This is the peer reviewd version of the followng article:

Syntactic diversity and language learnability / Crisma, P., Guardiano, C., Longobardi, G.. - In: STUDI E SAGGI LINGUISTICI. - ISSN 2281-9142. - 58:2(2020), pp. 99-130. [10.4454/ssl.v58i2.265]

Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

11/11/2026 20:26

(Article begins on next page)

SL

Syntactic diversity and language learnability

Paola Crisma, Cristina Guardiano, Giuseppe Longobardi

Abstract

We propose a preliminary model of a practical parameter setting procedure that aims at bridging the gap between descriptive and explanatory adequacy. We present a list of questions which can successfully set 94 binary parameters in 69 languages drawn from several different families using positive evidence only. Our proposal can be cast within a minimalist model of the language faculty, assuming an underspecified universal grammar and a rich network of implications among parameters. We argue that the workload of parameter setting can be significantly reduced by means of two assumptions: first by positing that only parameters with a positive value are set; second, by showing that parameters can be set exclusively on the basis of a core subset of positive evidence, which we call the *Restricted List*. We suggest that a model with these properties qualifies as a plausible framework for language acquisition studies, and also lends itself to be applied to closed corpora, such as those available as the sole sources for diachronic studies.

Keywords: learnability, parameter setting, underspecified Universal Grammar.

1. *Introduction*

In classical generative linguistics (Chomsky, 1957; 1964; 1965), a parallelism was proposed between the task of the linguist and that of the language learner: in both cases, the task is the correct reconstruction of the target grammar (adult *I-language*: Chomsky, 1986) that generates a certain amount of observable linguistic data. However, it is clear that the two procedures are different, both in terms of the input data and of mental states.

In this article, we present a crosslinguistically applicable procedure for discovering the target grammar from a *corpus* of syntactic data under conditions resembling those faced by a first-language learner: that the *corpus* must contain only *positive* and qualitatively *restricted* evidence. We argue that, to successfully work out such a procedure, it is necessary to identify for most parameters a default state, not to be set from experience; but it is *not* necessary to assume

Received: February 2020
Accepted: May 2020

that parameters are present as an innate list at the initial state S_0 of the Language Faculty[1], which can be presumed to be as minimal as possible.

The existence of language diversity will be the crucial problem of this discussion. The two most salient properties of human syntax studied in decades of structuralist and formal linguistics, namely constituent structure and constituent displacement, have now been reduced to a unified and computationally simple mechanism, regarded as a species-invariant property of the language faculty: Merge (Hauser, Chomsky and Fitch, 2002). This successful development can be regarded as the best example of 'genuine explanation' in linguistics[2]: for, at the same time, it provides an effective *descriptive* framework for grammars, makes an *explanatory* claim about the *LAD* (children parse inputs through Merge and only acquire grammars based on it), and also assumes a structure *simple* enough to be attributed to a single crucial step of the *evolution* toward biolinguistically modern humans[3].

However, language is obviously not as species-invariant as some other cognitive capacities; thus, I-languages cannot simply be the reflex of the innate initial state of the mind S_0, perhaps with progressively visible effects determined by maturation processes of other capacities, up to the relatively steady final state S_S; they are also shaped by culturally transmitted and environmentally triggered diversity.

Thus, to be explanatorily adequate (Chomsky, 1964), linguistic theory must be able to account for the selection of different grammars. For this reason, diversity is central in addressing learnability issues. Linguists must explain how a language learner can eventually converge on the correct grammar of the target language, notwithstanding the amount of alternatives. Not doing so essentially amounts to putting aside explanatory adequacy altogether. This point has been most lucidly described by Lightfoot:

> Generativists nowadays describe "parametric differences" between the grammars of, say, Japanese and Navaho, but they rarely mention how the parameters would be set for the particular grammars of these languages: what the triggering experience would need to be for the Japanese and Navaho child. Worse, if one tries to tease out the implicit assumptions about the trigger, they sometimes include exotic or negative data. (Lightfoot, 1989: 323)

[1] Sometimes called Language Acquisition Device (*LAD*) or Universal Grammar (*UG*).

[2] As suggested by N. Chomsky in personal correspondence with one of the authors, May 2019.

[3] Thus, such a notion of 'genuine explanation' seems to subsume the classical success levels of descriptive and explanatory adequacy, as well as of evolutionary adequacy, a term used in Longobardi (2003) to cover some of the main concerns of minimalist linguistics.

Showing that a system encoding diversity is learnable should in fact be a necessary pre-requisite to accepting it as a cognitively realistic representation of any specific I-language.

2. *Learnability, universals and diversity*

2.1. *Parameters and learnability*

A first hypothesis about the selection of a grammar from the Primary Linguistic Data (*PLD*) was based on the idea that learners could perform some *a posteriori* arithmetic evaluation of the fitness of each of several assumed grammatical hypotheses compatible with the data and with universal constraints (Chomsky, 1957). Since the number of grammatical hypotheses satisfying this condition was normally very high (often infinite), evaluation procedures were abandoned as non-realistic and inefficient, and replaced by Principles and Parameters (*P&P*) theories (Chomsky, 1981).

Heuristically, this model has been extremely productive, generating a large number of data-driven analyses of the most diverse languages, and isolating plenty of points of abstract discrete contrasts among such languages. Parametric theories have been very successful in describing grammatical diversity as a system of abstract binary choices predefined by the Language Faculty and in addressing typological (e.g. Kayne, 2000; Baker, 2001), diachronic (Lightfoot, 1979; 2006; Clark and Roberts, 1993; Roberts, 2007) and even taxonomic/phylogenetic (Longobardi *et al*., 2013; Ceolin *et al*., 2020) issues in formal linguistics. It remains to be shown, however, if they can really address the 'logical problem of language acquisition' (Hornstein and Lightfoot, 1981), a challenge posed by language variability even in strongly nativist frameworks.

In principle, the postulation of an innate system of open parameters, made available by Universal Grammar (*UG*), makes language learning pretty straightforward, breaking it down just to the individual setting of these parameters. For this reason the *P&P* model was quickly adopted in language acquisition studies[4], bringing about a series of stimulating research topics, such as for example: the question of whether parameters can be mis-set in the course of acquisition, and how a mis-set parameter could then be re-set

[4] For a brief outline of parameters in language acquisition studies, see in particular Thornton and Tesan (2007) but also Fodor (2001) and de Villiers (2001).

to converge with the target grammar[5]; the notion of deterministic 'triggers' or 'cues'[6] as opposed to statistical models[7]; the pervasive problem of ambiguous input, i.e. portions of the *PLD* that are compatible with more than one grammar[8]; the question of whether parameter values display set-theoretical relations, and the consequent formulation of the Subset Principle[9]; the necessity of reconciling the *P&P* model with the obvious fact that language acquisition is not instantaneous, as the mere 'switch' of a series of parameters may predict (see the debate on the Continuity Hypothesis[10] and the Maturation of *UG*[11]). All this led to the formulation of various hypotheses, some of which have some relation with our discussion.

As discussed at length in Fodor and Sakas (2017), however, a really plausible and established parameter-setting model has not been implemented yet, for several reasons. First, as noted by Fodor and Sakas, various highly influential learning models (Clark, 1992; Gibson and Wexler, 1994; Yang, 1999; 2002) still assume that whole grammars rather than single parameters are evaluated against the linguistic evidence; thus, they avoid the problem of dealing with the intricate dependencies and interactions among parameters (see also § 3.3) at the cost of hugely increasing the computational load of parameter setting. Second, all these models, including Fodor and Sakas (2017) and Sakas, Yang and Berwick (2017), make the assumption that parameters[12], as an extensional list, are part of S_0. But, as Fodor and Sakas (2017: 267) admit: «How burdensome that is, and how plausible it is from an evolutionary point of view, remains to be determined». Third, none of these models is based on a realistically meaningful number of parameters: even Sakas, Yang and Berwick (2017: 393), who claim that they «move beyond toy grammars[13] and provide a large-scale study of parameter setting in a linguistically complex domain» formulate the proposal on a set of 13 parameters only[14].

[5] Hyams (1986).

[6] Gibson and Wexler (1994), Fodor (1998), Roeper (1999), Lightfoot (1989), Clark and Roberts (1993).

[7] Yang (2002; 2004).

[8] Fodor (1998) and subsequent developments.

[9] Manzini and Wexler (1987).

[10] On this, see in particular Crain and Thornton (2015) and references cited.

[11] Defended in particular by K. Wexler, see Borer and Wexler (1987; 1992) and Wexler (1994).

[12] Or parameter values in the form of the 'treelets' of Fodor (1998).

[13] Where the 'toy grammar' was for example the three-parameter model of Gibson and Wexler (1994).

[14] Called «a modest collection» in Fodor and Sakas (2017).

The present work is meant to be a further step in this line of research, but it takes a different perspective: instead of testing the plausibility of the *P&P* model on the basis of a more or less haphazard collection of parameters, it takes a real-life collection of parameters that describe the variation in a circumscribed syntactic domain, and, even without presenting, at this point, a full operational learning algorithm, tests their settability on the basis of primary evidence. Analyzing our collection of parameters from the perspective of the language learner, we focus precisely on triggers, to determine how much of the information encoded in those parameters is in fact realistically provided to the child by the *PLD*.

Our starting point is a collection of 94 parameters which, abstracting away from purely lexical and phonetic idiosyncrasy[15], aims at a near-exhaustive description of the diversity observed in the mental grammars of nominal structures across 69 languages from 13 different historical families. We put together these 94 parameters using the customary tools of speakers' grammaticality judgements, including direct negative evidence and sometimes rare and complex constructions: a Table with the states of these parameters, along with a list of questions to elicit the triggers used to set them, is found in the on-line Supplementary Materials ((https://github.com/CristinaGuardiano/Parameter-setting-Questions, private access, upon request). This apparatus achieves a good deal of observational and descriptive adequacy in the specified domain.

Thus, the empirical basis of our study differs in at least three important respects from the other works with comparable goals mentioned above.

The first is the order of magnitude of our tools for encoding syntactic variation, compared to previous studies: Sakas, Yang and Berwick's (2017: 393) dataset consists of 3072[16] artificial languages generated by 13 binary parameters that encode differences observed in natural languages. Our system, with its 94 parameters, is almost one order of magnitude higher.

Second, our parameters are all drawn from a single module of grammar: this maximizes the possibility of finding and explicitly laying down the pervasive implicational system that characterizes syntactic diversity and has a large impact on the burden of parameter setting.

Third, our parameter system is entirely empirically motivated by the diversity observed in real, rather than artificial, languages.

[15] I.e. it excludes nano-parameters in Biberauer and Roberts (2017) terminology.

[16] The total number of languages generated by a set of 13 binary parameters is in principle 8192 (that is, 2^{13}), but, as the authors explain (Sakas, Yang and Berwick, 2017: 399), there are constraints on some parameters that reduce the resulting language set.

Against this background, we demonstrate that all our parameters are settable from unambiguous positive triggers; also, we argue that one of the two alternatives of each parametric choice does not need to be set at all, but simply corresponds to the unchanged form of S_0.

2.2. *How rich is* UG*?*

As noticed, in classical *P&P* models, it was assumed that parameters are all present at the initial state of language acquisition S_0 in the form of an extensional finite list of every possible point of variation in human grammars. This can be regarded as a 'preformistic' view of cognitive variability, adapting the term from 17th-18th century biology. This view was generally accepted in the 1980s and 1990s, when the common assumption was that the whole set of parameters of *UG* could amount to 20, 30 or maybe 40 items[17].

This hypothesis has proved increasingly untenable in the face of empirical work on syntactic diversity. Our empirical collection contributes to a clear demonstration of this failure: so far, the observable variation in just the module of nominal syntax has required the formulation of at least 94 parameters, even excluding nano-parameters; moreover, the investigation of the same domain in new languages is likely to require the addition of some new parameters to the inventory. This result neatly proves a point which has begun to emerge over the years: a comprehensive list of parameters extended to other domains and to other languages will quickly add up to hundreds or thousands of parameters, dwarfing the initial estimates.

This situation poses with great force the question of how plausible it is to maintain that the initial state S_0 of the faculty of language may consist of a preformistic list of parameters. With such a model, the mind of every speaker should start with several hundreds or thousands of open parametric choices[18]. Many works have argued that a minimalist framework must devise a model of variation that derives parameters from a much smaller set of primitives. Especially work by Boeckx and Leivadá (2014), Lightfoot (2017) and Longobardi (2017) has stressed counterarguments to the classical *P&P* model; accordingly, suggestions for replacing the extensional

[17] See for instance Fodor and Sakas (2017) and references cited.

[18] Most of which would be irrelevant at any stage of their life, since they turn out to be [–] or [0], in our descriptive terms, see § 4.2.

list of parameters with an intensional definition have been advanced since Longobardi's (2005) constructivist approach[19] and especially in Biberauer's (2019) neo-emergentist theory.

In the present work, we cast our hypotheses within a framework compatible with a radically underspecified theory of the initial state. Ideally, S_0 should consist of few and general invariant Principles of *UG*, e.g. Merge, perhaps locality constraints, and some tight externalization and mapping conditions. We assume that such conditions are restrictive enough to allow quite limited forms of parametric variation (see § 3.2)

A full proof of this model goes necessarily beyond the space limits of this article, but we show here that it is possible to envisage a simple parameter setting model that goes back to the origins of *P&P* (discovers the target grammar through a finite number of data questions), but does not commit itself to the assumption that the list of parameters is in fact part of *UG*.

Actually, our collection of parameters and the proof that they are settable in so many languages under restricted realistic conditions will also be a start for investigating what is left in S_0 once the information simply retrievable from the *PLD* is factored out: namely, the sum of *UG* and third-factor (Chomsky, 2005) effects. This is in line with the program advocated by Lightfoot:

> If the trigger or the "primary linguistic data" (*PLD*) were rich and well-organized, correspondingly less information would be needed in *UG*, and vice versa. (Lightfoot, 1989: 323)

3. *Some relevant properties of a parameter system*

First of all, we need to define the term parameter as will be used here: by it, we simply mean any point of minimal (binary) choice ultimately responsible for a set of observable syntactic differences between two languages. Such differences will be called *manifestations* of each parameter.

The analysis of the parameters in our sample, and further observation of parametric variation as described in the literature, highlights some properties that any adequate theory of parameters should take into account.

[19] See Gianollo, Guardiano and Longobardi (2008), Longobardi (2014; 2017).

3.1. *Clustering*

Already in the earliest work on parameters, for example in Rizzi (1980; 1982), Taraldsen (1978) and Chomsky (1981), it was noted that many of them could be associated with a cluster of co-varying surface manifestations, with different degrees of saliency[20]. This co-variation is in principle deduced from a single abstract point of structural diversity (i.e. it is not just an additive list of typological properties).

Over the years, a few cases of apparent macro-parametric clusters turned out to hold only in part and were therefore reformulated as hierarchies of smaller-scale parameters (micro- and meso-parameters in the terms proposed by Biberauer and Roberts, 2017 and Roberts, 2019; see also Manzini, 2019). However, several parameters in our sample still retain a robust cluster of co-varying surface manifestations.

Notice also that it is not necessarily the case that the surface manifestations of a parameter strictly co-vary. In theory, every parameter as such corresponds to one abstract structural choice, whose different manifestations should all consistently follow from general principles; but various factors intrinsic to the structure of grammar often produce predictable deviations from this ideal pattern. As an example, take parameter *FGP* [± *grammaticalized Person*] of our collection; the value [+], i.e. the grammaticalization of the feature Person, is manifested by the presence of Person inflection on verbs but also by the existence of visible expletive subjects[21], as in French; conversely, languages with [– *grammaticalized Person*] like Japanese, lack both. But if the Person inflection on verbs is particularly robust, a language may be a null subject language, and as such will exclude overt expletives, like Italian[22]. Thus, the actual cooccurence in the same language of all the potential manifestations of a parameter will often be a coincidence, depending on the interactions with other syntactic parameters or even morpholexical accidents of languages.

Especially macro- and meso-parameters affect different categories virtually by definition, so they are most exposed to interacting with category-internal variability (other syntactic parameters and also morpho-

[20] See Huang (1982), Rizzi (1986), Kuroda (1988), Pollock (1989), Fassi-Fehri (1993), Longobardi (1994; 2014), Zanuttini (1997), Kayne (2000), Biberauer (2008) among many others.

[21] Among other manifestations, see § 4.4.

[22] Which is still [+ *grammaticalized Person*] by virtue of at least the Person agreement on the verb. Parameter setting is expounded in § 4.

phonological idiosyncrasies): going back to *FGP*, a language grammaticalizing Person may show Person agreement on various categories, such as verbs, pronouns, reflexives, sometimes predicate nouns (as in Dravidian). Each of these constructions can by itself be a trigger (cue or *p-expression*[23] of that parameter), but in different languages a different subset of them is realized for general or more idiosyncratic reasons: for example, in IE predicate nouns do not carry Person morphology, in Mainland Scandinavian even the verb is normally deprived of it, in Slavic languages the same reflexive may be bound by an antecedent irrespectively of its Person specification.

3.2. *Schemata*

An observation made in Longobardi (2005) is that most parameters belong into a small number of types each identified by their general format. These similarities in format crosscut the specific applications of each parameter to a particular feature, or category (the latter understood as a set of lexically cooccurring features). These recurrent formats have been called *schemata* in Longobardi (2005) and subsequent work and should describe the domains in which invariant conditions of *UG* and third-factor fail to apply. Importantly, parameters of the same schemata are likely to share similar properties with respect to the way they are set (cf. § 5.2).

The schemata of possible variation fall into at least three main types. Some schemata have directly to do with the presence of certain formal features in the grammar. So, for α a feature or a set of features lexically associated with a functional category, the first type of schema interrogates about its occurence:

(1) a. is α available in language L?
b. is α grammaticalized in language L?

'Available' means here that a feature or a category can be used in a certain language (though not in others) as a choice by speakers, without being obligatory in a grammatically defined context. For instance, the licensing of a postnominal *non-prepositional* genitive in German or Greek as opposed to

[23] Clark and Roberts's (1993: 317) definition of *p*(*arameter*)*-expression* is reported here: «A sentence σ expresses a parameter p_i just in case a grammar must have p_i set to a definite value in order to assign a well-formed representation to σ».

English and most Romance is accounted for by an availability parameter in our sample (*GFO* [± *GenO*])[24].

'Grammaticalized', instead, is understood as 'obligatorily valued in a set of grammatically definable contexts'. Number has this characteristic in English or Italian nominal arguments, though not in Chinese (see *FGN* [± *grammaticalized Number*]).

Other schemata govern the combination of (sets of) features with each other:

(2) a. is α associated with phonological content in language L?
b. is α associated with interpretive content in language L?
c. is α associated with weak morphosyntactic realization (i.e. is it a clitic with designated hosts) in language L?
d. is α associated with some other feature or set of features β on the same functional item in language L?

The most typical instantiation of (2a) is an empty category, licensed in a language, though not in others: in our sample, Ibero-Romance differs from Italian or French in the possibility for a null noun to be licensed by a definite article in the presence of any modifier (*DNN* [± *null N-licensing article*]). In other cases, just a single feature may be phonologically absent from the realization of a head, such as Number from nouns in French (*FNN* [± *Number on N*]).

The semantic counterpart of (2a) is (2b). Some category or single feature in some language may, in certain contexts, lack mapping to the interpretive component; the most typical case is exemplified by expletive pronouns and articles, as e.g. in our parameters *NEX* [± *proper names in D*], *PEX* [± *personal proper names in D*], *FEX* [± *partial personal proper names in D*], all about the distribution of articles with proper names, which differs in languages as close as Italo-Romance varieties.

(2c) is represented by parameters distinguishing between languages with and without clitic articles or possessives, whose position is determined by that of a designated host, e.g. the head noun or certain adjectives for enclitic definite morphemes of Romanian or Bulgarian (*DCN* [± *article-checking N*]).

A transparent instantiation of (2d) is the difference between Italian and French with respect to possessive pronouns: French *mon* appears to combine

[24] See Crisma, Guardiano and Longobardi (in press).

two features, definiteness and 1st Person genitive, which occur on different items in Italian, *il* and *mio*[25].

Finally, some schemata govern the realization of relationships between (sets of) features not cooccuring in the same item:

(3) a. is α morphologically spread to positions where it is not interpreted in language L?
b. does α trigger overt movement (internal Merge) in language L?

Schema (3a) asks if agreement in the value of a feature holds between two categories (with the feature actually being interpreted in just one position): for instance, *FSN* [± *Number spread to N*] governs the difference between languages like Basque, in which Number is only expressed on the determiner and not on the head noun, and the rest of European languages, a distinction with a lot of indirect consequences (Crisma and Longobardi, in press, based on Delfitto and Schroten, 1991).

Finally, (3b) is based on the assumption that long distance relations between two positions (e.g. the scope- and θ-position of an operator phrase) are determined universally (so, there is no variation in probing by a category), but, as has been known since Huang (1982), such relations can be accompanied or not by overt displacement of the category involved. This variation schema seems one of the most commonly encountered in the study of cross-linguistic differences, *NWD* [± *weak Person*] (based on Longobardi, 1994; Crisma and Longobardi, in press, and governing raising of proper names to D, among many other consequences) is the first of a long series of such parameters in our sample[26].

In sum, such a system identifies few possible formats for the variability of syntactic relations/operations, determining a number of potential binary choices, which become actualized only on encountering overt evidence for them in the *PLD*. Thus, an interesting property of schemata is that they *describe* how tightly the possible variability appears to be constrained. A natural question to be addressed in further work is then how much of this constrained variability can be derived as a consequence of *UG* and freedom left by

[25] This schema takes inspiration from Sportiche's (1986) comparison of pronominal, anaphoric and bound-variable features in English and Japanese. Longobardi (2014) also discusses how the features of *no* and *any* in English may appear as specified on one and the same item (*nessuno/ningún*) in Italian/Spanish.

[26] The formulation of this schema does not commit us to claiming that all movement must be feature-driven.

how much can be ascribed to third-factor (Chomsky, 2005) effects or to the very nature of the *PLD*.

3.3. *Implicational structure*

In many cases, one of the two states of a parameter will be predictable from the states of other parameters (Baker, 2001; Guardiano and Longobardi, 2017; Roberts, 2019, among others).

To give an idea of the scope of this phenomenon, notice that 94 parameters should theoretically generate 2^{94} languages, but, given the implicational structure, the actual number of languages generated by our parameters is much lower, in fact. Using the procedure developed in Bortolussi *et al.* (2011), Ceolin *et al.* (2020) calculated that the number of languages generated by the first 30 of our 94 parameters is 152˙448 possible languages (~2^{17}) instead of 2^{30}. Beyond this threshold, a precise calculation becomes heavy in terms of computational resources, therefore it is more practical to proceed by projections. Given that the number of implications grows when one considers the other parameters in the list, it is probable that the cardinality of languages generated by the full 94 parameters is in the order of 2^{40}-2^{50}; this would still however be unmanageably huge from the point of view of language learners, if they were at all concerned with this kind of estimates, as in a model evaluating a large number of whole grammars at the same time. The implicational structure has important positive consequences for a system that sets parameters one by one, though, because it makes a very large proportion of potential choices completely irrelevant for the setting procedure (see § 4.2).

64

4. *Parameter setting*

4.1. *Parameters and* PLD

The parameters in our collection were initially formulated relying on *manifestations* which also include *un*grammaticality judgements, which can be obtained by linguists working with native consultants. However, the most restrictive hypothesis about language learners is that they have no access to negative evidence, direct or indirect[27]: the model presented here aims at testing this restrictive condition.

[27] See in particular Lightfoot (1989: 323-324) and Lasnik and Lidz (2017: 247).

We show that at least one value of each parameter in our collection can be unambiguously associated with positive evidence available in the *PLD*, i.e. with visible triggers or *p-expressions* (in Clark and Roberts' (1993) sense).

In the concrete, to do so, we associated each parameter with a list of one or more YES/NO questions with the following property:

(4) Each question must ask about the occurrence of (a set of) observable grammatical patterns/properties, and therefore have the logical form of a simple existential question:
'Does a (set of) structure(s)/interpretation(s) so-and-so occur in language L?'

The purpose of these questions is precisely that of identifying the possible *p-expressions* for each parameter. From the point of view of the language learner, it is only those *p-expressions* that matter, the 'structure(s) so-and-so' encountered in the *PLD*, not the questions, which have been reconstructed a posteriori from the linguists' experience.

Since each question must be answerable only on the basis of positive evidence, two corollaries follow:

(5) a. the question cannot have the form of a negative existential:
'Is a structure/interpretation so-and-so impossible/unattested/ungrammatical in language L?'[28]

b. the question cannot have the form of a universal assertion, which would amount to a hidden negative existential:
'Is a structure/interpretation so-and-so obligatory (i.e. present in all relevant contexts) in language L?'

This, of course, does not mean that answers to questions of the type in (5) are never part of the full set of manifestations for a given parameter. In fact, they are, and constitute fundamental evidence in the investigation of the linguistic competence of adult native speaker. However, under the most restrictive approach, they are not available or usable in language acquisition.

In the Supplementary Materials we present 94 sets of questions, one set for each parameter: at least one question in each set conforms to the

[28] Excluding this kind of question is equivalent to treating as irrelevant/unattainable the answer NO to questions of the form (4), thus ruling out the relevance of indirect negative evidence for the language learner.

condition in (4). This can be shown to be sufficient to warrant the settability of the parameters in the 69 languages of our sample from positive evidence only.

4.2. *Downsizing the setting task*

As mentioned in § 2.2, our model does not need to attribute an extensional list of alternatives to the initial state. Rather, we view 'parameter setting' as the addition of structure to S_0 (constrained by the schemata in (1)-(3)) only when positive evidence requires it, i.e. when such addition is needed to parse an utterance which contains a *p-expression* of a parameter. In this model, actual parameter setting in the course of language acquisition can be outlined as in (6):

(6) a. positive evidence adds the relevant structure to the mental grammar, conventionally indicated as [+ *parameter* P]
b. S_S results in a string of [+ *parameter P*]'s being added to S_0

Thus, parameters are additions to the initial state of the mind, i.e. only one state (conventionally coded here as [+]) needs to be set from empirical evidence. Therefore, if no relevant manifestation for [+ *parameter P*] is present in the data, the grammar does not change: no *p-expression* is encountered, no structure is added to the grammar. In this model, the expression '*parameter P* has the value [–]' is a useful metaphor for linguistic description and linguistic comparison but has no reality in the mental grammar at any stage (but see § 5.2). Thus, in our system, [–] is the default[29] value for *parameter P* in the sense that, recasting Biberauer's (2019: 60) words, it 'literally requires the acquirer to do nothing'.

[29] However, the default state for a given parameter does not necessarily coincide with it being statistically frequent (typologically 'unmarked'). Consider for example the fact that [– *parameter P*] may coincide with a seemingly less economical derivation, as in the case of *NWD* [± weak person], where the [–] value characterizes languages that overtly attract referential nominal material to the D area. Take also *FSN* [± *number spread to N*]: + *FSN* languages have (at least some class of) nouns displaying variable number morphology; on the contrary, languages which do not set the value [+] for this parameter exhibit indeclinable nouns. In our language sample, languages with a [+] value (i.e. the non-default) are by far the majority (62 out of 69). A more distributed language sample, that presented in Haspelmath (2013), still sees a predominance of languages that have obligatory plural marking on (at least some class of) nouns over those which don't (173 out of 291). Note however that what is counted as 'nominal plurality' in Haspelmath (2013) does not entirely coincide with [± *number spread to N*]. For a discussion of default states, see § 5.2.

Another factor that obviously reduces the acquirer's burden is the implicational structure discussed in § 3.3, which results in some states being predictable from other states. We label a predictable state [0 *parameter P*], and again [0 *parameter P*] is a descriptive tool with no reality in the mental grammar.

In sum:

(7) +/–/0 are conventional symbols:
[+] indicates that a structure is added to the mental grammar
[–] indicates that a structure is not part of the mental grammar
[0] indicates states predictable from other states

As a result, particular grammars (I-languages) are reduced to a small subset of the overall number of parameters used by linguists for describing broad-scale variability across languages. Ultimately, then, an I-language can be seen as a string of [+]s, all added to the mental grammar on the basis of positive evidence.

This produces a remarkable reduction of the workload needed for parameter setting: in our dataset there are 6486 (94x69) states to set, but in fact 2175 turn out to be [–] and 2925 to be [0], i.e. they have no reality for the language learner: as a result, there are only 1386 [+] that require positive evidence, an average of 20 per language.

4.3. *The Restricted List*

The clustering structure of parameters (see § 3.1) predicts three types of linguistic facts:

i. the grammaticality of some core manifestations;
ii. the grammaticality of some uncommon[30] manifestations;
iii. the ungrammaticality of some other strings/interpretations.

As discussed in § 4.1, our model assumes that negative evidence is not available for the language learner, therefore facts of type iii are not part of the *PLD*. But, what about facts of type ii? Lightfoot (1989) contends that only a robust and structurally simple subset of ~~the language learner's experience~~ acts as ~~triggers~~; in his proposal, only cues from unembedded sentences should be

PLD
triggering experience

[30] Or 'exotic' in Lightfoot's (1989: 323) terms.

taken into account for language acquisition (degree-0 learnability). Without engaging in a detailed evaluation of Lightfoot's formal hypothesis for distinguishing facts of type (i) and (ii)[31], we fully subscribe to his intuition that facts of type (ii) are not used in parameter setting, on a par with facts of type (iii)[32]. The parallelism between uncommon manifestations and negative evidence is well captured in Baker's (1989) commentary on Lightfoot (1989):

> Whereas it is now widely recognized that most negative data about a language are best viewed as "nonprimary" in nature, it is less often appreciated that at least some positive data *need to be accorded the same status* [emphasis ours, P.C., C.G., G.L.]. That is, certain types of acceptable sentences occur so rarely if at all in normal language use that their acceptability is more properly viewed as a result of language acquisition than as an input to it. (Baker, 1989: 334)

The perusal of the list of questions targeting our parameters' manifestations makes it immediately obvious that not all the *p-expressions* of a given parameter have the same saliency: some of the structures represented in the YES/NO questions associated with that parameter may be uncommon because of their structural complexity, or require elaborate pragmatic situations to be used felicitously. This is the kind of evidence that can be constructed by a linguist when eliciting grammaticality judgments from a (trained) native speaker, but may be hard for a learner to come across in a *corpus* of *PLD*[33]; it is therefore 'nonprimary' in Baker's sense. Thus, for a model to be learnable, and in particular for all the learners to converge on the same grammar, it is necessary that no parameter be set solely on the basis of this kind of nonprimary evidence: some core manifestation will set the parameter, while the more complex and uncommon manifestations will be predictable from this setting.

For this reason, we propose a further subsetting operation, illustrated in Figure 1. For each parameter, we try to isolate those *p-expressions* that, because of their structural or pragmatic characteristics, could be classed as nonprimary, and separate them from those which, conversely, could be considered the core primary evidence for that parameter. This latter set is what we call the *Restricted List* for each parameter, namely a subset containing only its core *p-expressions*.

[31] In its formulation, degree-0 learnability is anyway not easily applicable to the nominal domain.

[32] Cf. Clark and Roberts (1993: 303, fn. 3) for a similar approach.

[33] Or for a linguist working on a normal *corpus* of utterances, for that matter.

si può diminuire lo spazio qui in alto? In questo modo (8b), che è finito a pag. 18, tornerebbe sotto (8a) e ci starebbe meglio

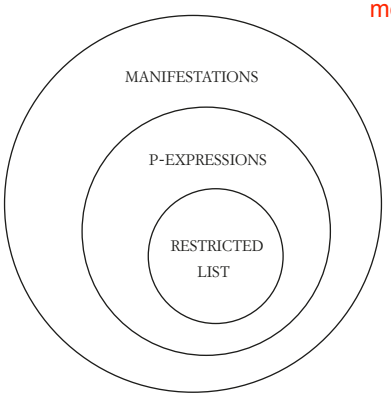

Figure 1. *Subsets of manifestations*.

This model, which we illustrate in § 4.4, has an important consequence for a learner and for the linguist: identifying just one manifestation in the *Restricted List* per parameter will set the parameter, while nonprimary manifestations[34] will follow from that setting[35]. This makes an immediate prediction: if a parameter has only one *p-expression*, this *p-expression* cannot be of the nonprimary type.

At this point, however, our individuation of the relevant *Restricted List* for each parameter is still intuitive and is not the result of a precise metric. We are fully aware that devising measurable criteria for sorting out *p-expressions* that belong in the *Restricted List* from those that can be safely labelled 'nonprimary' is a fundamental part in working out a full-fledged learnability algorithm, and we set this goal as ~~one of the~~ most urgent ones for future research work.

a

4.4. *The setting task: An example*

In sum, our model of parameter setting is based on the following tenets:

(8) a. only positive evidence is used to set parameters;

[34] As well as intuitions about the ungrammaticality of some constructions.

[35] Note that if one assumes that only the *Restricted List* provides evidence to the language learner for parameter setting, only the questions targeting those manifestations need be subject to the strictures posed in (4), i.e. need technically to search for *p-expressions*; the questions targeting 'nonprimary' manifestations are only used by the linguist, and therefore can (but do not have to) be of the type in (5) (i.e. admit negative evidence).

aggiunger spaziatura, a meno che non si riesca a portare (8b) sulla pagina precedente

b. no parameter is set on the basis of nonprimary patterns/structures.

Furthermore, it is compatible with an underspecified *UG*: S_0 does not need to contain an extensional list of parameters.

Now for concreteness, consider the parameter *FGP* [± *grammaticalized Person*], presented in Figure 2, which encodes whether the language has a formal feature 'Person' to be valued in certain contexts. The parameter has various manifestations, discussed in Crisma and Longobardi (in press); however, only three of them form the *Restricted List*: when the language learner encounters one of these three manifestations (which correspond to a YES answer to the questions), s/he adds the grammaticalized feature Person to her/his grammar[36].

As discussed in § 3.1, parameters are abstract properties that may have various observable manifestations, and it is not always the case that all the manifestations are encountered in all languages; therefore it is possible that different languages set *FGP* on the basis of different *p-expressions* in the *Restricted List*, and this is what we observe in our dataset: most of the languages that are [+ *grammaticalized Person*] set the parameter on the basis of the manifestation corresponding to *FGP_Qa*. However, two languages of the set, Norwegian and Danish, have no person agreement on the verb (a situation likely connected to an independent morphological parameter blocking exponence of ϕ-features on tensed verbs), and add [+ *grammaticalized Person*] on the basis of *FGP_Qb*.

The languages that did not answer YES to any of the questions in the *Restricted List* for *FGP* (Korean, Japanese, Mandarin, Cantonese in our dataset) are classed as [– *grammaticalized Person*] as the default option[37]. We predict that this state will be consistently reflected in *all* the other manifestations assigned to this parameter, in other words, that *none* of the other questions associated to *FGP* will receive a YES answer in these languages. This is indeed the case.

[36] In other terms, the parameter *FGP* is set to [+].

[37] Again, this notational device is used only for comparative purposes with no reality in language learning: we assume that the Korean/Japanese/Mandarin/Cantonese child, *not* encountering the relevant *p-expressions*, simply does not add the feature Person to his/her grammar.

Manifestations

p-expressions

Restricted list

FGP_Qa	Is there agreement in person between an argument and a verb?
FGP_Qb	Are there overt expletive pronouns in subject function?
FGP_Qc	Are there overt resumptive pronouns in object function?
FGP_Qd	Pick a 3rd person pronoun that can occur with no coreferential item in the sentence. Can it also occur as a variable bound by a quantified antecedent like *no-one/everyone*?
FGP_Qe	In argument position, the language displays the following order pattern: [(Determiner) Adjective Noun] (*some/those/the/Ø*) *poor/young children* are waiting outside Does the language also display the following pattern(s) in argument position? [Personal pronoun Adjective] *we poor/young* are waiting outside
FGP_Qf	Are there nominals that in non-argument position can occur bare but in argument position require a morpheme expressing some of the following functional meanings: person, gender, number, definiteness, countability, specificity?
FGP_Qg	Are argument proper names introduced by a functional morpheme that is not required in other (non-argument) occurrences?
FGP_Qh	Pick an element *x* (for instance a possessive, an adjective meaning 'other' 'same/even' or 'unique', the noun itself, etc.). Do nominals with a cardinal numeral before *x* always receive indefinite interpretation and nominals with a cardinal numeral after *x* receive definite interpretation?

Figure 2. *Manifestations for parameter* FGP [± grammaticalized person].

5. *Further issues*

5.1. *Ambiguity in the* PLD

The various attempts at devising a learning algorithm based on the *P&P* model for syntax, briefly referred to in § 2.1, recognize ambiguous data as a fundamental problem that any model has to deal with[38]. Ambiguous data are a consequence of the fact that even the simplest strings of an E-language are normally the expression of more than one parameter: the interaction among some of them may conspire to produce the same surface string in different languages, even if they have opposite values[39].

As an illustration of this point, consider the following noun phrase from four different languages:

(9)	a.	*lup-ul*	Romanian
	b.	*vŭlk-ŭt*	Bulgarian
	c.	*úlfur-inn*	Icelandic
	d.	*otso-a*	Basque
		wolf-art	
		"The wolf."	

In all the four cases, the head noun is followed by a morpheme that could be labelled 'an article'[40]. Notwithstanding this surface similarity, the four languages all differ in their nominal syntax, as is immediately apparent as soon as the same nominals are modified by an adjective:

(10)	a.	*lup-ul*	negru	Romanian
		wolf-art	black	
	b.	*cherni-yat*	*vŭlk*	Bulgarian
		black-art	wolf	
	c.	*svarti*	*úlfur-inn*	Icelandic
		black	wolf-art	

corsivo

[38] See in particular Fodor and Sakas (2017) and Sakas, Yang and Berwick (2017), and references cited.

[39] The typical example discussed in the literature (e.g. Gibson and Wexler, 1994 and Fodor, 1998) is that of SVO sentences, that are generated by a basic VO grammar, like that of English, but also by an OV V2 grammar, as in German.

[40] This label can group the four cases only in a very loose sense. In particular, the Icelandic morpheme, though it marks definiteness, is different from the article of Romanian and Bulgarian because of its structural position. As for the Basque clitic, it does not mark definiteness but argumenthood (see Crisma and Longobardi, in press).

d. *otso* *beltz-a* BASQUE
wolf black-ART
"The black wolf."

Thus, because of its ambiguity, the question 'Does the sequence Noun-Article occur in language L?' is not associated with any parameter in our collection, even if the sequence is obviously very robustly represented in the *PLD*. The sequence can in fact result from the cliticization of the article on the noun raised to the D position, as in Romanian and Bulgarian; from the cliticization of a definiteness affix on a noun lower in the structure, as in Icelandic; from the fact the whole DP is head-final, as in Basque. Therefore, we do not consider the sequence Noun-Article a *p-expression* that the language learner can use to set any parameter.

Other very common structures, however, do provide unambiguous triggers, such as the noun phrases modified by an adjective in (10). Equally clear are the examples in (11), where the phrase-final position of *asko* 'many' unambiguously singles out Basque as having head-final DPs, as opposed to Romanian, Bulgarian and Icelandic[41]:

(11) a. *mulți lupi* (*negri*) ROMANIAN
b. *mnogo* (*cherni*) *vŭltsi* BULGARIAN
c. *margir* (*svartir*) *úlfar* ICELANDIC
d. *otso* (*beltz*) *asko* BASQUE
"Many (black) wolfes."

The questions identifying the relevant *p-expressions* to set the parameters in our collection ~~are~~ of this type: they single out unambiguous evidence, but are still confined to relatively simple and common utterances, which could be good candidates for membership in the relevant Restricted Lists. target sequences

5.2. *Default states and the* PLD

The parameter-setting model described in § 4 assumes that a parameter is added to the mental grammar ([+ *parameter P*]) only if positive evidence requires it in order to be parsed; if no such evidence is available in

[41] We assume that head-finality in Basque is generated via overt movement of the complement of D to its Spec (see CRISMA and LONGOBARDI, in press, in KAYNE's 1994 framework).

the *PLD*, then the structure does not become part of the mental grammar at all[42]. Patterns such as the one in (11) may constitute a challenge for this model, for two reasons.

First, the model postulates that evidence plays no role in the case of one of the two values, the one called [– *parameter P*]: this state, from the point of view of the language learner, is not parameter setting, it is nothing at all. But in the case of (11), which manifests DP-headedness, it seems groundless to suppose that one property, say 'DP is head-initial' is set on the basis of the evidence (11a-c) and added to the relevant I-languages, while the reverse, 'DP is head-final', is not, with the Basque learner ignoring (11d).

Second, if one were adamant in maintaining that for all parametric choices one of the two options must be the default state [–] and the other one the [+], set on the basis of the *PLD*, it would be very hard to decide which one is which on a principled basis, it is almost like the flip of a coin[43].

In sum, this parameter has two properties not contemplated by our model so far:

(12) a. unambiguous p-expressions for each of the two possible values are robustly attested[44];
b. there is no obvious default value for the parameter by independent criteria.

As a matter of fact, from the learner's perspective, facts like those in (11) and (12) are still consistent with a setting model such as the one presented in (6): the learner encounters manifestations for 'DP is head-initial' or 'DP is head-final', the two possible values for a parameter ultimately responsible for DP headedness[45], adds the relevant feature to the grammar accordingly; which is the default state for the parameter is not an issue for the language learner. However, these facts are a problem for the linguist aiming at a consistent description of syntactic variation in terms of [+]s and [–]s, which may be a necessity in case the system is implemented in a full-fledged computational algorithm or in a tool for language comparison[46]. We are aware of this inconsistency in the model's architecture, and we suggest that a promising direction for investigating it may be taking into account that parameters

[42] Intended purely as a comparative linguist's notational device with no status for the language learner, see § 4.2.

[43] See the Questions relative to Parameter *NØD* in the Supplementary Materials.

[44] Obviously, not for both values in the same language.

[45] *NØD* in our collection.

[46] As e.g. in Ceolin *et al.* (2020).

that do not seem to have an obvious default value are headedness parameters: this category has been shown to be special in other respects, for it seems to be sensitive to prosodic cues and to be set very early in acquisition[47]. We leave this topic for future research.

Note that headedness parameters pose a problem for the model's architecture because of the *conjunction* of (12a) and (12b). A parameter with property (12a) alone is not a problem, for our model only requires that each parameter is associated with positive evidence that set [+ *parameter P*], but makes no prediction as to the kind of data associated with [– *parameter P*], the notational metaphor used by linguist for structures/features that are *not* part of the I-language in question. Actually, parameters with property (12a) but not (12b) are rather common in our dataset, and for certain purposes this property is more a blessing than a problem. Consider in fact that finding positive evidence for [– *parameter P*] can be of great practical value to the linguist working with native informants, and, crucially, with closed *corpora*: in fact, encountering evidence for [– *parameter P*] means that the search of *p-expressions* for [+ *parameter P*] can be terminated. Therefore, we call *Stop Questions* the questions meant to identify *p-expressions* for [– *parameter P*][48]. However, there is no need to assume that *Stop Questions* have any part in the child's acquisition process, because, different from what happens with headedness parameters, for these parameters it is rather straightforward to motivate the identification of [+] and [–] values, as we will show directly.

For example, one of the core manifestations to set parameter *FGN* [± *grammaticalized Number*][49] to [+] is the presence of visible agreement in Number between a singular/non-singular nominal argument and the verb. However, in languages that are [– *grammaticalized Number*] there is positive evidence that [± *grammaticalized Number*] cannot be set to [+]: in these languages the same noun phrases are systematically associated to both singular and plural interpretations. Similarly, the most salient manifestation of parameter *FSN* [± *Number spread to N*][50] is the presence of systematic number exponence on the head noun and not just on determiners; languages which set the value [–] for this parameter use the same

[47] See Biberauer (2019: 52-53) and references cited.

[48] Where available, in the Supplementary Materials they are listed after the questions that set each parameter to [+].

[49] Which was mentioned above as an example of schema (1b): 'is α grammaticalized?'.

[50] Which was mentioned above as an example of schema (3a): 'is α morphologically spread to positions where it is not interpreted?'.

head nouns with both singular and plural determiners. In both cases, it is natural to assume that the [+] value is associated with the presence (vs absence) of visible morphological marking and contrasts: upon encountering the relevant morphemes in the *PLD*, the learner adds [+ *grammaticalized Number*] or [+ *Number spread to N*] to his/her grammar[51]. The existence of positive evidence for [–], which amounts to morphological identity of singular and plural forms, can naturally be taken to play no role in language learning.

Finally, there are parameters which straightforwardly identify a default value, because they do not seem to have an easily conceivable *Stop Question*: that is, they only have *p-expressions* for one value, which then will be [+], with the other being the default value. Many parameters responsible for the presence/absence of overt movement (schema (3b)), along the model inaugurated by Huang (1981), seem to belong to this group[52]. Often, only one of the two states of the choice has unambiguous manifestations. The examination of empirical cases in our sample suggests that the default value is the one producing overt movement (e.g. N raising over adjectives: see Bernstein, 1991; 1993; Crisma, 1991; 1996; Valois, 1991 and the subsequent debate); the lack of corresponding movement in other languages (presumably the possibility of replacing them by a covert Long-Distance relationship) represents the marked value, requiring positive evidence in the data.

Actually, the availability (or not) of *Stop Questions* is not a property of each single parameter, but rather of parameter types, largely corresponding to the schemata (1)-(3). For example, among those that we have not already exemplified above, the parameters corresponding to schema (1a), do not seem to have a conceivable *Stop Question* for principled reasons: being an *available* category can only be a positive specification, because something unavailable simply does not have a *p-expression*, essentially by definition; for instance, among the manifestations which set the value [+] of parameter *FRC* [± *finite relative clauses*] is the possibility, in a language, for the predicate of relative clauses to bear morphology specific to verbs. Yet, the

[51] Note also that, in the case of grammaticalization parameters (schema (1b)), it is simply unfeasible to attribute the grammaticalization of *all* possible formal features to the initial state S_0, with positive evidence used by learners to eliminate all those features that are irrelevant in each particular language, see for example parameter FGT/[± *grammaticalized temporality*] in the Supplementary Materials.

[52] With the notable exception of headedness parameters, discussed above.

presence of such a structure is not incompatible with other types of realization of relative clauses (e.g. nominalization strategies). Hence, the presence in a language of relative clauses not realized as finite clauses, by itself does not provide evidence for [– *FRC*]. Another clear case is that of two features *associated* on the same functional item of the lexicon (schema (2d)): there is good evidence that they may also occur separately in the same language (e.g. Spanish *mi* 'my' combines the possessive features present in the same language on *mio/mia* and the definite ones of *el/la*).

If all these preliminary considerations prove correct, they may lead to a principled understanding of the notion of default parametric values.

5.3. *Minimalist speculations*

A full minimalist reduction of the parameter apparatus to structures emerging only in the course of the learning path from S_0 to S_S requires two further steps, that we are pursuing in ongoing research.

One is a deductive reduction of schemata to ~~simpler axioms, i.e. to representing just~~ the restricted space of variation left free by simple principles on Merge and its mapping to the interface levels, provided by *UG* and perhaps third-factor conditions.

The other crucial step consists of not leaving it to *UG* to stipulate the long and apparently idiosyncratic list of implications or hierarchies among such parameters, which are listed in Table A in the Supplementary Materials, because at S_0 the parameter themselves would not yet be there: then, the implicational structure must also be derived from general principles of the initial state of the mind or the third factor.

As noted, these topics exceed the goals of this article, but the empirical success in formulating a realistic setting procedure, as achieved here, corroborates the assumed model of parameters which provides the testing ground for these remaining theoretical issues.

6. *Conclusions*

In this work, we outline a parametric model of syntactic diversity and learnability that approximates to a practical discovery procedure for grammars (parameter values). We prove its success across a large number of real-world languages.

Our parameter setting procedure relies on the use of positive evidence only, therefore it constitutes a plausible framework for child language acquisition studies; furthermore, it only uses a core subset of the positive evidence, the one which is presumably more common in the *PLD* (the *Restricted List*).

Our proposal achieves these results on account of two assumptions. First, we drastically reduce the number of parameters to be set in each individual language, since we argue that only one state of each parameter, coded as [+], needs to be set from the *PLD*. The other state, coded as [–], corresponds to the original state of the system before exposure to environmental evidence, i.e. it does not exist in the minds of speakers. Second, we further downsize the number of parameters to be set, showing that most parameters are simply irrelevant (non-existing choices, coded as [0]) in each given language, owing to the pervasive implicational structure of possible syntactic variability.

Finally, we draw attention to the fact that our model does not require the postulation of a predefined extensional list of open choices at the initial state of the mind. As such, it is compatible with a radically underspecified theory of the language faculty.

Acknowledgments

The authors have been partly supported by MIUR PRIN 2017K3NHHY *Models of language variation and change: new evidence from language contact* (P. Crisma and C. Guardiano) and by the ERC Advanced Grant 295733 *LanGeLin* (G. Longobardi and C. Guardiano). The data discussed in the paper and presented in the Supplementary Materials have been collected by the *LanGeLin* project, and are the property of the *LanGeLin* database. Preliminary versions of this material have been presented at conferences at Arizona State University (June 2019), UCLA (November 2019), and Universität Göttingen (December 2019). We thank the audiences of these conferences, in particular Judy Bernstein, David Goldstein, Hilda Koopman, Pam Munro, Anna Roussou, Stavros Skopeteas, Tim Stowell; we also thank the two anonymous referees for SSL. This work is the result of collaboration among the three authors in all its parts; for official purposes, P. Crisma takes responsibility for § 2; C. Guardiano for § 4; G. Longobardi for § 3. The authors take joint responsibility for §§ 1, 5 and 6.

References

Baker, C.L. (1989), *Some observations on degree of learnability*, in «Behavioral and Brain Sciences», 12, 2, pp. 334-335.

Baker, M.C. (2001), *The Atoms of Language: The Mind's Hidden Rules of Grammar*, Basic Books, New York.

Bernstein, J. (1991), *DPs in French and Walloon, evidence for parametric variation in nominal head movement*, in «Probus», 3, 2, pp. 101-126.

Bernstein, J. (1993), *Topics in the Syntax of Nominal Structure across Romance*, PhD dissertation, City University of New York, New York.

Biberauer, Th. (2008, ed.), *The Limits of Syntactic Variation*, John Benjamins, Amsterdam.

Biberauer, Th. (2019), *Factors 2 and 3: Towards a principled approac*h, in «Catalan Journal of Linguistics» (special issue ed. by Galliego, A. and Ott, D.), pp. 45-88 [available online at *https://doi.org/10.5565/rev/catjl.219*, accessed on 01.08.2020].

Biberauer, Th. and Roberts, I.G. (2017), *Parameter setting*, in Ledgeway, A. and Roberts, I.G. (2017, eds.), *The Cambridge Handbook of Historical Syntax*, Cambridge University Press, Cambridge, pp. 134-162.

Boeckx, C. and Leivadá, E. (2014), *On the particulars of universal grammar: Implications for acquisition*, in «Language Sciences», 46 B, pp. 189-198.

Borer, H. and Wexler, K. (1987), *The maturation of syntax*, in «Studies in Theoretical Psycholinguistics», 4, pp. 123-172.

Borer, H. and Wexler, K. (1992), *Bi-unique relations and the maturation of grammatical principles*, in «Natural Language and Linguistic Theory», 10, pp. 147-189.

Bortolussi, L. Longobardi, G., Guardiano, C. and Sgarro, A. (2011), *How many possible languages are there?*, in Bel-Enguix, G., Dahl, V. and Jiménez-López, M.D. (2011, eds.), *Biology, Computation and Linguistics*, IOS Press, Amsterdam, pp. 168-179.

Ceolin, A., Guardiano, C., Irimia, M.-A. and Longobardi, G. (2020), *Formal syntax and deep history*, in «Frontiers in Psychology» 11: 488871 [doi: 10.3389/fpsyg.2020.488871].

Chomsky, N. (1957), *Syntactic Structures*, Mouton de Gruyter, The Hague.

Chomsky, N. (1964), *Current Issues in Linguistic Theory*, Mouton de Gruyter, The Hague.

Chomsky, N. (1965), *Aspects of the Theory of Syntax*, MIT Press, Cambridge (MA).

Chomsky, N. (1981), *Lectures on Government and Binding*, Foris, Dordrecht.

Chomsky, N. (1986), *Knowledge of Language*, Praeger, New York.

Chomsky, N. (2005), *Three factors in language design*, in «Linguistic Inquiry», 36, 1, pp. 1-22.

Clark, R. (1992), *The selection of syntactic knowledge*, in «Language Acquisition», 2, 2, pp. 83-149.

Clark, R. and Roberts, I.G. (1993), *A computational model of language learnability and language change*, in «Linguistic Inquiry», 24, 2, pp. 299-345.

Crain, S. and Thornton, R. (2015), *3rd year grammar*, in Gallego, Á. and Ott, D. (2015, eds.), *50 Years Later: Reflections on Chomsky's* Aspects, MIT Working Papers in Linguistics, Cambridge (MA), pp. 71-83.

Crisma, P. (1991), *Functional Categories inside the Noun Phrase. A Study on the Distribution of Nominal Modifiers*, Tesi di laurea, Università di Venezia.

Crisma, P. (1996), *On the configurational nature of adjectival modification*, in Zagona, K. (1996, ed.), *Grammatical Theory and Romance Languages* (Current Issues in Linguistic Theory, 133), John Benjamins, Amsterdam / Philadelphia, pp. 59-71.

Crisma, P., Guardiano, C. and Longobardi, G. (in press), *Toward a unified theory of Case form and Case meaning*, in Anagnostopoulou, E., Mertyris, D. and Sevdali, Ch. (in press, eds.), *The Place of Case in Grammar*, Oxford University Press, Oxford.

Crisma, P. and Longobardi, G. (in press), *The parametric space associated with D*, in Armoskaite, S. and Wiltschko, M. (in press, eds.), *The Oxford Handbook of Determiners*, Oxford University Press, Oxford.

Delfitto, D. and Schroten, J. (1991), *Bare plurals and the number affix in DP*, in «Probus», 3, 2, pp. 155-185.

Fassi Fehri, A. (1993), *Issues in the Structure of Arabic Clauses and Words*, Kluwer, Dordrecht.

Fodor, J.D. (1998), *Unambiguous triggers*, in «Linguistic Inquiry», 29, 1, pp. 1-36.

Fodor, J.D. (2001), *Setting syntactic parameters*, in Baltin, M. and Collins, C. (2001, eds.), *The Handbook of Contemporary Syntactic Theory*, Blackwell, Oxford, pp. 730-767.

Fodor, J.D. and Sakas W.G. (2017), *Learnability*, in Roberts, I.G. (2017, ed.), *The Oxford Handbook of Universal Grammar*, Oxford University Press, Oxford, pp. 249-269.

Gianollo, C., Guardiano, C. and Longobardi, G. (2008), *Three fundamental issues in parametric linguistics*, in Biberauer, Th. (2008, ed.), *The Limits of Syntactic Variation*, John Benjamins, Amsterdam, pp. 109-141.

Gibson, E. and Wexler, K. (1994), *Triggers*, in «Linguistic Inquiry», 25, 3, pp. 407-454.

Guardiano, C. and Longobardi, G. (2017), *Parameter theory and parametric comparison*, in Roberts, I.G. (2017, ed.), *The Oxford Handbook of Universal Grammar*, Oxford University Press, Oxford, pp. 377-398.

Haspelmath, M. (2013), *Occurrence of nominal plurality*, in Dryer, M.S. and Haspelmath, M. (2013, eds.), *The World Atlas of Language Structures Online*, Max Planck Institute for Evolutionary Anthropology, Leipzig [available online at *http://wals.info/chapter/34*, accessed on 01.08.2020].

Hauser, M.D., Chomsky, N. and Fitch, T.W. (2002), *The faculty of language: What is it, who has it, and how did it evolve?*, in «Science», 298, pp. 1569-1579.

Hornstein, N. and Lightfoot, D. (1981, eds.), *Explanation in Linguistics: The Logical Problem of Language Acquisition*, Longman, London.

Huang, C.-T.J. (1982), *Logical Relations in Chinese and the Theory of Grammar*, PhD Dissertation, MIT.

Hyams, N. (1986), *Language Acquisition and the Theory of Parameters*, Reidel Publishing Company, Dordrecht.

Kayne, R.S. (1994), *The Antisymmetry of Syntax*, MIT Press, Cambridge (MA).

Kayne, R.S. (2000), *Parameters and Universals*, Oxford University Press, New York.

Kuroda, S.-Y. (1988), *Whether we agree or not: A comparative syntax of English and Japanese*, in «Linguisticae Investigationes», 12, pp. 1-47.

Lasnik, H. and Lidz, J. (2017), *The argument from the poverty of the stimulus*, in Roberts, I. (2017, ed.), *The Oxford Handbook of Universal Grammar*, Oxford University Press, Oxford, pp. 221-248.

Lightfoot, D. (1979), *Principles of Diachronic Syntax*, Cambridge University Press, New York / London.

Lightfoot, D. (1989), *The child's trigger experience: Degree-0 learnability*, in «Behavioral and Brain Sciences», 12, 2, pp. 321-334.

Lightfoot, D. (2006), *How new languages emerge*, Cambridge University Press, Cambridge.

Lightfoot, D. (2017), *Discovering new variable properties without parameters*, in «Linguistic Analysis» (special issue ed. by Karimi, S. and Piattelli-Palmarini, M.), 41, 3-4, pp. 409-444.

Longobardi, G. (1994), *Reference and proper names: A theory of N-Movement in syntax and logical form*, in «Linguistic Inquiry», 25, 4, pp. 609-665.

Longobardi, G. (2003), *Methods in parametric linguistics and cognitive history*, in «Linguistic Variation Yearbook», 3, pp. 101-138.

Longobardi, G. (2005), *A Minimalist Program for Parametric Linguistics?*, in Broekhuis, H., Corver, N., Huybregts, M., Kleinhenz, U. and Koster, J. (2005, eds.), *Organizing Grammar: Linguistic Studies for Henk van Riemsdijk*, Mouton de Gruyter, Berlin / New York, pp. 407-414.

Longobardi, G. (2014), *Theory and experiment in Parametric Minimalism: The case of Romance negation*, in Pensalfini, R., Turpin, M. and Guillemin, D. (2014, eds.), *Language Description Informed by Theory*, John Benjamins, Amsterdam / New York, pp. 217-261.

Longobardi, G. (2017), *Principles, parameters and schemata. A radically underspecified UG*, in «Linguistic Analysis» (special issue ed. by Karimi, S. and Piattelli-Palmarini, M.), 41, 3-4, pp. 517-557.

Longobardi, G., Guardiano, C., Silvestri, G., Boattini, A. and Ceolin, A. (2013), *Toward a syntactic phylogeny of modern Indo-European languages*, in «Journal of Historical Linguistics», 3, pp. 122-152.

Manzini, M.R. (2019), *Parameters and the design of the Language Faculty. Northern Italian partial null subjects*, in «Evolutionary Linguistic Theory», 1, 1, pp. 23-56.

Manzini, R.M. and Wexler, K. (1987), *Parameters, binding theory, and learnability*, in «Linguistic Inquiry», 18, 3, pp. 413-444.

Pollock, J.-Y. (1989), *Verb movement, UG, and the structure of IP*, in «Linguistic Inquiry», 20, pp. 365-424.

Rizzi, L. (1980), *Violations of the* wh *island constraint in Italian and the Subjacency Condition*, in «Journal of Italian Linguistics», 5, pp. 157-195.

Rizzi, L. (1982), *Issues in Italian Syntax*, Foris, Dordrecht.

Rizzi, L. (1986), *Null Objects in Italian and the theory of* pro, in «Linguistic Inquiry», 17, 3, pp. 501-557.

Roberts, I.G. (2007), *Diachronic Syntax*, Oxford University Press, Oxford.

Roberts, I.G. (2019), *Parameter Hierarchies and Universal Grammar*, Oxford University Press, Oxford.

Roeper, Th. (1999), *Universal Bilingualism*, in «Language and Cognition», 2, pp. 169-186.

Sakas, W.G., Yang, C. and Berwick, R.C. (2017), *Parameter setting is feasible*, in «Linguistic Analysis» (special issue ed. by Karimi, S. and Piattelli-Palmarini, M.), 41, 3-4, pp. 391-408.

Sportiche, D. (1986), *ZIBUN*, in «Linguistic Inquiry», 17.2, pp. 369-374.

Taraldsen, K.T. (1978), *On the NIC, vacuous application, and the that-t filter*, Ms. MIT, Indiana University Linguistics Club, Bloomington (IN).

Thornton, R. and Tesan, D. (2007), *Categorical acquisition: Parameter-setting in Universal Grammar*, in «Biolinguistics», 1, pp. 49-98.

Valois, D. (1991), *The internal syntax of DP and adjective placement in French and English*, in Shearer, T. (1991, ed.), *Proceedings of NELS 21*, Graduate Student Linguistic Association, University of Massachusetts, Amherst, pp. 367-381.

de Villiers, J. (2001), *Continuity and modularity in language acquisition and research*, in «Annual Review of Language Acquisition», 1, pp. 1-64.

Wexler, K. (1994), *Optional infinitives, head movement and the economy of derivations*, in Lightfoot, D. and Hornstein, N. (1994, eds.), *Verb Movement*, Cambridge University Press, Cambridge, pp. 305-350.

Yang, Ch. (1999), *Unordered merge and its linearization*, in «Syntax», 2, pp. 38-64.

Yang, Ch. (2002), *Knowledge and learning in natural language*, Oxford University Press, Oxford.

Yang, Ch. (2004), *Universal Grammar, statistics, or both*, in «Trends in Cognitive Sciences», 8, 10, pp. 451-456.

Zanuttini, R. (1997), *Negation and Clausal Structure: A Comparative Study of Romance Languages*, Oxford University Press, Oxford.

Paola Crisma
Dipartimento di Studi Umanistici
Università di Trieste
Androna Campo Marzio 10
34123 Trieste (Italy)
paola.crisma@units.it

Cristina Guardiano
Dipartimento di Comunicazione ed Economia
Università di Modena e Reggio Emilia
Viale Allegri 9
42121 Reggio Emilia (Italy)
cristina.guardiano@unimore.it

Giuseppe Longobardi
Department of Language and Linguistic Science
University of York
Vanbrugh College
YO10 5DD Heslington (United Kingdom)
giuseppe.longobardi@york.ac.uk