

# A Biolinguistic Interpretation of Parameters and Language Acquisition

**Tiao-yuan Mao**

Massachusetts Institute of Technology, USA

## *Abstract*

The Principles-and-Parameters framework, the essential construct of the Generative Grammar, is endowed with new interpretation in the current Minimalist Program of biolinguistic paradigm, about which scholars have different views. From the illustration of the relation between the logical problem of language acquisition and parameters to the examination of different proposals for parameter-setting, the theoretical and empirical consideration of the status of parameters turns out to be clear. Meanwhile, along with the re-pondering of Principles in the current syntax theory, an integrative project for parameter-setting is put forward tentatively in terms of language acquisition, with reference to the theoretical achievements and tendencies in the current biolinguistic framework.

*Keywords: biolinguistics, language acquisition, feature, parameter, double-stage proposal*

## **1. Introduction**

Parameter setting is an essential step for children to acquire their mother tongues, or for children and adults to acquire the second language (L2), the third language (L3), etc. Different proposals probe into the biological growth or maturation of children's first language (Radford, 1990; Wexler, 1991, a.o.), or the acquisition of L2, L3, Ln by adults (White, 2003; Slabakova, 2016, a.o.). The proposals concerned are closely related to the definition of parameters in the theoretical linguistic frameworks.

In the last few decades or so, the exploration of biological properties of language has shifted from the concentration of ontogenetic language acquisition to serious consideration of phylogenetic language acquisition. The aim of a theory of language

is to explain not only how children can biologically acquire grammars, but also how our species could evolutionarily acquire human languages (Stroik & Putnam, 2013). The tradition, devoted to the exploration of biological properties of human language, has originated “from Chomsky’s (1959) review of Skinner’s *Verbal Behavior* and his publication of *The Aspects of the Theory of Syntax* in 1965” (Boeckx, 2006, pp. 16-17), especially, “the publication of the *Biological Foundations of Languages* (Lenneberg, 1967) founded the modern study of biology of language” (Berwick & Chomsky, 2016, p. 96). Following this move, generativists have committed themselves to examine how the language faculty evolved, such as Lewontin (1998), Jenkins (2000), Hauser et al. (2002, 2014), Lowenthal and Lefebvre (2014), Berwick and Chomsky (2016), a.o.

That the exploration of biological properties of human languages has turned into the focus results from “the language faculty as an organic or a cognitive system” (Chomsky, 2005, 2010a, 2010b, 2016; Friederici et al., 2017), mainly probing into “the shared properties between human language and other organisms by adopting the naturalistic methodology” (McGilvray, 2013, pp. 28-30). Moreover, biolinguistics tries to explore language evolution and acquisition outside language, and examine or define the scope of Universal Grammar (UG) beyond UG. This proposal renews the Principles and Parameters framework that commits to dissolve the tension between descriptive and explanative adequacy (Chomsky, 2005, p. 8; 2007a, p. 2), and endows the Principles and Parameters framework with a new mission, which, of course, excites the hot debate over the framework in the theoretical linguistic circle.

Such being the case, it is obligatory to examine the process and reasons for the transmutation of parameters, together with the explanation of reduction of principles when necessary, trying to elucidate the issues brought about by the revision of connotations of principles and parameters in the current Minimalist Program of the biolinguistic paradigm. This would be conducive to the formation of a tentative and integrative project for the parameter-setting in terms of language acquisition and diversity.

The paper is organized as follows: the second part introduces the relation between the parameter-setting and theoretical linguistic theories, such as the logic problem of language acquisition; the third part touches upon the characteristics and classification of parameters; the fourth part analyzes the different proposals about how to set the parameters, and evaluates the merits and demerits of current proposals of the parameter-setting; and the fifth part brings about my tentative proposal.

## 2. The Logical Problem of Language Acquisition and Parameters

As we know, young children often hear less sufficient messages while acquiring their mother tongues or other languages simultaneously. More specifically, what they hear contains performance errors, such as slips of tongue, false starts, etc. However, it is true for children with immature cognitive capacity and poor linguistic inputs to succeed in the acquisition of L1, or being a competent bilingual. How it is possible for children to

acquire their languages remarkably quickly in such situation has posed a puzzle for a long time. Besides, adults also show sophisticated subconscious knowledge about their languages that they cannot be taught (cf. White, 1985, 2003). In other words, the speakers possess subtle knowledge (competence) that they cannot articulate. For example,

- (1) a. \*Is this desk which \_\_\_ in the classroom is yellow?  
b. Is this desk which is in the classroom \_\_\_ yellow?

Children and adults could both make a correct judgment of (1a) as false, even though they are not taught explicitly that (1a) commits the violation of structure-dependent rule.

Chomsky (1980, p. 34) dubbed it as “the poverty of stimulus” (POS), or as “Plato’s Problem” (Chomsky, 1986), which says: how can we know so much given the limited experience? Put it in another way, “people attain knowledge of the structure of their language for which no evidence is available in the data to which they are exposed as children” (Hornstein & Lightfoot, 1981).

The above argumentation reveals a gap between what acquirers master about languages and the available information of languages. It is the gap that justifies the logical problem of language acquisition, and the innateness of linguistic knowledge—UG, with reference to Descartes’ discovery of people’s innate ability to understand novel concepts even without exposure to them. Based on this hypothesis, we can understand why young children fully acquire the basic linguistic knowledge at an early age, and adults are able to make subconscious grammatical judgment, even if both of them are not exposed to all structures of languages.

As the earliest UG theory assumes, “UG provides a format for rule systems and an evaluation metric that assigned a ‘value’ to each generative procedure of the proper format” (Chomsky, 1995, p. 24). Taking advantage of this endowment, children acquiring their mother tongue make use of the internal “abductive mechanism” (cf. Mao & Zeng, 2011), to restrict the hypotheses that are relevant to the rules. More specifically, the grammatical rules activated in the brain, perhaps only one, are consistent with the kinds of rules available on the primary linguistic data (PLD). The language acquisition problem is hence solved. However, the tension between descriptive adequacy and explanatory adequacy would be intensified, for to reach a detailed description of language rules in terms of language acquisition means multiplication of the possible rules. Gradually, the contents of UG turn out to be richer, which goes against the rapidity and consensus of children’s acquisition of their mother tongues.

To remove the theoretical redundancy, the practical method is to abstract the important principles from the rules, and then the operations of grammatical rules fall under the guideline of UG Principles. Thus, the system of grammatical rules is replaced by UG Principles, and the theoretical assumption of grammar becomes concise with the advent of the “Principles & Parameters framework” (Chomsky, 1981). Under this framework, UG Principles are defined as a kind of invariant, implicit knowledge shared by all human

beings, connecting with the “switchbox” of Parameters. For example, English is a head-initial language, while Japanese is a head-final one.

- (2) a. Taro bought the book.  
 b. Taroo-ga    ano hon-o    kat-ta.  
     Taro-NOM   that book-ACC buy-PAST  
     ‘Taro bought that book.’

In (2), the position of heads is determined by the on/ off of head parameter, viz., [+head initial] for English in (2a) and [-head initial] for Japanese in (2b). Thus, whether a language is head-initial or head-final is regarded as a parameter that is either on or off. The attractiveness of this type of parameter turns out to be that the number of grammatical consequences or “parametric clustering” attached to a single on/off switch, made it plausible that bunches of switches at most, say ten, twenty, or thirty, would suffice to capture the grammars children attained.

Based on the implicit knowledge and switchboard metaphor, children are able to acquire their languages quickly, and make very accurate judgment of certain linguistic phenomena, such as anaphora and the parasitic gap. At the same time, children’s acquisition of the first language is to “fix a language from the limited variety available in principle” (Berwick & Chomsky, 2011, p. 28). For adults, it could be possible to fix the parameter on the basis of L1 and UG. As a result, the first specific proposal comes into being. It probes into how acquirers, both L1 and L<sub>n</sub>, attain their individual languages via the parameter-setting. Along this line, several different proposals of the parameter-setting occur.

### 3. Characteristics of Parameters

As evidenced by the above discussion, the parameter-setting is thought to be an optimal solution to demonstrate how children and adults acquire their languages. Although the theoretical linguistic theories have been updated several times, the purpose of parameter-setting has remained the same in terms of language acquisition, with several salient points worthy of clarification.

#### 3.1 Principles or position-oriented

In the Principle and Parameter framework, the firstly-proposed parameters are “grammatical parameters”, viz., “parameters within the statements of the general principles that shape natural language syntax”, but “not localized in the lexicon per se” (Baker, 2008, p. 353). In other words, shortly after the first introduction of parameter into the theoretical linguistic field, the parameters are “macro-parameters” rather than “lexical parameters” or “micro-parameters” which locates the syntactic variations within the lexicon as Borer (1984) demonstrated.

The reason why the grammatical parameters pop out at the early period of the Government and Binding theory would attribute to the analysis of Pro-drop languages by Chomsky (1981) and Rizzi (1982). For instance,

- (3) a.  $\emptyset$  vendrá. (Spanish)  
 $\emptyset$  will-come.3sg  
 b.\*  $\emptyset$  will come. (English)

In Spanish (3a), the subject is optional, while in English (3b) it is not. Thus, the parametric variation in the absence or presence of subject captures the differences of Spanish and English. The Null Subject Parameter (NSP) is then proposed. Baker (2008, p. 351) called NSP a “medioparameter” because “the NSP is medium-sized in its formulation and its effects.” More specifically, the former property means that it is as a general rule of grammar, not as the features of a lexical item or variation in a core principle of grammar; the latter refers to the fact that the NSP is to account for a cluster of properties, including subject omission, inversion, and the absence of complementizer-trace effects. Therefore, the parametric choice might have an impact on the grammar of a language, triggering a clustering of syntactic properties, just as Chomsky (1981, p. 6) commented: “In a tightly integrated theory with fairly rich internal structure, change in a single parameter may have complex effects, with proliferating consequences in various parts of the grammar.” The basic idea lurking behind this statement could be that “the language faculty favors the alignment of parametric values of related parameters” (Roberts, 2016, p. 83).

Meanwhile, the locus of parametric variations could also be concerned with a particular module of UG, such as Theta Theory, Case Theory, etc. or related to the choice of grammatical levels, such as D-Structure, S-Structure, Logical Form on which a grammatical operation applied (Eguren et al., 2016). The difference in English and Chinese Wh-movement is a good example to demonstrate the parametric variation in the Logical Form; namely, Wh-movement in English takes place overtly, while Chinese Wh-movement occurs covertly in the Logical Form.

These two types of parametric variations agree with Chomsky’s (1981, p. 6) explanation: “Each of the systems of (1) (subcomponents of rule system of grammar) and (2) (subsystems of principles) is based on principles with certain possibilities of parametric variation.” With all of the above in mind, it is not difficult for us to have a general idea of the characteristics of parameters in the Principles and Parameters framework. That is to say, the early parameters are grammatical parameters, closely connected with the Principles, operations and different grammatical levels. More importantly, the parameters or parametric values can directly date back to UG, and the parameters imply sets of formal properties that clustered together.

For L1 or L<sub>n</sub> acquirers, to attain a language means to set or determinate the loci where parametric variations happen, switching on or off certain parameters, with reference to

inputs available in the learning environment. No matter how complex the target languages are or how poorly the inputs turn out to be, the language acquirers would achieve complex knowledge of target languages beyond the inputs and reach the end state.

Even though the early parametric proposal is conducive to the shape of grammars, for example, it helps to decide the head-directionality, or realize the linearization of syntactic entities via fixing their order, etc., yet as researchers pointed out, the clustering properties of parameters have not borne out completely.

Moreover, with more research exploring the potential Principles and Parameters, more and more Principles are created from different languages. And UG faces the embarrassing situation again, that is, too rich in contents. Originally, the formation of UG Principles is to eliminate the theoretical redundancy that the system of grammatical rules incurs; yet they are facing the same problems now. In this case, the simplification of UG Principles turns out to be an urgent task during the implementation of the biolinguistic paradigm. Accordingly, it is inevitable to redesign the proposal for the parameter-setting.

Actually, some potential ideas did sprout out to address the language learnability in a different way in the early Principles-and-Parameters framework. For example, to solve the theory internal problem about the clustering properties of parameters, we might say that the clusters of properties have been conceived as sets of different grammatical properties that are all meant to derive from one abstract property (cf. Eguren, 2016). Or, as Borer (1984, p. 29) argued:

...by reiterating the conceptual advantage of a system that reduces all interlanguage variation to the properties of the inflectional system (in the lexicon-the author notes). The inventory of inflectional rules and of grammatical formatives in any given language is idiosyncratic and learned on the basis of input data. If all interlanguage variation is attributable to that system, the burden of learning is placed exactly on that component of grammar for which there is strong evidence of learning: the vocabulary and its idiosyncratic properties. We no longer have to assume that the data to which the child is exposed bear directly on universal principles, nor do we have to assume that the child actively selects between competing grammatical systems.

Although the proposal was not taken into consideration seriously at that time, it marked a potential possibility for the subsequent change in the design of parameters.

### **3.2 Functional feature or third factors-oriented**

Jacob (1977, p. 1165) argues that “divergence and specialization of mammals resulted from mutations altering regulatory circuits rather than chemical structures.” That is to say, the minor changes in regulatory mechanism can bring about huge differences in phenotype. Meanwhile, as Sherman (2007, p. 1873) points out: “the Metazoan phyla, all having similar genomes, are nonetheless so distinct because they utilize specific combinations of developmental programs.” In this sense, it is the regulatory mechanism

and organizational principle that lead to the phenotypic differences during the evolution of organisms. It is the very similarities between biological and linguistic evolution that provide the theoretical support for Chomsky's reformulation of the Principles-and-Parameters framework. The minor changes in parameters might lead to "the diversity of language through the interaction between the invariant principles and the choice of parameters" (Chomsky, 2010b, p. 49).

More importantly, the neurological research of brain shows that "the evolution of human language, at least in part, could be an automatic but adaptive consequence of increased absolute brain size" (Striedter, 2006, p. 10). This indicates that some slight or minor mutation could result in the rewiring of brain (Berwick & Chomsky, 2011, p. 27). Recursive Merge, a core linguistic principle assumed in the Minimalist Program, takes shape in a small group of hominids in East Africa (cf. Hauser et al., 2002; Chomsky, 2010a), and endows individuals with the internal generative system for the construction of thoughts (language of thought). Then, it is possible that applying the recursive Merge to the conceptual atoms—lexical items in the lexicon, will generate discrete, endless hierarchical structures. Merge becomes a unique recursive principle (Freidin, 2014, p. 144). Presumably, the occurrence of Merge would have created the generative procedure about 75,000 years ago (Hauser et al., 2002; Chomsky, 2016), which takes human concepts (lexical items) as the computational atoms.

Chomsky (2007a, p. 7) assumes that if Merge comes from other system, namely, not unique for language, then there must be a genetic instruction to use Merge to form structured linguistic expressions, satisfying the interface conditions. In other words, the real intention to define Merge as UG Principle is that there exists an innate neural mechanism, which is formed during evolution. This idea brings Chomsky's early view in our sight again, that is, "acquisition of language involves not just a few years of experience and millions of years of evolution, but also principles of neural organization that may be even more deeply grounded in physical law" (Chomsky, 1965, p. 59).

Furthermore, this assumption implies that Merge is possibly the language properties decided by genes. In this sense, it is not difficult to understand "why the recursive Merge must be a part of the genetic endowment and the initial state of the language faculty" (Freidin, 2014, p. 145). At the same time, that Chomsky regards Merge as the (at least) unique UG Principle accounts for a forceful response to Empiricism's questioning, for Empiricism criticizes the basic assumption of nativism, that is, arguing "against the existence of internal specific linguistic knowledge" (Carr, 2006, p. 333). Currently, Chomsky takes Merge, a kind of physical law shared by all species, as the UG Principle, which positively demonstrates that the criticism from Empiricism is not on the right ground; instead, it exactly explains the feasibility for assuming the existence of the internal linguistic knowledge, no matter whether the internal UG constituents could only be Merge after the drastic simplification of UG contents, and whether it is possible to falsify the assumption with the development of scientific and linguistic research. Also, this view is consistent with Chomsky's latest comment: "UG is not to be confused with

descriptive generalizations about language such as Joseph Greenberg's universals" (Chomsky, 2013, p. 35).

Therefore, it could be reasonable to assume that there are non-trivial analogies between the evolutionary developmental biology (Evo-Devo) and language. In this sense, it is commonly and plausibly assumed that "the emergence of language was a core element in this sudden and dramatic transformation" (Berwick & Chomsky, 2011, p. 20). Furthermore, language functions as one component of the human capacity that could be studied seriously. As a result, the research that is purely linguistic actually falls within the scope of biolinguistics. Following this tradition, the research of language, which is regarded as the mental components of humans (language faculty), is a kind of biological research of language. This practice brings about the issues that correspond to form and function, ontogeny and phylogeny in biology and other related fields, such as "what language knowledge is, how the children language develops and how human language evolves" (Jenkins, 2013, p. 4). To integrate the linguistic issues with the biolinguistic research seriously, some conditions should be met as McGilvray (2013, p. 44) argued:

...the issue of acquisition could be seen as sufficiently well addressed began with the introduction of the Principles and Parameters model, and with the naturalistic scientists' effort to simplify cashed out in a minimalist program that was also now capable of addressing "third factor" contributions to evolution and ontogenesis...

If these conditions are satisfied, the possibility of accommodation to biology could be realized. Accordingly, it is unavoidable to revise the principles and parameters in the current biolinguistic paradigm.

Taking the biological and neurological findings as a guide, the Minimalist Program is firmly established as a research program, which aims to implement the Principles-and-Parameters framework, in which UG is not rich and highly structured any more. The Principles in the former theoretical framework have undergone remarkable reduction. The core principle could only include Merge (Chomsky, 2012, p. 13). The reason for the reformulation of UG principle is to verify "UG is a simple and elegant theory, with fundamental principles that have an intuitive character and broad generality" (Chomsky, 1995, p. 29). Hence, it is not difficult for us to understand "why the Principles-and-Parameters framework removes the major conceptual barrier to the study of evolution of language" (Chomsky, 2007b, p. 19), and "the serious conceptual barriers to the implementation of principled explanation" (Chomsky, 2008, p. 2).

The reformulation of the Principles has huge influence on the design of parameters. For instance, the new move has catalyzed the major shift from probing into classical parameters to looking for microparameters, since it is not practical to narrow down the parametric variation to the core Principle-Merge. This tendency, according to Baker (2008, p. 351), follows "the methodology and reasoning championed by Richard Kayne

(2005), and uses the same reasoning as classical parameter theory, but applies it at a higher level of magnification". The local differences in the grammars of closely related languages, including dialects, come into focus. Along this line, macro-parametric differences among languages are mainly thought to be the result of the aggregation of the microparameters (with Baker's insistency on macro-parameters as a complementary one). To some degree, this approach has proven to be "very fertile, creating a steady stream of results" (Baker 2008, p. 351).

Chomsky (2005, p. 6) pointed out: "if the language faculty shares some general properties with other biological systems, then it is possible to seek out the basic factors that influence the ontogenetic acquisition." These factors include "genetic endowment, linguistic experience and language-independent principles (principles of third factors—the author notes)" (Chomsky, 2005, p. 6). The cross-linguistic diversities are then not defined by the UG Principle, but attributed to the interaction of several factors. Put another way, "the notion is a taxonomic artifact emerging from (i) the formal features of certain heads are underspecified by UG, (ii) the Primary Linguistic Data, and (iii) general markedness conditions" (Roberts, 2016, p. 179). That is to say, the language variation could result from the interaction among the functional features, linguistic data and computational efficiency of the third factor principles, such as c-command, intervention and locality, among others.

The idea that parameters emerge from the interaction is actually an updated version of Borer's (1984) early idea of language learnability. As discussed in the previous section, Borer proposed the parameter-setting as the acquisition of inflectional lexical properties. The simple reason is that lexical learning is the first and indispensable step for the language acquisition. Provided that an acquirer obtains the inflectional properties of lexical items, the morpho-syntactic mapping will become successful, and to acquire a language comes to natural end. Biberauer (2008, p. 25) clarifies the idea as follows:

'learning vocabulary' does not just entail learning the idiosyncratic component of language... presumably at the same time, also what the phonotactic and morphological properties of the language in question are (i.e., two types of system-defining knowledge); it also entails acquiring the parametric profile of the language, fixing the unvalued parameters of  $S_0$ .

Given the properties of lexical items in the lexicon are so important for linguistic computation and language acquisition, Chomsky (1995) incorporated Borer's early idea into the Minimalist Program. Lexicon becomes a core assumption in the minimalist theorizing. The "Borer-Chomsky Conjecture" (Baker, 2008, p. 353) comes into being as stated in (4), attracting a lot of attention and popularizing in linguistic sub-fields.

(4) The Borer-Chomsky Conjecture (BCC)

All parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon.

From this conjecture, we can see that UG assumes an inventory of functional features in addition to Merge, and the loci of parametric variations are narrowed down to the formal features on the functional heads. In other words, the specific properties of the language-specific rules or principles might be abstracted as the atomic syntax-driven features, stored in the UG inventory of functional features. Obviously, the parameters implemented by the formal features are not a part of UG any more. The classical parameters, “a kind of parameterized principle” (Boeckx, 2016, p. 75), give way to the micro-parameters, which are confined to the formal features that are uninterpretable at the interface. UG thus turns out to be concise and elegant. In this case, the evolution of the language faculty can be addressed properly, since the language-specific principles are reduced to language itself, and the functional features, as micro-components which manifest the idiosyncrasies of language, are put into the UG inventory of features, functioning as the essential syntax-driven constructs.

However, the insight that parameters are reduced to the functional features on the lexical items (lexical features) poses a problem for the language acquisition. It is because UG does not make clear how a particular language makes use of which part of them, and the only UG Principle, Merge, cannot provide any clues for how to select them. Then, it might be a myth for how children and adults acquire certain languages. For this, Chomsky (1998, p. 13) extended a clear explanation:

On these (fairly conventional) assumptions, acquiring a language involves at least selection of the features [F] (A subset of F made available by UG), construction of lexical items **Lex**, and refinement of  $C_{HL}$  (the computational procedure for human language) in one of the possible ways-parameter setting.

This statement goes hand-in-hand with the “Borer-Chomsky Conjecture”, for the key points for language acquisition is attributed to the selection of formal features. Specifically, after the lexicon for the individual language is established by one-time selection of [F] from F into the individual lexicon,  $C_{HL}$  can get access to [F] to generate expressions. Therefore, it is very crucial for acquirer to obtain the formal properties of lexical items.

## **4. Different Proposals of Parameter-Setting and the Evaluation**

Given that the functional features have played an important role in the language acquisition and the realization of language diversity, the theoretical linguistic circle has spared no efforts to initiate several trials to explain “Plato’s Problem”.

### **4.1 Core-syntax proposal for parameter**

In the classic Principles-and-Parameters framework, the Parameters are mostly grammatical parameters, attributed to UG and connected with the Principles as “the

switchboard metaphor” suggested, or set in different grammatical levels and operations, with clustering properties or cascade effect. Clearly, the parametric variations locate the operations within the core syntax. Roberts (2012, 2016) also proceeded in this line. A proper example is Hale’s (1983) Configurationality Parameter (CP), as stated as follows:

(5) The Configurationality Parameter

- a. In configurational languages, the Projection Principle holds of the pair (LS-lexical structure, PS-phrase structure).
- b. In non-configurational languages, the Projection Principle holds of LS alone.

On the basis of (5), CP can distinguish English from Warlpiri, a configurational and a non-configurational language respectively. According to Hale (1983), CP is established in terms of Chomsky’s (1981, p. 29) “Projection Principle: the subcategorization properties of lexical items are projected at each syntactic level, D-structure, S-structure and LF”. Therefore, in the configurational languages, the Projection Principle sets up a structurally isomorphic connection between LS arguments and PS nominal expressions, while in the non-configurational ones, the CP will not propose any relation between LS and PS.

Other proposals set the parameters within the core syntax including Saito and Fukui’s (1998) view on parametric variation caused by Merge, Baker and Collins’s (2006) proposal for parameterization in terms of minimal link condition, and Chierchia’s (1998) parametric project in the semantic component like Nominal Mapping Parameter, among others. It seems that the limitation of parametric variation in the core syntax could neatly cover empirical phenomena, and file a good solution to language acquisition.

However, with the further development of the theoretical linguistic inquiry, the proposal might turn against the primary target of the theoretical linguistic inquiry.

Boeckx (2011, p. 209) argued that UG does not include parameters in nature, and parameters are a conceptual tool. In other words, parameters are “logically indispensable anywhere, including the core syntax, yet do not make any sense biologically” (Boeckx, 2014, p. 158). The conceptual or theoretical reason or the argumentation, as we know, might date back to Kayne’s analysis of macro-parameters: “some of properties of the language faculty are too deeply built in to be possible loci of variation” (quoted from Boeckx, 2016, p. 71). It is because “limiting syntactic parameters to features of functional heads is also intended to exclude the possibility that there could be a syntactic parameter that is a feature of no element of the lexicon at all” (Kayne, 2005). Put it simply, if we completely follow a bottom-up derivational syntax, it is not reasonable for us to get a (macro) parameter that has no root in the lexicon. Possibly, in Kayne’s (2005) view, macro-parameters might be decomposed into arrays of microparameters.

At the empirical level, as Boeckx (2016, p. 71) summarized, the work under the classic Principles-and-Parameters framework reveals that “macro-parameters leak in

Sapir's words". The early estimation that limited number of parameters, plus the cascade effect, could capture the grammars children acquire, cannot cover the empirical linguistic phenomena. It is because the clustering effect does not automatically follow in some situations, as Newmeyer (2004, 2005) argued. Instead, to handle the language acquisition problem completely, we need more independent or specific parameters. It seems unavoidable that the additive effects of some number of micro-parameters could break down the macro-parametric differences.

## 4.2 Pre-syntax proposals for parameter

While reviewing the significance of the (classic or new) Principles-and-Parameters framework, Chomsky (2007b, p. 19) commented "the Principles-and-Parameters framework denotes a kind of shift of perspective, enabling us to address the issues of explanatory adequacy and transcending explanatory adequacy". Obviously, what he said mirrors the process of parameter transmutation.

In the Minimalist Program, parameters are confined to the formal features, viz., attributing parameters to the formal features of functional heads in the lexicon. The reason why the concept of parameter has undergone the innovation is due to the reduction of UG Principles to Merge, and the shifting from the top-down to the bottom-up model. The classic Parameters gradually fade out from the center, and "Borer-Chomsky Conjecture" comes into focus. Specifically, the parameter-setting means to assemble the functional features, or clustering of features (cf. Rizzi, 2009), into the functional lexical items, and the syntactic computations take the lexical items as the atomic constituents, realizing the morphological pattern of individual language via the simplest syntactic operations. Accordingly, it is very essential for the parameter-setting "to assemble features into lexical items" (Chomsky, 2008, p. 2). The complexity and diversity of language are confined to the lexicon. Thus, the lexicon becomes the locus of parametric variation (Chomsky, 2007b, p. 25).

With the shift from macro- to micro-parameters, many syntacticians have practiced this idea, mainly "for more empirical than explanatory reasons" (Boeckx, 2016, p. 72). Such as Bobaljik and Thráinsson (1998), Fortuny (2008), a.o., which favor the "Bundling Parameter":

### (6) Bundling Parameters

Given two lexical features  $f_1$  and  $f_2$ , drawn from a universal repertoire (UG), does a given language  $L$  project  $f_1$  and  $f_2$  as a bundle or do  $f_1$  and  $f_2$ , function as distinct heads in syntax?

(Boeckx, 2016, p. 72)

However, Boeckx (2014, 2016) questioned the nature of the lexical bundling operation. As Boeckx (2016, p. 72) argued: "bundling is nothing but Merge operating in the pre-syntax lexicon." That is to say, "like Merge, 'bundle' operates on lexical features;

it groups them into a syntactically combinable (mergeable) set; ‘bundle’ builds syntactic structures (feature trees)” (Boeckx, 2016, p. 72). In other words, “bundle” equals Merge. Bundling parameters are syntactic parameters, and fall within the characterization of Snyder’s “constructive parameter” (Boeckx, 2016, p. 94). It seems that the bundling parameters function as a structural blocks and the formation of bundling parameters relies on bundling operations.

Eguren et al. (2016) thought that Chinese Analyticity Parameter (Huang, 2010) is a case of bundling parameter. According to Huang (2010), Modern Chinese differs from English in “exhibiting high analyticity over a full range of lexical and functional categories”. Namely, the analyticity of Modern Chinese lexical items scatters at the levels of lexical, functional and argument structures. However, different properties seem to cluster together in Modern Chinese; the properties are all manifestations of a single macro-parameter (Eguren et al., 2016, p. 17).

Meanwhile, macro-parameters presuppose an active lexicon, allowing the application of derivational operations, which is different from the traditional assumption of lexicon, a list of lexical entries. Then, the lexical parameters turn into syntactic parameters (as Baker, 2008 argued), nullifying the alleged explanatory superiority of Borer’s insight of locating all parameters in the lexicon, suffering from the same critiques (Boeckx, 2016, p. 72).

Therefore, based on Chomsky’s (2001) Uniformity Principle, Boeckx (2016) suggested a Strong Uniformity Thesis, i.e., Principles of narrow syntax are not subject to parameterization; nor are they affected by lexical parameters. This view shuts down the possibility of parameterization in the core syntax or lexicon via removing the pre-syntactic lexical parameter. Rather, the locus of variation is confined to the margins of the narrow syntax as stated in (7), agreeing with Berwick and Chomsky’s (2011, p. 37) idea: “parameterization and diversity would be mostly, possibly entirely, restricted to externalization.”

(7) Locus of variation

All “parameters” reduce to the realization of options.

This hypothesis sounds intriguing, and also agrees with the latest thought in the current biolinguistic research. We will extend our critique on the hypothesis later on.

### 4.3 Post-syntax proposals for parameter

As we discussed in the previous section, Boeckx strongly suggested the parametric variation to be confined to externalization at PF level. It could be attributed to Chomsky’s (2005, 2007b, a.o.) emphasis on the exploration of the relation between the language diversity and the third factor principles.

The above theoretical considerations trigger the post-syntactic proposal for the parameter-setting. If UG is not highly structured and rich, “diversity or variety of

language would be reduced to ancillary mapping involved in the externalization” (Chomsky, 2007b, p. 25). It means that a third way is added for the parameter-setting, for the variety of language phenotype would also directly attribute to the restriction of the third-factor principles in terms of computation efficiency. The externalization of syntactic products does not involve the genetic evolution, rather, “a process of applying the current cognitive ability to solve the questions” (Chomsky, 2010a, p. 121). In detail, the externalization of parameters would depend on “the currently evolved core computational system and the Sensory-Motor (SM) system existing for millions of years” (Chomsky, 2010b, p. 60). The former supplies the structural expressions and the latter provides the place where linear externalization is realized according to the morphological requirement of individual languages (cf. Chomsky, 2007b). In the process of externalization, the morphological and phonological computations function as the practical operators, which transform the syntactic entities that are interpretable in the SM system. These entities, regulated by such constraints as filters, optimal computation, etc., are subject to further readjustments according to the morphological requirement of individual languages. The parameters of individual languages are then fixed. In this sense, the parametric variation among languages will be measured by the external factors.

## 5. Double-Stage Proposal

If all the proposals sound reasonable to some degree, the next question is how we can integrate different proposals of parametric variation. In other words, it is important to reconcile the relation among functional features, Merge, the language-specific principles and the third factor principles.

As discussed above, in addition to invalidating the parameterized principles, Boeckx also doubted the feasibility of the bundling parameter in the lexicon, viz., questioning about the pre-syntax parameters. This argumentation might be based on the progress that the Minimalists have realized the possibility in the separation of generality from universality. Most importantly, the narrow syntax is at the heart of the language faculty, and not subject to parameterization, nor is it affected by lexical parameter (Boeckx, 2011, p. 209). It might be true because it is unanimously recognized in the current biolinguistic research that UG would not include parameters in nature, and parameter is a conceptual tool. The bundling lexical parameters were basically motivated by the empirical requirements, and it changes lexical parameter into syntactic parameter, against “Borer-Chomsky Conjecture”.

However, we would not think all of the demerits of lexical parameters mentioned above might be true, though some of critiques seem to make sense from the perspective of current biolinguistic research.

Although the functional features on the functional heads bundled together function as syntactic parameters, they are the essential constructs for syntactic operations. More

features bundled or assembled into lexical items could facilitate the economic syntactic computations, which are in accord with the intuitive empirical evidence, viz., a converged syntactic computation presupposes the conceptual atoms - lexical items with all necessary features assembled, with the features giving corresponding instructions to the syntactic operations (cf. Mao & Meng, 2016, a. o.). Meanwhile, it is impossible to retrieve the necessary features again in the lexicon during the syntactic computation. In this sense, even though the final patterns of parameters fixed on the basis of the bundling features in lexical items look like the same as those of the classic syntactic parameters, yet the theoretical considerations are totally different. The former adopts the “bottom-up” perspective, taking the parameter-setting as the selection of features from UG feature inventory and assembling them into lexical items; the latter abides by the classical “top-down” model and emphasizes the internal properties of principles, holding that the parameter-setting is the deductive process of the parameterization of principles.

Currently, the bilingualistic research gains momentum, and the feature theory in the “bottom-up” model remains the effective tool for generativists to interpret linguistic phenomena, especially in the fields of syntax and language acquisition. This might be attributed to the fact that the pre-syntax lexical parameters are not against the implementation of bilingualistic spirit.

Then, the functional features bundled into the heads would not entail the macro syntactic parameters, even though they might subsume hierarchical feature trees (Harley & Ritter, 2002). It is because a single feature can realize the syntactic effect, as the Cartographic approach presents a fine-grained articulation of syntactic structures based on the basic tenet: one feature, one head (Mao & Meng, 2016, a.o.). Even though some of features bundled together motivate the syntactic operations, it is not clear whether the bundling operation could be regarded as a kind of syntactic operation, or as the morphophonological computation in the lexicon as Kiparsky (1982) argued. The bundling features mean to get the necessary atomic items ready for the syntactic operations in the narrow syntax, or the assembly of the features to the lexical items just prepares for the subsequent syntactic computation. The assembly or bundling processing is closely related with lexical parameters, rather than syntactic parameters, since to assemble one or more functional features into the lexical items is equal to set the lexical micro-parameters. It might be the reason why Kayne (2005) and Roberts and Holmberg (2009) propose the macro-parameter as the aggregate of the micro-parameter. In this sense, to assemble or bundle features for the lexical parameter is more flexible than those of purely syntactic parameter, since the former confines the loci of variation to the lexicon, rather than at all levels of syntactic computations.

More essentially, the empirical motivation for the parameter-setting can date back to the first manifestation of lexical parameter by Borer (1984). Along this line, to acquire the inflectional properties seems quite sensible for language acquisition from an empirical perspective. Crucially, as Friederici et al. (2017, p. 714) pointed out that

“to demonstrate that experience alone without prior constraints provided by UG is insufficient for language acquisition needs ample empirical evidence from infant studies as well as theoretical modeling.” Further, L2 or Ln acquisition provides solid empirical evidence that the recognition and assembly of functional features into lexical items are the primary step to reach the final state of language acquisition (Mao, 2016). In this case, it seems that we cannot say that we do not need the lexical parameters because they are connected with the empirical research. Rather, we’d better admit that the lexical parameters could function as a kind of empirical evidence to demonstrate how language acquisition happens. This could be the real purpose to set up the Principles-and-Parameters framework.

As for Modern Chinese Analyticity Parameters as Huang (2010) delineated, for us, Huang actually assumed that it might be reasonable for Modern Chinese to follow what Kayne (2005) and Roberts & Holmberg (2009) argued: the macro-parameters as aggregates of the micro-parameters with correlating values. That is to say, Modern Chinese would not lack lexical syntax, against Boeckx’s (2016, p. 72) judgment: “‘highly analytic’ languages like Modern Chinese lack l(exical)-syntax.”

Currently, the diversity of language phenotypes would be attributable to the function of the third factor principles. It is because Chomsky attempts to absorb the viewpoint of externalization of parameters in the post-syntax stage. However, he does not mean to abandon the lexical parameter, namely, “attributing the parameters to the functional elements in the lexical items” (Chomsky, 2014, p. 12). On the other hand, to ignore totally the deductive consequences of parameters which are advocated in the “top-down” model is against the generally-accepted view, “integration of the theoretical perspectives of top-down and bottom-up” (Chomsky, 2007a, p. 26).

Proposed as it is, the recursive Merge needs the lexical items assembled with necessary features for the sake of syntactic computation on the one hand; on the other, the lexical items as the conceptual atoms feed the subsequent semantic interpretation or thinking in the Conceptual-Intentional (C-I) system. During the syntactic computation, the functional features assembled into the lexical items would send out the instructions, and drive the syntactic computation via projecting the corresponding heads according to the morphological requirements of individual languages. When certain syntactic computation is completed, the diversity of language phenotype would be expressed in the syntactic structures to some degree. Here, the parameters are still attributable to the functional elements in the lexicon; and the narrow syntax transfers the hierarchically-structured representations to the SM system for further phonological interpretation. However, it is possible that the syntactic structures that the narrow syntax transfers to the SM system would not meet the linear morphological requirements of target languages, because “Merge leaves elements unordered” (Chomsky, 2014, p. 7). In this case, further computation in the SM system is needed, for the adjustment of unordered elements in SM do not affect the mapping between the narrow syntax and the C-I system.

Continuously, during the externalization of the representations of syntactic computation into the SM system, the structural expressions would undergo further processing, fully representing the structural expressions into the linear structures which meet the morphological requirements of the individual languages in the post-syntactic phonological level. In other words, with the aid of general cognitive operations, the post-syntactic parameters are fixed, and the diversity of language phenotype is realized. In this sense, it seems to be an optimal proposal in the construction of a unified parameter-setting project through the integration of pre-syntactic micro lexical parameter and the post-syntactic phonetic externalization operations in the SM system.

The integrative proposal might earn us more theoretical advantages. Specially, on the one hand, based on UG inventory of features, the subset of features, which are supposed to express the diversity of language phenotype, are selected in the lexicon and assembled into the lexical items for the syntactic computation (also a kind of syntactic acquisition). This clearly demonstrates the deductive consequences of parameters in a top-down way; on the other, in the post-syntactic morphological and phonological level, the syntactic representations, in order to meet the linear morphological requirements of specific languages, would undergo further adjustment of word orders. This indicates that parameters are set on the basis of a bottom-up induction or generalization. If the analysis is on the right track, the double-stage project of parameter-setting will provide a more coherent and systematic parameter theory for the language acquisition. However, it is not very clear whether the integrative project would bring any theoretical redundancies. We will use the following examples to illustrate the parametric differences between English and Chinese Wh-movement:

- (8) a. What do you like?  
 b. 你 喜欢 什么?  
     ni xihuan shenme  
     you like what

(8) typically reflects the differences between Chinese and English Wh- movement. Specifically, Wh-word is fronted at the beginning of Wh-questions in English, while remaining in-situ in Chinese Wh-questions. In this case, how can a Chinese-English learner or an English-Chinese learner acquire Wh-movement in the target languages? Here, we will only focus on the former situation because it might be significant for the study of English learning in China. In our proposal, the acquisition of the cross-linguistic variation between English and Chinese Wh-movement can be attributed to the feature assembly in the pre-syntactic lexicon and the post-syntactic phonological processing in the SM system. Especially the former, because the functional features assembled onto the lexical items functioned as instructions for the subsequent syntactic computations. Let us explain the derivation or acquisition process in detail.

Firstly, the Chinese-English learners could be able to differentiate the various

interpretations of English and Chinese Wh-words (more details see Mao & Dai, 2015), then pick up in the lexicon the lexical items like  $\{\{\text{what, DO, you, like, } v^*\}, T, C\}$  (DO is treated as an auxiliary here), and assemble  $[uQ][uWh][EPP]$ , and  $[uPhi][uTense][EPP]$  into  $\{C\}$ ,  $[uPhi][EPP]$  into  $\{v^*\}$ . Mapping of  $[+Present]$  into  $\{DO\}$ . At the same time, setting the features  $[+Wh][+Q][+Singular]$  into  $\{\text{what}\}$  (putting aside Case here). When T merges with C, it inherits the  $[uPhi]$  from C. When the feature assembly is ready, the derivation or (invisible) acquisition process could proceed with the instructions from the formal features assembled on the lexical items as follows:

- (9) a.  $[_{v^*P} \text{you } v^*_{[uPhi][EPP]} \text{-like } [_{VP} \text{like what }_{[+Q][+Wh][+Singular]}]]$   
 b.  $[_{v^*P} \text{What}_{i[+Q][+Wh][+Singular]} \text{you } v^*_{[uPhi][EPP]} \text{-like}[_{VP} \text{like what}_{i[+Q][+Wh][+Singular]}]]$   
 c.  $[_{CP} C_{[uQ][uWh][EPP], [uPhi], [uTense][EPP]} [_{TP} T\text{-DO}_{[+Present][uPhi]} [_{v^*P} \text{What}_{i[+Q][+Wh][+Singular]} \text{you } v^*_{[uPhi][EPP]} \text{-like}[_{VP} \text{like what}_{i[+Q][+Wh][+Singular]}]]]]$   
 d.  $[_{CP} \text{What}_{i[+Q][+Wh][+Singular]} \text{do}_{j[+Present][uPhi]} \text{-} C_{[uQ][uWh][EPP], [uPhi], [uTense][EPP]} [_{TP} T\text{-}\emptyset_j_{[+Present][uPhi]} [_{v^*P} \text{you}_{i[+2nd][+Plural]} v^*_{[uPhi][EPP]} \text{-like}[_{VP} \text{like what}_{i[+Q][+Wh][+Singular]}]]]]$

The derivations from (a)-(d) illustrates what the pre-syntactic parameter denotes, i.e., the formal features assembled on the lexical items have driven the generation of hierarchical structures. Furthermore, in (d), the in-situ Wh-word is only used for semantic interpretation, such as “for which thing x, you like the thing x” when (d) is transferred into C-I system, but it cannot stand long in terms of normal English linearization in the SM system, because the normal English Wh-question sentence does not keep the in-situ copy of Wh-word at the surface. In this case, when the narrow syntax transfers (d) into the SM system for the phonological interpretation, the in-situ Wh-word is deleted according to Chomsky (2007, p. 21), “at the sensorimotor side, only one of the two identical syntactic objects is pronounced, typically the structurally most salient occurrence”. The in-situ “what” is thus deleted because it is not structurally salient. The post-syntactic phonological parameter is set. If this idea is on the right track, the double-stage proposal seems to be able to realize the integration of the theoretical perspective of top-down and bottom-up as we clarified before.

## 6. Conclusion

We examine the different proposals for parameterization in terms of language acquisition and linguistic diversity. The loci of parameters are located either in different modules/components of language, such as in the computational and semantic components, the pre-syntax lexicon, even in the third-factor conditions, or in the externalization process. If the computational results of syntactic component are subject to the interfaces for evaluation or linearization, it seems possible to attribute the loci of variation to the initial preparation for syntactic computation in the lexicon, or to the externalization of syntactic results in SM. It is because the structures computed by

Merge are unordered in the narrow syntax. If this is right, the double-stage proposal of parameterization might be on the right track. But the validity of this proposal needs more empirical and theoretical evidence to prove it, especially, evidence from language acquisition.

## Notes

- 1 We adopt the capitalized “P” here to indicate the innate assumption of principles and parameters of language in the Government and Binding Theory before the formulation of the Minimalist Program.
- 2 Kayne (2005, a.o.) promoted this idea by showing what it counts in the comparative syntax paradigm.
- 3 As for Chierchia’s (1998) semantic parameters, language can vary with respect to the predicative or argumental nature of their nouns: some languages merely possess predicative nouns (denoting properties), which must obligatorily combine with a determiner in argumental positions; like Chinese, nouns are argumental (names of kinds) and can thus function as arguments on their own, i.e., without a DP projection. Along this line, different proposals for the variation of DP include Kayne’s visible vs. invisible hypothesis, namely, all languages have DP, some have visible DP and others have invisible DP; and externalization hypothesis, confining the variations to realization of DP or not in PF level.

## References

- Baker, M., & Collins, C. (2006). Linkers and the internal structure of vP. *Natural Language & Linguistic Theory*, 24, 307-354.
- Berwick, C., & Chomsky, N. (2011). The biolinguistic program: The current state of its evolution and development. In A. Di Sciullo & C. Boeckx (Eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty* (pp. 19-41). New York: Oxford University Press.
- Berwick, R., & Chomsky, N. (2016). *Why only us*. Cambridge, MA: MIT Press.
- Biberauer, T. (2008). *The limits of syntactic variation*. Amsterdam / Philadelphia: John Benjamins.
- Bobaljik, J., & Thrainsson, H. (1998). Two heads aren’t always better than one. *Syntax*, 1(1), 37-71.
- Boeckx, C. (2006). *Linguistic minimalism, origins, concepts, methods and aims*. Oxford: Oxford University Press.
- Boeckx, C. (2011). Approaching parameters from below. In A. Di Sciullo & C. Boeckx (Eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty* (pp. 205-221). New York: Oxford University Press.
- Boeckx, C. (2014). What principles and parameters got wrong. In M. Carme Picallo (Ed.), *Linguistic variation in the minimalist framework* (pp. 155-178). Oxford: Oxford University Press.
- Boeckx, C. (2016). Considerations pertaining to the nature of logodiversity. In L. Eguren, O. Fernandez-Soriano, & A. Mendikoetxea (Eds.), *Rethinking parameters* (pp. 64-104). Oxford:

- Oxford University Press.
- Borer, H. (1984). *Parametric syntax, case studies in Semitic and Romance languages*. Dordrecht: Foris Publications.
- Carr, P. (2006). Philosophy of linguistics. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed.) (pp. 331-337). Oxford: Elsevier.
- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, 6, 339-405.
- Chomsky, N. (1959). Review of B. F. Skinner's *Verbal Behavior*. *Language*, 35, 26-58.
- Chomsky, N. (1965/2015). *Aspects of the theory of syntax* (50th Anniversary Edition). Cambridge, MA: MIT Press.
- Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. New York: Praeger.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, N. (1998/2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step: Essays in syntax in honor of Howard Lasnik* (pp. 89-155). Cambridge, MA: MIT Press.
- Chomsky, N. (2001). *Beyond explanatory adequacy* (MIT Occasional Papers in Linguistics, 20). Cambridge, MA: MIT Working Papers in Linguistics.
- Chomsky, N. (2005). Three factors in the language design. *Linguistic Inquiry*, 36, 1-22.
- Chomsky, N. (2007a). Approaching UG from below. In U. Sauerland & H.-M. Gärtner, (Eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics* (pp. 1-29). New York: Mouton de Gruyter.
- Chomsky, N. (2007b). Of minds and language. *Biolinguistics*, 1, 9-27.
- Chomsky, N. (2008). On phases. In R. Freidin, C. Otero, & M. Zubizarreta (Eds.), *Foundational issues in linguistic theory, essays in honor of Jean-Roger Vergnaud* (pp. 133-166). Cambridge, MA: MIT Press.
- Chomsky, N. (2010a). The biolinguistic program: Where does it stand today? *Linguistic Science*, 2, 113-123.
- Chomsky, N. (2010b). Some simple evo devo theses: How true might they be for language? In R. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language, biolinguistic perspective* (pp. 45-62). New York: Cambridge University Press.
- Chomsky, N. (2012). Poverty of stimulus: Unfinished business. *Studies in Chinese Linguistics*, 33, 3-16.
- Chomsky, N. (2013). Problems of projection. *Lingua*, (130), 33-49.
- Chomsky, N. (2014). Minimal recursion: Exploring the prospects. In T. Roeper & M. Speas (Eds.), *Recursion: Complexity in cognition* (pp.1-18). New York: Springer.
- Chomsky, N. (2016). *What kind of creatures are we?* New York: Columbia University Press.
- Eguren, L., Fernandez-Soriano, O., & Mendikoetxea, A. (2016). Introduction. In L. Eguren, O. Fernandez-Soriano, & A. Mendikoetxea (Eds.), *Rethinking parameters* (1-45). Oxford: Oxford University Press.

- Fortuny, J. (2008). *The emergence of structure in syntax*. Amsterdam: John Benjamins.
- Freidin, R. (2014). Recursion in generative grammar. In F. Lowenthal & L. Lefebvre (Eds.), *Language and recursion* (pp. 139-147). New York: Springer.
- Friederici, A., Chomsky, N., Berwick, R., Moro, A., & Bolhuis, J. (2017). Language, mind and brain. *Nature*, (1), 713-722.
- Hale, K. (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory*, (1), 5-47.
- Harley, H., & Ritter, E. (2002). Structuring the bundle: A universal morphosyntactic feature geometry. In H. Weise & H. Simon (Eds.), *Pronouns-grammar and representation* (pp. 23-39). Amsterdam: John Benjamins.
- Hauser, D., Chomsky, N., & Fitch, W. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, (298), 1569-1579.
- Hornstein, N., & Lightfoot, D. (1981). *Explanation in linguistics: The logical problem of language acquisition*. London: Longman.
- Huang, J. (2010). Macro- and micro-variations and parametric theory: Principles-and-Parameters and Minimalism. Manuscript, Harvard University.
- Jacob, F. (1977). Evolution and tinkering. *Science*, (196), 1161-1166.
- Jenkins, L. (2000). *Biolinguistics: Exploring the biology of language*. New York: Cambridge University Press.
- Jenkins, L. (2013). Biolinguistics: A historical perspective. In C. Boeckx & K. Grohmann (Eds.), *The Cambridge handbook of biolinguistics* (pp. 4-11). New York: Cambridge University Press.
- Kayne, R. (2005). *Movement and silence*. Oxford: Oxford University Press.
- Kiparsky, P. (1982). Lexical morphology and phonology. In Linguistic Society of Korea (Ed.), *Linguistics in the morning calm: Selected papers from SICOL (1981)* (pp. 3-91). Seoul: Hanshin.
- Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.
- Lewontin, R. (1998). The evolution of cognition: Questions we will never answer. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science* (pp. 107-132). Cambridge, MA: MIT Press.
- Lowenthal, F., & Lefebvre, L. (2014). *Language and recursion*. New York: Springer.
- Mao, T. (2016). *Chinese EFL learners' acquisition of Comp* (Doctoral dissertation). Beijing Foreign Studies University, Beijing.
- Mao, T., & Dai, M. (2015). A conceptual reformulation of parameters in L2 Feature Reassembly Hypothesis. *Modern Foreign Languages*, (5), 667-677.
- Mao, T., & Meng, F. (2016). The cartographic project of the generative enterprise. *Language and Linguistics*, (17), 916-936.
- Mao, T., & Zeng, F. (2011). On transmutation and unification of the concept of abduction: The implications for Pragmatics. *Journal of Foreign Language*, (6), 40-47.
- McGilvray, J. (2013). The philosophical foundations of biolinguistics. In C. Boeckx & K. Grohmann (Eds.), *The Cambridge handbook of biolinguistics* (pp. 22-46). New York: Cambridge University Press.

- Newmeyer, F. (2004). Against a parameter-setting approach to language variation. *Linguistic Variation Yearbook*, (4), 181-234.
- Newmeyer, F. (2005). *Possible and probable languages*. Oxford: Oxford University Press.
- Radford, A. (1990). *Syntactic theory and the acquisition of English syntax*. Oxford: Basil Blackwell.
- Rizzi, L. (1982). Wh-Movement and the Null Subject Parameter. In L. Rizzi (Ed.), *Issues in Italian syntax*. Dordrecht: Foris.
- Roberts, I. (2012). Macroparameters and minimalism: A programme for comparative research. In C. Galves et al. (Eds.), *Parameter theory and linguistic change* (pp. 320-335). Oxford: Oxford University Press.
- Roberts, I. (2016). Some remarks on parameter hierarchies. In L. Eguren, O. Fernandez-Soriano, & A. Mendikoetxea (Eds.), *Rethinking parameters* (pp. 170-199). Oxford: Oxford University Press.
- Roberts, I., & Holmberg, A. (2009). Introduction: Parameters in minimalist theory. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan (Eds.), *Parametric variation: Null subjects in minimalist theory* (pp.1-57). Cambridge: Cambridge University Press.
- Saito, M., & Fukui, N. (1998). Order in phrase structure and movement. *Linguistic Inquiry*, 29(3), 439-474.
- Sherman, Y. (2007). Universal genome in the origin of metazoan. *Cell Cycle*, 6(15), 1873-1877.
- Slabakova, R. (2016). *Second language acquisition*. Oxford: Oxford University Press.
- Striedter, G. (2006). Precis of principles of brain evolution. *Behavioral and Brain Sciences*, 29, 1-36.
- Stroik, T., & Putnam, M. (2013). *The structural design of language*. New York: Cambridge University Press.
- Wexler, K. (1991). On the arguments from the poverty of the stimulus. In A. Kasher (Ed.), *The Chomskyan turn*. Oxford: Blackwell.
- White, L. (1985). Is there a “logical problem” of second language acquisition? *TESL Canada Journal*, 2, 29-41.
- White, L. (2003). *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press.

**(Copy editing: Alexander Brandt)**

### **About the author**

Tiao-yuan Mao (tiaomarc@mit.edu) currently works as a postdoctoral fellow in the Department of Linguistics and Philosophy of Massachusetts Institute of Technology (MIT). He earned his doctoral degree from Beijing Foreign Studies University. His research interests cover syntax, language acquisition (both the first and second language acquisition), semantics-pragmatics interface, language teaching, esp., with keen interest in the biolinguistics and neurolinguistics. He has hosted several scientific research projects sponsored by the Ministry of Education of the PRC or other Provincial Social Science

Foundation, such as ‘Study of the Formalization of Pragmatic Reasoning’. Most of his research papers have been published on several prestigious linguistic journals both at home and abroad.