all varieties proposed to have both types of present perfect. This includes modern Standard Dutch, where the construction is not attested. However, I make the novel claim that the lack of perfect doubling constructions in that variety is only apparent and results from a parameter which blocks the spell out of the embedded participial form of *have*. This proposal is favourable from a Minimalist perspective by localizing variation away from the computational system and LF. Moreover, I show that, far from being stipulative, it is empirically supported by several parallel verbal constructions in modern Standard Dutch where a covert, structurally second head has also been posited, like perfect passives (e.g. *het boek is verkocht* (*geworden*); van Bart et al. 1998) which feature only one overt auxiliary. Indeed, further empirical support for this proposal is that varieties which attest perfect doubling, like HVDs, show a strong tendency for perfect passives with two overt auxiliaries. In sum, the entire analysis shows that variation in West Germanic *have*-participial constructions can be reduced to parameters governing merger and spell-out.

4. Implications for grammaticalization. This account builds upon raising approaches to grammaticalization (Postma 1995, Roberts and Roussou 1999, Ijbema 2002). Notably, Ijbema (2002) proposes a raising approach to the further grammaticalization of Germanic perfects, positing that this involves a shift from T(Anterior)/Asp_perfect to T(Past) in Cinque’s universal hierarchy of functional projections. Here, following van Craenenbroeck and van Koppen’s (2017) account of Dutch perception verbs, I argue for a Wurmbrandian raising approach to the grammaticalization of *have* in West Germanic. This account has a number of advantages over Ijbema’s account. Firstly, it makes a raising approach to the grammaticalization of *have* viable given the broader advantages of a Wurmbrandian-based approach to a Cinquian one, none the least the reduced structural architecture required. Secondly, the empirical coverage of a raising approach is increased: whereas Ijbema’s account covers just the further grammaticalization of present perfect *have*’s, the current account also incorporates the resultatives and eventives, tracing its entire grammaticalization path.

Rebracketing (Gliederungsverschiebung), Minimize Structure, and the Early Merge Principle.

Helmut Weiß, Goethe-University Frankfurt a.M.

Reanalysis was firstly introduced as a systematic means to explain syntactic change by the Neogrammarians. It was presumably Holzmann (1875: 483) who proposed for the first time that the complementizer dass ‘that’ emerged in this way, i.e. that the pre-OHG demonstrative pronoun thaz that referred cataphorically to a following main clause was eventually reanalyzed as part (i.e. complementizer) of this second clause (which then was reanalyzed as subordinated). Reanalysis as originally conceived by the Neogrammarians and subsequently adopted by many linguists (Harris & Campbell 1995) consists of two separate, though often combined processes: rebracketing (boundary shift or loss) and relabeling (recategorization).

My talk (cf. Weiß, t.a.) is concerned exclusively with the shift of boundaries between sentences, that is, with cases where a lexical item originally belonging to the first sentence is reanalyzed as belonging to the second one. The main purpose of the talk is threefold: I will demonstrate, firstly, that rebracketing (i.e. the shift or loss of sentence boundaries) exists and occurs relatively frequently (what is sometimes questioned, cf. Whitman 2012); secondly, that it is not reducible to other types of syntactic change, but constitutes a separate type of its own; and, thirdly (and most importantly), that it is subject to a hitherto unnoted condition, namely that it is restricted to cases where the sentence whose boundary is shifted, is already subordinated. Therefore, the prototypical case for this kind of boundary shift – the development of the complementizer that/dass – can be excluded for principal reasons. A consequence of this condition on rebracketing is that the assumption that hypotactic structures developed from paratactic ones, which is widespread among functional linguists (cf. Hopper & Traugott 2003: ch. 7), loses considerable empirical evidence.

I will argue that sentence boundaries can be crossed or shifted in the process of rebracketing only if the item that is to be reanalyzed as part of the following sentence is a head that selects the respective sentence or a phrase that contains the respective sentence. Therefore, rebracketing can have the forms as given in (1a, b):

1. a. \([XP \alpha [CP \ldots]] > [CP \alpha \ldots]\]
   b. \([XP \alpha [YP(\beta)[CP \ldots]] > [CP \alpha(\beta)\ldots]\]

The development of the temporal complementizers bis ‘until’ – out of the respective preposition (2a-c) – and seitdem ‘since’ – out of the construction seit dem mal dass ‘since the time that’ (3a-d) – are concrete examples for (1a, b).

2. a. \([PP \text{bis} [CP \text{daz} \ldots]] \rightarrow [CP \text{[C° bis] \ldots}]\]
   b. ich will bî dir bestân, biz daz ich nû geleben mac (R. von Ems, Barlaam, V. 5727)
      ‘I will stand by you as long as I may live’
   c. daz her alle tage sîchte / sine uenie also ofte / biz ime der sweiz abe ran (Ägidius 295)
      ‘that he all days searched his genuflection so often until the sweat run down on him’
I will present and discuss several further cases mainly (but not exclusively) from the history of German (e.g., weil, als, seit) to empirically corroborate my claim. In addition, I will try to specify the selectional relation holding between the head and the sentence and discuss why it seems to be a necessary precondition. This condition on rebracketing follows from the fact that otherwise, the respective clause would be a phase in the Minimalist sense (Chomsky 2005, Citko 2014) and no longer accessible for operations from outside (thus excluding rebracketing). Additionally, I will argue that rebracketing (as other kinds of syntactic changes, cf. van Gelderen 2004) is economy-driven in that it minimizes the structure and it is motivated by the Principle of Early Merge that applies when no further movement follows.


Affiliation:
Prof. Dr. Helmut Weiß
Johann Wolfgang Goethe-Universität Frankfurt am Main
Institut für Linguistik
Postfach: 21
Norbert-Wollheim-Platz 1
D-60629 Frankfurt am Main
http://www.uni-frankfurt.de/59458210/Historische-Sprachwissenschaft
Paratactic Negation in the History of German

Elisabeth Witzenhausen, Ghent University

In this talk, I want to present novel data of a corpus study on paratactic negation in Middle High German (MHG). I show that van der Wouden's (1997) theory of paratactic negation together with data from the Referenzkorpus Mittelhochdeutsch (ReM) favor an analysis of the old preverbal negative marker as a spell out of an affective or non-veridical feature of C in subordinate contexts.

**Theoretical background**
Van der Wouden (1997) identifies two types of paratactic negation: an element of negative import, meaning sentential negators such as *not* or verbs such as *deny, doubt or forbid*, (i) triggers the occurrence of one or more negative morphemes in their complement clause or (ii) the negative element selects a special type of complementizer that may or may not be homophonic to a negation operator. At first glance, the instances of paratactic negation in (1) and (2) are of type (i), as there is no 'special type of complementizer' (ibid. 196).

(1) Nikolaus von Straßburg: Predigten, early 14th century

_Er zwiflete oc h daz vinser here (Jesus) (Christus) ni*

He doubted even that our Lord Jesus Christ NEG

_geborn wree von einer mege
grown were by a maid

'He even doubted that the Lord Jesus Christ was born by a maid.'

(2) St. Pauler Predigten, early 13th century

_wer sotl nv zwivel si ne sin alle heilich di mit dem plyte_

Who shall now doubt they NEG are all sacred who with the blood

_des almehtigen gotes werdent besprenget_

the allmighty god.GEN are splashed

'Who shall doubt that they are all holy who were splashed with the blood of the allmighty god.'

In contrast to this intuition, Breitbarth (2009) analyses preverbal *ne/en* in West Germanic languages as the head of PolP, which she identifies as Laka's (1990) negative complementizer, arguing for an upwards reanalysis of the former negative marker from containing [pol:neg] to a [pol:] feature only. Petrova (2018) argues for MHG that verbs with negative import such as *deny, forbid or doubt* in affirmative matrix clauses license *dass*-clauses with _V_\text{end} order, while overtly negated matrix predicates trigger V2 clauses with preverbal *ne/en* and proposes that the preverbal *ne/en* is reanalysed as an exponent of the Old High German complementizer *ni*.

**Data**

The results from my corpus investigation of the verbs *vermiden* ('to avoid'), *zwivelen* ('to doubt'), *lougenen* ('to deny'), *bewar* ('to prevent') and *ruwen* ('to regret') confirm Petrova's (2018) distribution of *ne/en* in V2 and *niht* in V\text{end} clauses, but also draw a more fine grained picture. Only around 1/3 of the complement clauses even contain a negative element (either single *ne/en* or single *niht*). Among them, verb final *daz* complement clauses ('that'-clause) such as (1) with the adverbial *niht* are very sparsely attested. V2 clauses with preverbal *ne/en* as the only negative marking are the most common type of paratactic complement clause. Investigating other contexts for single preverbal *ne/en* in the ReM\textsuperscript{1}, it becomes clear that single *ne/en* as paratactic negation almost exclusively appears in V2 clauses (175 of 178 clauses). Most of these clause (n=159) are adverbial clauses, such as (3) and only a few complement clauses (n=12).

(3) Gottfried von Straßburg: Tristan, early 13th century

_daz mich dehein ander wip iemer von dir gescheide_

that me any different women ever from you separate

---

\textsuperscript{1} While the results for complement clauses were collected by searching for semantically negative verbs, I additionally took a random sample of 20% (n= 3929) of all clauses containing a negative particle in the ReM corpus (n=19645) and manually searched for clauses containing paratactic negation with single *ne/en.*
Therefore, the analysis has to capture four main observations: (a) if semantically dependent clauses containing paratactic negation appear with single ne/en, they always include verb movement; (b) these dependent clauses almost exclusively have some negative element in their associated/main clause; (c) the assumed distribution between adversative predicates triggering one type of paratactic negation while negated (adversative) predicates trigger another type (Wallage 2005, Petrova 2018) cannot be sustained; (d) generally, paratactic negation is optional but seems more frequent in adverbial clauses when exeptional or adversative semantics are involved.

**Analysis** The analysis distinguishes adverbial clauses with non-negative ne/en and paratactic negation in complement clauses to negative or negated matrix predicates. For complement clauses, earlier approaches (cf. van der Wurff 1999, Wallage 2005 and Breitbarth 2009) as well as the clear distribution and syntactic behavior will be taken as starting point to discuss different accounts explaining the behavior of ne/en in contrast to niht in clauses with paratactic negation. I will argue along the lines of Breitbarth 2009 that ne/en can be analyzed as an exponent of the head of PolP. As Breitbarth takes PolP as part of the CP layer at the interface between clauses, I reinterpret PolP as an [affect] feature on C which causes ne/en to be spelled out on the finite verb. I claim that negative or negated matrix predicates optionally select a C[affect] or PolP, as in (4). I will discuss whether it is reasonable to assume a semantic contribution of paratactic negation in complement clauses. The use of niht and Modern German nicht as a paratactic negation marker is argued to be a later development and to arise when V2 word order in subordinate clauses became more restricted (Reis 1997).

Adverbial clauses such as (3) to their associated clause via a discourse head as in (5) (Cinque 2008, Koster 2000, De Vries 2009, Haegeman and Greco 2018) which also show ne/en will be argued to have a different structure. In adverbial clauses, ne/en can be shown to function as a discourse particle similar to present-day German adverbial connectives, hence clearly adding meaning to the clause it appears in.

![Diagram](image)

Selected References


