

Handout 2: Formal Syntax and Language Change

Elly van Gelderen; Salzburg March 2017

In this chapter, I sketch a formal generative syntactic perspective on language change. I first review the tension between generative grammar and historical linguistics, in both directions. After discussing the role of Universal Grammar, I present the major advances from the 1950s to the present in generative grammar and how they have an impact on diachronic studies. Initially, there is an emphasis on (catastrophic) reanalyses but recently this has shifted to gradual change. I describe work in this area involving features. Finally, I summarize the major contributions and list some unsolved questions.

keywords: cartography, features, generative, modal, phrase structure

1 Introduction

In this chapter, a description is given of some of the work in diachronic formal syntax. Because the field is over 50 years old, an overview of some of the changes it has undergone is necessary. Formal linguistics is often opposed to functional linguistics, with the main difference being that language is seen in terms of communication in the latter but not the former. I will focus on formal syntax and will not discuss the differences very much (see Newmeyer 1998 for an excellent account). Although there are other formal models (e.g. Lexical Functional Grammar¹ and Head Driven Phrase Structure Grammar), I will only discuss the generative approach.

This chapter is organized as follows. Section 2 starts out with some general remarks on the relationship between historical linguistics and generative linguistics and, because Universal Grammar is crucial for a generative account of language change, I also provide some background in how ideas regarding Universal Grammar have shifted. In section 3, I provide more detail on the various generative models, i.e. phrase structure and

¹ See Vincent (2001) for how Lexical Functional Grammar is very well suited to account for grammaticalization and, I assume, cyclical change as well.

transformations, Principles and Parameters, and Minimalism, and what role they play in accounting for historical linguistics within the generative framework. In section 4, I examine two areas of current importance in more detail, features and cartography. Section 5 considers the major contributions of generative grammar to historical linguistics. It also sketches unsolved questions.

2 Generative Grammar and Historical Linguistics

In this introduction, I will first turn to how generative grammar regards historical work and what it sees as the causes of change. I also provide an overview of how the notion of Universal Grammar, crucial to generative grammar since its start, has developed.

2.1 Generative Grammar and Historical Linguistics

There has always been an inherent tension between generative syntax and historical linguistics. I first sketch what the reasons for the tension are (see also van Gelderen 2011b: 45-47) and I then outline what is seen as the major cause for historical change in this model.

In his own writings, for instance, Chomsky has never been interested in language change. The exception is a chapter in Chomsky & Halle's (1968) *The Sound Pattern of English*, which is a description of four stages of English vowels. This inclusion was most likely due to Halle's interests, as evidenced in Halle (1962). Since Chomsky has set the agenda for generative linguists for almost 60 years, pursuing historical linguistics has therefore been 'less popular'.

Formal, generative syntax has typically relied on introspective data, i.e. asking a native speaker for grammaticality judgments because of its emphasis on the internalized grammar of a speaker. For historical periods, such a method of data gathering is obviously impossible and, as a result, historical generative linguistics has often suffered. The use of corpora (and now google-searches) has also been controversial in mainstream generative syntax up to recently. Generative grammar places much emphasis on the distinction between competence and performance, or more recently I(nternal)- and E(xternal)-language (see Chomsky 1986). However, finding a pattern in a (spoken) corpus shows that there is something systematic going on: repeatedly finding *shouldof*

and *shoulda* in contemporary texts (as well as those from the 15th century) indicates that something interesting is happening with modals and perfect auxiliaries, especially if the participle is often replaced by the simple past, as in (1a), from a regular internet search and by an infinitive, as in (1b), in earlier English.

- (1) a. I should **of knew** this was too good to be true.
b. *There xuld not a be do so mykele.*
‘There shouldn’t have been done so much.’ (Margaret Paston a1469)

Similarly, if, as in (2), first (and second) person pronouns are ‘repeated’ (to be adjacent to the finite verb) more than third person ones, this indicates an important difference between those pronouns in the language.

- (2) Which **i** perhaps **i could** adapt, changing my code.

Such differences cannot be attributed to performance but must be determined by the I-language.

Since the 1990s, a group of generative linguists has worked on the creation of parsed corpora (see <http://www.ling.upenn.edu/histcorpora/> and also Kroch, this volume). This has resulted in much better descriptions of changes in the word order (e.g. work by Pintzuk, Haerberli, Taylor, van Kemenade and others), changes in *do*-support (e.g. Kroch and Ecay), Adverb Placement (Haerberli, van Kemenade, and Los), and pro drop (Walkden).

If there exists this tension in the nature of the data used by historical linguists and the type of data that is preferred by generativists, what are the practical consequences for generative historical linguists? One consequence is that mainstream historical linguists often see generative work as data-poor, over-theoretical, and not very insightful. This can be seen from the impact, or rather lack thereof, of generative linguists on mainstream historical linguistics conferences. For instance, conferences such as ICHL (International Conference on Historical Linguistics) and ICEHL (International Conference on English Historical Linguistics) will have generative papers and occasional plenary addresses

using that framework, but generative grammar is a minor framework. Thus, ICHL 20 in Osaka, Japan in 2011 was organized around workshops but did not have a specific workshop on generative grammar and no plenary speaker, and ICEHL 17 in Zurich, Switzerland in 2012 had no generative plenary speakers and included only a handful of generative papers.

Another consequence of the uneasy relationship between generative grammar and historical linguistics – the reverse of the first - is that generative conferences and journals do not see historical linguistics as a crucial component to their enterprise of understanding the faculty of language. Two recent, very influential generative conferences had only two papers devoted to historical linguistics: NELS (North East Linguistic Society) 42, held in Toronto in 2011, included two papers with a diachronic focus in the 67 papers that appeared on the final program and GLOW (Generative Linguists of the Old World) 35, held in Potsdam in 2012, included no historical ones in the 32 papers of the general session.

There has, however, always been a group of generative linguists interested in historical change, arguing that such change gives a special insight into the innate language faculty. Work by Closs (later Closs Traugott), King, Kiparsky, Klima, Lakoff, and Lightfoot testifies to that. Starting in 1990 in York, generative historical linguists have come together through the DIGS (Diachronic Generative Syntax) conference which is devoted entirely to diachronic generative syntax. These linguists have used written sources and have, for the last 20 years, been using various corpora. Many of the DIGS conferences had selected papers appear, e.g. as Battye & Roberts (1995), van Kemenade & Vincent (1997), Pintzuk, Tsoulas, & Warner (2000), and Crisma & Longobardi (2009).

As for causes of change, the emphasis in generative grammar is on one main cause, namely reanalysis during the acquisition process. If this new form is subsequently accepted by the speech community (and this acceptance is not something generative grammar has a lot to say about), the form survives. There are two approaches to the reason behind reanalysis, namely (a) the view that reanalysis is triggered by external events and (b) the view that emphasizes internally motivated change. The first approach has led to a search for cues (Lightfoot 2006 and Westergaard 2009) and the second has

made people look for general patterns of change trying to see if they could be based in the language faculty (e.g. Roberts & Roussou 2003 and van Gelderen 2011a).

Research that emphasizes cues examines how much input a child needs to reset a parameter. According to Lightfoot, "children scan their linguistic environment for structural cues" (2006: 32). These cues are very specific (2006: 82-83), e.g. the *Ce* cue makes a child pay attention to empty complementizers and the *CP[wh-* cue triggers the realization that *wh*-elements move. Under the cue approach, change comes from the outside, i.e. it is triggered by variable data. The challenge for this type of an account is to determine the external triggers. Lightfoot estimates that, in order to reset a word order setting, 30% of main clauses have to show the new word order.

The research that emphasizes internal factors as reasons for change faces the challenge to find the principles that children use to economize the input. This approach focuses on unidirectional change such as full phrases being reanalyzed as heads and prepositions as complementizers. On the basis of recurring cyclical change, van Gelderen (2004; 2011a) formulates a number of principles that can be seen at work in how language is reanalyzed by the language learner. This approach also looks at changes in the input due to external factors.

This difference about the cause of change has also led to a controversy on the role of grammaticalization and unidirectionality. Is grammaticalization real or epiphenomenal? Roberts & Roussou (2003: 2), very much in the spirit of Newmeyer, state that "grammaticalization is a regular case of parameter change ... [and] epiphenomenal." Others, e.g. van Gelderen (2004; 2011a), have argued that the unidirectional patterns that are shown by grammaticalization can be explained internally, namely through the child reanalyzing the input in a certain way.

A last issue is whether change is gradual or abrupt. Most functionalist explanations assume it is gradual whereas many formal accounts think it is abrupt in the internalized grammar of the speaker. As we'll see in 2.2, early generative approaches emphasize a catastrophic reanalysis of both the underlying representation and the rules applying to them. Lightfoot, for instance, argues that the category change of modals is an abrupt one from *V* to *AUX* and, as we'll also see in 2.2, as is the change from impersonal to personal verbs (the verb *lician* changing in meaning from 'please' to 'like'). With the shift to parametric

parameters, it becomes possible to think of gradual change through reanalysis as well (e.g. Roberts 2009 and van Gelderen 2009).

2.2 *Universal Grammar*

In the 1950s, Chomsky's generative grammar introduces an alternative to the then current behaviorist models of language acquisition and the structuralist models of language description. Chomsky focuses not on the structures present in the language/outside world but on the mind of the language learner/user. The input to language learning is seen as poor (the 'poverty of the stimulus' argument): speakers know so much more than what they have evidence for in the input. How is this possible? The answer to this problem is Universal Grammar (hence UG), the initial state of the language faculty, a biologically innate organ. UG helps the learner make sense of the data and build an internal grammar. In the 1960s (e.g. Chomsky 1965), UG consists of substantive universals, concerning universal categories (V, N, etc) and phonological features, and formal universals relating to the nature of rules. The internalized system that is the result is very language-specific, as we'll see.

In the 1970s (e.g. Chomsky 1973), UG comes to be seen as consisting of Principles (true in all languages) and Parameters (choices to be made depending on the language). An example of a principle would be the Subadjacency Principle (Chomsky 1973) that constrains an element from moving too far and an example of a parameter would be the Headedness Parameter, needing to be set as initial or last depending on the language a child encounters: head-initial in Arabic and head-final in Urdu. I have not been able to find the first formulation in the literature of the headedness parameter. It seems that idea was in the air, certainly influenced by Joseph Greenberg's work. As we will see, this period results in a lot of work on parameter resetting, i.e. a younger generation will set its parameter in a different way from the older generation.

Currently, i.e. from the middle of the 1990s on, the role of parameters is considered much less important than it was in the 1980s. Parameters now consist of choices of feature specifications as the child acquires a lexicon (Chomsky 2004; 2007). All parameters are lexical and determine linearization; therefore, they account for the variety of languages. Baker, while disagreeing with this view of parameters, calls this the

Borer-Chomsky-Conjecture (2008: 156): "All parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon." This has spurred some work on the role of features in language change (e.g. Breitbarth 2012, Roberts 2009, van Gelderen 2008). The set of features that are available to the learner is determined by UG. Even as early as 1965, Chomsky says that "semantic features ..., are presumably drawn from a universal 'alphabet'" (1965: 142, although Chomsky continues that "little is known about this today").

A final shift in current thinking is to deemphasize the role of Universal Grammar in favor of what are called third factors. One of the reasons to do so is the evolutionary time UG had to develop. If language arose in humans between 100,000 and 150,000 years ago, a very specific Universal Grammar would not have had the time to develop. The factors not specific to language are therefore preferred and referred to as 'third factors'. The three factors are given in (3).

- (3) **Three Factors:** "(1) genetic endowment, which sets limits on the attainable languages, thereby making language acquisition possible; (2) external data, converted to the experience that selects one or another language within a narrow range; (3) principles not specific to FL [the Faculty of Language]. Some of the third factor principles have the flavor of the constraints that enter into all facets of growth and evolution.... Among these are principles of efficient computation". (Chomsky 2007: 3)

The third factor marks a new emphasis but is somewhat related to the first factor. The third factor is favored above the language-specific first one (for reasons of simplicity) and can be divided into several types, including principles of efficient computation, which are "of particular significance in determining the nature of attainable languages" (Chomsky 2005: 6). Economy Principles are probably also part of more general cognitive principles, thus reducing the role of Universal Grammar even more.

Unfortunately, the third factors are not well defined and have been invoked to account for a number of phenomena, e.g. pro-drop (Sigurðsson 2011), phrase structure (Medeiros 2012), and language change (van Gelderen 2011a). Constraints on word

learning, such as the shape over color bias (Landau et al 1992), would also be third factor, but this then becomes very general. Like UG before it, third factor reasons would remain stable and not responsible in language change and would in fact constrain change.

Having sketched the general notion of UG, we'll now turn to the details of sentential derivations throughout the last 50 years and their impact on diachronic data.

3 From the 1950s to the present

Chomsky's work in the 1950s inspired a formal approach to linguistics. Early on, the emphasis is on phrase structure rules and transformations. In 3.1, I show how these changes are relevant to historical syntax and very briefly discuss generative historical phonology. In 3.2, I discuss the Principles & Parameters stage of the 1970s through the mid 1990s. In 3.3, I show how the emphasis shifts to the inventory of functional categories and features and to third factors from the late 1990s to the present. The delineation between the stages isn't always straightforward, however. I start with an excursion to generative historical phonology as that was influential for concepts such as transparency and reanalysis, core concepts for syntactic change.

3.1 Early Generative Grammar: PS rules and transformations

The early generative phonology is fairly abstract with lots of rules. Most sound change is seen as change in the phonological rules, either by rule loss, addition, and restructuring/simplification. The underlying representation of, for instance, 'write' may be /rajt/, as in (4), with a raising rule deriving the phonetic representation [rʌjt], as in a number of varieties of Canadian English. Because the language also has a voicing rule (before a vowel), the voicing rule can 'bleed' the raising one since the latter occurs only before voiceless consonants. It does so in, for instance, the word 'writer'. The observation of Canadian raising was first made in Joos (1942) and (4) is adapted from a discussion in Chomsky & Halle (1968: 341-342).

(4)	underlying	/rajt/	'write'	/rayt-ər/	'writer'
	a.	voicing	not applicable	raydər	
	b.	raising	rʌjt	can no longer apply	

phonetic

[rʌjt]

[rajder]

If (4a) and (4b) were ordered differently, raising could apply before voicing with [rʌjtər] as result. Kiparsky (1965) and King (1969) discuss many such cases where rules end up being ordered to apply to as many inputs as possible.

Kiparsky also formulates an opacity principle that if too many rules are needed to derive the phonetic form the child reanalyzes the underlying form. This Opacity Principle becomes very important in diachronic syntax, especially in the work of Lightfoot (e.g. 1974 and 1979) who puts it as a Transparency Principle.

As mentioned in section 1, the early grammars are very language-specific. I now turn to the mechanisms needed to derive a sentence. Early on, there are phrase structure rules, as in (5), and transformations, as in (6) for the passive.

(5) a. Sentence --> NP + VP

b. VP --> Verb + NP (Chomsky 1957 : 27)

(6) If S1 is a grammatical sentence of the form

NP1 - Aux - V - NP2, then the corresponding string of the form

NP2 - Aux + be + en - V - by + NP1

is also a grammatical sentence. (Chomsky 1957: 43)

These derive sentences in a very language-specific way. Some of the transformations are ordered and a similar interest in rule reorderings appears as in phonology.

Using these transformations and phrase structure rules, Closs (1965), Kiparsky (1965), Lakoff (1968), Closs Traugott (1972), and Lightfoot (1974; 1979) show how languages change. Early work on syntactic change examines changes in word order, *do*-support, relative pronouns, modals, complementizers, and subjunctives. For instance, Haiman (1974) examines Verb-second word order in Germanic in terms of the interaction of the rule of fronting and expletive insertion; Bever & Langendoen (1971) provide a description of the changes in relative pronouns in the history of English and advocate looking not just at the interacting rules but at performance systems as well, e.g.

perceptual complexity; and Lakoff (1968) examines changes in complementation, especially in Latin, emphasizing that the universal nature of the complementation system.

Closs (later Traugott) (1965) presents a groundbreaking study of English modal verbs and other auxiliaries. This work is the basis of much later work on auxiliaries, both in terms of the data as well as the analysis. She examines the phrase structure rules of Old English, argues that the AUX is different in Old English, and suggests ways to account for the differences. Her formulation of the Modern English Phrase Structure rules is as in (7) and those of Old English as in (8). MV stands for main verb, Vt for a transitive verb, Vi for an intransitive one (with `move` indicating a movement verb), PrP and PP for Present and Past Participle, “env” for “environment”.

- (7) S → NP VP
 VP → AUX MV
 MV → V - NP
 AUX → T (M) (have - PP) (be-PrP) (Closs 1965: 404)
- (8) S → NP - VP
 VP → MV + AUX
 MV → $\left. \begin{array}{l} \text{NP - Vt} \\ \text{Vi} \end{array} \right\}$
 AUX → $\left. \left\{ \begin{array}{l} \text{PP - habb, in env. Vt and Vi-move} \\ \text{PP - wes, in env. Vi} \\ \text{PrP - BE} \end{array} \right\} \right\} \text{ (Inf M) T}$
 (Closs 1965: 407-408)

Closs argues that modals are a class separate from main verbs and accounts for the Old English *have-be* auxiliary-split. The main differences between Modern English (7) and Old English (8) are (a) the order of the main verb (MV) and the auxiliary (AUX) and (b) that, in addition to the modal (M), only one other auxiliary can appear. That complementarity is shown by having the AUX in (8) have the brace brackets around the *habb* `have`, *wesan* `be` and BE. Modern English can have a total of one modal and three auxiliaries, as in (9).

- (9) the ligaments will have been being unknowingly put under pressure due to the masking effects of the medication (Sean Anderson's find on <http://languagelog ldc.upenn.edu/nll/?p=898>).

Lightfoot (1974: 234), also focusing on modals, formulates the phrase structure rules for the modern English modals in fairly similar ways but argues that Old English modals “all behave exactly like ordinary, complement-taking verbs”. Thus, the main concern for Lightfoot is whether modals are main verbs or auxiliaries and this remains a huge debate for a while. He argues they are not full verbs in Modern English, due to a “radical change in the deep structure” (1974: 234). With this statement, he goes against the work of Lakoff (1968) who argues against any changes in the underlying rules. Modals have come to haunt historical linguists and much subsequent ink has been used. I will come back to them in section 4.

The most important work of this period may be Lightfoot (1979) who emphasizes the role of the language learner and formulates the Transparency Principle, as in (10), which is quite similar to Opacity mentioned in section 3.1.

- (10) “[D]erivations may be only of a limited degree of complexity” (1979: 344).

The description of the phenomena he discusses in the book are still couched in terms of phrase structure and transformations, e.g. restructuring of phrase structure rules, and recategorization of modals, quantifiers, and infinitives. Transparency requires the latter categories to be analyzed differently after having become opaque as verbs, adjectives, and prepositional phrases, respectively.

The motivation in much of this early work on historical syntax lies in testing certain aspects of the generative model. For instance, Allen (1977: 1) justifies her study ‘Topics in English Diachronic Syntax’ as follows: “[t]he complementizer has become a focal point of the so-called Extended Standard Theory, as developed by Chomsky in his works from around 1970 to the present ... Because of this, the history of the system of complementation in English is of great potential interest ...”. In this period, there is no

mention of grammaticalization (even though Closs Traugott's data present prototypical examples of grammaticalization).

3.2 *Principles & Parameters Approach (P&P)*

In the mid 1970s and 1980s, generative grammar moves beyond phrase structure rules and transformations. Starting in the 1970s, insights into phrase structure from Chomsky (1970) and Jackendoff (1977) replace rules such as (5) with X-bar theory, applicable cross-linguistically. And after Ross (1967) 'discovers' **islands**, domains from which movement cannot take place, rules such as (6) are replaced by 'move alpha' (= move anything anywhere). Such rules are applicable in any language. The consequence of all of this is a system that invites cross-linguistic comparison and, as a result, Universal Grammar comes to be seen in the late 1970s and early 1980s as consisting of Principles (true in all languages) and Parameters (choices to be made depending on the language).

Allen (1977) tests some of these more general principles on clausal extraction and on what constrains preposition stranding in Old English. She proposes, for instance, that there are two types of relative clauses in Old English, one with and one without movement. More importantly for the theoretical literature are her data on *that*-trace as grammatical in Old English, as (11) shows, taken from Allen (1977). In (11), the complement after the verb 'say' is a clause from which the *wh*-word *hwaet* has been extracted.

- (11) *Ac hwaet saegst ðu ðonne ðaet ~~hwaet~~ sie forcuðre*
But what say you then that -- be wickeder
ðonne sio ungesceadwisnes?
than be foolishness

'But what do you say is wickeder than foolishness?' (Boethius 36.8, from Allen 1977: 122)

The '*that*-trace filter' of Chomsky and Lasnik (1977) had excluded these constructions for Modern English, as in (12). The discovery of (11) showed that the internal grammar has become more restricted in this area.

(12) *Who did you say that ~~who~~ loves Bill?

What I have crossed through in (11) and (12) are often represented as traces, i.e. traces of moved elements, and hence the name the 'that-trace filter'. The data in (11) together with similar observations from other languages spurred much work on the environments in which traces, and empty elements in general, could appear. This becomes relevant to what a universal principle is and what a parameter.

Important work in this period is also done on word order and headedness (V-initial or V-final) and that leads to an examination of the inventory and order of the functional categories. Canale (1978) argues Old English changes around 1200 from an OV to VO language. Van Kemenade (1987) examines the word order in Old English and devises a phrase structure with a V-final VP and an I(nflection) position outside of the S(entence). She identifies two changes in the history of English, a change from OV to VO around 1200, echoing Canale, and a change in the position of I.

Around this time, many linguists start looking at the category of I and C cross-linguistically. Many of the languages examined are, however, Indo-European. Pintzuk (1991) looks at the position of I in Old English, Roberts (1993) at Verb-movement, auxiliaries, and I in French and English, Kiparsky at V-to-C movement in early Indo-European and Germanic, Eyþórrson (1995) at early Germanic Verb-movement, and Los (2000) looks at the status of infinitives and the position of *to*.

Thus, in the 1990s, the topics that are popular are the modals, the auxiliary *do*, verb-movement, and the infinitival marker *to*; complementizers, demonstratives, pro drop, and articles also become of interest. Changes in modals, demonstratives, complementizers and infinitive markers are all prime instances of grammaticalization. With the shift towards an emphasis on functional categories in the late 1980s, grammaticalization could be discussed in structural terms, as Roberts (1993) and van Gelderen (1993) do. Roberts (1993) doesn't mention grammaticalization and there is of course no account for the regularities of the changes, i.e. volition verbs grammaticalize as future and spatial prepositions turn into temporal and causal ones. Van Gelderen

discusses grammaticalization but as something to be responded to by the learner, not as set in motion by the language learner as she does in later work (2011a).

Battye & Roberts' (1995) introduction to one of the DIGS volumes reviews the diachronic work done within the P&P framework and note some of the main issues: parametric differences in verb-movement, both V-to-I and V-to-C, in pro-drop, and in clitics. The title of that volume includes 'Clause Structure' and that is very apt since much of the work around this time is on the formulating the clausal skeleton of CP, IP, and VP. Van Kemenade & Vincent (1997) entitle a later DIGS volume 'Parameters of Morphosyntactic Change' and here we can see the beginning of the shift towards Minimalism with its emphasis on features. Most of the papers continue to be more in the P&P framework, as does the volume edited by Pintzuk, Tsoulas, and Warner in 2000. Typical topics are null arguments, negatives and polarity, position of subjects, the change to VO order, and verb-movement. Thus, historical generative syntax continued (and continues) to be concerned with clause structure and embraces Minimalism slowly.

3.3 *Minimalist Generative grammar*

Current generative historical linguists engage in research that is either heavily based in quantitative corpus data and continue to investigate word order shifts and patterns (both OV/VO and Verb-movement), have started to consider changes in alignment and argument structure (mainly with *psych*-verbs), are working on information structure and word order, or work on feature parameters. I will focus on the latter but will introduce the other areas first.

Some notable results using quantitative corpus data have been made in recent years. One of the first to examine Middle English syntax using the parsed corpus is Kroch & Taylor (2000) who study the supposedly immediate, i.e. catastrophic change around 1200 from OV to VO and show that this particular change is gradual, starting in Old English and continuing into Middle English. They take the position of I(nflection) into account and distinguish three stages, i.e. I-final and OV, I-medial and OV, and I-medial and VO, and discuss in depth the percentage of the clauses in early Middle English texts are still I-final. Since rightward movement of objects is prevalent in Germanic, it is hard to find if an observed VO order is really VO or OV with a moved object. As is usual in

such cases, Kroch & Taylor therefore consider postverbal pronouns (unlikely to have moved rightwards) and other light elements. This study shows a much more complex interplay of various movement factors and a language that is truly ambiguous for the language learner and therefore prone to reanalysis.

Other historical (parsed) corpora have appeared or are appearing and spurring much work among generative (and non-generative) linguists. For instance, there is the Tycho Brahe parsed corpus of historical Portuguese, o corpus do Português, the Corpus del Español, the Regensburg Russian Diachronic Corpus, and a Hungarian corpus is under construction.

Changes in alignment type have been studied in Indo-European and in Austronesian. For instance, Aldridge (2011) argues Austronesian is ergative at one stage and splits into languages that are split-ergative and languages that are accusative. As for argument structure, there has always been an interest in the change from impersonal to personal constructions. Lightfoot (1979) uses early work by van der Gaaf (1904) to argue that, due to the loss of morphological case, the argument roles are opaque and a reanalysis takes place with the experiencer role as subject. Allen (1995) questions this relationship between the loss of case and the change to personal constructions and argues instead for a gradual spread of a particular lexical frame. Since Hale & Keyser's (2003) views where argument structure, i.e. thematic structure, is represented as a vP-shell, this model has been used to account for changes in verb particle constructions (Elenbaas 2007) and a loss of intransitives (van Gelderen 2011c). Reintges (1997; 2005) investigates the passive and aspect in the history of Egyptian in this framework.

Recently, there has been a shift towards looking at discourse and information structure. Van Kemenade (1987) recognizes that *wh*-elements and negatives in initial, pre-verb position trigger absolute V-second but that, with topicalized elements, the subject pronoun can follow resulting in V-third, as (13) shows.

- (13) *Das þing we habbað be him gewritene ...*
these things we have about him written

'These things we have written about him.' (Chronicle E, 1087, 143, van Kemenade 1987: 110)

This early work provides the seed to the current shift by van Kemenade and others, e.g. Los (2012), to connecting word order and information structure. This construction also leads Kroch & Taylor (1997) to investigate the regional variation in the Verb-third of (13) and Speyer (2008) revisits the issue and argues that the choice between verb-second and verb-third is determined by a requirement to avoid two focused elements next to each other.

I'll now turn to feature-based accounts. As mentioned in 2.2, in the Minimalist Program (Chomsky 1995), all parameters are encoded in the lexicon and this has as a consequence that linguistic variation and change is related to the morphological properties/features of the lexical items. Lexical items have features that vary across languages, but can be divided into phi-features (number, person, and gender), case features (dependent marking of nominals by the T(ense) and light verb *v*), and EPP features (the features that are responsible for movement of the grammatical subject).

The inventory of features is innate (see e.g. Chomsky 1965: 142 and 2007: 6) and the child has to choose which of these are relevant in the language to be acquired. The number and kind of features is still debated. Cinque and Rizzi (2008) discuss the question of the number of functional categories. There are 32 in Cinque (1999: 130), and around 40 in Kayne (2005). Cinque and Rizzi, using Heine & Kuteva's (2002) work on grammaticalization, come up with 400 features that are targets in Heine & Kuteva. Benincà & Munaro (2010: 6-7) note that syntax has reached the detail of phonological features. Differences among the inventory of features account for changes in prepositions, complementizers, auxiliaries, demonstratives and the like.

According to Roberts (2009), we need "an appropriate feature system which breaks down major categories (N, V, etc) into smaller ones (count noun, transitive verb, etc.) up to a fairly fine-grained level. In the case of functional categories, which will be the major concern in what follows, we will be dealing with categories such Modal, which can be divided into epistemic, alethic, deontic, etc."

So, let's assume that the inventory of features is given and that the semantic and phonological interfaces require that their input is legible, this being responsible for the various formal features. A model of how syntax interfaces with the performance systems is given in Figure 1. The Sensorimotor (SM) and Conceptual-Intentional (CI) systems are

external to the syntax but interact with the syntax through the interfaces PHON and SEM, corresponding to PF and LF in older frameworks.

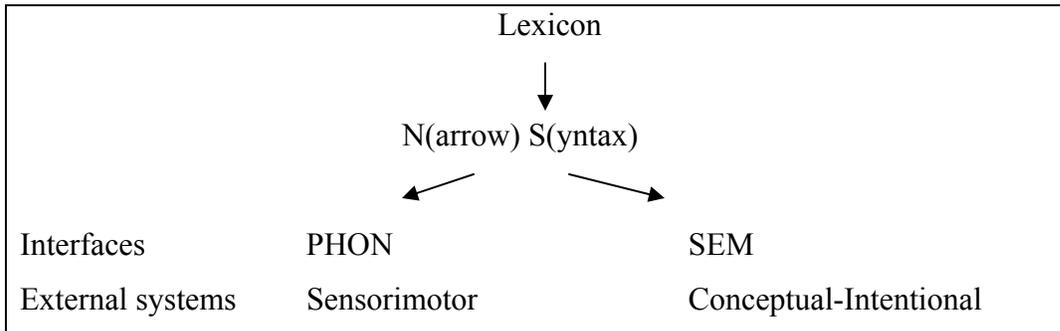
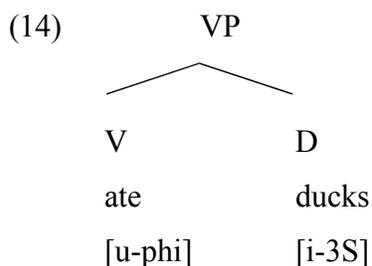


Figure 1: Interfaces

Phonological features are obviously not needed at CI and semantic features are not legible at SM. Most formal features in Minimalism therefore come in two kinds: interpretable at CI and uninterpretable at CI. The child has to learn this as well and may analyze features that are interpretable in the older generation as uninterpretable. An example of this is uninterpretable verbal agreement that derives in many languages from the interpretable features on nouns and pronouns. I'll go into a Minimalist derivation a little more to show the various features.

In Chomsky (1995, 2004, etc), the phrase structures of (5) are abandoned in favor of a general rule Merge. Merge combines two bundles of features taken from the lexicon and would look like (13), with only the phi-features marked². The [i-3S] in (14) stands for interpretable third person and singular number features on the noun and the [u-phi] for uninterpretable, and as yet unvalued, person and number features on the verb.



² Note that I am simplifying. The phi-features should be on a light verb.

In (14), we see that the features of the verb are uninterpretable and those on the noun interpretable. As mentioned, the reason for this distinction is whether they are relevant at the semantic interface or not. The features of the verb will get valued before being transferred to PHON. Figure 2 provides the interpretable and uninterpretable features of the noun *airplane* and the verb *build*.

	<i>airplane</i>	<i>build</i>
interpretable:	[nominal] [3 person] [non-human]	[verbal] [assign accusative]
uninterpretable:	[case]	[phi]

Figure 2: Interpretable and uninterpretable features of *airplane* and *build* (adapted from Chomsky 1995: 278)

To (14), we add functional categories, such as T and C, each with their own sets of features. The question for a bottom-to-top approach, as in (14), is how the universal sequence is achieved where C is the highest functional projection, followed by T(ense), and ASP(ect) and various other functional categories. The cartographic approach (Rizzi 1997; Cinque 1999) provides an answer through a Functional Hierarchy, as in (15), for a subset of cases.

- (15) T_{past} T_{fut} Moodir Modnec Modpos ASP_{hab} ASP_{prep} ASP_{freq}
once then perhaps necessarily possibly usually again often
 (from Cinque 1999: 107)

Although Chomsky (2002) says that cartography is not incompatible with minimalism, I personally see a clash between a derivational approach, as in (13), and a representational approach, as in (15).

To finish, I add a few examples of feature changes but without much justification provided here. As well-known, demonstratives can be reanalyzed as articles, complementizers, and copulas. This is shown in Figure 3 as a change in features. The child acquiring English might have a demonstrative in its lexicon with location ('loc')

and person and number features. These features are interpretable and the demonstrative can be on its own. The child then also uses the demonstrative with fewer features.

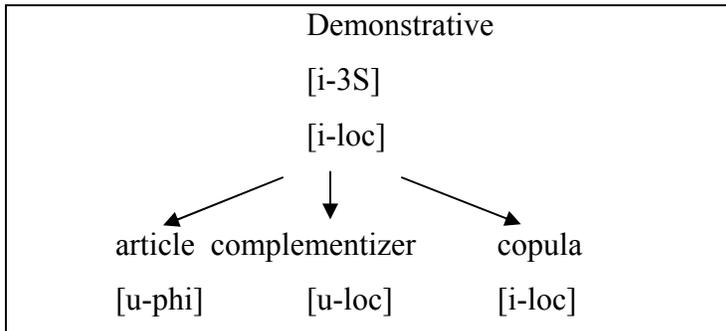


Figure 3: Reanalyses of the demonstrative pronoun

Also well-known is that adpositions, shown in Figure 4, can be reanalyzed as verbs, complementizers, and case (see Heine & Kuteva 2002 for an overview). Again, this represents use of the features in a different way.

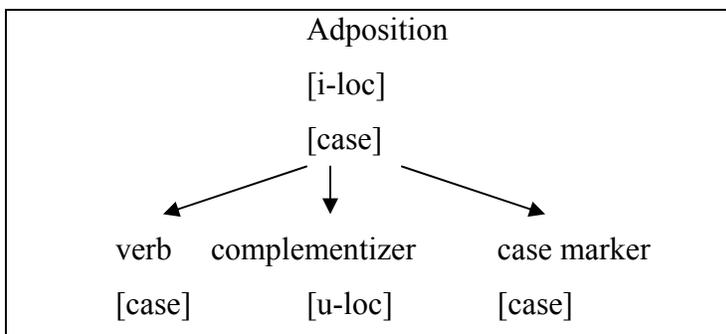


Figure 4: Reanalyses of the adposition

The shift in current Minimalism from syntactic parameters to lexical parameters, by means of feature choices, makes it easier to translate the insights from the grammaticalization literature into generative grammar. What remains unclear is the question of how to constrain the vast number of possible features. Connected to this is the role of cartography (see Salvi 2005 and Danckaert 2012 for diachronic work on the cartography of the left periphery)

4 Major contributions of formal syntax and unsolved questions

Generative linguistics focuses on language change as instigated by the language learners faculty of language and not by communicative, functional needs. This has numerous

advantages. I will list a few in this section and I also outline a few unresolved questions. I won't consider disadvantages of which there are also many.

One advantage over purely functional approaches is that there are numerous instances where a communicative breakdown does not explain change. For instance, why would reflexive pronouns be introduced for first and second person, as happens in the history of English and in other languages. Old English uses personal pronouns as reflexives, as in (16), and there is no functional reason to add *self* to first and second person as starts to happen from late Old English on.

- (16) *Wit unc wið hronfixas werian þohton*
we us-DAT against whales defend thought
'We intended to defend ourselves against the whales.' Beo 540

Formal syntax, by focusing on internal mechanisms, also finds patterns in the data that would be missed if one just looked at communication, e.g. the *that*-trace phenomena of (11). Claire Bowers (p.c.) suggests that generative approaches are useful in explaining changes that result in system reorganization, as in (4) above. Functional approaches have a hard time with such changes.

Some challenges that I have mentioned in previous sections are the need to make the 'third' factor more precise and to restrict the role of features. Chomsky talks of 'third factor' principles of computational efficiency but these are not well-defined. In Minimalism, a major role is played by features, semantic and grammatical, and their power needs to be restricted. Finally, the tension between a Minimalist model of trees without labels and highly specific structures, as in (15), needs to be understood better (see van Gelderen to appear).

5 Conclusions

Historical linguistics and generative grammar have had an uneasy relationship. Proponents of generative grammar love data that are elicited from native speakers and that is impossible in historical linguistics. Generative linguists were initially very skeptical about corpus data because they would not be representative of the underlying

structure. However, the use of corpora, so prevalent in historical linguistics, is gradually becoming accepted over just the use of native speaker judgments.

I have showed how some of the changes in emphasis of the generative framework have impacted historical generative linguistics. I have identified an interest in word order and verb-movement especially from the 1980s to the present, and interest in argument structure, a recent interest in information structure and currently in features and cartography.

Abbreviations are to be completed

DIGS

I

ICEHL

ICHL

OV

P&P

VO

References are to be completed

Aldridge, Edith 2011. Antipassive in Austronesian Alignment Change. *Grammatical Change: Origins, Nature, Outcomes*. Dianne Jonas, John Whitman, Andrew Garrett, eds, 331-345. Oxford University Press.

Allen, Cynthia 1977. Topics in Diachronic English Syntax. UMASS Diss.

Allen, Cynthia 1995.

Baker, Mark 2008. *The syntax of Agreement and Concord*. Cambridge: Cambridge University Press.

Battlori, Montserrat & Lluïsa Hernanz (eds) 2011. *Generative Diachronic Syntax: Word Order and Information Structure*. Special issue of the *Catalan Journal of Linguistics*.

Battye & Roberts, Ian 1995. Clause Structure and Language Change.

Bever, Thomas & Terry Langendoen 1971. A Dynamic Model of the Evolution of Language. *Linguistic Inquiry* 2.4: 433-463.

- Breitbarth, Anne 2012. The Development of Conditional Modal Verbs. DIGS 14, Lisbon, July.
- Canale, Michael 1978. Word order change in Old English. McGill PhD.
- Chomsky, Noam 1986. *Knowledge of Language*. New York: Praeger.
- Chomsky, Noam & Halle, Morris 1968. *The Sound Pattern of English*. Cambridge: MIT Press.
- Closs, Elizabeth 1965. Diachronic Syntax and Generative grammar. *Language* 41.3: 402-415.
- Eyþórrson, Thorhalldur 1995. Verbal Syntax in the Early Germanic Languages. Cornell Dissertation
- Gelderen, Elly van 1993. *The Rise of Functional Categories*. Amsterdam: Benjamins.
- Gelderen, Elly van 2008. Where did Late Merge go? Grammaticalization as Feature Economy. *Studia Linguistica*: 287-300.
- Gelderen, Elly van 2011a. *The Linguistic Cycle*. OUP.
- Gelderen, Elly van 2011b. Grammaticalization and generative grammar. In Narogg & Heine, 43-55. OUP.
- Gelderen, Elly van 2011c. Valency changes in the history of English”, *Journal of Historical Linguistics* 1.1: 106-143.
- Gelderen, Elly van to appear. *Clause Structure*. Cambridge: Cambridge University Press.
- Haiman, John 1974. *Targets and Syntactic Change*. The Hague: Mouton.
- Halle, Morris 1962. *Phonology in Generative Grammar*. Chicago: Word.
- Joos, Martin 1942. A phonological dilemma in Canadian English. *Language* 18.2. 141-144.
- Kemenade, Ans van 1987. Syntactic Case and Morphological Case in the History of English. Dordrecht: Foris.
- Kemenade, Ans van & Nigel Vincent (eds) 1997. *Parameters of Morphosyntactic Change*. CUP.
- Kiparsky, Paul 1965. Phonological Change. MIT PhD.
- Kiparsky, Paul 1996. Indo-European Origins of Germanic Syntax. In Ian Roberts and Adrian Battye (eds.), *Clause Structure and Language Change*, 140-167. Oxford: Oxford University Press.

- Klima, Edward 1965. *Studies in Diachronic Transformational Syntax*. Harvard PhD.
- Lakoff, Robin 1968. *Abstract Syntax and Latin Complementation*. MIT Press.
- Lightfoot 1974,
- Lightfoot, David 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- Lightfoot, David 2006. 'Cuing a New Grammar'. In Ans van Kemenade & Bettelou Los *The Handbook of The History of English*, 24-44. Oxford: Blackwell.
- Lightfoot, David 2006. *How new languages emerge*. CUP.
- Los, Bettelou 2000. *Infinitival Complementation in Old and Middle English*. Free University, Amsterdam PhD.
- Newmeyer, Frederick 1998. *Language Form and Language Function*. Cambridge: MIT Press.
- Pintzuk, Susan 1991. *Phrase Structures in Competition*. University of Pennsylvania Dissertation.
- Reintges, Chris 1997; 2005 in Kiss
- Roberts, Ian 1993. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer.
- Roberts, Ian 2009. Semantic Bleaching, Grammaticalisation and the Clausal Hierarchy. In Elizabeth Traugott & Graeme Trousdale (eds), *Gradience, Gradualness and Grammaticalization*, ** Amsterdam: John Benjamins.
- Roberts, Ian & Roussou, Anna 2003. *Syntactic Change*. Cambridge: CUP.
- Salvi 2005 (in Kiss)
- Steele, Susan et al 1981. *An Encyclopedia of Aux*. MIT Press.
- Traugott, Elizabeth 1972. *A History of English Syntax*. New York, Holt, Rinehart and Winston.
- Vincent, Nigel 2001. LFG as a model of syntactic change. In Miriam Butt & Tracy Holloway King (eds), *Time over Matter*, 1-42. Stanford: CSLI.
- Meillet, Antoine. (1921 [1912]) "L'évolution des formes grammaticales". In: *Linguistique historique et linguistique générale* (ed. Édouard Champion). Paris: Librairie Ancienne Honoré Champion, 130-148.

Further reading

- Biberauer, Theresa (ed) 2008. *The Limits of Syntactic Variation*. Benjamins.
- Breitbarth, Anne et al (eds) 2010. *Continuity and Change in Grammar*. Benjamins.
- Crisma, Paola & Giuseppe Longobardi (eds) 2009. *Historical Syntax and Linguistic Theory*. OUP.
- Kiss, Katalin (ed) 2005. *Universal Grammar and the Reconstruction of Ancient Languages*. Berlin: Mouton de Gruyter.
- Longobardi, Giuseppe 2001. Formal syntax, diachronic Minimalism, and etymology: the history of French *chez*. *Linguistic Inquiry* 32, 275–302.
- Hinterhölzl, Roland & Svetlana Petrova 2009. *Information Structure and Language Change*. Berlin: Mouton de Gruyter.