

1

Micro-change and macro-change in diachronic syntax

ERIC MATHIEU AND ROBERT TRUSWELL

1.1 Syntactic theory and syntactic change

Syntactic theory has come a long way in the past couple of decades. Despite superficial disagreements, almost every researcher in every major framework is now committed to a conception of a grammar as a large, language-particular, and open-ended population of lexical items (richly specified correspondences between sets of phonological, syntactic, and semantic properties), related by a small, universal set of syntactic operations (see, e.g., Chomsky 1995, Steedman 2000, Bresnan 2001, Sag 2012). Gone are the phrase-structural or transformational rules of Chomsky (1957) or Gazdar *et al.* (1985); the parameters of Chomsky (1981) are on their way out, too, except insofar as they can be reduced to properties of individual lexical items (e.g. Borer 1984) or perhaps to properties of interfaces between syntax and neighbouring domains (Berwick and Chomsky 2011).

By and large, though, this pared-down syntactic theory has to hold itself accountable to the same range of data as ever. In particular, the papers in this volume are concerned with a series of logical and empirical problems relating to syntactic change. Studying syntactic change can help us test and refine our synchronic syntactic theories; it can also remove some of the empirical burden from synchronic theories, by providing a diachronic basis for attested facts, whether at the level of individual languages or typological generalizations. Finally, it can provide a new source of insight into related matters such as language variation or acquisition.

All of these issues are daunting, because modern syntactic theory gives us so little to work with. If a grammar is just a handful of universal operations or relations, plus a lexicon, then all syntactic change must ultimately reduce to lexical change (see also Walkden 2012). More specifically, most of the researchers in this volume adopt the feature-based approach to syntactic relations from Chomsky (1995, 2000, 2008). For those researchers, in most cases, syntactic change must reduce to a change in the featural specification of one or more lexical items. That really is all there is: other, apparently different

modes of explanation (for example, claims that a certain constituent used to move to the left periphery, but now merges there directly, e.g. Roberts and Roussou 2003; or that a certain constituent now moves to a new landing site) are just abridged descriptions of changes in particular configurations of lexical items with certain, interacting featural specifications.

This is the most basic type of diachronic syntactic research: reduce changes in surface syntactic patterns to a series (minimally two) of synchronic ‘snapshots’ of the underlying grammars, conceived of as sets of lexical items which interact to yield the surface patterns. Many papers in this volume engage in this kind of detailed analysis. As well as its own intrinsic interest, this endeavour feeds into synchronic syntactic theorizing: minimally, every postulated synchronic ‘snapshot’ of a grammar G carries with it the implicit claim that G is a possible human grammar: we cannot postulate illegitimate synchronic states, just to get ourselves more smoothly from A to B . Although this sounds trivial, it needn’t be: to take two familiar examples, van Gelderen’s (2004) analyses of certain cyclical changes in terms of tendencies for specifiers to become heads, and for interpretable features to become uninterpretable, if accurate, commit us to a syntactic architecture in which specifiers can be distinguished from heads, and uninterpretable features from interpretable. Many varieties of Minimalism make the distinctions that van Gelderen needs; other theories may not. In this way, a theory of syntactic change implies a theory of syntax.

At one step further removed from synchronic syntactic theory, snapshot-based analyses of particular syntactic changes, when aggregated, imply notions of ‘possible (and probable) syntactic change’. Given a series of snapshots $\langle G_1, \dots, G_n \rangle$, we can infer a crude analogue (to be refined below) of a series of grammatical changes, as sets of paired grammars $\{\langle G_1, G_2 \rangle, \langle G_2, G_3 \rangle, \dots, \langle G_{n-1}, G_n \rangle\}$. It is possible, in principle, to aggregate such sets of changes, and take them to imply a roadmap of sorts through the space of possible grammars: grammars like G_i tend to develop into grammars like G_j or G_k , but not grammars like G_g or G_h . Once we have enough pairs of snapshots, we may expect to feel confident making such statements.

Something like the above model is visible in several of the most lively current approaches to grammatical change, including grammaticalization theory (Heine and Kuteva 2002) and cyclical change (van Gelderen 2011). In fact, though, no one currently works in exactly the way sketched in the preceding paragraph. Once we understand why, a range of further diachronic questions open up which point to the real challenge, and real interest, posed by trying to reconstruct a theory of syntactic change within such a minimal architecture.

Firstly, for all the architectural minimalism, grammars remain extremely complex objects. We are all told in LING101 that the average college student knows 30,000 words, or maybe 60,000, or 100,000. On the lexicalist approach to

syntax espoused above, that means that a grammar consists of tens of thousands of statements about correspondences between bits of phonology and semantics with bundles of syntactic features. Any change from G_1 to G_2 is doomed to be a hapax legomenon when G_1 and G_2 are objects of such complexity.

We can reduce the problem somewhat, by making a few further assumptions about syntactic theory. For instance, the exoskeletal approach (Marantz 1997, Borer 2005*a,b*), which pairs category-neutral roots with c-commanding categorizing heads, means that the majority of those tens of thousands of lexical items have no syntactic features, and, *a fortiori*, can only participate in syntactic change by gaining syntactic features. That makes the syntactic hapax legomenon problem an order of magnitude smaller, but still quite real.

We could take a step further, if we remember that a grammar is just a collection of lexical items. Rather than comparing entire grammars, we could extract pairs of ‘corresponding’ lexical items, and note any changes. Rather than seeing syntactic change as a pair $\langle G_1, G_2 \rangle$, we would thereby construct a more fine-grained object, a set of pairs of sets of syntactic features associated with corresponding lexical items $\{\langle LI_{11}, LI_{12} \rangle, \langle LI_{21}, LI_{22} \rangle, \dots, \langle LI_{n1}, LI_{n2} \rangle\}$. Regularities in syntactic change can be established at this finer grain: maybe LI_{11} can correspond to LI_{12} in an immediately subsequent snapshot, but never LI_{13} ; or perhaps if LI_{11} corresponds to LI_{12} in a subsequent snapshot, then LI_{21} must correspond to LI_{22} . Recurring pathways of change boil down to statements of the form: if grammar G_1 contains LI_{11} , then with greater than chance frequency, grammar G_2 will contain either LI_{11} or LI_{12} . And in the ideal case, possible and likely long-term syntactic changes would be nothing more than the transitive closure of these ‘instantaneous’ paired snapshots. As Lightfoot (2002) has vigorously argued, the notion of ‘correspondence’ is far from unproblematic. However, Walkden (2012) shows that phonological and semantic similarities, as well as independently established properties of probable syntactic changes, have significant heuristic value.

All of this can be recast as a theory of reanalysis in acquisition: statements about possible and probable changes, about contingently related sets of changes, and so on, can be seen as statements of the form ‘If speaker S_1 has a grammar containing LI_{11} , and if S_2 learns from S_1 , then S_2 may/will/will never induce a grammar containing LI_{12} .’

At this point, we have something approaching a formally coherent theory of syntactic change, with at least the hope of avoiding the hapax legomenon problem and having some predictive power. (Of course, the more restrictive our theory of syntactic features, the more restrictive our theory of change in the featural composition of lexical items will independently be, but this is one of the more contentious areas of current synchronic syntactic theory, and we prefer to avoid wading in). However, such a theory would be far from satisfactory to any

specialist in syntactic change. Here are just a few examples of problems that cannot yet be analysed in the above, snapshot-based approach.

- It is not straightforward to state when a language has undergone a syntactic change, because of synchronic variation.
- Many syntactic changes play out over centuries, or even millennia.
- Complex, long-term changes form recurrent ‘pathways’, often composed of series of subsidiary changes.

Many of these facts cannot even be stated in the vocabulary developed above, without doing violence to the presumed cognitive basis of these groups of corresponding lexical items, rooted in the acquisition process. To give an example, imagine an exceptionally well-behaved cycle, in which Stage 1 always progressed to Stage 2, and Stage 2 to Stage 3. We can describe this using the ideas we have developed, and have done so above. However, such a cycle would put language learners in an awkward position: they can innovate and create a Stage 1 or Stage 2 grammar, but only if they promise that some subsequent generation will go further and create a Stage 3 grammar.

Of course, we must not say things like this, but the empirical phenomena will not go away. And perhaps perfectly well-behaved cycles do not exist, with some languages getting stuck at Stage 1 or Stage 2, and others going somewhere else altogether, but the phenomenon of long-term change is real enough: it is not an accident that researchers see Jespersen’s Cycle, for example, as a set of changes with an internal cohesion over and above the usual, despite the fact that we can take dozens of generations to get from the start to the end.

It is now common to take the first two of the problems enumerated above as indicative of interactions between multiple grammars. This has its origin in the seminal results of Kroch (1989), which introduced the Constant Rate Hypothesis, and derived the CRH from a model of grammar competition. Again, given the limited resources at our disposal, it makes sense to make grammar competition do as much for us as possible.

As for the third problem, that of recurring pathways of change, one possibility is to see this as evidence for biases in syntactic acquisition. Here, too, there is a link to multiple grammars. As described above, a complex, cyclical change can be decomposed into a series of ‘instantaneous’, simple changes. If transitions like $\langle LI_{11}, LI_{12} \rangle$, $\langle LI_{12}, LI_{13} \rangle$, and $\langle LI_{13}, LI_{14} \rangle$ are all very likely (either likely to happen in the first place or likely to spread through a population), then the result will be a long-term pathway of change which is still likely to be followed. The goal at this point is to explain why these changes are very likely to occur, rather than others. One very common answer is that certain biases in acquisition entail that syntactic change is not a free-for-all.

This is our area, then. We start small, obsessing about the internal structure of lexical items, the set of features which can be lexically associated with a head, and so on; but basic considerations of empirical adequacy force us to consider population dynamics, interactions between multiple grammars, and acquisition biases. Neither half is complete or satisfying without the other: we have seen that individual lexicalist grammars alone lead to incomplete theories of change, but the larger scale considerations only have any predictive teeth if they are related to well-defined basic objects. The micro and the macro: the two sides of a modern syntactic theory of change.

1.2 Overview of the volume

We have not attempted to group these papers into parts: the major themes are too intertwined, and any attempt to draw boundaries just leads to artificiality. However, the chapters at the beginning of the volume are ‘more macro’: they concern gradual change, among individuals and among larger populations, often over long periods of time. As the volume progresses, we focus increasingly on the micro-changes: particular proposals about particular groups of features or lexical items, in a couple of languages, over shorter time periods.

The two styles of analysis cannot easily be separated: macro-level work builds on concrete diachronic grammatical analysis; concrete fine-grained diachronic work gains an extra dimension of interest from being considered in the big picture. Although diachronic syntax is far from a predictive science, we are not scared as a discipline of asking why. In other words, although we cannot predict what will happen next in any individual case, we can, and should, attempt to describe what will tend to happen. And given the uniformitarian perspective of most diachronic syntactic research, every micro-level piece of research here is an implicit example of the kind of thing that tends to happen, grist to the mill of macro-level theorizing.

We begin with a series of chapters that explore the relationship between acquisition and pathways of gradual change in nonuniform populations of linguistic agents. **Cournane** focuses on the relationship between acquisition biases and the typology of change. She distinguishes two ways in which acquisition pathways can relate to diachronic pathways: they can align (root modals develop before epistemic modals in acquisition; epistemic modals often arise from root modals by a process of grammaticalization), or they can oppose (monoclausal structures with parentheses are acquired before subordinating multiclausal structures, but multiclausal structures can grammaticalize into monoclausal structures with auxiliary and main verb). Cournane proposes that

the two patterns relate to two different areas of grammar, and so to two different acquisition tendencies: alignment results from semantic overextension during acquisition, while opposition between acquisition pathways and diachronic pathways results from conservativity biases during morphosyntactic acquisition.

Gisborne and Truswell share a focus on acquisition biases, but with a view to explaining typological particularities, rather than universals. Their focus is on parallel syntactic evolution, the establishment of a recurring pathway within a language family which is only rarely attested outside that family. For example, interrogative forms and demonstrative nominals repeatedly develop into phrasal headed relative markers in Indo-European languages, but not in other languages. Gisborne and Truswell claim that this can serve as a new source of evidence about acquisition biases: parallel changes in the syntax of cognate lexical items suggest that syntactic change, and so the acquisition of syntax, is biased.

This typological perspective is shared by **Whitman and Ono**, who take up the longstanding idea (Givón 1975, Aristar 1991) that crosscategorical word order generalizations, like many of Greenberg’s (1963) implicational universals, can be attributed to recurring patterns of change. Whitman and Ono perform a clustering analysis on the properties described in the *World Atlas of Language Structures Online* (Dryer and Haspelmath 2013), in order to reveal robust patterns of co-occurrence between word order parameters. Whitman and Ono show that the only robust crosscategorical generalizations to emerge (other than those restricted to unmarked values) concern the order of a head and its arguments. These can be reduced to ‘relabelling’ instances of grammaticalization (for instance, prepositions developing from verbs and maintaining the relative position of the verb’s complement). A relatively restricted and well-attested set of recurring reanalyses can therefore be taken as the motor driving crosscategorical word order generalizations.

The next two papers address the range of effects that can arise from language contact, maintaining the above focus on increases in complexity arising from interactions among multiple grammars, but in a somewhat different form.

É. Kiss shows through an analysis of changes in the Hungarian verbal complex that contact can lead to pervasive changes in syntax. She demonstrates specifically that the Hungarian tense system became more complex under Turkic influence, and then simplified under subsequent Slavic influence, raising questions about structural conditions for borrowing of such systems. É. Kiss argues that the evolution of complex tenses started with a micro-change: the reanalysis of the feature content of a verbal suffix. This step initiated further processes of reanalysis, analogical extension, and abstraction, as a consequence of which the tense system inherited from Uralic, distinguishing only past and nonpast, developed into a complex system marking both tense and viewpoint aspect.

A longstanding obsession in diachronic syntax has been the S-curve as a model of the time course of linguistic change (Osgood and Sebeok 1954, Weinreich *et al.* 1968, Bailey 1973, Altmann *et al.* 1983, Kroch 1989, Blythe and Croft 2012). Three papers here investigate an alternative pattern, in which a change initially appears to move away from its eventual end-state. **Postma** and **Bacovcin** both address the dynamics of failed changes, where a change gets underway but is never fully actualized, from apparently incompatible perspectives. Postma develops his earlier work (Postma 2010) tying failed changes to successful changes, presenting a reconceptualization of Kroch’s Constant Rate Hypothesis, expanded to include a *transient state* (reflecting the failed change) between the initial and final grammars in a competition process.

Bacovcin argues, based on the diachrony of a short-lived Middle English pattern in which a ditransitive verb phrase is realized as *V to-NP NP*, that at least some failed changes are an emergent result of interactions between two changes which go to completion. On Bacovcin’s analysis, the *V to-NP NP* pattern emerged as a result of a reanalysis of *to* as a dative marker, and disappeared as a consequence of a second reanalysis, a couple of centuries later, of the recipient in recipient–theme orders as accusative-marked. The *V to-NP NP* order spread with the first reanalysis, and disappeared as the second reanalysis spread.

Meanwhile, **Troberg and Burnett** show that the development of verb-framed modern Romance languages from satellite-framed Latin does not reduce to an incremental accumulation of verb-framed constructions. Rather, Medieval French, *en route* to the verb-framed behaviour of modern French, passed through a stage in which it exhibited even more satellite-framed behaviour than Latin. This is not a U-shaped curve in the same sense as Bacovcin, then: Troberg and Burnett are not concerned with fluctuations in frequency over time. Rather, they are concerned with trajectories between two idealized grammar ‘types’. They analyse the unusual progression from Latin to Medieval French to Modern French as a series of two reanalyses: first, place prefixes (which can be supplemented with a null path morpheme) are reanalysed as Path heads, leading to the innovation of resultative and verb–particle constructions. Secondly, the Path head is reanalysed as an inseparable part of the verb, leading to the loss of the whole range of satellite-framed constructions.

Lavidas aims to explain the appearance of an unusual pattern of accusative-marked subjects of finite predicates in Early Byzantine Greek. The predicates in question are usually mediopassive verbs; Lavidas claims that the unusual, and quite short-lived, option of an accusative subject of such verbs arose as a result of developments in the case system: Accusative case was initially interpretable (in opposition to partitive) and would subsequently become reanalysed as a structural case marking objects; meanwhile, the voice morphology is argued to show a complementary change: initially arising as a result of agreement with the accusative subject, once accusative case became

structural, voice morphology became inherently determined. The unusual intermediate system arose during the transition between the two more standard systems.

By now, we are approaching the ‘micro’ level of diachronic syntax. The meat of Bacovcin’s analysis (if not the implications for the modelling of change) consists of small changes in the English case system; and the analyses of Troberg and Burnett, and Lavidas, are based on equally small changes, in the French system of path/place prefixes and prepositions and the Greek aspectual system respectively. Much of the current fascination of diachronic syntax comes from the large-scale effects arising from such small lexical changes: the relationship between surface phenomena and underlying analyses is far from straightforward, as has been emphasized repeatedly since Andersen (1973) and Lightfoot (1979). One example of the complexity of this relationship, currently the subject of intense analysis, concerns the range of fronting operations in Germanic and Romance. It has become increasingly clear that the by-now well-established descriptive label ‘verb-second’ covers multiple possible featural specifications for T and C heads, and beyond. At this stage, even the synchronic typology is not definitively established, and it so happens that many ‘old’ languages (including Old French, Old English, and Old High German) are of particular interest in refining that typology. **Labelle and Hirschbühler** concentrate on the related phenomenon of Stylistic Fronting, attested in modern insular North Germanic varieties and described in Old French by Mathieu (2006). Labelle and Hirschbühler dispute this analysis, claiming that the construction described by Mathieu does not share certain properties with North Germanic Stylistic Fronting, and actually does not reflect a single syntactic structure, but rather a range of structures differentiated by patterns of word order, derived by phrasal movement of VP, possibly in combination with short scrambling of the object. There is thus not one Leftward Stylistic Displacement construction but different constructions allowed by the grammar.

Salvesen and Walkden focus on the distinction between symmetric and asymmetric V2 languages. In contrast to modern German, the V2 language *par excellence*, Old French and Old English show a number of V3 and other orders. This has led to different analyses being proposed for Old French and Old English: some researchers aim to maximize the similarity to German, by postulating V-to-C movement in all cases, together with housekeeping measures to account for V3 orders; other analyses place the verb in these languages lower, for example in T⁰, leaving more scope for multiple constituents to appear on its left. Moreover, the range of V3 constructions in Old French and Old English is not identical: Old English is largely as predicted by this description, while Old French V3 orders are more limited in scope, and largely reduce to the constructions which Labelle and Hirschbühler address in their chapter. Because of this, it is unclear whether a unified analysis of Old French and Old English is

possible, or even desirable. Salvesen and Walkden’s contribution is to examine the distribution of embedded V2. If V in these languages does not raise past T, then embedded V2 should be quite unremarkable. If, on the other hand, V has raised to C, then we are inclined to treat embedded V2 as an embedded root phenomenon, of the sort first described by Hooper and Thompson (1973): only compatible with certain classes of matrix predicate. Salvesen and Walkden show that the latter prediction better fits the corpus data, generating an expectation that V3 orders can be explained elsewhere, for instance within a more articulated syntax of the left periphery.

Finally, **Light** uses current ideas about the syntax of V2 in Germanic to argue that demonstrative pronouns across historical Germanic varieties are interpreted as contrastive. The argument runs as follows: many Germanic varieties allow fronting of contrastive elements, but only German allows fronting to preverbal position of noncontrastive elements other than the subject (a fact which Light links to the wider range of possibilities for scrambling to the *Mittelfeld* in German). However, 16th century Bible translations in Early New High German, Early Modern English, and Icelandic all contain instances of fronting of object demonstrative pronouns to preverbal position. Light takes this as evidence that object demonstratives are interpreted as contrastive in these languages, as the only preverbal noncontrastive elements are subjects. She links this to a technical characterization of contrast, building on the pragmatic notion of *subinformativity* (Gast 2010).

Ecay and Tamminga develop a new method for relating surface phenomena to underlying grammatical patterns, similar in its scope and outlook to the Constant Rate Effect of Kroch (1989). The logic of the Constant Rate Effect is that, if two surface phenomena are linked to a common underlying mechanism, the rates of change of the two phenomena across time should be similarly linked. Ecay and Tamminga’s new diagnostic relies instead on *persistence*, or non-independence of sequential utterances as a consequence of priming-like effects. If persistence can be demonstrated between two surface phenomena, that could indicate that they have underlying elements in common. Ecay and Tamminga apply this logic to the analysis of Middle English negation, which surfaced in three variants: *ne*, *ne . . . not*, and *not*. Competing accounts have analysed this surface alternation either in terms of two possibly co-occurring underlying atoms, or three independent forms. The evidence from patterns of persistence supports the three-atom model.

At the most micro level, diachronic syntax involves tracking developments in the syntactic composition of a single form and its descendants. For example, **Danckaert** deals with a typical problem in diachronic syntax: there are two candidate sources in an ancestral language for a later construction, and the task is to distinguish which is the actual source. His study involves the Latin sources of Romance analytic passives. Late Latin had two analytic passive forms: the

infectum (e.g. *amatus sum*), and the *perfectum* (e.g. *amatus fui*). It has typically been assumed that the *infectum* is the source of the modern Romance analytic passive (e.g. French *Je suis aimé* ‘I am loved’), because of the formal continuity between *sum* and *suis*. However, Danckaert shows that word order facts support the opposite conclusion: forms like *fui*, along with most auxiliaries, tend towards preverbal position in Late Latin, while *sum* remains largely postverbal. This suggests that largely preverbal *fui* is the source of the modern Romance construction, while the *sum* form came to behave syntactically more like the ancestors of the future marker. This in turn raises a number of questions about how the modern paradigms emerged from these formally distinct ancestors.

Courtney discusses the Dutch inflectional paradigm, with a focus on complementizer agreement. Competing syntactic and postsyntactic analyses of complementizer agreement exist in the literature; Courtney’s contribution is to argue that the two types of analysis are each appropriate to different dialects, and that it is plausible to construe the synchronic microvariation between syntactic and postsyntactic agreement patterns as a reflex of gradual grammaticalization of the agreement relation, in the sense of recent Minimalist approaches to grammaticalization such as that of Roberts and Roussou (2003).

Finally, **Jędrzejowski** tracks the directional change of a single lexical item, German *versprechen* ‘promise’. As in many languages, *versprechen* is polysemous: its best-known interpretation is as a control verb, with the speaker committed to a state of affairs described in a complement clause. However, it also can have a raising interpretation, typically with inanimate subjects like *das Wetter* ‘the weather’, an interpretation without speaker commitment, and other changes in argument structure. Jędrzejowski uses well-established grammaticalization mechanisms to establish a pathway in which the raising interpretation of *versprechen* emerges first in constructions embedding a nominal, rather than clausal complement.

By now, we have come a long way from the very general patterns of change in the earlier chapters. But the point of modern diachronic syntax is that you can’t do one without the other. Specific long-term time courses of change must emerge from discrete, local, intergenerational changes in the lexical specification of grammatical features. No paper in this collection is far from a reduction of syntactic change to ‘imperfect’ lexical acquisition; no paper concentrates on a phenomenon so small as to have no broader architectural implications. The counterpoint between small, discrete changes and large-scale, emergent, multigenerational diachronic phenomena is the core of this new discipline.

Bibliography

- Altmann, Gabriel, von Buttlar, H, Rott, W, and Strauß, U (1983). A law of change in language. In *Historical Linguistics* (ed. B. Brainard), pp. 104–115. Studienverlag Brockmeyer, Bochum.
- Andersen, Henning (1973). Abductive and deductive change. *Language*, **49**, 765–793.
- Aristar, Anthony (1991). On diachronic sources and synchronic patterns: An investigation into the origin of linguistic universals. *Language*, **67**, 1–33.
- Bailey, Charles-James (1973). *Variation and Linguistic Theory*. Center for Applied Linguistics, Arlington, VA.
- Berwick, Robert and Chomsky, Noam (2011). The biolinguistic program: The current state of its evolution and development. In *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty* (ed. A. M. Di Sciullo and C. Boeckx), pp. 19–41. Oxford University Press, Oxford.
- Blythe, Richard and Croft, William (2012). S-curves and the mechanisms of propagation in language change. *Language*, **88**, 269–304.
- Borer, Hagit (1984). *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Foris, Dordrecht.
- Borer, Hagit (2005a). *Structuring Sense. Volume 1: In Name Only*. Oxford University Press, Oxford.
- Borer, Hagit (2005b). *Structuring Sense. Volume 2: The Normal Course of Events*. Oxford University Press, Oxford.
- Bresnan, Joan (2001). *Lexical-Functional Syntax*. Blackwell, Oxford.
- Chomsky, Noam (1957). *Syntactic Structures*. Mouton, The Hague.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Foris, Dordrecht.
- Chomsky, Noam (1995). *The Minimalist Program*. MIT Press, Cambridge, MA.
- Chomsky, Noam (2000). Minimalist inquiries: The framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik* (ed. R. Martin, D. Michaels, and J. Uriagereka), pp. 89–115. MIT Press, Cambridge, MA.
- Chomsky, Noam (2008). On phases. In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud* (ed. R. Freidin, C. Otero, and M. L. Zubizarreta), pp. 133–166. MIT Press, Cambridge, MA.
- Dryer, Matthew and Haspelmath, Martin (2013). The World Atlas of Language Structures online. Available online at <http://wals.info>, accessed on 24 April, 2016.
- Gast, Volker (2010). Contrastive topics and distributed foci as instances of sub-informativity. In *Comparative and Contrastive Studies of Information Structure* (ed. C. Breul and E. Göbbel), pp. 15–50. John Benjamins,

- Amsterdam.
- Gazdar, Gerald, Klein, Ewan, Pullum, Geoffrey, and Sag, Ivan (1985). *Generalized Phrase Structure Grammar*. Blackwell, Oxford.
- Givón, Talmy (1975). Serial verbs and syntactic change: Niger-Congo. In *Word Order and Word Order Change* (ed. C. Li), pp. 47–112. University of Texas Press, Austin, TX.
- Greenberg, Joseph (1963). Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language* (ed. J. Greenberg), pp. 73–113. MIT Press, Cambridge, MA.
- Heine, Bernd and Kuteva, Tania (2002). *World Lexicon of Grammaticalization*. Cambridge University Press, Cambridge.
- Hooper, Joan and Thompson, Sandra (1973). On the applicability of root transformations. *Linguistic Inquiry*, **4**, 465–497.
- Kroch, Anthony (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change*, **1**, 199–244.
- Lightfoot, David (1979). *Principles of Diachronic Syntax*. Cambridge University Press, Cambridge.
- Lightfoot, David (2002). Myths and the prehistory of grammars. *Journal of Linguistics*, **38**, 113–136.
- Marantz, Alec (1997). No escape from syntax: Don’t try morphological analysis in the privacy of your own lexicon. In *University of Pennsylvania Working Papers in Linguistics* (ed. A. Dimitriadis, L. Siegel, C. Surek-Clark, and A. Williams), pp. 201–225.
- Mathieu, Eric (2006). Stylistic fronting in Old French. *Probus*, **18**, 219–266.
- Osgood, Charles and Sebeok, Thomas (1954). *Psycholinguistics: A Survey of Theory and Research Problems*. Waverly Press, Baltimore, MD.
- Postma, Gertjan (2010). The impact of failed changes. In *Continuity and Change in Grammar* (ed. A. Breitbarth, C. Lucas, S. Watts, and D. Willis), pp. 269–302. John Benjamins, Amsterdam.
- Roberts, Ian and Roussou, Anna (2003). *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge University Press, Cambridge.
- Sag, Ivan (2012). Sign-Based Construction Grammar: An informal synopsis. In *Sign-Based Construction Grammar* (ed. H. Boas and I. Sag), pp. 69–202. CSLI, Stanford, CA.
- Steedman, Mark (2000). *The Syntactic Process*. MIT Press, Cambridge, MA.
- van Gelderen, Elly (2004). *Grammaticalization as Economy*. John Benjamins, Amsterdam.
- van Gelderen, Elly (2011). *The Linguistic Cycle*. Oxford University Press, Oxford.
- Walkden, George (2012). *Syntactic Reconstruction and Proto-Germanic*. Ph.D. thesis, University of Cambridge.
- Weinreich, Uriel, Labov, William, and Herzog, Marvin (1968). Empirical foundations for a theory of language change. In *Directions for Historical*

Linguistics (ed. W. Lehmann and Y. Malkiel), pp. 95–195. University of Texas Press, Austin, TX.