

TYOLOGICAL GENERALIZATIONS AND
THE LOCUS OF THEIR EXPLANATION

TETSUHARU MORIYA
*Kanazawa University**

Possible and Probable Languages: A Generative Perspective on Linguistic Typology, by Frederick J. Newmeyer, Oxford University Press, Oxford, 2005, x+278pp.

Keywords: generative grammar, functionalism, typological generalization, parameter

1. Introduction

It is generally agreed that languages do not differ without limit; languages vary only within the limited range of possibilities and these possibilities are specified by implicational or distributional universals. Even though one language superficially seems to differ from another language to a great extent, the difference should be within the bounds of theoretically specified range. To give a well-known example, some languages have SVO while other languages have SOV as a basic word order. The difference in the basic word order correlates with other syntactic characteristics. Greenberg (1963) has listed 45 universals of language based on observations of thirty individual languages. The majority of them are implicational and 25 of them are related to word order. In those universals, Greenberg (1963) proposed three basic word orders, VSO, SVO and SOV, and argued that these basic word orders correlate with the orders of other elements in the

* I am indebted to Kaoru Horie, who gave me valuable advice in writing this review article. I would also like to thank two anonymous reviewers for valuable comments. This work is partly supported by a Grant-in-Aid for Scientific Research (C) 16520234 from the Japan Society of Promotion of Science.

grammar.¹ For example, VSO languages are always prepositional while most SOV languages are postpositional. Concerning the structure of noun phrases, if in SOV languages the genitive follows the modifying noun, the adjective also follows the noun.

Generative grammarians attribute the typological generalizations to the innate Universal Grammar (UG). UG is assumed to have some mechanisms which give rise to cross-linguistic variations. One representative example of the mechanisms is given by parameters in the principles-and-parameters approach. For instance, Chomsky (1981) proposes that UG contains a pro-drop parameter, the positive value of which produces the clustering of following properties:

- (1) (i) missing subject
- (ii) free inversion in simple sentences
- (iii) "long *wh*-movement" of subject
- (iv) empty resumptive pronouns in embedded clause
- (v) apparent violations of the *[*that-t*] filter

(Chomsky (1981: 240))

Pro-drop languages like Italian have all the properties in (1) while non-drop languages like English and French do not. The difference is supposed to derive from different settings of the pro-drop parameter.

As another example, the head-first/ head-last parameter has been posited to explain the word order universals. If a language selects the head-first value, the heads generally precede their complements. On the other hand, in the head-last languages the heads generally follow their complements. This parameter is intended to capture the kind of typological generalizations Greenberg (1963) presented. The pro-drop parameter not only explains typological generalizations, but also plays a role in eliminating the phrase structure component:

- (2) [T]he Phrase Structure Component can be entirely eliminated, apart from certain parameters of X-bar theory: e.g. does the Head precede its Complements, as in English-like languages, so that we have the constructions N-Complement, V-Complement, A-Complement, P-Complement; or does it follow them, as in

¹ These universals are not without problems. For example, Hawkins (1983: Ch. 2) points out that SVO word order fails to predict significant co-occurrence regularities, and argues that a verb-based typology is not valid. Hawkins (1983) refined the universals based on the data of 350 languages.

Japanese-like languages, so that we have the corresponding constructions Complement-N, Complement-V, Complement-A, Complement-P? (Chomsky (1986: 82–83))

This passage shows that in the principles-and-parameters approach, explanations by parameters are assumed to be theoretically superior to those by rule systems.

While generative linguists explain the typological generalizations with recourse to formal concepts like parameters, functionalists explain such generalizations with language external factors. For example, in Hawkins (1990, 1994), word order universals are explained by efficiency of parsing mechanism. The main claim of these studies is that “linear ordering seems to be determined by the rapidity with which the immediate constituents of syntactic groupings can be produced and recognized on-line” (Hawkins (1990: 225)).

As Hawkins (2004) states, it has been quite difficult either for formalists or functionalists to evaluate the respective studies:

- (3) These two research traditions, the formal and the functional, use different methodologies and they formulate different kinds of linguistic generalizations, and this makes it hard for an uncommitted observer to assess their respective arguments.

(Hawkins (2004: xi))

In bridging the gap between the two research traditions, the book under review is quite valuable. *Possible and Probable Languages* (henceforth PPL) has three main claims. The first claim is on the role of UG in explaining typological generalizations. It first reviews the attempts to explain typological universals and then argues that UG is not responsible for explaining language variations: UG has to specify what possible languages are, but not what probable languages should be. Such variations, PPL claims, should be explained by performance theory. The second claim is on the distinction between knowledge of language and language use. PPL argues that although UG is not responsible for typological generalizations, that does not mean formal or structural principles are totally dispensed with and replaced by functional principles. In other words, the distinction between competence and performance should be maintained. The last claim is that although typological generalizations are, on the whole, functionally motivated, the link between grammatical constructs and functional motivations is extremely indirect. PPL denies the view that properties of particular grammars are directly motivated by functional properties.

For formalists, this book gives a good opportunity to review developments

of their theoretical principles to assess whether they have been successful in explaining typological generalizations. In particular, PPL is critical of the Minimalist Program (MP), as stated in the following remark: “The parametric program continues to this day, though it is perhaps less evident how to handle typological generalizations within the Minimalist Program than within GB” (p. 72).

For functionalists, this book gives a useful overview of formal approaches to typological generalizations since 1960’s. It is usually quite difficult to cover all the generative linguistic literature reviewed in PPL to understand what methodologies generative linguists have used and what generalizations they consider important. PPL is valuable also for functionalists because it is a critical review of the functionalist approaches as well. To prove the correctness of the position that grammatical elements and functional motivations are not directly linked, PPL attacks the opposite position, which is called “atomistic functionalism” (AF). Under AF, PPL subsumes such studies as Haiman (1983), Givón (1983), Dik (1989), Lakoff (1987), and Hopper (1987, 1988).² In this sense, PPL is a continuation of Newmeyer (1998), which points out that the basic principles of generative grammar provide compelling accounts of “phenomena such as prototype effects, grammaticalization, the grounding of formal structure in external pressure, and so on—phenomena that few generativists have, in the past, even thought worthy of consideration” (Newmeyer (1998: 7)). Therefore, it is important for functionalists to assess the validity of arguments against AF in PPL.

The organization of this review article is as follows. The following section gives a synopsis of contents of the book. Section 3 examines arguments against AF and shows that the link between grammatical elements and language-external motivations is more direct than PPL assumes. The final section gives an overall evaluation of the book and concludes the review.

2. Overview

This book is made up of 5 chapters. Chapter 1 sets the stage for the overall discussion, elaborating on the problems of identifying and explaining possible and probable languages. Chapter 2 reviews studies on the param-

² Several studies of Optimality Theory (OT) are classified into AF, although OT is a generative approach.

eterization of principles of UG, the principal means of generative theory to capture typological generalizations. Chapter 3 criticizes the parametric approach and argues that it is not the job of UG to account for typological generalizations. Chapter 4 maintains that although UG is not responsible for explaining typological generalizations, performance factors alone cannot explain the grammatical properties of language. Chapter 5 discusses the link between grammatical principles and functional motivations. PPL admits grammatical elements are largely motivated by extra-linguistic factors. However, the link between grammatical properties and functional motivations is claimed to be only indirect.

2.1. Possible and Probable Languages

In Chapter 1, PPL first shows that characterizing possible human languages has been not only the goal of generative grammar from its beginning as indicated by Chomsky (1962: 536–537) and Chomsky (1965: 27) but also the concern of functional linguists as indicated, for example, by Croft (1990: 34) and Shibatani and Bynon (1995: 19). Thus, generative grammarians and functionalists share their goal of linguistic theory but differ in their view of the language faculty.

Identifying what is possible in human language has several difficulties. First, it is not possible to decide which features are possible and which are impossible in a theory-independent way and therefore it is impossible to evaluate the hypotheses theory-independently. Second, we cannot conclude the impossibility of a grammatical feature simply because we cannot find it in a limited sample of languages. Therefore it is important to distinguish between necessarily universal features and incidentally universal features derived from extra-grammatical principles. PPL then argues that difficulties in explaining what is possible in human language lie in “so little agreement on the precise nature of the human language faculty” (p. 8). Not all the relevant faculties can be attributed to UG. For example, impossibility of multiple embedding is not considered a UG property but can be explained by the parsing difficulties. But having a plausible functional explanation is not necessarily evidence for supporting the functionalists’ view, either. The structural dependence on grammatical rules and extraction phenomena such as subadjacency are such examples.

Based on these arguments, PPL takes “a cautious approach to attributing some particular feature of grammar to UG” (p. 14). First, PPL assumes that UG determines “grammatical architecture in its broad lines” (p. 14), but does not give any specific proposal on the precise nature of the grammati-

cal architecture. Second, it assumes that certain more specific universals of grammar are solely determined by UG without any influence of extragrammatical factors, but the majority of the putative universals will prove not to be universal with more cross-linguistic evidence. Finally, although some constraint such as subadjacency is obviously ascribed to UG, it will not assume that “innateness is the default hypothesis for any grammatical principle” (p. 15).

On the problem of identifying what is probable in language, PPL observes that the discussion of the English auxiliary in Chomsky (1965) set the tone for subsequent generative work and made up the dominant idea that UG is responsible for specifying probable languages. Then both functionalists and formalists are interested in identifying probable languages, but the methodologies are different. The methodological differences are illustrated by focusing on an extended exchange in the 1980s between Bernard Comrie and Peter Coopmans and several problems of the standard generative approach to typological generalizations are preliminarily pointed out.

2.2. Criticism of Parameterized Principles

Chapter 2 reviews a series of studies dealing with typological variations by generative linguists since the 1960s. The main focus of this chapter is the parameter-based approach since the 1980s. Chapter 3 then criticizes the approach both on theoretical and factual grounds. Based on these criticisms, PPL tries to support an alternative view of linguistic theory as in (5), which contrasts with the standard view as summarized in (4):

- (4) Some central features of (currently predominant) linguistic theory:
 - a. Principles of Universal Grammar (or, more recently, a set of functional projections provided by UG), which have
 - b. Different parameter settings for different languages (thereby accounting for language-particular differences).
 - c. By means of ([4]a) and ([4]b), typological variation is accounted for.
 - d. A residue of marked (language-particular) morphosyntactic properties. (p. 73)
- (5) Some central features of an alternative way of looking at linguistic theory:
 - a. Unparameterized principles of Universal Grammar.
 - b. Language-particular rules constrained by these UG principles.
 - c. Extragrammatical principles that account for typological vari-

ation.

(p. 73)

Under the view (5), language-particular differences are accounted for by differences in language-particular rules, and the locus of accounting for typological generalizations is transferred from competence to performance.

This subsection introduces some of these arguments focusing on several features of parametric approaches.

2.2.1. Simplicity

One characteristic of the parametric approach on which it is generally regarded as theoretically better than rule-based explanations is the descriptive simplicity. Take the Head Parameter, for example. It can be stated rather simply as in (6):

- (6) HEAD PARAMETER—Complements are consistently to the left or to the right of the head.
- a. HEAD-LEFT (English, Swahili, ...)
 - b. HEAD-RIGHT (Japanese, Lakhota, ...) (p. 43)

On the other hand, in a rule-based model, it would be necessary to specify phrase structure rules for each category and that would give the impression that rule-based explanations are cumbersome and complex.

PPL argues against this view on the ground that “parameters are motivated only to the extent that they lead overall to more formal simplicity” (p. 77). If the number of parameters in a parameter-based model is almost the same as the number of rules in a rule-based model, then “clearly nothing is gained opting for parameters” (p. 77). An example of this situation is an analysis of the differences between adjective-noun ordering in English and French. Cinque (1994) argues that the contrasts shown in (7) to (9) result from the parametric differences as in (10):

- (7) a. un **gros** ballon **rouge**
 b. a **big red** ballon
- (8) a. un tissue **anglais cher**
 b. an **expensive English** fabric
- (9) a. an **old** friend (= friend who is aged or friend for a long time)
 b. une **vieille** amie (= friend for a long time)
 c. une amie **vieille** (= friend who is aged)
- (10) a. French has postnominal adjectives (as in [7]a) because of a parametric difference with English that allows N-movement in the former language, but not in the latter.
 b. *Cher* has scope over *anglais* in ([8]a) because French has a

parametric difference with English that triggers movement of an N-ADJ constituent.

- c. In ([9]), the two positions for *vieille* in French, but only one for *old* in English, result from a parametric difference regarding the feature attraction possibilities of functional categories in the two languages. (p. 77)

PPL points out that in such an account, the word “parameter” is virtually a synonym of the word “rule.” Another example is English word order. In English, orders other than SVO are possible as shown in (11):

- (11) a. The last lecture Mary hated. (OSV)
- b. Drink the whole bottle, John never would. (VOS)
- c. Away ran John. (VS) (p. 78)

A rule-based model will derive the word orders in (11) by positing optional movement rules which operate on an underlying SVO order. In a parameter-based model, it will also be necessary to posit additional parameters or regard the orders in (11) as “marked periphery.” PPL argues neither alternative leads to overall formal simplicity.

2.2.2. Smallness of Number

The number of parameters is assumed to be smaller than that of rules. According to Lightfoot (1999: 259), for example, it amounts to thirty or forty. Also, the smallness of number is considered to reduce the amount of learning. But PPL points out that “[c]ertainly hundreds have been proposed since the notion was introduced around 1980” (p. 81). For example, even if we assume that there is one parameter for each functional category based on the assumption that only functional categories are parameterized (Fukui (1988: 267)), the number of proposed functional categories itself is quite large. Cinque (1999) posits thirty two functional heads in the IP domain alone. CP is also considered to have more than a dozen projections (Rizzi (1997)). Also, as more and more publications focus on “micro-parametric differences” (p. 68) to distinguish between language and dialectal variations, the number of parameters inevitably increases.

Transferring parameters into lexicon does not reduce the potential number of parameters and “the statement that some language makes the parametric choice that lexical item L licenses functional projection P is indistinguishable from the statement that there is a language-particular rule involving L that specifies P” (p. 83). Therefore, there is no reason to conclude that the number of parameters is greatly smaller than that of rules.

2.2.3. Predicting Clustering of Typological Properties

In the principles-and-parameters approach, value was placed on predicting a clustering of unexpected set of typological properties. With the positive setting of the Null Subject Parameter, null subject languages such as Italian and Spanish are predicted to manifest the properties listed in (1) above. Pollock (1989) explains different behaviors of VP-adverbs, clausal negation, and floated quantifiers in English and French by positing a lexical verb movement to a higher inflectional position in French while prohibiting the movement in English. Hale (1982, 1983) argues that properties of non-configurational languages are predicted by a single parameter setting. Snyder (2001) has proposed a Compounding Parameter, which predicts the availability of a range of complex predicate constructions should pattern closely with the availability of root compounding.³

Concerning the prediction, PPL first points out that many studies on parameters and clustering deal with a single language or several related languages. In that case, PPL argues that rule-based models can also predict clustering of properties, as illustrated by the proposals in Chomsky (1957: 39):

- (12) a. AUX → TNS (M) (*have + en*) (*be + ing*)
 b. affix element + verbal element → verbal element + affixal element (p. 87)

These two rules predict various behaviors of English auxiliaries although they do not predict anything with respect to other languages.

PPL then shows that the Null Subject Parameter, the best-studied parameter in the principles-and-parameters approach, underpredicts the class of possible languages. Chomsky (1982) proposes that the empty subject is the small “pro” and it is licensed by rich inflection on AGR. Rizzi (1982) further assumes that the pronominal INFL in null subject languages is either referential or non-referential and that only the referential INFL licenses the null-subject. In spite of the revision, existence of languages that have null subjects but not subject inversion such as Brazilian Portuguese and Chinese cannot be predicted. Furthermore, Jaeggli and Safir (1989)’s proposal that null subjects are permitted in languages with morphologically uniform para-

³ Snyder (2001), which PPL mentions very briefly, has two noteworthy characteristics. First, he adds child language acquisition as a novel source of evidence to get over the limitations of a purely comparative approach. Secondly, he explicitly defends the existence of a syntactic parameter in the classical sense of Chomsky (1981).

digms is refuted by the facts in Old French (Roberts (1993)), Swedish and Russian (Speas (1994)), Brazilian Portuguese (Rohrbacher (1994)). More significantly, works after Jaeggli and Safir (1989) all omit the discussion of rich clustering of typological properties.

PPL also throws doubt on the robustness of clustering predicted by the Head Parameter. Citing the data from Dryer (1991), PPL observes "SVO languages are in general intermediate in their properties between those in which the verb is on the right margin or on the left margin" (p. 94). Therefore a simple head parameter cannot explain the difference between SVO and V-initial languages.

PPL assumes that a parametric theory would receive support if unexpected correlations of properties were predicted by the abstract parameters, but concludes that "two decades of intensive research has failed to reveal the existence of the hoped-for correlations" (p. 98).

2.2.4. Lexical Parameter Hypothesis

In the Minimalist Program (MP), the locus of parametric variation is mainly shifted to the heads of functional categories. The idea originates in the Functional Parameterization Hypothesis in Fukui (1988). In the early MP, Chomsky (1995) posits that functional heads could have either strong or weak features. Strong features are visible at PF if not checked so that functional heads with strong features need to undergo overt movement to have the strong feature checked. On the other hand, functional heads with weak features need not undergo overt movement since weak features are invisible at PF. Typological variations, then, will be attributed to different combinations of feature values on functional heads. In recent MP work, the strong-weak feature distinction is given up and movement is triggered to eliminate uninterpretable features.

The main problem of the MP approach pointed out in PPL is that as the locus of parametric variation is moved from the entire grammar to functional categories in the lexicon, it has become all but impossible to predict clustering of typological properties:

- (13) The original vision of parameters was an extremely attractive one, in that the set of their settings was conceived of as a checklist for a language as a whole. But the Lexical Parameterization Hypothesis (LPH), the idea that values of a parameter are associated not with particular grammars, but with particular lexical items has put an end to this vision. (p. 95)

Thus it is puzzling for Newmeyer that the shift of the locus of parametric

variation is “often portrayed as a major step forward” (p. 95). Furthermore, as typological variations concerning head directionality and word order are attributed to PF, we have the problem of determining which variation belongs where and of capturing correlations of the variations in the two domains.

2.2.5. The Irrelevance of Typology for Grammatical Theory

PPL further argues that “grammars do not encode typological generalizations either directly or indirectly” (p. 104) by showing that typologically significant generalizations are fundamentally different from linguistically significant generalizations.

First, PPL shows that typological generalizations do not derive from UG principles. The fact that V-final languages tend to lack *wh*-movements and to have final question particles is explained by the lack of a specifier for COMP (Fukui (1986)), quantifier movement (Kim (1990)), the parametric choice of a final question particle or *wh*-movement (Cheng (1997)). However PPL throws doubt on these analyses since they are highly stipulative as the explanation of typological generalizations and cannot explain, for example, why the languages which choose a final Q-particle/*wh*-in-situ are overwhelmingly OV. Furthermore, there are some typological generalizations which no imaginable parametric device can explain. One such example is the tendency that if in prepositional phrases complex modifiers on nouns are permitted, then in that language simpler modifiers are also permitted:

- (14) Proposition Noun Modifier Hierarchy (PrNMH; Hawkins (1983))
 If a language is prepositional, then if RelN then GenN, if GenN then AdjN, and if AdjN then DemN. (p. 107)

Unless generative grammarians appeal to relative structural complexity, it is difficult to incorporate the hierarchy. On the other hand, it is easy to explain this hierarchy by the functional pressure to reduce constituent recognition time as proposed in Hawkins (2004).

Next, typological generalizations are not always stated as D-structure generalizations. For example, the basic word order of German and Dutch are basically SOV in subordinate clauses and SVO in main clauses. Through detailed analyses of generative grammarians, it is generally agreed that the basic word order of these languages is SOV and the SVO order is derived from the basic word order. Nevertheless German and Dutch have many VO correlates.

Thirdly, the cross-linguistically rarer phenomena are often characterized more simply than the more common ones. The case in point is preposi-

tion-stranding. Although preposition-stranding is cross-linguistically rare, a language with preposition-stranding is theoretically characterized as simpler than the one without the operation in that all four lexical categories are specified as proper governors. On the other hand, in the language without preposition-stranding, N, V and A are proper governors but P is not. An additional rule or principle is necessary to explain why P alone is excluded.

Finally, the relation between linguistic knowledge and typological generalizations is examined. Typological generalizations are usually stochastic and children acquiring a language are not exposed to cross-linguistic generalizations. PPL then claims that typological generalizations are “not part of knowledge of language at all” (p. 118).

Thus, PPL concludes that typological generalizations belong to the domain of E-language and that performance theory should be responsible for those generalizations. As an example of a performance theory, a processing-based explanation of cross-linguistic variation by Hawkins (2004) is described in detail at the end of this chapter.

2.3. Distinction between Language Knowledge and Language Use

Although typological generalizations are claimed not to be part of language knowledge and they are functionally motivated, in Chapter 3, PPL supports the distinction between language knowledge and language use. That is, although the role of formal grammar is diminished, performance-based explanations cannot replace the formal principles. In Newmeyer’s words, “Saussure (and Chomsky after him) were correct in distinguishing language knowledge from language use” (p. 128). This chapter supports the classical Saussurean position.

In the usage-based model, which derived from generative semantics and is shared by most cognitive linguists, no sharp distinction is made between linguistic and non-linguistic knowledge. PPL gives several reasons that the usage-based model seems appealing to some linguists. First, a lot of “incontrovertible evidence” has piled up “that grammars are shaped in part by performance considerations” (p. 130). Publications such as Hopper and Thompson (1980), Bybee (1985), Heine and Claudi (1986), and Croft (1990) demonstrate the point. Secondly, it has also become evident that “language users are sensitive to the frequency of grammatical forms” (p. 130). For example, Lightfoot (1991) makes crucial use of frequency effects in explaining the change of basic word order in English. Thirdly, great discrepancy between sentences generated by formal grammars and actual utterances also reinforces the skepticism about generative models. A lot of studies have

shown that in actual speech, the majority of sentences fail to express the full propositional structure and this leads functionalists to conclude that full argument structure is of no great significance in linguistic theory. Finally the rise of the connectionist approach also contributes to the popularity of the usage-based modes.

PPL refutes these arguments in favor of the usage-based model. First, it points out that generative approaches are not different from functional approaches in assuming that much of grammatical structure is motivated by functional pressure. This is clear from Chomsky's remark cited below:

- (15) Surely there are significant connections between structure and function; this is not and has never been in doubt.... Searle argues that "it is quite reasonable to suppose that the needs of communication influenced the structure" of language, as it evolved in human prehistory. I agree. (Chomsky (1975: 56–58))

Thus, the issue is "not whether grammars have functional motivation, but where and how much, and the centrality of focusing on this motivation in one's research program" (p. 136). Secondly, as any theory involving modeling on a network is now called connectionist, the connectionist approach itself is not necessarily non-modular and this approach has not given remarkable results in dealing with grammatical facts. Thirdly, PPL provides several pieces of evidence that full argument structure is mentally represented. For example, elliptical sentences have the same characteristics as non-elliptical ones with respect to anaphors:

(16) Who does John_i want to shave?

- (17) a. Himself_i
 b. Him_j
 c. *Myself
 d. *Him_i

(p. 145)

In response to (16), answers (17a, b) are possible while (17c, d) are not. This is totally parallel with the full sentences in (18):

- (18) a. John_i wants to shave himself_i.
 b. John_i wants to shave him_j.
 c. John_i wants to shave me.
 d. *John_i wants to shave myself.
 e. *John_i wants to shave him_j.

(p. 145)

This parallelism shows that when speakers use sentence fragments as in (17), they have full representation of the argument structure. The same is true of complementizer choice and the internal structure of noun phrases. For another example, studies on child speech indicate that chil-

dren know more about grammar than is expressed by their utterances. The fact that little-used sentences or introspective data can elicit a stable pattern of response also supports the point.

If the distinction between competence and performance is strictly maintained, then it is natural that grammar itself is not suitable for communication because it is only one among various factors that make communication possible. Therefore, to find some generalizations that grammar cannot cover will not be in itself the refutation of formal grammar. Furthermore PPL points out that the arguments based on “needs” or “usefulness” are not predictive. Any grammatical element can be given some functional motivations while not all “useful” concepts are grammaticalized. For example, to distinguish inclusive first person plural and exclusive first person plural is certainly “useful,” but such distinction is not grammaticalized in every language. Thus, “[w]e need a theory, which is now lacking, of why some seemingly needed features result in grammatical coding and some do not” (p. 154).

PPL also argues against the use of stochastic data as a reflection of individual grammar. For one thing, corpus data is the composite of sentences and utterances produced by different individuals and stochastic data hide variations among them. Also, the probability of using some forms can be affected by real-world knowledge. *Walk* is used more often as an intransitive verb than as a transitive verb because walking alone is a more common activity than walking something else. Furthermore, applying stochastic approaches to syntax is problematic because syntactic variations are often related with meaning differences. Therefore, the probabilistic data may reflect the difference of meaning to be conveyed, not some inherent property of grammar.

In summary, PPL claims that grammatical principles are mentally represented and functional motivations are not registered in the innate language faculty although much of grammar is functionally motivated.

2.4. The Relation between Form and Function

Chapter 3 has shown that formal grammar is irrelevant for typological generalizations and the locus of explanation is shifted to functional pressures. Chapter 4 then claims that although functional elements affect the distribution of formal elements, the distinction between linguistic knowledge and use should be rigidly maintained. Then, the next question is how direct the linkage between functional pressures and the distribution of grammatical elements motivated by those pressures.

In Chapter 5, PPL examines two opposing positions on the relation:

- (19) Atomistic functionalism (AF): There is direct linkage between properties of particular grammars and functional motivations for those properties. (p. 174)
- (20) Holistic functionalism (HF): There is no direct linkage between external functions and grammatical properties. The influence of the former on the latter is played out in language use and acquisition and (therefore) language change and is manifested only typologically. (p. 175)

This chapter is devoted to refuting AF and supporting HF.

At the beginning, two problems in deciding between AF and HF are indicated. The first problem is “overavailability” of functional explanation. Candidates for functional motivations are so diverse and numerous that any element in formal grammar can be given some functional motivations. Even the transformational rules in the pre-GB period were claimed to have functional motivations by Langacker (1974), for instance. The second problem is the tenuous relation between the results of psychological experiments and cross-linguistic tendencies. Although it has become evident through psycholinguistic experiments that isomorphic structures are more readily interpretable, it is not clear whether structures which are easier to parse are cross-linguistically more common. For example, Kirby (1998) shows that parallel function relatives, where the grammatical role of the moved element and that of the relative clause itself are the same as in (21a), are easier to parse than non-parallel function relatives as in (21b), although the former type does not cross-linguistically predominate nor does the latter type typologically implicate the former type.

- (21) a. The woman who is walking in the hallway is my friend.
 b. The woman who John knows is walking in the hallway.

(p. 181)

PPL then suggests that language change is a testing ground for deciding between the two positions because changes are “more concrete and easier to study than more abstract properties of grammars” (p. 183).

On this testing ground, PPL tries to argue that diachronic change gives no support to AF because many instances of diachronic change are not necessarily regarded as maximization of functionality of extra-linguistic properties. Drawing on Weinreich, Labov, and Herzog (1968), PPL identi-

fies three stages in the process of language change:⁴

- (22) a. innovation (the first appearance of the change)
- b. propagation (the adoption of the change by the speech community)
- c. retention (the transmission of the change from grammar to grammar in successive generations) (p. 184)

In the retention stage, functional motivations are obviously irrelevant. Today's English has SVO word order not because subjects are cognitively more prominent than objects but simply because the grammars of the parents and peers of language acquiring children have that property. This conventionality itself has a functional motivation since it reflects "mental routinization and social convention" (p. 185). Taking the structural change of genitive noun phrases in English as another example, PPL also argues that although various kinds of functional motivations affect structures of a language in the course of development, these structures cease to reflect the functional pressures as the language further changes. Functional pressures are also irrelevant in the propagation stage because "the mechanisms of propagation are social, not linguistic" (p. 187).

On the other hand, Newmeyer admits that a lot of structures are functionally motivated in the innovation stage. One example is the loss of dative and genitive cases governed by verbs, which are replaced by a structural accusative case in some Germanic languages like English. This change reflects the tendency to increase the form-meaning transparency. There are also parsing-motivated changes, in which language becomes more "harmonic" to make parsing easier. One example is reanalysis of verbs as prepositions in harmony with development of VO orders as described in Aristar (1991), Givón (1971), Heine and Reh (1984), and Vennemann (1973).

Nevertheless, PPL claims that not all innovations make languages easier. In some cases, applicability of a rule can narrow down, as il-

⁴ It should be noted that a leading functionalist William Croft also recognizes the three stages in language change:

- (i) [A] theory of language change must explain why languages do not change in many ways, sometimes over many generations of speakers ... [A] theory of language change must distinguish the two processes of change, that is, it must distinguish altered replication from differential replication. To use the terminology more typically found in linguistics, the two processes are INNOVATION or actuation—the creation of novel forms in the language—and PROPAGATION or diffusion (or, conversely, loss of those forms in the language.

(Croft (2000: 4–5); emphasis original)

illustrated by the change of applicability of V-to-I movement in English history. This change cannot be regarded as maximizing functionality. In other cases, innovations are apparently in the direction of decreasing functionality. Newmeyer considers the development of French sentential negation as an example of this dysfunctional innovation. In negative sentences in Old French, the negative particle *ne* was reinforced by independent nouns such as *point* “dot, point,” *mie* “crumb,” *gote* “drop,” and in the course of time *pas* “step” was recruited as the sole reinforcer and has been virtually compulsory since the seventeenth century. For Newmeyer, this process cannot be regarded as a process leading to increased functionality and “[p]resumably, an AF-oriented theory would predict that French would have left things alone” (p. 189). Furthermore, innovations induced by language contact are even more remotely functional. The most conspicuous example is the system of English word stress. In Old English, stress was placed on the first syllable except for the words with certain prefixes. But in Modern English, the stress rule has become more complicated due to borrowings from French and Latin. The words from Old French have a stress on the final syllable and those from Latin have a stress on the penultimate syllable. As a result, it is difficult to provide English word-stress rules with functional motivations.

PPL opposes AF on other grounds, too. The first argument is “overgeneralization” of the functional motivations. Principles that were originally motivated by functional pressures are sometimes overgeneralized to sentences with no processing difficulties. For example, (23a) and (23b) are almost the same in complexity, but only (23a) contains a constraint violation and hence is ungrammatical:

- (23) a. *John tried for Mary to get along well with.
 b. John is too snobbish for Mary to get along well with.

(p. 192)

Secondly, principles with good functional motivations sometimes have dysfunctional consequences. For example, parsing considerations can motivate the condition that traces of movement must be lexically governed. But this condition blocks the extraction of subjects:

- (24) a. *Who_i do [sic] think [e_i that e_i saw Fay]?
 b. *Who_i do you wonder [e_i how [e_i solved the problem]]?

(p. 193)

It is not difficult to imagine the case where we want to ask the identity of the subordinate subject. So this is certainly an undesirable result. But the negative effect of this principle does not affect the functioning of languages

because they have means to get around it. In English, the complementizer is adjusted so that the extraction will be licensed:

(25) Who do you think saw Fay? (p. 193)

Thus, PPL asserts that no rule or constraint in itself is functionally motivated but functionality is relevant only in the total system. Thirdly, PPL raises the problem of competing motivations. It observes “the existence of competing motivations threatens to render functional motivations vacuous” (p. 195). The point of the claim is that if one language has property X motivated by some functional explanation α and another language has property Y motivated by another functional explanation β , where X and Y are incompatible, the explanation of why α motivates X and β motivates Y will be necessarily circular.

Furthermore, as another instantiation of AF, functionally based Optimality Theory (FOT) is criticized because paring each constraint with an external functional motivation “locates the form-function interplay in the mental grammar itself, rather than seeing the response of form to function as emerging from language use and acquisition” (p. 225).

Based on these arguments PPL concludes that HF is better motivated than AF. But some of the arguments seem to have problems. The next section will examine some of them.

3. Discussion

PPL is a critique of generative approaches, especially the parameter-based one, to typological generalizations. Firstly, it has shown that the parametric approach has failed to explain typological generalizations and claims that such generalizations are beyond the scope of formal grammar. It then criticizes functional approaches to typological generalizations and claims that the relation between functional motivations and grammatical elements is indirect.

The first point is criticized by Roberts and Holmberg (2005), which is a critique of Newmeyer (2004). They propose that Newmeyer’s arguments against the principles-and-parameters model of cross-linguistic variation “are based on misunderstandings either of theory or of data, are conceptually misconceived, illogical or simply false” (Roberts and Holmberg (2005: 538)), examining Newmeyer’s arguments against the utility of parameters one by one and demonstrating an example of the efficacy of the parameter-based approach. Newmeyer (2006) rebukes Roberts and Holmberg (2005) pointing out deficiencies of their arguments both on factual and theoretical

grounds. To the reviewer, Newmeyer's arguments seem more persuasive than those of Roberts and Holmberg (2005).⁵ Furthermore, as pointed out by PPL, since MP localizes parametric variations in functional categories in the lexicon, it has become quite complicated to characterize cross-linguistic variations by feature strength or interpretability. To the reviewer, the first point seems convincing.

But as to the second point, it is not clear to the reviewer whether the case against AF is legitimately made. Thus, this section focuses on the second point and will show that there are some cases in which functional factors are more relevant than Newmeyer claims and argue that it is necessary to consider how functional factors interact with non-functional ones.

The organization of this section is as follows. In 3.1, the criticisms focusing on diachronic aspects are examined and we will show that some of the arguments are not convincing. In 3.2, we will argue that competing motivations are not in themselves a problem to AF.

3.1. Functional Motivations in Language Change

PPL's arguments against AF are based on considerations of language change and are summarized in two points:

- (26) a. Since functional factors are irrelevant in the propagation and the retention stages, synchronically these factors do not motivate grammatical structures. In other words, "the forces that bring a construction into a language are not necessarily the same ones that keep it there" (p. 185).
- b. Although functional factors are relevant in the innovation stage, they do not motivate all innovations. In fact, some innovations may have dysfunctional consequences.

We will first argue that at the synchronic level, there are cases where functional elements seem to be relevant in motivating grammatical forms. If they are synchronically relevant, then PPL's first claim (26a) is too strong. It is necessary to consider how functional factors interact with non-functional ones to keep a construction as it is. We will also argue that in some of the cases regarded as showing dysfunctional innovations, we can find cases which seem to have some functional motivations. If they prove

⁵ Roberts and Holmberg (2005) defend parameters in the principles-and-parameters model. It is not clear why they do not defend the parametric approach in the MP and what they think about Newmeyer's arguments against the MP approach.

to be functionally motivated, then PPL's second claim (26b) is also too strong. It is also necessary to consider how and when functional factors can motivate innovations collaborating with other factors.

3.1.1. Functional Pressures at the Synchronic Level

As mentioned in 2.4, PPL examines the development of the genitive noun phrase in English to illustrate irrelevance of functional motivations at the retention stage. As English basic order shifted from SOV in OE to SVO in ME, the structure of genitive noun phrases also shifted from Gen-N in OE to N-Gen in ME. But Gen-N order has revived in Modern English and the two structures coexist. PPL points out that the two orders are maintained in Modern English not simply because Gen-N is used for animate possessors and N-Gen is used for inanimate possessors. Such a functional factor is only one of many factors such as convention, weight, and purely structural pressure. For example, the animacy factor cannot explain the pair in (27):

- (27) a. the table's leg
 b. the mother of the lawyer (p. 186)

PPL then claims that the relation between grammatical rules or principles and functional pressures is quite indirect:

- (28) The point is that languages are filled with structures that arose in the course of history to respond to some functional pressure, but, as the language as a whole changed, ceased to be very good responses to that original pressure. Such facts are challenging to any theory like AF, in which the sentences of a language are said to be a synchronic product of constraints that must be functionally motivated. (p. 187)

However, what the examples show is not that functional motivations are irrelevant, but that a functional motivation can compete with non-functional factors. If we can predict in a principled way when the functional motivation is given a priority over other factors, then the criticism (28) will not hold. Also, we can find cases in which a functional motivation is more readily observable as there are not so many competing factors. Below we will examine such cases.

Givón (1989) proposes an iconic principle concerning a correlation between code quantity and mental effort:

- (29) Code—Quantity and Mental Effort:
 The more mental effort is used in processing information, the more coding material will be used in representing the information in language. (Givón (1989: 106))

With the scale of referential predictability in (30), this principle explains the coding differences as illustrated in (31):

- (30) **most predictable/continuous**
 a. zero anaphora
 b. unstressed pronoun
 c. stressed pronoun
 d. definite noun
 e. restrictively modified definite noun
least predictable/continuous (Givón (1989: 105))
- (31) a. John came in, [Ø] looked around and [Ø] gasped.
 b. John talked to Bill; then *he* left. (=John left).
 c. John talked to Bill; then **he** left (=Bill left).
 d. John came in and paused. *The woman* got up.
 e. *The tall woman* remained in the room. Later on, *the man who had been hiding under the couch* emerged.
 (Givón (1989: 106))

Horie (2000) argues that principle (29) is also applicable to the coding difference between core and oblique argument positions in Japanese and Korean. Horie (2000) points out that Japanese sentence nominalizers *no* and *koto* are both acceptable in argument positions as in (32a) but only *koto* is acceptable in oblique positions as in (32b):

- (32) a. [Kare-ga amerika-ni it-ta **no/koto**]-o
 he-NOM America-to go-PAST NMLZR-ACC
 sitte-masi-ta ka?
 know:GER-POL-PAST-Q
 'Did you know that he went to the US?'
- b. Sono hito-wa [eigo-o osieru-**{koto/*?no}**]-de
 that person-TOP English-ACC teach-NMLZR-INSTR
 seikei-o tatete-iru.
 living-ACC make:GER-exist
 'That person makes his living by teaching English.'
 (Horie (2000: 90–91))

This is because while a nominalized clause occurring in unmarked core argument positions such as subject and direct object does not require particular mental effort in processing and thus requires only minimal linguistic coding, a nominalized clause occurring in marked oblique argument positions requires more processing effort and thus more explicit coding of the oblique nominalized clause.

Kuno (1983) points out that a difference exists between *-(r)enai* and *-koto*

ga dekinai, both of which express negation of ability:

- (33) a. Shibafu-ni-wa hai-renai
 lawn-to-Top enter-not:able
 b. Shibafu-ni-wa hairu-koto-ga-dekinai
 lawn-to-Top enter-Comp-Nom-not:able
 ‘You cannot enter the lawn.’

Kuno (1983: 151) notes that (33b) is more natural as a notice sign because it implies that the prohibition is imposed by some external authority while (33a) indicates lack of ability of someone who might try to enter the lawn. It is possible to explain the contrast by (29) because the longer form *-koto ga dekinai* expresses the external prohibition while the shorter form expresses the more direct internal inability of the subject.⁶

These examples show that functional motivations for grammatical forms are available at the synchronic level. It is important to consider how these functional factors interact with non-functional ones.

3.1.2. Functional Motivations of “Dysfunctional” Innovations

As briefly discussed in 2.4, PPL cites the development of French sentential negation as an example of dysfunctional innovations. The point is that there is no functional reason for *pas* to reinforce and finally replace the negator *ne*. But we must draw attention to the fact that English parallels this pattern of development. Horn (1989: 455) shows the parallelism clearly as in (34):

- | | |
|---|---|
| (34) Old French: Jeo ne dis
Modern French
(standard): Je ne dis pas
Modern French
(colloq.): Je dis pas | Old English: Ic ne secge
Middle
English: Ic ne seye not
Early Modern
English: I say not |
|---|---|

In English, too, the negator *ne* was reinforced by *not* in Middle English. In fact, Jespersen (1917: 5) gives a functional account of why negation needs

⁶ One of the anonymous reviewers pointed out that to prove the validity of arguments based on iconicity, it is necessary to show that explanations based on frequency do not work. It is true that explanations based on iconicity sometimes do not seem to work as in the case of anti-causatives pointed out by the anonymous reviewer. Also, Haspelmath (2007) argues that most of the core phenomena for which iconicity seems responsible are explained by frequency principles and the economy principle. However, we cannot still exclude the possibility that iconicity is a factor influencing frequency at the innovation stage (Bybee (1988: 359)). Furthermore, it is not clear how frequency can explain such phenomena as illustrated in (32) and (33).

to be strengthened:

- (35) The negative notion, which is logically very important, is ... made to be accentually subordinate to some other notion; and as this happens constantly, the negative gradually becomes a mere proclitic syllable (or even less than a syllable) prefixed to some other word. The incongruity between the notional importance and the formal insignificance of the negative may then cause the speaker to add something to make the sense perfectly clear to the hearer.

Then, from the Jespersen's point of view, the development of French negation is in no way dysfunctional.

Another example PPL cites as a dysfunctional innovation is borrowing. But it is worth noting that not all borrowed forms are counter-functional. A conspicuous example of such cases is borrowing of pronouns in English. In Old English, the third person plural pronouns were *hīe* (nominative), *hiera* (genitive), *him* (dative) and *hīe* (accusative). On the other hand, the third person singular feminine pronouns were *hēo* (nominative), *hiere* (genitive), *hiere* (dative) and *hīe* (accusative). Also the dative of the third person singular masculine and neuter pronoun was *him*. In that situation, borrowing *they*, *their* and *them* from Old Norse had a merit of disambiguation, as pointed out below:

- (36) Possibly the Scandinavian words were felt to be less subject to confusion with forms of the singular.

(Baugh and Cable (1993: 100))

This case shows borrowing is not necessarily dysfunctional and the problem is to specify conditions that affect the prominence of functional motivations among competing functional or structural factors.

3.1.3. Grammaticalization

Grammaticalization is a process in which a lexical unit or structure assumes a grammatical function. It has been studied from both historical and synchronic perspectives. In the former perspective, grammaticalization is a subset of linguistic changes; in the latter perspective, it is seen as "primarily a syntactic, discourse pragmatic phenomenon, to be studied from the point of view of fluid patterns of language use" (Hopper and Traugott (2003: 2)). Functionalists show more interest in this phenomenon than generativists and many of them claim that grammaticalization challenges generative grammar. For example, Traugott and König (1991: 189) observe "[t]he study of grammaticalization challenges the concept of a sharp divide

between langue and parole, and focuses on the interaction of the two.” If that is the case, then grammaticalization presents a challenge to the position of PPL on language change. But Newmeyer (1998: Ch. 5) denies the existence of grammaticalization as a distinct grammatical phenomenon:

- (37) Far from calling for a “new theoretical paradigm,” grammaticalization appears to be no more than a cover term for a conjunction of familiar developments from different spheres of language, none of which require or entail any of the other.

(Newmeyer (1998: 295))

The reviewer agrees with Newmeyer that grammaticalization is an epiphenomenon involving various components of independent changes. But that does not exclude the possibility that some components of such changes have functional motivations. For example, Traugott and Dasher (2002) argue that the main mechanism of semantic change is subjectification and intersubjectification, the processes in which meanings become grounded in the speaker and the speaker-hearer relation, respectively. These processes are claimed to be not only relevant to grammaticalization but also typical of semantic change in general. Specifically, Traugott and Dasher give detailed analyses of the developments of modal verbs, adverbials with discourse marker function, performative verbs and constructions, and social deictics. More research in this direction will make it clear how and to what extent functional principles contribute to the innovation stage of language change.

3.2. Competing Motivations

As noted in section 2.4, PPL claims that the existence of competing motivations renders functional explanations vacuous because the explanation of differing functional motivations will necessarily be circular. PPL does not illustrate the point with actual examples. Below we will examine a set of examples to argue that it is not necessarily the case that the explanation will be circular.

English modal auxiliaries have several distinct meanings. For example, according to Oxford Advanced Learner’s Dictionary (7th edition), *may* is “used to say that something is possible” as in *He may have missed the train* or “used to ask for or give permission” as in *May I come in?* The same is true of *must*, which is “used to say that something is likely or logical” as in *You must be hungry after all that walking* or “used to say that something is necessary or very important (sometimes involving a rule or a law)” as in *Cars must not park in front of the entrance*. The former uses expressing possibility or likelihood are referred to as *epistemic* while the latter uses

expressing permission or obligation are referred to as *deontic* or *root*. The same string of words can express both epistemic and deontic meanings:

- (38) a. John *must* be home by ten; Mother won't let him stay out any later.
 b. John *must* be home already; I see his coat.

(Sweetser (1990: 49))

In (38a) *must* is used to express obligation ascribed to John while *must* in (37b) expresses assessment of probability based on perceived evidence. Sweetser (1990) emphasizes the universality of such ambiguity:

- (39) This ambiguity is not peculiar to English; indeed, there is an evident crosslinguistic tendency for lexical items to be ambiguous between these two sets of senses. Many unrelated languages (Indo-European, Semitic, Philippine, Dravidian, Mayan, and Finno-Ugric, among others) are alike in having some set of predicates which carry both the root and epistemic modal meanings as English modal verbs do. (Sweetser (1990: 49))

In Old Japanese, there were several auxiliaries which encoded both epistemic and deontic meanings as shown in (40):

(40)	Deontic	Epistemic
- <i>beshi</i>	Obligation, Intention	Certainty
- <i>mu</i>	Intention	Probability
- <i>mashi</i>	Wish	Counterfactual conjecture

(Horie (1997: 440))

However, modal auxiliaries in Modern Japanese are quite different from those in Old Japanese. We can point out two main characteristics of their diachronic development. First, only a few Old Japanese auxiliaries survived into Modern Japanese. For example, *meri* and *nari*, which mainly express evidentiality in Old Japanese, do not exist in Modern Japanese. Those survived, such as *beki-da* (<-*beshi*), underwent semantic narrowing (Kondo (2000: 479)). In Modern Japanese, *beki-da* is used mostly as a deontic modal expression and more often it is used as a nominal-modifier form. As another example, *-yoo* and *-daroo* (<-*mu*) mostly function as sentence-final predicates and are rarely used as prenominal modifiers (Kondo (2000: 479)).

The second characteristic is emergence of periphrastic modal auxiliaries. From around the 17th century, new periphrastic modal auxiliaries began to appear as if they complement decreasing Old Japanese modal auxiliaries. Although English also has periphrastic modal auxiliaries such as *have to* and *be going to*, the number is quite small and they are usually ambiguous as non-periphrastic auxiliaries are. The following list shows

first attested dates of periphrastic modal auxiliaries based on Kitahara et al. (2000):

- (41) *nakere-ba-nara-nai* (obligation): 1638
ni-chigai-nai (certainty): 1734
temo-ii (permission): 1833
kamo-shire-nai (probability): around the end of 16th century

Consequently, in Modern Japanese, there are few instances of modal auxiliaries exhibiting deontic-epistemic ambiguity, which is not the case in English.

We can interpret the difference as deriving from different functional motivations. While English chose economy over efficiency, Japanese chose efficiency sacrificing economy. To avoid a circular explanation, we must attribute the difference in choice to some other property. One candidate of the property is the difference in source categories. While English modal auxiliaries evolved from main verbs, modal auxiliaries in Old Japanese derived from parts of verbal inflection and periphrastic modal auxiliaries are composed of stative predicates. Sweetser (1990) argues that the shift from deontic to epistemic meaning in English is metaphoric extension of force-dynamic properties. To be more specific, Sweetser (1990) proposes that deontic meanings are extended to epistemic meanings through a metaphorical process whereby the logic of the external (sociophysical) world is applied to that of the internal mental world:

- (42) Thus, we view our reasoning processes as being subject to compulsions, obligations, and other modalities, just as our real-world actions are subject to modalities of the same sort.

(Sweetser (1990: 50))

On the other hand, as Onoe (2001: 459) points out, no force dynamics is involved in the development of Japanese modal auxiliaries because modal auxiliaries in Old Japanese, such as *-u* and *-yoo*, did not derive from main verbs. In this case the difference in source categories affects the availability of some developmental patterns and leads to differing priorities of functional motivations.⁷

If the analysis proves to be appropriate, it will show that the existence of competing motivations itself does not threaten functional explanations in gen-

⁷ Moriya and Horie (2004) show that Korean modal auxiliaries are the same as Japanese ones in these respects. However, as one anonymous reviewer points out, it is necessary to investigate many other languages to prove that the difference in source categories causes the difference in modal auxiliary systems.

eral.

4. Conclusion

Together with Newmeyer (1998), PPL is a quite informative overview and comparison of generative grammar and functional approaches. Reading it and referring to numerous papers and books cited in it, you will have a bird's eye view of researches concerning typological generalizations in both formal and functional approaches. The claim that explaining typological generalizations is beyond the scope of UG is noteworthy as an overall evaluation of generative approaches to typological generalizations. On the other hand, we have argued that some of the arguments against AF are too strong. In some cases, functional motivations seem more relevant than Newmeyer claims. Therefore, it is necessary to explicate how functional motivations interact with grammatical elements. This statement will not contradict with what Newmeyer himself mentioned: "[w]e need a theory, which is lacking, of why some seemingly needed features result in grammatical coding and some do not" (p. 154).

Lastly, the reviewer noticed only two typographical errors. In the sentence "[s]o, Manning (2002a) observes that Pollard and Sag (1994) consider sentence (57) grammatical, but they star (57b)" on p. 161, (57) should be replaced with (57a). The example (29a) on p. 193, which is cited as (24a) in this review article, lacks the main subject *you*. The smallness of the number of typographical errors also reflects the high quality of this book.

REFERENCES

- Aristar, Anthony R. (1991) "On Diachronic Sources and Synchronic Patterns: An Investigation into the Origin of Linguistic Universals," *Language* 67, 1–33.
- Baugh, Albert C. and Thomas Cable (1993) *A History of the English Language*, 4th ed., Routledge, London.
- Bybee, Joan (1985) *Morphology: A Study of the Relation between Meaning and Form*, John Benjamins, Amsterdam.
- Bybee, Joan (1988) "The Diachronic Dimension in Explanation," *Explaining Language Universals*, ed. by John A. Hawkins, 350–379, Basil Blackwell, Cambridge, MA.
- Cheng, Lisa L.-S. (1997) *On the Typology of Wh-Questions*, Garland, New York.
- Chomsky, Noam (1957) *Syntactic Structures*, Mouton, The Hague.
- Chomsky, Noam (1962) "Explanatory Models in Linguistics," *Logic, Methodology,*

- and *Philosophy of Science: Proceedings of the 1960 International Congress*, ed. by Ernst Nagel, Patrick Suppes and Alfred Tarski, 528–550, Stanford University Press, Stanford.
- Chomsky, Noam (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.
- Chomsky, Noam (1975) *Reflections on Language*, Pantheon, New York.
- Chomsky, Noam (1981) *Lectures on Government and Binding*, Foris, Dordrecht.
- Chomsky, Noam (1982) *Some Concepts and Consequences of the Theory of Government and Binding*, MIT Press, Cambridge, MA.
- Chomsky, Noam (1986) *Knowledge of Language: Its Nature, Origin, and Use*, Praeger, New York.
- Chomsky, Noam (1995) *The Minimalist Program*, MIT Press, Cambridge, MA.
- Cinque, Guglielmo (1994) “On the Evidence for Partial N Movement in the Romance DP,” *Paths Towards Universal Grammar*, ed. by Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi and Raffaella Zanuttini, 85–110, Georgetown University Press, Washington, DC.
- Cinque, Guglielmo (1999) *Adverbs and Functional Heads: A Cross-Linguistic Perspective*, Oxford University Press, Oxford.
- Croft, William (1990) *Typology and Universals*, Cambridge University Press, Cambridge.
- Croft, William (2000) *Explaining Language Change: An Evolutionary Approach*, Longman, London.
- Dik, Simon, C. (1989) *The Theory of Functional Grammar, Part 1: The Structure of the Clause*, Foris, Dordrecht.
- Dryer, Matthew S (1991) “SVO Languages and the OV:VO Typology,” *Journal of Linguistics* 27, 443–482.
- Fukui, Naoki (1986) *A Theory of Categorical Projection and Its Applications*, Doctoral dissertation, MIT.
- Fukui, Naoki (1988) “Deriving the Differences between English and Japanese: A Case Study in Parametric Syntax,” *English Linguistics* 5, 249–270.
- Givón, Talmy (1971) “Historical Syntax and Synchronic Morphology: An Archaeologist’s Field Trip,” *CLS* 7, 394–415.
- Givón, Talmy (1983) “Topic Continuity in Discourse: An Introduction,” *Topic Continuity in Discourse: A Quantitative Cross-Language Survey*, ed. by Talmy Givón, 1–42, John Benjamins, Amsterdam.
- Givón, Talmy (1989) *Mind, Code and Context: Essays in Pragmatics*, Lawrence Erlbaum, Hillsdale, NJ.
- Greenberg, Joseph H. (1963) “Some Universals in Language with Special Reference to the Order of Meaningful Elements,” *Universals of Language*, ed. by Joseph H. Greenberg, 73–113, MIT Press, Cambridge, MA.
- Haiman, John (1983) “Iconic and Economic Motivation,” *Language* 59, 781–819.
- Hale, Kenneth (1982) “Preliminary Remarks on Configurationality,” *NELS* 12, 86–96.
- Hale, Kenneth (1983) “Warlpiri and the Grammar of Nonconfigurational Languages,” *Natural Language and Linguistic Theory* 1, 5–47.

- Haspelmath, Martin (2006) "Parametric Versus Functional Explanations of Syntactic Universals," ms., Max Planck Institute.
- Haspelmath, Martin (2007) "Frequency vs. Iconicity in Explaining Grammatical Asymmetries," ms., Max Planck Institute.
- Hawkins, John A. (1983) *Word Order Universals*, Academic Press, New York.
- Hawkins, John A. (1990) "A Parsing Theory of Word Order Universals," *Linguistic Inquiry* 21, 223–261.
- Hawkins, John A. (1994) *A Performance Theory of Order and Constituency*, Cambridge University Press, Cambridge.
- Hawkins, John A. (2004) *Efficiency and Complexity in Grammars*, Oxford University Press, Oxford.
- Heine, Bernd and Ulrike Claudi (1986) *On the Rise of Grammatical Categories: Some Examples from Maa*, Dieter Reimer, Berlin.
- Heine, Bernd and Mechthild Reh (1984) *Grammaticalization and Reanalysis in African Languages*, Helmut Buske Verlag, Hamburg.
- Hopper, Paul (1987) "Emergent Grammar," *BLS* 13, 139–157.
- Hopper, Paul (1988) "Emergent Grammar and the Apriori Grammar Postulate," *Linguistics in Context: Connecting Observation and Understanding*, ed. by Deborah Tannen, 117–134, Ablex, Norwood, NJ.
- Hopper, Paul and Sandra A. Thompson (1980) "Transitivity in Grammar and Discourse," *Language* 56, 251–299.
- Hopper, Paul and Elizabeth C. Traugott (2003) *Grammaticalization*, 2nd ed., Cambridge University Press.
- Horie, Kaoru (1997) "Form-Meaning Interaction in Diachrony: A Case Study from Japanese," *English Linguistics* 14, 428–449.
- Horie, Kaoru (2000) "Core-Oblique Distinction and Nominalizer Choice in Japanese and Korean," *Studies in Language* 24, 77–102.
- Horn, Laurence R. (1989) *A Natural History of Negation*, University of Chicago Press, Chicago.
- Jaeggli, Osvaldo and Kenneth J. Safir (1989) "The Null Subject Parameter and Parametric Theory," *The Null Subject Parameter*, ed. by Osvaldo Jaeggli and Kenneth J. Safir, 1–44, Reidel, Dordrecht.
- Jespersen, Otto (1917) *Negation in English and Other Languages*, A.F. Høst, Copenhagen.
- Kim, Soowon (1990) *Chain Scope and Quantificational Structure*, Doctoral dissertation, Brandeis University.
- Kirby, Simon (1998) *Function, Selection and Innateness: The Emergence of Language Universals*, Oxford University Press, Oxford.
- Kitahara, Yasuo et al. (2000) *Nihon Kokugo Daijiten*, Dai 2 Han (Great Japanese Language Dictionary, 2nd ed.), Shogakukan, Tokyo.
- Kondo, Yasuhiro (2000) *Nihongo Kijutsu Bunpo-no Riron* (The Theory of Japanese Descriptive Grammar), Hituzi Syobo, Tokyo.
- Kuno, Susumu (1983) *Shin Nihon Bunpo Kenkyu* (A New Study of Japanese Grammar), Taishukan, Tokyo.

- Lakoff, George (1987) *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*, University of Chicago Press, Chicago.
- Langacker, Ronald W. (1974) "Movement Rules in Functional Perspective," *Language* 50, 630–664.
- Lightfoot, David W. (1991) *How to Set Parameters: Arguments from Language Change*, MIT Press, Cambridge, MA.
- Lightfoot, David W. (1999) *The Development of Language: Acquisition, Change, and Evolution*, Blackwell, Oxford.
- Moriya, Tetsuharu and Kaoru Horie (2004) "Periphrastic Modal Expressions in Korean and Japanese: A Cognitive and Socio-cultural Perspective," *Inquiries into Korean Linguistics I*, ed. by Sang-Oak Lee et al. 67–76, Thaeaksa, Seoul.
- Newmeyer, Frederick J. (1998) *Language Form and Language Function*, MIT Press, Cambridge, MA.
- Newmeyer, Frederick J. (2004) "Against a Parameter-Setting Approach to Typological Variation," *Linguistic Variation Yearbook* 4, 181–234.
- Newmeyer, Frederick J. (2006) "A Rejoinder to "On the Role of Parameters in Universal Grammar: A Reply to Newmeyer by Ian Roberts and Anders Holmberg,"" ms., University of Washington.
- Onoe, Keisuke (2001) *Bunpo to Imi* Vol. 1 (Grammar and Meaning, Vol. 1), Kurosio, Tokyo.
- Pollock, Jean-Yves (1989) "Verb Movement, Universal Grammar, and the Structure of IP," *Linguistic Inquiry* 20, 365–424.
- Rizzi, Luigi (1982) *Issues in Italian Syntax*, Foris, Dordrecht.
- Rizzi, Luigi (1997) "The Fine Structure of Left Periphery," *Elements of Grammar: Handbook of Generative Syntax*, ed. by Liliane Haegeman, 281–337, Kluwer, Dordrecht.
- Roberts, Ian (1993) *Verbs and Diachronic Syntax: A Comparative History of English and French*, Kluwer, Dordrecht.
- Roberts, Ian and Anders Holmberg (2005) "On the Role of Parameters in Universal Grammar: A Reply to Newmeyer," *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, ed. by Hans Broekhuis et al., 538–553, Mouton de Gruyter, Berlin.
- Rohrbacher, Bernhard (1994) *The Germanic VO Languages and the Full Paradigm: A Theory of V to I Raising*, Doctoral dissertation, University of Massachusetts, Amherst.
- Shibatani, Masayoshi and Theodora Bynon (1995) "Approaches to Language Typology: A Conspectus," *Approaches to Language Typology*, ed. by Masayoshi Shibatani and Theodora Bynon, 1–25, Clarendon Press, Oxford.
- Snyder, William. (2001) "On the Nature of Syntactic Variation: Evidence from Complex Predicates and Complex Word-Formation," *Language* 77, 324–342.
- Speas, Margaret J (1994) "Null Arguments in a Theory of Economy of Projection," *University of Massachusetts Occasional Papers* 17, 179–208.
- Sweetser, Eve (1990) *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*, Cambridge University Press, Cambridge.

- Traugott, Elizabeth C. and Richard B. Dasher (2002) *Regularity in Semantic Change*, Cambridge University Press, Cambridge.
- Traugott, Elizabeth C. and Ekkehard König (1991) "The Semantics-Pragmatics of Grammaticalization Revisited," *Approaches to Grammaticalization* Vol. 1, ed. by Elizabeth C. Traugott and Bernd Heine, 189–218, John Benjamins, Amsterdam.
- Vennemann, Theo (1973) "Explanation in Syntax," *Syntax and Semantics* 2, ed. by John P. Kimball, 1–50, Seminar Press, New York.
- Weinreich, Uriel, William Labov and Marvin I. Herzerg (1968) "Empirical Foundations for a Theory of Language Change," *Directions for Historical Linguistics*, ed. By Winfred P. Lehmann and Yakov Malkiel, 95–188, University of Texas Press, Austin.

Faculty of Education
Kanazawa University
Kakuma, Kanazawa-shi
Ishikawa 920–1192
e-mail: moriya33@kenroku.kanazawa-u.ac.jp