

## Historical Syntax and the Generative Paradigm

Review article of

David LIGHTFOOT 2006. *How New Languages Emerge*. Cambridge: Cambridge University Press.

Ian ROBERTS 2007. *Diachronic Syntax*. Oxford: Oxford University Press.

Chiara Gianollo

### 1. Introduction

The two monographs under review are recent contributions by eminent scholars in historical linguistics, both written with the intent of summarizing the main results and the new perspectives offered by decades of formal studies in diachronic syntax. Treating the two books in parallel, however, calls for a proviso, as they are explicitly conceived for two quite different kinds of audience.

Lighfoot explores the contribution of generative historical syntax to the study of change within human communities, by focusing on the development of what the author believes to be “a more sophisticated analysis of history and change” than what has been proposed by evolutionary and developmental biologists and political historians (Lightfoot 2006: viii). His intent is to make it accessible to scholars with interests at the interface with linguistics (anthropologists, sociologists, psychologists, neuroscientists).

On the other hand, Roberts has designed a comprehensive textbook introducing the discipline of diachronic syntax to students in linguistics, who approach for the first time the formal study of syntactic change and, possibly, generative syntactic theory itself.

The modes of exposition of the two books vary accordingly, as do the choice of arguments and the space allotted to many topics they share. In discussing the import of the two volumes under review, it is of course important to keep in mind the authors' basic difference of intent. However, such a difference cannot prevent one from noticing a significant similarity which hopefully makes a parallel treatment of the two monographs worth pursuing: both are essentially motivated by the major growth witnessed by diachronic syntax over the past few decades. This growth, unexpected as it was at the beginning of the Seventies, has transformed the field into one of the liveliest areas

of contemporary linguistic theory, which has obtained a place of its own in many undergraduate and graduate curricula in linguistics. In addition, it is constantly interacting with many other neighboring disciplines, thereby offering valuable insights for research on cultural variation and change. An overview of the progressing efforts of diachronically-minded generative linguists can be grasped from a number of volumes which collect the work discussed during the biennial DIGS (Diachronic Generative Syntax) meetings (Battye & Roberts 1995, Van Kemenade & Vincent 1997, Pintzuk, Tsoulas & Warner 2000, Lightfoot 2002, Battlori, Hernanz, Picallo & Roca 2005, Jonas *to appear*, Crisma & Longobardi *to appear*).

Apparently both Lightfoot and Roberts felt that, as in any quickly-developing field, the moment had come to offer a synthesis, highlighting, on the one hand, the contribution of historical syntax to linguistic theory itself and to other connected scientific areas, but also pinpointing the priorities and challenges ahead.

The exposition will proceed as follows: first I will summarize the basic lines of a generative approach to diachronic syntax, which represent the common background for the two studies (section 2). Then, after giving a short overview of each volume (section 3), I will concentrate on some issues which appear to be particularly salient in evaluating the past and future role of historical studies of syntax and which receive an insightful and often innovative treatment in the two books (section 4). Sometimes the two authors will be shown to hold quite different positions, and the comparison will target issues which combined synchronic and diachronic research may hopefully enlighten.

## 2. *Diachronic syntax in a generative perspective*

The central tenet in both books, strongly in compliance with the Chomskyan view about language, is that language change is best understood as change in individuals, in the internal system of grammatical knowledge arising through first-language acquisition (internal, I-language). Such connection between diachronic change and language acquisition was first cast in generative terms by David Lightfoot's (1979) influential monograph, *Principles of diachronic syntax*. Language change came to be seen as an inherent possibility arising, given certain circumstances, during the process of acquisition of grammatical structures, currently interpreted as a process of parameter setting. A new generation may converge over a parameter

value different from that of the previous generation, once the triggering experience, i.e. the corpus of primary data children are exposed to (the external, E-language), changes significantly from one generation to the other.

Lightfoot (1979) viewed language change as “a function of chance and necessity” (a formula which deliberately echoed Monod’s (1970) famous essay on genetic transmission). Chance resides in oscillations in the triggering experience, yielded either by the existence of earlier changes in the grammar or by extra-grammatical factors (contact-induced borrowing, massive imperfect second-language acquisition of the target grammar, desire for expressivity and consequent variation in frequency of a given construction). Necessity is induced by a series of universal, biological characteristics of human language, namely “that the grammar should not allow excessive opacity, that surface strings should be processed with minimal perceptual difficulty, and that generations should maintain mutual comprehensibility” (Lightfoot 1979: 396). In sum, language change is interpreted as triggered by local causes, either internal or external to the grammatical system, and carried out during the stage of language acquisition, a process which at the same time ensures conservativity (mutual intelligibility), elimination of opacity, and obedience to innate restrictions on the form of Universal Grammar.

For the first time, the understanding of language change came to be strictly tied to the elaboration of restrictive theories of grammar, which could define the limits of variation and, thus, the possible outcome of change given a certain triggering experience. In 1979, Lightfoot held the opinion that the secondary role of the study of diachronic syntax in historical linguistic research was “a function of inadequate theories of *synchronic* syntax on the part of neogrammarians, American structuralists and transformational generative grammarians alike” (Lightfoot 1979: vii). In this respect, he considered work on word-order implicational universals originating from Greenberg’s (1963) research to be extremely promising. And, in fact, the real takeoff of generative diachronic linguistics has been determined by the rise of the Principles&Parameters approach to syntactic variation, since its first formulation in Chomsky’s *Pisa Lectures* (Chomsky 1981).

The Principles&Parameters framework offers to historical linguistics the possibility of operating with an extremely powerful tool in the investigation of grammatical change, and, in particular, in the explanation of its ‘bumpiness’, i.e. the frequent occurrence of clusters of changes appearing simultaneously in a given language, which

were considered already by Lightfoot (1979: 402) in a pre-parametric era as “various surface realizations of a single change in the abstract grammar”.

A parametric analysis of syntax, based on a deeply deductive theory of language variation, helps capture the connection between different co-occurring superficial changes at least in two ways. First, it often allows embracing various apparently scattered grammatical changes under only one, abstract point of variation, thereby favoring explanatory approaches of language change based on the study of acquisitional mechanisms. Second, it offers a way to handle clusters of distinct parametric changes by proposing a theory of the complex interdependencies existing among parameter values, which might trigger a chain-shift effect in the language.

The Principles&Parameters framework thus represents a theoretical model which can account both for clusterings of synchronic properties of grammar (i.e. implicational universals), and for clusterings of diachronic properties.

Historical linguists working in a generative perspective are convinced that all major kinds of syntactic change can be reframed in terms of parameter-resetting operations, which appear to be “the principal explanatory mechanism in diachronic syntax” (Roberts 2007: 121). At the same time, they believe that a better understanding of change in terms of parameter resetting has to be considered fundamental in order to answer more general questions at the synchronic level of explanation, such as, among others, the nature of parameters and parametric networks, the mechanisms of first-language acquisition and the nature of the triggering data, the existence of default values, and the level of grammatical variability in linguistic communities.

The interpretation of language change as a particular instance of parameter setting, i.e. as an acquisition-driven phenomenon, leads generative historical linguistics to refute any theory of change which appeals to explanations involving more than one generation. Syntax is seen as an essentially conservative, inert module, guaranteeing the overall convergence of the newborn grammars with the parental ones found in ideally ‘normal’ situations of language transmission (i.e. with no substantial mutation in the primary data). Syntactic change does not arise, unless it is locally caused; in Keenan’s original formulation of the Inertia Principle, “Things stay as they are unless acted on by an outside force or decay” (Keenan 1994: 2). In Lightfoot’s terms, “structural change in I-languages is contingent, resulting from changes in the grammars or in the use of grammars of earlier generations

that have the effect of changing the availability of grammatical cues” (Lightfoot 2006: 164).

History cannot transcend speakers; grammars, being constructed by individuals, cannot retain “racial memories” (Lightfoot 1979: 391), and there is no space in the theory for explanatory notions such as diachronic universals or long-term teleological changes. Lightfoot entertains quite a radical view on this point, which is critically discussed by Roberts in his book, as we will see below (4.3).

No principles of history and, thus, no predictive theory of change can be formulated according to such premises. Lass, in his (1980) monograph, held the extreme view that language change would in principle not be subject to explanation, as it is not deterministic, hence unpredictable. According to Lightfoot, grammatical change can be explained insofar as it is considered to be a change of some properties of the individual’s grammar from one generation to another. What cannot be generally subject to a principled explanation is the primitive change in the linguistic environment, in the E-language (a similar point is made by Lass 1987, 1997): “Grammatical, structural changes need explanation; but there is no theory of why trigger experiences should change, except insofar as they change as a result of earlier structural changes” (Lightfoot 1999: 207). Once extra-grammatical factors enter the picture, “Clio is free to play idly with her water clock” (*ib.*).

The radical disruption with a long historicist tradition brought about by generativist thinking on language change is patent; it has entailed a drastic revision of many traditional treatments of long-standing issues in the study of language change, and has prompted, during the past decades, a lively debate, which is conspicuously reflected in the two volumes under review here. After a short general overview of their respective content (3.1 and 3.2), I will concentrate on two main aspects which appear to be particularly challenging for the generative approach to the study of diachronic syntax: the search for local causes in processes of change and the connected theory of learnability (4.1 and 4.2), and the necessity of dealing with long-term processes in the history of languages (4.3).

### *3. Overview of the two volumes*

#### *3.1 Lightfoot (2006)*

David Lightfoot is not new to the task of offering linguists and the broader scientific community comprehensive monographs deve-

loping a wide-ranging integrated approach to the problem of language change (cf. e.g. Lightfoot 1982, 1991, 1999, and, with Stephen Anderson, 2002). With his latest book, he is particularly concerned with the ‘outside world’, in two different senses.

First, from the point of view of the intended audience, the book is especially devoted to non-linguists, to “people who have thought a little about language but who do not necessarily work on syntactic theory, who have no concern whether syntax is minimalist or cognitive or unificationalist or systemic, but who might be interested in the implications of that work for understanding how new systems can develop” (Lightfoot 2006: viii). For this reason, technicalities are kept to a minimum and various sections sum up the history of core theoretical problems in the study of language history.

Secondly, from a more substantial point of view, Lightfoot’s main interest shifts from the internal-language dimension to the mutual relationship between change at the internal-language level and in the external language, two dimensions feeding each other in a dynamic interaction which has not yet been satisfactorily studied.

The book contains eight chapters, some of which represent the expansion and the revision of previous works by the author, especially Lightfoot (1991) and (1999), in light of new research.

The first chapter introduces basic concepts underlying the cognitive approach to the study of language change, which is pursued in the rest of the volume; it represents a clear and synthetic overview of the role of historical linguistics within the biolinguistic framework. In particular, the reader is pointed to the core problem of generativist research on change and variation, namely the fact that, once it is acknowledged that there exists a human “language organ”, which is genetically transmitted and invariant within the species, the origin of the actual variation among the observed languages of the world becomes a real paradox. In order to solve it, it has to be posited that the possibility of variation is biologically based, and that human language capacity is actuated by a range of ‘phenotypical shapes’; however, the observed variation cannot have a biological basis, rather it must be determined by environmental factors, which become crucial during the stage of language acquisition. The child, in building her mental grammar, her internal I-language, is guided by innate principles to develop grammatical structures which are recognized within the corpus of primary linguistic data. These data come from the external E-language, which is a function of the use of I-grammars by individuals in the child’s environment. The emergence of new internal systems of linguistic

knowledge must be linked to a shift occurring in the external primary corpus, and vice versa. The central chapters of the book are, in fact, devoted to the definition of a distinction, both in the causes and in the actuation, between two different types of linguistic change: while E-language changes affect the primary corpus for language acquisition, possibly preparing I-language changes, the latter are formal changes in a new generation's mental grammar. Thus, E-language, which depends on the various speakers' use of their internal linguistic knowledge, is a constantly changing entity, inherently in flux. On the contrary, I-languages usually remain stable during the adult age: they only change from one generation to another. It is the interplay between internal and external languages which gives rise to new grammatical systems.

The second chapter is devoted to a short survey of traditional approaches to language change, starting with the Comparative Method in the Nineteenth Century and discussing the structure of Nineteenth-Century historical explanations. The third chapter represents an expansion of the first chapter with respect to the notions of I-language and the poverty-of-stimulus argument in support of a universal innate basis for linguistic knowledge. Issues concerning learnability theories and the nature of acquisitional mechanisms – a topic I will come back to in 4.1 – are illustrated in the fourth chapter. The fifth chapter is concerned with an account of syntactic change at the I-language dimension, and presents evidence from the history of English modal verbs and from the study of verb movement in connection to morphology. In the sixth chapter the triggering mechanism of structural changes is investigated, relating it to the change in the use, at the level of E-language, of pre-existing constructions; here the empirical material is represented by a study of the split genitive construction from Old to Middle and Early Modern English, and of the shift in the verb-object order in the history of various Germanic languages. The seventh chapter deals with the accelerated process of new language creation taking place in the case of creoles and signed languages. The concluding chapter takes up again the classical concerns of historical linguistics discussed in the second one and surveys some new perspectives raised by the cognitive approach to language change, with particular attention to the issue of reconstruction. According to Lightfoot, the factor of chance intrinsic to diachronic processes, which makes every predictive theory of change impossible, also has the effect of severely restricting the possibility of reconstructing proto-grammars.

### 3.2 Roberts (2007)

Ian Roberts, with his textbook, has tried to satisfy a profound need of many linguistics departments, that of introducing the generative study of syntactic variation and change to a public which is supposed to have no previous knowledge of formal syntax nor of historical linguistics. Although the author explicitly warns the reader that the book is not intended to be an introduction to syntactic theory or a manual for syntactic analysis, he succeeds in making his textbook self-sufficient. The task is accomplished thanks to the intelligent choice and clear exposition of a limited but significant number of points of syntactic variation, which are investigated in the history of various languages and tackled from a variety of perspectives in different sections of the book.

The accessibility of the text is substantially helped, especially from a didactic point of view, by the use of separate boxes for the illustration of more technical aspects of the theory, by the presence of a useful glossary, and by the thorough and well-organized index of subjects. Each chapter is enriched by a concluding section containing detailed suggestions for further reading, where each reference is briefly presented in its principal achievements, constituting in this way also a valuable survey of different approaches to phenomena mentioned in the discussion and a guide to recent developments.

The first chapter introduces the Principles&Parameters approach to syntactic variation in its Minimalist manifestation and then focuses on the group of six parameters whose effects in various stages of different languages will be used as the main empirical material throughout the book: the existence of null categories in subject position, the movement of the lexical verb to a functional position known as V-to-T movement, the possible successive movement of the verb instantiating the phenomenon of Verb Second, aspects of the syntax of negation (negative concord), the movement of interrogative operators (Wh-movement), and finally the much-debated head-complement parameter. Parameters are presented first in their synchronic dimension, by adopting a truly comparative perspective which covers data from a rich sample of languages (it is not frequent to find a book on formal syntax with a four-page Index of Languages); then follows an analysis of witnessed changes relative to the value of each parameter, conducted with a thorough discussion of references on the subject.

The second chapter is concerned with different types of syntactic change, such as reanalysis, grammaticalization, and change in argument structure, complementation, and word order. All receive an explanation in terms of parameter resetting. The section on reanalysis



is particularly interesting, as it takes the author to address two crucial problems for any theory of change, termed by Roberts the ‘Regress problem’ and the ‘Chicken-and-Egg problem’ (Roberts 2007: 125-127). These are both related to the abductive nature of grammatical change, and refer respectively to the paradox of language acquisition (a new generation abduces a different grammar from what has generated the primary corpus it has been exposed to) and to the problem concerning the causal relationship between two correlated changes.

The role of first-language acquisition in language change is the subject of the third chapter, which introduces a Principles&Parameters approach to the dynamics of acquisition and discusses the nature of triggers, the role of morphology, the definition of structural simplicity and its importance in guiding reanalyses, and the significance, in this respect, of a theory of markedness of parameter values. This chapter is a noteworthy original contribution, blending the exposition of some firmly established points in generative acquisitional research with the illustration of novel perspectives on the analysis of syntactic change stemming from the author’s recent research. I will comment on some of these aspects in 4.2 and 4.3.

The discussion of the dynamics of syntactic change, the ‘transition problem’, in the fourth chapter is the occasion to introduce the issue created by an I-language perspective on language change with respect to the apparent gradualness in the diffusion of innovation. Roberts reviews the mechanisms of lexical diffusion, which brings the author to the concept of microparametric variation and change, the role of formal optionality in causing apparent gradualness, the competing-grammars approach. Then he presents a major debate in historical studies, that concerning the notion of ‘drift’, which I discuss in 4.3. The concluding section of the chapter is dedicated to the topic of syntactic reconstruction. The fifth and last chapter deals with the role of contact in causing syntactic change: in the first part, the effects of massive second-language acquisition and, more in general, of linguistic substrata are introduced, whereas the creation of creoles and sign languages are the subject of the second part.

#### *4 Challenges for generative diachronic syntax*

##### *4.1 Acquisition and the nature of the triggering experience*

In the fourth chapter of his latest book, Lightfoot incisively summarizes his original approach to learnability, which he has been developing since the late Eighties (Lightfoot 1989). Lightfoot considers

learnability theories based on evaluation metrics to be psychologically implausible, as they rely too heavily on the process of grammar comparison over finite corpora of unanalyzed sentences, i.e. on E-languages. Models such as Gibson & Wexler's (1994) Triggering Learning Algorithm, an error-driven, serial device, presuppose, in fact, the child's access to a stored data set, formed by E-language elements, and their batch processing, yielding serious feasibility problems caused by the "exponential re-explosion" (Fodor 2001: 736) of workload necessary to evaluate parametric systems of a plausible extension (Gibson & Wexler's simulation involved just three parameters). A major feature of natural languages, already forcefully pointed out by Gibson & Wexler themselves (but see also previous observations by Clark 1989), is represented by pervasive ambiguity in the input sentences constituting the triggering experience, which in principle enables the child to generate more than one grammar for given sentence types. Once ambiguity generates an error in setting one parameter, this error may cause, in turn, successive input to receive a wrong interpretation, leading to an unrecoverable situation given these kinds of models of acquisition.

Lightfoot argues that this aporia can be avoided once one adopts a perspective on learnability genuinely centered on I-language, according to which the real triggering experience is not to be recognized in sentences or word-strings, E-language elements, but rather in pieces of structure, pure I-language entities. The same strategy is followed by Fodor in proposing the Structural Triggers Learner model (Fodor 1998 and subsequent work, especially Fodor 2001 and Fodor & Sakas 2001): children do not learn from ambiguous input, because they are guided by innate structural triggers called 'treelets', pieces of structure adopted in the grammar only if successful in parsing strings coming from the external input. Each input sentence becomes parametrically unambiguous once it receives a complete structural description, thus each structurally represented sentence will represent an unambiguous trigger once triggers are supposed to be pieces of structure. If a surface string may receive different structural descriptions, it will be disregarded during the process of parameter setting: the learner, within this model, is able to detect parametric ambiguity and to beware of it, learning only from unambiguous evidence (see Fodor 1998: 23-27 for a discussion of cases of subset languages, where the learner might be thought to be compelled, in fact, to learn from ambiguous evidence).

Similarly, Lightfoot (1997, 1999, 2006: 77-86) proposes a model based on structural triggers, called "cues", pieces of structure provided by Universal Grammar that guide the child to the mental repre-

sentations of sentences coming from the E-language. Such sentences will be considered to express a cue only if they unambiguously need that piece of structure in order to be analyzed. So, for instance, the child looks for a cue like  $_1V$ , i.e. inflected verbs in the Inflection/Tense position: the instantiation of this cue in a sufficient number of E-language utterances (e.g. verb-initial interrogative sentences or negative sentences in Modern French) will imply for the child the assumption, in her I-grammar, of a V-to-I (in current Minimalist framework, V-to-T) movement. Cues/treelets are “global” triggers in Gibson and Wexler’s (1994: 409) sense: they are unambiguous triggers for a given parameter value in every language.

A further hypothesis couched in Lightfoot’s model of acquisition concerns the syntactic space which is relevant for the process of parameter setting: Lightfoot (1991) introduces the idea that children look for cues only in simple structures, unembedded domains, such as simple clauses and the top of embedded clauses (complementizers and subjects). The “degree-0” learnability hypothesis stems mainly from diachronic observations, clearly summarized in Lightfoot (2006: 123-136). Here the author discusses the change from OV to VO order in the history of English, connecting it to Verb-Second properties of main clauses, which increase the number of VO orders in the E-language corpus, and also to the reanalysis of modals and to the rise of the periphrastic *do* construction. The different rate of change in main and embedded clauses, gradual in the former but later in time and much more rapid in the latter, is considered to be evidence of the fact that children, in changing their grammar from OV to VO, only considered unembedded evidence, which had gradually shifted to a prevalence of superficial VO orders due to accidental variation in the use of E-language by the linguistic community; as soon as the change happened, the new value was generalized to embedded domains as well (for a discussion of different analyses of this change, and for reasons for skepticism about its suddenness in embedded domains, see Roberts 2007: 175-198 and cited references).

Within a cue-based approach, the notion itself of parameter comes to be cast into a different perspective: in fact, Lightfoot states that “there is no need for an independent notion of parameters” (Lightfoot 2006: 78); children, in processing data coming from the primary corpus they are exposed to, scan the environment for cues: “cues that are realized only in certain grammars are the points of variation between grammars” (*ib.*). Roberts (2007: 242-245) confronts this view, and raises some critiques, the most important of which, in my opinion, concerns the unrestricted nature of the cue-based approach: “if there

is no independent notion of cues, then we have no way of specifying the class of possible parameters, and hence the range along which languages may differ, synchronically or diachronically” (Roberts 2007: 244). If the format of cues is not explicitly defined and shown to obey restrictions following from properties of Universal Grammar, a cue-based approach runs into the risk of overloading the genetic endowment, which would contain a potentially infinite list of pieces of tree structures to be matched with the mental representations of external data during the process of language acquisition. Moreover, it would fail to offer a principled account for the interactions among parameter values which are a pervasive characteristic of natural languages. It has to be added, though, that the same methodological difficulties may arise within a parametric approach, if parameters are not subject to a critique concerning their possible format and their hierarchical organization (see Gianollo, Guardiano & Longobardi *to appear* for some remarks and a proposal concerning a restrictive theory of the form of parameters). The study of language change may prove to be decisive in yielding better formalizations of the space of grammatical variation, whose success, as we have seen, is strictly connected to an investigation of the problem of parametric expression in the Primary Linguistic Data (henceforth, PLD).

Beyond the issue of the relevant format for triggers, historical concerns challenge classical learnability theories also when it comes to account for what is traditionally referred to as the logical problem of language change, also termed the “Regress problem” by Roberts (2007: 126): the paradox to be explained consists of the fact that, when language change happens, the trigger experience produced by a generation which has acquired a given grammar is not sufficient for the next generation to converge on exactly the same grammar. Thus, a new grammar has to be ascribed to a mutation in the corpus generated by speakers of the ‘older’ grammar, immune to the grammatical change. Such primitive mutation is assumed not to be grammatical in nature; it arises from variation in the actual use of the pre-existing system, which may change the next generation’s primary experience to a threshold level which triggers a new grammatical system.

In order to cope with these facts in elaborating a learnability theory, the learning algorithm to be assumed cannot be strictly deterministic, but rather “weakly deterministic” (Roberts 2007: 231) and allow for parametric change to take place (this is why Probably Approximately Correct Algorithms of language learning, such as those discussed e.g. by Clark & Roberts 1993 and Niyogi 2006, have been proposed).

Given a weakly deterministic model, the factor assuring substantial convergence with the parent grammar is a drive towards conservativity inherent to the language acquisition device, formulated as the Inertia Principle (see discussion in Roberts 2007: 231-232), which blocks changes not necessarily triggered by local causes in the primary corpus. In the presence of an adequate trigger (P-expression in Roberts' terms, cue in Lightfoot's), the value attributed to a given parameter by a new generation converges on that of the previous one; otherwise, once the trigger for a given value has become less robust, either because of the intervention of extra-syntactic factors or as the consequence of an independent syntactic change, it paves the way towards grammatical change. But then, how to define the 'strength' threshold of a trigger, i.e. the extent to which the learner can be said to find unambiguous parametric expression in the PLD? Possible answers to this core question are presented at length, in constant confrontation with Lightfoot's positions expressed over the years, by Roberts (2007), in his discussion of grammaticalization and, especially, reanalysis. We will address this issue in the following paragraph.

#### *4.2 Economy of representation and diachrony*

Grammatical change is abductive in nature and, as such, open to 'error': given the principles of UG and a primary corpus, the child abduces a grammar, which may or may not converge on the system which has yielded the primary corpus itself. The reason for this, as discussed in the previous paragraph, must lie in the variable *use* of grammars within a linguistic community, which may reduce the strength of some triggers and result in parameter change. Roberts (2007) argues that change in parameter values is generally associated with a process of reanalysis of a given syntactic pattern, a process which involves Move and Agree operations affecting functional heads (whose feature set is assumed, since Borer 1984, to be the locus of parametric variation). Reanalysis happens when a given string surfacing in the PLD is assigned by the learner a structural representation different from the one which has generated it. The conditions causing reanalysis are investigated by Roberts (2007) in his second chapter: Roberts connects reanalysis to the notion of *transparency* introduced by Lightfoot (1979), which is however innovatively interpreted in terms of economy of representation.

The "Transparency Principle" invoked by Lightfoot (1979: 98-115, 121-141) in his study of the changing categorization of English modals had been conceived as a UG principle prompting the learner's "therapeutic" reaction, in terms of reanalysis, in case of accumulated

*opacity* in the primary data. The Transparency Principle “requires derivations to be minimally complex and initial, underlying structures to be ‘close’ to their respective surface structures” (Lightfoot 1979: 121). Thus, the clustering of exceptional features on the class of English modal verbs caused them to be opaque as main verbs and led the learner to their new categorization as auxiliaries.

Roberts, at least since Roberts (1993), has attempted to reach a formal characterization of the notion of opacity, by relating it to that of *simplicity*, in the wake of Lightfoot’s original formulation of the Transparency Principle. He connects opacity to ambiguity and defines the latter by referring to Clark & Roberts’ (1993) notion of parameter expression and P-ambiguity. Strong P-ambiguity, whereby a string may express both values of a given parameter, is considered to be linked to reanalysis. Strong P-ambiguity may arise as a consequence of changes in other modules (phonology, semantics) or may be caused by other syntactic changes.

In developing the notion of “Diachronic Reanalysis”, Roberts (1993: 153-160) considers structural ambiguity arising from opacity to be addressed by the learner by appealing to a notion of structural simplicity: in the presence of competing structural representations for a given string, the learner would opt for the simplest one. He defines simplicity in terms of number of links within a syntactic chain, and proposes that the notion plays a role in acquisition by means of what he calls the “Least Effort Strategy”, according to which “[r]epresentations assigned to sentences of the input to acquisition should be such that they contain the set of the shortest possible chains (consistent with (a) principles of grammars, (b) other aspects of the trigger experience)” (Roberts 1993: 156). Reanalyses guided by the Least Effort Strategy may create the conditions for a subsequent parametric change, by reducing, and ultimately removing, structural evidence for the older parameter setting.

Roberts (*ib.*) is very resolute in setting the Least Effort Strategy apart from similar claims made by Chomsky (1989, reprinted with minor revisions as the second chapter of Chomsky 1995), who, according to a research agenda which will become, in its mature form, the Minimalist Program, considers principles of simplicity in derivation and representation to be part of UG in general, and not just an acquisitional strategy to deal with ambiguity. Nonetheless, the notion of Diachronic Reanalysis and the connected proposal concerning the Least Effort Strategy are strongly criticized by Lightfoot (1999: 216-220) as “an attempt to explain some changes entirely through UG, independently of changes in trigger experiences” (Lightfoot

1999: 216), and, as such, suggesting a deterministic view of history. According to Lightfoot, Diachronic Reanalyses occur where a grammatical change has already taken place and, thus, have no real explanatory power: the real explanation lies in shifts in the input, which are the only cause for parametric changes. In his view, the Least Effort Strategy, and his own Transparency Principle, represent a wrong turn towards the search for “endogenous” tendencies towards “optimization” (Lightfoot 1999: 218), for an explanation of grammatical changes motivated uniquely by internal factors, such as economy.

Roberts (2007: 132) partially answers this critique by defending the usefulness of a notion of Diachronic Reanalysis in highlighting the role of structural ambiguity in parametric change. In his discussion of grammatical changes arising through reanalysis, he also shows that his approach does not entail abstracting away from actual changes in the triggering experience: on the contrary, in order to attain a real explanation of the change, it is necessary to single out a local cause which prevented reanalysis in a previous generation and motivates its appearance in the new one. For instance, in his discussion of the development of the French question particle *ti*, he argues that the trigger to reanalysis comes from a previous change in the phonological system, although he acknowledges that this interpretation, while solving the problem for syntax, shifts the burden of explanation to phonology (see Roberts 1993: 220-224 and 2007: 129-132). However, Roberts does not deny that higher-order cross-linguistic principles might be active in guiding acquisition and, therefore, change.

Roberts (2007) recasts the Least Effort Strategy in terms of a “simplicity preference”, according to which “reanalysis is motivated by a general preference on the part of language acquirers to assign the simplest possible structural representations to the strings they hear” (Roberts 2007: 131). He further elaborates on the notion of simplicity by proposing a “simplicity metric”, drawn from recent joint work with Anna Roussou on grammaticalization (Roberts & Roussou 2003). Roberts & Roussou (2003: 200 ff.), while noticing that, in principle, many different approaches to syntactic complexity may be proposed (based on number of nodes, branching nodes, traces, chain links -cf. Roberts 1993-, symbols or features), favor a feature-counting approach, following the proposal made by Chomsky & Halle (1968) for treating complexity in phonological systems. A simpler representation for a given string will be that containing fewer formal features. Roberts and Roussou further assimilate complexity to *markedness*; their approach is discussed by Roberts (2007) in his third chapter, where he connects this proposal to the similar one

made by Chomsky & Halle (1968) for the distinctive-feature system in phonology. Accordingly, an asymmetry is recognized in parameter values: “the unmarked value of a parameter determines a grammar which generates simpler structures than those generated by the marked value” (Roberts 2007: 254).

As features are responsible for movement, that is, a Probe causing movement will be more complex than one not causing it, it is possible to develop a markedness hierarchy for grammatical operations, along the lines of Roberts & Roussou (2003: 210-213): Move > Agree > neither.

Marked parameter settings are associated with opaque, relatively complex, constructions, e.g. constructions involving movement; this explains, for instance, the mechanism of grammaticalization, which creates exponents of functional categories out of lexical elements: given certain co-occurring circumstances, such elements are reanalyzed as items directly merged in the functional position (and not copied there by means of a costlier movement operation from the lexical layer).

Also, morphology is considered to be associated with markedness and, as a consequence, with complexity. Roberts (2007: 264) suggests that “[i]f a formal feature of a category C is inflectionally expressed, then C is associated with a marked parameter value”; as a consequence, morphological loss is recognized as a decisive factor in parameter change, yielding an unmarked setting.

This formulation, however, should not be taken as an instantiation of the isomorphic view of syntax and inflectional morphology convincingly criticized by Anderson (2002). Morphology may act as a cue for movement, and thus its loss may perturbate the primary data and weaken the parametric expression. This does not entail, however, as argued by Lightfoot (2002; see also the other contributors to the same volume), that there be a two-way relationship between morphology and syntax, whereby there is no movement where there is poor morphology (for a similar remark, see Lightfoot 2006: 106).

Following Chomsky & Halle (1968), Roberts’ markedness theory does not posit the existence of a single unmarked value for a given parameter; markedness has to be contextually determined, in light of parametric interactions, and markedness reversals might take place. This helps avoid an impending paradox: once a preference for unmarked systems is built into the learning theory, one would suppose grammatical change to be directional, only from marked to unmarked values. As this is obviously not the case, Roberts develops a system whereby changes from unmarked to marked are possible, once



what he calls a “markedness-induced harmony” (Roberts 2007: 275) is at work. Accordingly, the markedness hypothesis comes to be referred not just to individual parameters, but to systems of parameters: “the markedness of a particular parameter will depend on the values assumed by other parameters in a given system” (Roberts 2007: 273). Crucially, Roberts assumes that something like Vennemann’s (1974) and Hawkins’ (1983) cross-categorical harmony, inducing a preference for languages where there is a homogeneous ordering of constituents across phrasal categories, is active in grammatical systems: he states it in terms of a “preference for potential movement triggers to act together” (Roberts 2007: 194), harmonically. Thus, for instance, even if a given head having an EPP feature – causing movement – may represent a complex feature for a grammar, once this head will conform to other heads in the language in having such an EPP feature, its value will not be marked, because a preference for harmonic ordering will override the markedness of the movement operation.

We see here that Roberts recasts the typological notion of cross-categorical harmony as a preference for simple, i.e. relatively unmarked, grammars. In his discussion of the head-complement parameter, Roberts takes into consideration typological approaches to the problem, trying to relate implicational universals of the Greenbergian tradition to clustering effects created by single, deep parameters, or to the complex system of interactions existing among values of different parameters. In discussing the issue of cross-categorical harmony in the order of heads and their complements, he appears to be convinced of the fact that a strong theory of syntax should be seriously concerned with frequently observed cases of harmonic ordering, despite the obvious counterexamples. In particular, he proposes that harmonic effects cannot be the result of a unitary syntactic operation, but rather that of the interaction of distinct grammatical features, and that “the preference for ‘harmonic’ ordering may thus derive from an overriding tendency for independent parameters to conspire to produce a certain type of grammar”, in compliance with “a higher-order cross-linguistic principle” (Roberts 2007: 101-102). The concept of cross-categorical harmony has recently been revived in parametric terms by Roberts in work with Theresa Biberauer (Biberauer & Roberts 2005; Biberauer & Roberts *to appear*), together with that of parametric conspiracy.

According to Biberauer and Roberts (2005), effects of parametric harmony may arise as a symptom of a general drive towards simplification of structures which is active during language acquisition, a ‘least-effort’ strategy forcing reanalysis once primary data become opaque or ambiguous with respect to a given parameter value. There

is no internal tendency to mutation within syntax itself, and the trigger for the mechanism is considered to reside in the external language, in the linguistic community; however, as soon as ambiguity arises in the input, deep overarching grammatical principles of simplicity and harmony will affect the learner's reanalysis of the primary data.

Obviously, saying that a principle of simplicity guides learning and, as a consequence, syntactic change, does not amount to saying that change makes grammars simpler: besides the fact that reanalysis processes may cause complications in other areas of grammars (see the example taken from French grammar in Roberts 1993: 177-186), the result of parametric change is just another grammar with different properties, in no way simpler than those of the former system.

However, the interaction of simplicity with the notion of harmony in Roberts' model of change introduces into the theory a directional force able to conspire towards a syntactic 'type' over periods of time transcending generations. Roberts' model, therefore, entails a revision of 'orthodox' generative treatments of long-range historical phenomena, which are usually seen as epiphenomenal, non-directional clusterings of parametric properties. The controversial perspectives initiated by such an approach clearly appear from Lightfoot's and Roberts' discussions of the notion of 'drift', to which we turn in the next paragraph.

#### 4.3 *The notion of drift*

When Sapir introduced the notion of 'drift' in 1921, he considered it, as appropriately stressed by Lightfoot (2006: 37; cf. also 1979: 386 ff., 1999: 208 ff.), as an *explanandum*. Sapir considered drift in a language to be the "unconscious selection on the part of its speakers of those individual variations that are cumulative in some special directions" (Sapir 1921: 155). Drifts for Sapir represented a peculiar facet of variation, i.e. that part of variation which was not random, but perceived by hindsight as directed. Nonetheless, phenomena of drift were assumed to be due to specific local causes which could be isolated by linguistic analysis (see the analysis of four factors involved in the "drift away from *whom*" in Sapir's (1921) seventh chapter).

However, Lightfoot claims that, since the first applications of typological observations on word-order harmony to the study of language history (e.g. Lehmann 1974, Vennemann 1975), 'drift' became a sort of explanatory force, an *explanans*, within a theory of change rooted in a deterministic view of historical developments, accounting for them without looking for local causes. Basically following the histo-

ricist thinking of Nineteenth-Century linguistics, typologists started to investigate the existence of diachronic universals, of directional grammatical phenomena capable of spanning centuries and transforming one 'pure' linguistic type into another, following universally predictable historical paths and giving rise to intermediate, transitional stages represented by 'mixed' grammars. A particularly debated phenomenon in this area of study has been the relative order of constituents (subject/verb/object, noun/adjective, auxiliary/verb, and so on), i.e. what in parametric linguistics is known as the head-complement parameter. Typological approaches to these facts assume a cross-categorical coherence in the order of heads with respect to their complements, which is perturbed during the process of language change; accordingly, a 'mixed' system will be driven by an internal force along a predictable succession of stages to attain a new equilibrium.

Such long-spanning change is incompatible with an approach which recognizes the locus for change in the process of transmission of individual grammars, as most explicitly stressed in Lightfoot's (1979: 385-405) discussion of drift. Nonetheless, gradualness and discontinuity of change, as well as convergence effects are undeniable phenomena. As Roberts emphasizes (Roberts 2007: 348; cf. also 2001: 91), given the huge space of possible parametric variation created by a few binary parameters (thirty independent binary parameters generate  $2^{30}$  languages = 1,073,741,824), the fact that one can, in fact, observe linguistic types or diachronic tendencies is astonishing. This hints to the fact that change is less random than what might be supposed and calls for an explanation which goes beyond the individual. To adopt Lass' (1987) powerful simile, "speakers seem rather like Tolstoy's 'little men', caught up in great historical currents whose import they're unaware of, but who nevertheless play their 'ordained' parts in the larger design" (Lass 1987: 162).

Lightfoot (1979: 396, 402 ff.) readily acknowledged the existence of drag-chains in syntactic change, and, while judging that they had been overstated by typological research and misunderstood in their role, he considered them "of enormous importance for a theory of grammar" (Lightfoot 1979: 396). Furthermore, he was persuaded that advance in syntactic theory could represent an important step towards their understanding. For instance, he was confident in the significance of X-bar conventions for diachronic linguistics, as they constrained possible innovations by providing a restrictive theory of phrase structure and, in particular, could account for observed cross-categorical generalizations in the order of specifiers, heads, and complements (Lightfoot 1979: 402-403).

More recently, in his 1999 book, coming back, when evaluating Niyogi & Berwick's (1997) computational simulations of language change, to the issue of the existence of diachronic trajectories, Lightfoot states that "the explanation for long-term tendencies, if they emerge, will be a function of the architecture of UG and the learning procedure and of the way in which populations of speakers behave. In this way the historical tendencies will turn out to be epiphenomena, derived in an interesting fashion, not stipulated by brute force" (Lightfoot 1999: 225).

In his fourth chapter, Roberts (2007) discusses Lightfoot's critique to the typological interpretation of the notion of 'drift', embracing his perspective; nonetheless, being convinced that "something like Sapir's notion of drift is required on both empirical and conceptual grounds" (Roberts 2007: 350), he lays the basis for a parametrically oriented explanation of long-term trajectories by exploring the validity of the notion of 'parametric drift'. He singles out three main problems an explanation has to face, which can be summarized as follows: the causation problem (what causes purported linguistic 'cycles'), the directionality problem (which natural direction a given drift would have), and the incompatibility with a principle of Inertia which is supposed to cause conservativeness in grammatical systems.

Roberts proposes that drift, and, in general, implicational sequences of syntactic changes, which include also grammaticalization phenomena, be interpreted in parametric terms as a "cascade of changes, a kind of 'domino effect' in the parametric system, whereby an initial, exogenous change destabilizes the system and causes it to transit through a series of marked states until it eventually restabilizes as a relatively unmarked system again" (Roberts 2007: 341-342). Roberts interprets the gradual diffusion of change within the grammatical systems as obeying dynamics similar to those observed in the phenomenon of lexical diffusion: as sound changes gradually diffuse through the lexicon by involving one item at a time, in the same way parametric changes in the formal features of functional heads might gradually diffuse, head by head, in the set of functional categories.

Summarizing recent joint work with Theresa Biberauer (Biberauer & Roberts 2005), he offers an exemplification of 'parametric drift' by analyzing changes in the English verbal system as a cascade of parametric resetting operations spanning from the Fifteenth to the Seventeenth century, starting from the loss of verb movement to the C position and from there on involving each a change in the feature composition of elements merging in the T position (Roberts 2007: 351-356).

The two notions of *stability* of a system and *markedness* of states raise a series of crucial questions for an acquisition-based theory of change, and are at risk of bringing back into the picture the sort of long-term teleology refuted by generative historical research. First, if a system is unstable, i.e. at least dispreferred by some principles of UG, how is it possibly acquired as such by new learners? Why should it last for long periods, without being ‘normalized’ in the space of a few generations? It seems to me that the *actuation* problem, to use the classical terminology introduced by Weinreich, Labov, and Herzog’s (1968) seminal work, is there again.

Moreover, when successive cohorts of acquirers “reanalyse different aspects of the PLD which have been rendered marked by an earlier change” (Roberts 2007: 342), what drives the gradual choice of such aspects? Roberts’ idea with respect to this last point is that “each parameter change skews the PLD in such a way that the next is favoured” (Roberts 2007: 356). Causes, thus, are still local, within the triggering experience; however, they may not be induced by forces external to syntax, but rather be couched in the syntactic system itself, and appear one by one, each as the consequence of a previous parameter resetting operation. As for the causation problem, therefore, the immediate cause for the drift mechanism to be set in motion is considered to reside in the outside world, where even a small variation in the primary corpus for acquisition may trigger a process of language change; this, in turn, would entail a series of internally-caused shifts in the PLD triggering further parametric changes and causing the ‘cascade’ effect.

There is nothing predetermined in the observed sequences of change, insofar as they are never inevitable: they just proceed as long as no other factor intervenes. This is accounted for by revising the formulation of the principle of Inertia, according to which it does not necessarily entail stasis, but persistence in a given direction, unless external interference on the primary data occurs.

As concerns the directionality problem, a solution to it is proposed by invoking markedness considerations, which play an important role in Roberts’ line of reasoning: “certain areas of the parameter space attract grammatical systems, by being relatively unmarked” (Roberts 2007: 350). A drift would be directional as a result of its preference towards unmarked areas of the grammatical space, i.e. in virtue of a motivation built within UG itself. The process of change would not be completed within the individual, but would continue being pulled towards a “basin of attraction”.

It has to be noticed, however, that contemporary linguistic theory seems to be still far from the result of elaborating a solid theory of markedness, which may guide historical explanations.

Roberts is concerned with markedness at length in the third chapter of his book. The concept of markedness appears to be tightly connected, in his account, to that of structural simplicity, as we have seen in 4.2.

Lightfoot (2006) discusses markedness in connection to creoles studies in his seventh chapter. It is well known that an important aspect of Bickerton's (1984a and b) Language Bioprogram Hypothesis consisted in the promise of finding in the investigation of creole genesis special evidence concerning unmarked settings of UG parameters. More specifically, creoles would represent a collection of parameter settings coming in part from the superstrate language and, much more substantially, instantiating unmarked values innately predefined by UG, which would manifest themselves in response to an extremely impoverished triggering experience, such as that offered by pidgins. Lightfoot criticizes this hypothesis, offers an alternative explanation in terms of cue-based acquisition, and argues that nothing in principle forces one to think that "marked settings require access to more extensive experience, and perhaps to fairly exotic data, and that this is not available to first speakers of a creole. [...] One can easily imagine a marked setting being triggered by readily available data, even in the first forms of a creole" (Lightfoot 2006: 144). Most interestingly for our discussion here, he maintains that the only basis for postulating markedness values within UG is represented by arguments from the poverty of the stimulus, i.e. by the necessity of ranking parametric values in order to avoid resorting to negative evidence during acquisition. Markedness, under this analysis, results from the application of the Subset Principle, by which the learner chooses the most restrictive grammar which can generate all and only the structures found in the primary corpus: the smallest language derives from the least marked parameter setting. Once markedness is interpreted in this principled way, it can be shown that 'radical creoles', like Saramaccan, do display marked parametric values (Lightfoot 2006: 142-144). The determination of marked vs. unmarked parametric values on the basis of the Subset Principle is criticized by Roberts (2007: 257-261), who, nevertheless, acknowledges the explanatory power of markedness considerations based on the Subset Principle in cases of genuine formal optionality and in the diachronic process of restriction of function applying to given grammatical operations.

Despite the authors' attempts to lay the basis for a serious theory of markedness, the problem appears to remain open, together with the connected issues relative to a principled account of long-range diachronic phenomena. A renewed interest in such topics by generative syntacticians is, however, welcome. Although the study of syntactic persistencies and cross-generational tendencies in language change is a notoriously tricky field, also for objective difficulties due to the nature of our historical records for most languages, I feel that one must conclude with Lass that "[i]f the linguist ties himself down to the parochial, to the individual brain and its spatiotemporally bound knowledge, he imposes on himself a methodological and conceptual limitation that no other historian would let himself be constrained by" (Lass 1987: 157).

### *5. Conclusion*

The volumes under review are both successful attempts to bridge gaps, with Lightfoot focusing on the relationship with disciplines at the interface with linguistics in the study of human behavior, and Roberts being concerned with a too frequent dichotomy between formal and historical studies of language, and, especially, of syntax. In choosing to concentrate in depth on just a few aspects of Lightfoot's and Roberts' argumentation, I have run the risk of overlooking some other important contributions of their discussion to the current debate. However, I hope to have given a clear idea of the importance of operating at the edge of different branches of linguistics in order to forcefully address diachronic issues which might, in turn, contribute to a more general understanding of the human language faculty. To say it with Lightfoot, "a modern historical linguist needs to be a generalist and to understand many different subfields – grammatical theory, variation, acquisition, the use of grammars and discourse analysis, parsing and speech comprehension, textual analysis, and the external history of languages" (Lightfoot 2006: 6).

The volumes by Lightfoot and Roberts highlight the particularly urgent necessity of a better connection between acquisitional and historical studies. Roberts (2007: 225) regrets that "there is something of a sociological divide between linguists working on L1 acquisition and those working on diachronic syntax". This is especially unfortunate given the weight that notions such as that of simplicity and cross-categorial harmony are gaining in historical explanations, a weight which calls for a better investigation of the validity of such construals by means of observational studies of language acquisition.

Another point which clearly emerges from the two books concerns the significance of conclusions drawn from the observation of networks of parameters, and not just scattered points of variation: approaches to theories of cascade resettings or parametric conspiracies such as Roberts' have to be elaborated over large parametric systems, displaying their full range of interactions, which may be understated when studying only a few parameters at a time. Crucially, therefore, sound historical analyses of syntactic change should widen their scope to encompass full-fledged systems of parameters. In this respect, the application of the Modularized Global Parametrization strategy, first proposed by Longobardi (2003), to diachronic investigation seems to qualify as particularly promising in order to understand mechanisms of propagation of change within parametric networks. According to this method, an entire, coherent module of syntax – noun phrase structure in Longobardi's work – is analyzed in a significant number of languages, and parameters are formulated in order to capture minimal contrasts existing between any two varieties and to single out interactions among parametric features. In this way, it becomes easier to comprehend instances of co-variation and co-evolution.

In closing, let me observe a fact which might seem marginal to the discussion but, indeed, directly or indirectly enables many of the proposals sketched above, namely the dramatic improvement in the fine-grained analysis of data yielded by the existence of syntactically coded electronic corpora for different stages of the English language, the *York-Toronto-Helsinki Corpus of Old English*, the *Penn-Helsinki Parsed Corpus of Middle English*, and the *Penn-Helsinki Parsed Corpus of Early Modern English* (respectively, YCOE, PPCME2 and PPCEME in the references). In sharp contrast, the much more fragmentary picture of the syntactic history of languages for which such technical developments are yet to come is apparent.

*Address of the Authors:*

CHIARA GIANOLLO, Università di Trieste (gianollo@units.it)

*References*

- ANDERSON Stephen R. 2002. Syntax and Morphology are Different: Comments on Jonas' "Residual V-to-I". In LIGHTFOOT David (ed.). *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press. 271-275.



- ANDERSON Stephen R. & David LIGHTFOOT 2002. *The Language Organ: Linguistics as Cognitive Psychology*. Cambridge, UK: Cambridge University Press.
- BATTLORI Montserrat, Maria Lluïsa HERNANZ, Carme PICALLO & Francesc ROCA 2005 (eds.). *Grammaticalization and Parametric Change*. Oxford: Oxford University Press.
- BATTYE Adrian & Ian ROBERTS 1995 (eds.). *Clause Structure and Language Change*. Oxford: Oxford University Press.
- BIBERAUER Theresa & Ian ROBERTS 2005. Changing EPP-parameters in the history of English: accounting for variation and change. *English Language and Linguistics* 9.1. 5-46.
- BIBERAUER Theresa & Ian ROBERTS *to appear*. Cascading parameter changes: internally-driven change in Middle and Early Modern English. In EYTHÓRSSON, Thórhallur & Jan Terje FAARLUND (eds.). *Grammatical Change and Linguistic Theory: the Rosendal Papers*. Amsterdam: Benjamins.
- BICKERTON Derek 1984a. The language bioprogram hypothesis. *Behavioral and Brain Sciences* 7.2. 173-188.
- BICKERTON Derek 1984b. Creole is still king. Author's response to commentary. *Behavioral and Brain Sciences* 7.2. 212-218.
- BORER, Hagit 1984. *Parametric Syntax*. Dordrecht: Foris.
- CHOMSKY Noam 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- CHOMSKY Noam 1989. Some notes on economy of derivation and representation. *MITWPL* 10. 43-74.
- CHOMSKY Noam 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- CHOMSKY, Noam & Morris HALLE 1968. *The Sound Pattern of English*. New York: Harper and Row.
- CLARK, Robin 1989. On the relationship between the input data and parameter setting. *NELS* 19. 48-62.
- CLARK, Robin & Ian ROBERTS 1993. A Computational Model of Language Learnability and Language Change. *Linguistic Inquiry* 24.2. 299-345.
- CRISMA, Paola & Giuseppe LONGOBARDI *to appear* (eds.). *Historical Syntax and Linguistic Theory*. Oxford: Oxford University Press.
- FODOR, Janet D. 1998. Unambiguous triggers. *Linguistic Inquiry* 29.1. 1-36.
- FODOR, Janet D. 2001. Setting syntactic parameters. In BALTIN, Mark & Chris COLLINS (eds.). *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell. 730-767.
- FODOR, Janet D. & William G. SAKAS 2001. The structural triggers learner. In BERTELO, Stefano (ed.). *Language Acquisition and Learnability*, Cambridge, UK: Cambridge University Press. 172-233.
- GIANOLLO, Chiara, Cristina GUARDIANO & Giuseppe LONGOBARDI *to appear*. Three fundamental issues in parametric linguistics. In BIBERAUER, Theresa (ed.). *The Limits of Syntactic Variation*. Amsterdam: Benjamins.
- GIBSON Edward & Kenneth WEXLER 1994. Triggers. *Linguistic Inquiry* 25.3. 355-407.
- GREENBERG Joseph 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In GREENBERG, Joseph (ed.). *Universals of Language*. Cambridge, MA: MIT Press. 73-113.

- HAWKINS, John A. 1983. *Word order universals*. London: Academic Press.
- JONAS, Dianne *to appear* (ed.). *Grammatical Change: Origins, Nature, Outcomes*. Oxford: Oxford University Press.
- KEENAN Edward 1994. Creating anaphors. An historical study of the English reflexive pronouns. Los Angeles: UCLA. Ms.
- LASS Roger 1980. *On explaining language change*. Cambridge, UK: Cambridge University Press.
- LASS Roger 1987. Language, Speakers, History and Drift. In KOOPMAN, Willem, Frederike VAN DER LEEK, Olga FISCHER, Roger EATON (eds.). *Explanation and Linguistic Change*. Amsterdam: Benjamins. 151-176.
- LASS Roger 1997. *Historical linguistics and language change*. Cambridge, UK: Cambridge University Press.
- LEHMANN Winfred P. 1974. *Proto-Indo-European Syntax*. Austin: University of Texas Press.
- LIGHTFOOT David 1979. *Principles of Diachronic Syntax*. Cambridge, UK: Cambridge University Press.
- LIGHTFOOT David 1982. *The language lottery: Toward a biology of grammars*. Cambridge, MA: MIT Press.
- LIGHTFOOT David 1989. The child's trigger experience: degree-0 learnability. *Behavioral and Brain Sciences* 12.2. 321-334.
- LIGHTFOOT David 1991. *How to set parameters: arguments from language change*. Cambridge, MA: MIT Press.
- LIGHTFOOT David 1997. Catastrophic change and learning theory. *Lingua* 100. 171-92.
- LIGHTFOOT David 1999. *The development of language: acquisition, change and evolution*. Oxford: Blackwell.
- LIGHTFOOT David 2002 (ed.). *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press.
- LONGOBARDI Giuseppe 2003. Methods in Parametric Linguistics and Cognitive History. *Linguistic Variation Yearbook* 3. 101-138.
- MONOD Jacques 1970. *Le Hasard et la Nécessité: essai sur la philosophie naturelle de la biologie moderne*. Paris: Seuil.
- NIYOGI Partha 2006. *The Computational Nature of Language Learning and Evolution*. Cambridge, MA: MIT Press.
- NIYOGI Partha & Robert BERWICK 1997. A Dynamical Systems Model for Language Change. *Complex Systems* 11.161-204.
- PINTZUK, Susan, George TSOULAS & Anthony WARNER 2000 (eds.). *Diachronic Syntax. Models and Mechanisms*. Oxford: Oxford University Press.
- PCEME. *Penn-Helsinki Parsed Corpus of Early Modern English*. Maintained by Anthony KROCH & Beatrice SANTORINI - University of Pennsylvania. <http://www.ling.upenn.edu/hist-corpora>.
- PCME2. *Penn-Helsinki Parsed Corpus of Middle English*. Maintained by Anthony KROCH & Ann TAYLOR - University of Pennsylvania. <http://www.ling.upenn.edu/hist-corpora>.
- ROBERTS Ian 1993. *Verbs and Diachronic Syntax. A Comparative History of English and French*. Dordrecht: Kluwer.

- ROBERTS Ian 2001. Language change and learnability. In BERTOLO Stefano (ed.), *Parametric Linguistics and Learnability*. Cambridge, UK: Cambridge University Press. 81-125.
- ROBERTS Ian & Anna ROUSSOU 2003. *Syntactic Change: a Minimalist Approach to Grammaticalization*. Cambridge, UK: Cambridge University Press.
- SAPIR Edward 1921. *Language. An introduction to the study of speech*. New York: Harcourt Brace & Co.
- VAN KEMENADE, Ans & Nigel VINCENT 1997 (eds.). *Parameters of Morphosyntactic Change*. Cambridge, UK: Cambridge University Press.
- VENNEMANN Theo 1974. Topics, subjects and word order. From SXV to SVX via TVX. In ANDERSON John M. & Charles JONES (eds.), *Historical Linguistics. Proceedings of the First International Congress of Historical Linguistics*. Vol II. Amsterdam: North-Holland. 339-376.
- VENNEMANN Theo 1975. An explanation of drift. In Li Charles N. (ed.). *Word Order and Word Order Change*, Austin: University of Texas Press. 269-305.
- WEINREICH, Uriel, William LABOV, & Marvin HERZOG 1968. Empirical foundations for a theory of language change. In LEHMANN Winfred P. & Yakov MALKIEL (eds.). *Directions for historical linguistics: A symposium*. Austin: University of Texas Press. 95-188.
- YCOE. *York-Toronto-Helsinki Corpus of Old English*. Maintained by Ann TAYLOR, Anthony WARNER, Susan PINTZUK, Frank BETHS - University of York.  
<http://www-users.york.ac.uk/~lang22/YcoeHome1.htm>.