Chapter 2

Reconciling linguistic theories on comparative variation with an evolutionarily plausible language faculty

Kleanthes K. Grohmann
University of Cyprus, Cyprus Acquisition Team

Evelina Leivada
UiT The Arctic University of Norway, Cyprus Acquisition Team

This work aims to reconcile the atomic objects of study typically assumed within comparative variation studies with an evolutionarily plausible faculty of language. In the process, we formulate and address the *incompatibility problem*, the observation that studying comparative (micro)variation has progressively led to an evolutionarily implausible Universal Grammar. We identify a solution to this problem through arguing in favour of a so-called emergentist approach to some linguistic primitives. We then address the *granularity mismatch problem* and argue on the basis of this emergentist approach firstly, that linguistic and neurocognitive studies of language may be brought to the same level of granularity, and secondly, that specific insights from comparative variation can inform an evolutionarily plausible approach to human language.

1 Introduction

The topic of language variation and how it informs our study of the faculty of language (FL) together with its initial state are currently at the forefront of linguistic research (for latest overviews, see e.g. Hinzen 2014; Trettenbrein 2015;

Kleanthes K. Grohmann & Evelina Leivada. 2020. Reconciling linguistic theories on comparative variation with an evolutionarily plausible language faculty. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 25–42. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280629

Berwick & Chomsky 2016). As a matter of fact, the exploration of variation from a comparative, cross-linguistic perspective can be considered one of the very few topics which both linguists and cognitive neuroscientists agree merits further attention.

A representative perspective of the first area of research is that of generative linguist Noam Chomsky. When asked in a recent interview what the main advantages and/or reasons to study linguistic variation are, he reiterated a view that has been repeatedly explored in his work: In order to determine the capacity to use and understand language, we need to know "what options it permits" (Chomsky 2015). Put differently, if we want to understand FL and its initial state, Universal Grammar (UG), we must determine what structures UG is capable of generating. In the same vein, we should also determine what structures UG is *not* capable of generating as striking typological gaps across phylogenetically diverse languages call for explanations that can enrich our theory of language (see Biberauer, Holmberg & Roberts 2014 for a concrete example). From a linguistic perspective, we will call this the "insider" view.

To pursue the analogy, the perspective of cognitive neuroscientist Peter Hagoort can be described as the "outsider" view. Hagoort devoted part of his plenary talk at the 47th annual meeting of the Linguistic Society of Europe to how linguistics, once seen as a key player in the field of cognitive science, has seen its influence fade over the years (Hagoort 2014). This alienation directly relates to how linguists have presented their discoveries in the study of language variation. Often linguists have captured aspects of comparative variation through postulating primitives that they did not grow or derive in any sense, typically by assuming that a UG-encoded feature drives the relevant linguistic representation. Such postulations cannot be informative in the long run. Perhaps they can be successfully employed when one deals with some language A or B, but when the aim is broader (e.g., to approach our language-readiness and UG as its initial state), then such postulations are rather impeding progress.

In this context, the two most important questions to be addressed are (i) why this alienation across disciplines is happening and (ii) whether there is a remedy for this situation. The second question is the topic of §2. With respect to the first question, it seems that the reason is in part the way the topic of language variation has been approached over the last few years. More specifically, discussing comparative syntax and the way parametric models capture variation (see, for example, the recent collection of papers in Fábregas et al. 2015), Biberauer, Holmberg, Roberts & Sheehan (2014) argue that linguistic descriptions that have emerged since Chomsky (1981) have achieved an increasingly high level of

descriptive adequacy, but sacrificed explanatory adequacy due to the postulation of more and more entities in UG. In their words:

> Arguably, the direction that [principles & parameters] (P&P) theory has taken reflects the familiar tension between the exigencies of empirical description, which lead us to postulate ever more entities, and the need for explanation, which requires us to eliminate as many entities as possible. In other words, parametric descriptions as they have emerged in much recent work tend to sacrifice the explanatory power of parameters of Universal Grammar in order to achieve a high level of descriptive adequacy. (Biberauer, Holmberg, Roberts & Sheehan 2014: 104)

Describing linguistic data and formulating observations or generalisations over these data may then offer observational adequacy, possibly even descriptive adequacy, but not explanatory adequacy.

Although Biberauer, Holmberg, Roberts & Sheehan's point is well-taken, it is only a part of the issue at hand. Another part is presented by Yang (2004) when he writes that

> adult speakers, at the terminal state of language acquisition, *may retain multiple grammars, or more precisely, alternate parameter values*; these facts are fundamentally incompatible with the triggering model of acquisition […] *It is often suggested that the individual variation is incompatible with the Chomskyan generative program.* (Yang 2004: 50–51)

We can thus phrase the full problem as follows:

(1) *The incompatibility problem*: Studying microvariation has led to a model entailing an evolutionarily implausible UG/FL.

Put differently, we have managed to describe many linguistic structures across different languages, but now we have trouble explaining the ontology of the biological "structure" underlying their existence: UG. Given the short time scale typically assumed for evolution, the higher the degree of linguistic specificity encoded in UG, the more difficult the task of accounting for it in evolutionary terms.

Reconciling a bottom-up approach to UG and a resulting evolutionarily plausible FL with the findings from the literature on language variation has the potential to solve not only the incompatibility problem but also *Poeppel's problem*. More specifically, this reconciliation can overcome the granularity mismatch

considerations according to which linguistic and neuroscientific studies of language operate with objects of different granularity in a way that makes the construction of interdisciplinary bridges particularly difficult (cf. the granularity mismatch problem in Poeppel & Embick 2005). A bottom-up approach to UG entails a non-overarticulated UG which consists of a few computational principles (as Di Sciullo et al. 2010 have argued) only, leaving outside of this component many of the linguistic primitives that have been ascribed to it within comparative variation studies.

In this context, the next section discusses the importance of studying variation from a comparative, cross-linguistic perspective while at the same time maintaining a bottom-up approach to UG (i.e. an approach to UG from below that seeks to ascribe to it as little as possible, while maximizing the role of the other two factors in language design; Chomsky 2007). Pursuing a bottom-up vs. a top-down approach matters because depending on how much one ascribes to UG, the plausibility of the latter from an evolutionary perspective changes significantly. Our main aim is to offer the following solution to the incompatibility problem: An emergentist approach to some UG primitives can reconcile the Chomskyan generative program and the individual variation attested in reality. §3 then aims to offer a concrete demonstration of how relevant findings and primitives from the field of language variation can inform a biological approach to human language. §4 concludes and presents some suggestions for future work on this topic.

2 An emergentist approach to UG primitives

The second question that arose in the context of Hagoort's view on the interaction of linguistics with the larger field of cognitive science is whether there is a remedy for the observed decreased influence of linguistics. Hagoort (2014) offers five different directions for rectifying this issue. We apply some of these directions through pursuing an approach to UG primitives from below (Chomsky 2007), while at the same time retaining in our theory of FL some of the theoretical notions that pertain to the comparative variation literature. This combination has the potential of killing two birds with one stone, solving not only the incompatibility problem but also doing justice to the patterns of (micro)variation that are attested across languages in the following, two-step way:

I. Disentangling variation by teasing apart the different contributing factors which are responsible for deriving it in a way that does justice to sociolinguistic and psycho-/neurolinguistic aspects of language use, such as mono- vs. bilingual acquisition trajectories, the sociolinguistic status of the linguistic input, and the non-linguistic part of the environment.

II. Keeping UG primitives to a minimum in order to effectively comply with both minimalist principles and evolutionary constraints.

Point (I) has a second part that will not be addressed in this paper but that should be kept in mind nevertheless if the goal is to construct interdisciplinary bridges and overcome the granularity mismatch problem: Embedding the theory of language variation that emerges from step (I) into a "shared context of justification" (Hagoort 2014) by obtaining reliable data from different language groups, each of which may contribute its own characteristics towards deriving variation.[1] In practice, this would mean that careful elicitation of data should be followed by an attempt to interpret the data through *deriving* their properties rather than assuming that they are driven by a new, ad hoc postulated feature. If the aim is to understand FL rather than describe structure A in language B, then this process of interpretation should also be cautious to not rely on assumptions that are hard to sustain in the long run and quickly decompose under the light of interdisciplinary examination.

Talking about different contributing factors in (I) boils down to realising that variation across developmental paths of individuals that speak the same language can be the outcome of different modalities, environmental factors, non-linguistic features that affect linguistic development, and so on. For instance, research has shown that non-standard varieties allow for greater grammatical fluidity in a way that blurs the boundaries across different varieties. This, in turn, affects speakers' perceptions of whether a specific variant belongs to their linguistic repertoire or not (Cheshire & Stein 1997; Henry 2005). Another contributing factor is the trajectory of language acquisition and subsequent development, and the circumstances in which it takes place. For example, non-heritage speakers of a language may differ from heritage speakers of the *same* language with respect to the amount of variation attested in their repertoire (Montrul 2002; 2008; Lohndal & Westergaard 2016). The sociolinguistic status of the language(s) one is exposed to (the mono- vs. bilingual trajectory is in and of itself another factor that leads to variation) is yet another potential source of variation: In the case of non-standard varieties, speakers' perceptions about their native grammatical

[1] Hagoort (2014) argues that running sentences in one's head and consulting a colleague is fine for discovering interesting phenomena and possible explanations (the "context of discovery"), but it does not suffice as "the context of justification", due to innate confirmation biases and the fallibility of introspection. Thus, "to justify one's theory, empirical data have to be acquired and analysed according to the quantitative standards of the other fields of cognitive science". In the context of addressing the incompatibility problem, Hagoort's perspective is relevant because it shows how findings that may target points of grammatical (micro)variation should be analysed and interpreted.

variants are likely to be affected by their knowledge that many of their dialectal structures are considered unacceptable or "incorrect" by speakers of the standard variety (Henry 2005 for Belfast English; Leivada, Papadopoulou, Kambanaros, et al. 2017 for Cypriot Greek) in a way that enhances grammatical fluidity. Also, in those cases in which a standard variety co-exists with a structurally proximal, non-standard variety, the discreteness across grammatical variants at times fades away by the emergence of intermediate (Cornips 2006) or "diaglossic" speech repertoires (Auer 2005), resulting once more in a greater degree of variation (see also Rowe & Grohmann 2014 and relevant references cited for Cypriot Greek).

Understanding the multitude of faces that variation can acquire (for a more extensive overview, see Leivada 2015a) is of key importance when it comes to approaching UG primitives from an emergentist perspective. The reason is that cross-linguistic variation has long been described as part of UG, that is, deriving from UG parameters. Showing that patterns of variation are not as stabilised or uniform as the traditional UG parameters-account predicts opens the way for an emergentist approach to linguistic primitives that were traditionally viewed as part of UG. Understanding what terms like "stabilised" or "uniform" refer to in the present context requires shifting our attention to how variation *within* linguistic communities has been approached.

A crucial challenge for any approach to variation derives from the mainstream conception of the notion of "surface variation" (i.e. grammatical variation among speakers of the same language that is not the result of any acquired or developmental pathology) *within* a linguistic community. For example, Chomsky's idealised picture of a "completely homogeneous speech community" and an "ideal speaker-listener [...] who knows its language perfectly" (Chomsky 1965: 3) is often assumed together with the assumption that the so-called "linguistic genotype" is uniform across the species in the absence of severe and specific pathology (Anderson & Lightfoot 2000). Another related idea is that attained adult performance is "essentially homogeneous with that of the surrounding community", unless again a pathology is present (Anderson & Lightfoot 2000: 698). When translated into empirical terms, idealisations like these, although theoretically well-argued in their original context, paint a picture directly related to both Hagoort's and Poeppel's considerations. More specifically, by not doing justice to the patterns of surface variation that are attested in reality, theoretical linguistics may *lose* a significant part of its potential for interactions with fields that deal with recent sign language emergence, evolutionary linguistics, or sociolinguistics. Despite what the idealised picture suggests, variation can be found even in the absence of any pathology, even among speakers of the same language, and even within a native speaker who has passed the L1 acquisition period. The core

of this idea can be analysed across two dimensions, the linguistic dimension and the developmental one.

The developmental dimension refers to the fact that the presence of a severe and specific pathology is not a necessary condition for obtaining variation, even among neurotypical speakers of the same language. Individuals that share a diagnosis of cognitive disorder (or the absence of one) are not necessarily uniform in terms of their innate endowment: Individuals with a pathogenic variant of a gene can be impaired in a non-uniform fashion (variable expressivity), which may result in different cognitive phenotypes at times not reaching a cut-off point where the diagnosis of a specific pathology is possible. To demonstrate this with two examples, Fowler (1995) observes that there is tremendous variability with regard to language function in individuals with Down syndrome (variable expressivity). And it has also been observed that the existence of subsyndromal schizotypal traits in the general population is higher than average in first-degree relatives of patients with schizophrenia (Calkins et al. 2004). This led to the realisation that

> schizophrenia is not, despite its clinically important and reliable categorical diagnosis […], a binary phenotype (present, absent) with sudden disease onset. (Ettinger et al. 2014: 1)

In other words, some pathological characteristics might be present even if the cut-off point for reaching a diagnosis is not met – and, on the other hand, a diagnosis of schizophrenia might be reached, even if the pathological characteristics manifested among individuals with the same diagnosis are far from uniform. Together, these two examples suggest that it is equally plausible to expect that attained adult performance is not uniform among members of the same linguistic community in the absence of a pathology or in the presence of the same pathology.

With respect to the linguistic dimension, this is where factors related to non-standard varieties and inherent grammatical fluidity enter the picture. Evidently, not all linguistic communities are homogeneous, and in many cases this variation goes well beyond bi- or multilingualism. Similarly, in the case of recent language emergence de novo, as in the case of Al-Sayyid Bedouin Sign Language (ABSL) and other sign languages, fieldwork has shown that not only is the development of grammatical markers subject to environmental factors (e.g., time, distribution of speakers/signers, etc.), but also that great grammatical fluidity is attested at the various stages in the development of a language. In these recently emerged languages, points of variation ("parameters" in generative terms) are *not* fixed in

terms of their values, resulting in the realisation of alternate settings both within and across speakers (Washabaugh 1986; Sandler et al. 2011).

To mention a concrete example, consider the head-directionality parameter. S(ubject) O(bject) V(erb) is the prevalent word order among ABSL signers; this was, however, established as the prevalent order from the second generation of signers onwards only (Sandler et al. 2005), meaning that for some time the manifestations of this "parameter" were more fluid than what a stabilised parameter value would permit. Even more important is the fact that variation exists past the "stabilisation" point: Sandler et al. (2005: 2663) report the existence of some (S)VO patterns. As Leivada (2015a) argues in her discussion of ABSL, the fact that SOV patterns became robust in the second generation of speakers illustrates that variation is present when certain grammatical properties are still emerging. Fluctuating parameter values within a syntactic environment are incompatible with the idea that a parameter value is fixed past the terminal state of acquisition. Observing that this fluctuation exists in various cases, be it non-standard varieties or recently emerged grammars, is an indication that the head-directionality parameter "should indeed be better viewed as a surfacey decision that allows for varying realizations, rather than a fixed, deeply rooted syntactic parameter" (Leivada 2015a: 48). This does not mean that points of variation are unfixed and eventually culminate in an "anything goes" grammar, but it does mean that this surface decision is not (i) syntactic (i.e. Chomsky in recent work has explicitly recognized that variation between grammars is a matter of variable externalization; see Berwick & Chomsky 2011: 41), (ii) UG-encoded, or (iii) binary, as the classical parametric approach would suggest. Non-binarity is particularly evident in case of bidialectal speakers; their linguistic repertoire may include functionally equivalent variants (Kroch 1994) with *different* values that are alternatively realized in the *same* syntactic environment (Leivada, Papadopoulou & Pavlou 2017).

An emergentist approach to some linguistic primitives that were previously thought to be parts of UG will be able to reconcile the Chomskyan generative program (and especially UG, as one of its main pillars) with the patterns of variation that are attested in reality (see Yang's 2004 point mentioned earlier). Moreover, an emergentist approach will solve the incompatibility problem, as the number of linguistic primitives allocated to UG will be reduced. The notion of *emergent parameters* (Roberts & Holmberg 2010; Roberts 2012; Biberauer, Roberts & Sheehan 2014; Biberauer & Roberts 2017) is an important step in this direction. The central idea behind emergent parameters is that instead of postulating a richly specified parametric endowment as part of the initial state of our FL (UG; Chomsky 1981), parameters are derived (i.e. emergent) properties falling out of the interaction of

Chomsky's (2005) three factors in language design (Biberauer, Holmberg, Roberts & Sheehan 2014). In the context of emergent parameters in which UG does not provide a pre-specified "menu" of parametric choices, Biberauer, Roberts & Sheehan (2014) note that it is very important to provide independent motivation for the plausibility of the parameters that acquirers will postulate as well as for the sequence in which each point of variation should be considered. Here lies the solution to the incompatibility problem and a first step towards approaching the granularity mismatch problem.

With respect to the incompatibility problem, if the points of variation that are meaningful from a comparative (micro)variation perspective are treated as emergent properties, they are no longer translated as innately specified options. The consequence of this move is that UG would be considerably deflated and much easier to discuss from an evolutionary perspective. As Chomsky (2007) has very convincingly argued, for any given component or structure, the less attributed to structure-specific factors for determining the development of an organism, the more feasible the study of its evolution, hence the need for a bottom-up approach to UG.

In relation to the granularity mismatch problem, the important component of the "emergent parameters"-account lies in the element of *interaction*. As Biberauer, Roberts & Sheehan (2014) explicitly claim, it is the interaction of the second factor (linguistic input) and the third factor (non-language-specific principles of cognition) plus the language-readiness (provided by the first factor, UG); that delivers emergent parameters. To illustrate this with an example, let's return to the head-directionality parameter, which makes reference to the position of a head in relation to its dependents. Traditional accounts of grammar would describe Japanese as a head-final and English as a head-initial language, with the difference between the two explained in terms of the different value to which the head-directionality parameter is set. The typological preference given to harmonic orders (i.e. *consistent* head-initial or head-final patterns within a language; see Hawkins 2010) might also be taken to suggest that a UG-based head-directionality parameter is indeed operative and, once set, its effects are diffused across different syntactic environments.[2] Alternatively, one could argue that the realisation of the head in relation to its dependents does not boil down to setting a UG-based parameter. This latter approach should be preferred because it is compatible with the fact that variation *can* be attested past the "setting" state in the repertoire of a neurotypical, adult speaker who has fully acquired her language

[2] A reviewer points out that this is not assumed within the emergentist approach just outlined. Indeed, it is not and we do not embrace this explanation either; we only point out that it is an alternative explanation, which, however, should not be preferred, since it does not accommodate the patterns of variation that are attested.

(as suggested in the case of ABSL). If one chooses to approach this parameter as an emergent parameter, the interaction of this grammatical choice with principles of general cognitive architecture becomes meaningful. For example, why are harmonic orders preferred if they are not *imposed* by the setting of a predetermined parameter? Of course, an emergent parameter would also need to be "set" in order to reflect the options that are permitted in the adult grammar, but crucially by not being encoded in UG, its variable realizations within and across speakers of the same language (e.g., in the form of functionally equivalent variants; Leivada, Papadopoulou & Pavlou 2017) would not be a problem for our theory of UG and/or FL.

Roberts (2016b) suggests that these generalisation effects are related to the computational conservatism of the learning device. This is formally captured by his *input generalisation*: "There is a preference for a given feature of a functional head F to generalise to other functional heads G, H …" (cf. Roberts 2007: 275) – that is, to "maximise available features" (Biberauer & Roberts 2016; Roberts 2016b). This computational conservativism is a third factor principle. If so, preference for harmonic orders no longer amounts to a UG-wired principle or parameter, but to the way human memory or even learning more broadly works. It has been shown that sequence edges are particularly salient positions and facilitate learning in a way that gives rise to *either* word-initial *or* word-final processes much more often than otherwise (see, for example, Endress et al. 2009 on the prevalence of prefixing and suffixing across languages in comparison to the rarity of infixing). At the syntactic level, Dryer (1992) observes the following correlation with respect to generalisation effects in relation to the position of the Head on the basis of 434 languages: OV languages are mostly postpositional and VO languages are mostly prepositional. From Dryer's dataset, Hawkins (2010) calculated that the vast majority of languages (93%) are consistently OV-postpositional or VO-prepositional. Hawkins (2010) approaches harmonic word-orders in terms of third factor demands, and, more specifically, a processing preference that favours shorter processing domains. Evidently, the workings of comparative (micro)variation which deal with headedness patterns across typologically different languages can now be revisited and explained from a different perspective. This perspective involves the *interaction* of linguistic patterns with the driving forces of general cognition in a way that addresses Hagoort's considerations. With respect to the "messy" patterns of variation that just do not fit in the classical notion of a binary parameter, but that are just as uncontroversially there, an emergentist approach has the potential to cover these too. If parameters are emergent and allow for non-binary realizations, then the incompatibility that Yang (2004) correctly observes between these "messy" patterns and UG disappears.

Despite its theoretical and empirical benefits, this interaction may not solve the *granularity mismatch problem*. It may contribute to the construction of interdisciplinary bridges in some respects, but still a good portion of primitives may be left unmapped across disciplines. Put differently, even if parameters or other linguistic primitives are explained through an emergentist approach, this would not entail that the granularity mismatch problem has been solved. This could be due to the complicated nature of the task at hand; as Hornstein (2009: 156–157) argues, "the right theory of grammar will be one that has (roughly) the empirical coverage of [government-and-binding theory], *and* that 'solves' Plato's problem, Darwin's problem, *and* the granularity mismatch problem" (emphasis added).[3] In other words, given how polylithic both the problem and its solutions are, there can be no a priori guarantee of success. Despite recognising this possibility, the next section will follow Hagoort's (2014) suggestion to maximise the interdisciplinary contributions of linguistics within a larger cognitive (neuro)science environment. We endeavour to approach a constraint, which in the linguistics literature has been called "linguistic" or "syntactic" more often than not, in neurocognitive terms.

3 Levels of granularity: Anti-identity as a case study

Anti-identity has received many distinct names in the linguistics literature; consider, for example, the *obligatory contour principle* in phonology (Odden 1986), *identity avoidance* (van Riemsdijk 2008), *distinctness* (Richards 2010), *X-within-X recursion* (Arsenijević & Hinzen 2012). This is also the basis for *anti-locality* relations in syntax (Grohmann 2003, recently surveyed with additional references in Grohmann 2011). Regardless of the level of linguistic analysis at stake, anti-identity in general describes the absence of adjacent elements of the same category (e.g., [*XX] in syntax).

There are different ways to approach this phenomenon. In the linguistics literature, it has been approached in terms of a UG-imposed well-formedness ban that precludes the adjacency of same-category elements (see Richards 2010 for a more detailed discussion). This position would place the ban in UG, together with the configurations of categorial features that the ban is sensitive to. Alternatively, one could aim to keep UG at a minimum and see whether [*XX] can be shown to boil down to a general, cognitive principle. A first step in this direction

[3] According to Hornstein (2009), Darwin's problem refers to "the logical problem of language evolution", how language emerged in the species (see also Boeckx & Grohmann 2007 on the relation between Plato's problem and Darwin's problem).

is made by van Riemsdijk (2008) when he briefly argues that identity avoidance might be "a general principle of biological organization" (p. 242). If so, one expects to find its manifestations not only in language, but also in other domains of cognition.

Taking one step back, if this comparison across cognitive domains is fruitful, one would have successfully mapped an element that appears in the "parts list" (i.e. a list that enumerates concepts canonically used in the fields of study it represents; see Poeppel & Embick 2005) of two different disciplines. In more recent work, Poeppel (2012) talks about the *mapping problem*. In his words, the mapping problem "addresses the relation between the primitives of cognition (here speech, language) and neurobiology. Dealing with this mapping problem invites the development of linking hypotheses between the domains" (Poeppel 2012: 34). Developing these linking hypotheses is the only route to potentially solving the granularity mismatch problem. Returning now to the case at hand, linking hypotheses *can* be constructed for [*XX].

It seems to be true that humans do not like repetitions in general and that anti-identity in language is not the result of a linguistic ban but of a bias that finds application in other domains of human cognition too. Walter's (2007) biomechanical repetition avoidance hypothesis proposes a *physiological* motivation for this dislike: Repetition of articulatory gestures is relatively difficult, and this difficulty results in phonetic variation; that is, in [XX] it is likely that the two elements are not spelled out identically. We propose the term "novel information bias", which has a *cognitive* motivation: It refers to the well-demonstrated fact that subjects are unable to tokenise multiple adjacent instances of the same type (Treisman & Kanwisher 1998, Walter 2007) because of a general bias in the perceptual system to be more attentive to novel sensory information than to repeated information (Leivada 2017).

In the body of research by Kanwisher (1987 et seq.), *repetition blindness* has been described as the result of difficulties in detecting repeated tokens in rapid serial visual presentations of words. Another illustration is the *apparent motion illusion*: Identical stimuli flashed in different locations are largely perceived as a single moving stimulus; in other words, subjects show a clear preference for a representation of different tokens as one moving token (Vetter et al. 2012). What this means in the context of [*XX] is that talking about a general cognitive bias on anti-identity instead of a UG-wired linguistic constraint that bans [*XX] explains why a limited number of [XX] patterns do surface cross-linguistically (as shown in Leivada 2015b). In sum, the strong preference for anti-identity in language has to do with the way our brain computes types and tokens, and not with a syntactic ban on same-category embedding.

Overall, this approach to anti-identity can be extended to other UG primitives such as parameters or categorial features. In line with Poeppel & Embick's (2005) suggestion to "tak[e] linguistic categories seriously and us[e] them to investigate how the brain computes with such abstract categorical representations" (p. 107), this approach can lead to an evolutionarily plausible UG, while at the same time describing and accounting for the patterns of variation that one has to deal with in the field of comparative variation.

4 Outlook

The approach to UG primitives advocated in this work is still in its earliest stages. An important thing to keep in mind for future work is that deflating UG does not equal arguing against its existence. In other words, there can be a noticeable change in the way we treat UG primitives, without denying the existence of UG (for further discussion, see Roberts 2016a and many of the contributions to that volume). The second important note is that achieving the right levels of abstraction and representation in this effort is crucial: The more linguists abstain from postulating UG-encoded primitives that are very language-specific in nature, the more progress will be made in embedding findings from linguistics in a productively shared context of justification. Last, a third part of this type of approach that is worth mentioning is the conclusion reached in Biberauer, Roberts & Sheehan (2014): What were previously thought to be hard-wired properties of FL could actually reduce to emergent properties that feature the element of interaction among the different factors in language design.

Abbreviations

ABSL	Al-Sayyid Bedouin Sign Language	FL	faculty of language
		UG	Universal Grammar

Acknowledgements

We thank two anonymous reviewers for their helpful comments. KKG's contribution was partially supported by Leventis project 3411-61041 (University of Cyprus). EL acknowledges support from European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 746652.

References

Anderson, Stephen R. & David W. Lightfoot. 2000. The human language faculty as an organ. *Annual Review of Physiology* 62. 697–722. DOI: 10.1146/annurev.physiol.62.1.697.

Arsenijević, Boban & Wolfram Hinzen. 2012. On the absence of X-within-X recursion in human grammar. *Linguistic Inquiry* 43(3). 423–440. DOI: 10.1162/LING_a_00095.

Auer, Peter. 2005. Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In Nicole Delbecque, Johan van der Auwera & Dirk Geeraerts (eds.), *Perspectives on variation: Sociolinguistic, historical, comparative*, 7–42. Berlin: De Gruyter.

Berwick, Robert C. & Noam Chomsky. 2011. The biolinguistic program: The current state of its evolution. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 19–41. Oxford: Oxford University Press.

Berwick, Robert C. & Noam Chomsky. 2016. *Why only us: Language and evolution.* Cambridge, MA: MIT Press.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2). 169–225. DOI: 10.1162/LING_a_00153.

Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. 2014. Complexity in comparative syntax: The view from modern parametric theory. In Frederick J. Newmeyer & Laurel B. Preston (eds.), *Measuring grammatical complexity*, 103–127. Oxford: Oxford University Press.

Biberauer, Theresa & Ian Roberts. 2016. Emergent parameters and syntactic complexity: New perspectives. Paper presented at 'Notions of 'feature' in linguistic theory, Cross-theoretical and cross-linguistic perspectives' workshop at the 49th Annual Meeting of the SLE 2016, Naples.

Biberauer, Theresa & Ian Roberts. 2017. Parameter setting. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge handbook of historical syntax*, 134–161. Cambridge: Cambridge University Press.

Biberauer, Theresa, Ian Roberts & Michelle Sheehan. 2014. No-choice parameters and the limits of syntactic variation. In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics*, 46–55. Somerville, MA: Cascadilla Proceedings Project.

Boeckx, Cedric & Kleanthes K. Grohmann. 2007. The *Biolinguistics* manifesto. *Biolinguistics* 1. 1–8.

Calkins, Monica E., Clayton E. Curtis, William M. Grove & William G. Iacono. 2004. Multiple dimensions of schizotypy in first degree biological relatives of schizophrenia patients. *Schizophrenia Bulletin* 30(2). 317–325. DOI: 10.1093/oxfordjournals.schbul.a007081.

Cheshire, Jenny & Dieter Stein. 1997. *Taming the vernacular: From dialect to written language*. Harlow: Longman.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22. DOI: 10.1162/0024389052993655.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax–semantics*, 1–30. Berlin: De Gruyter.

Chomsky, Noam. 2015. An interview on linguistic variation with Noam Chomsky. *Isogloss* 1. 143–145. DOI: 10.5565/rev/isogloss.11.

Cornips, Leonie. 2006. Intermediate syntactic variants in a dialect-standard speech repertoire and relative acceptability. In Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky & Ralf Vogel (eds.), *Gradience in grammar: Generative perspectives*, 85–105. Oxford: Oxford University Press.

Di Sciullo, Anna Maria, Massimo Piattelli-Palmarini, Kenneth Wexler, Robert C. Berwick, Cedric Boeckx, Lyle Jenkins, Juan Uriagereka, Karin Stromswold, Lisa Lai-Shen Cheng, Heidi Harley, Andrew Wedel, James McGilvray, Elly van Gelderen & Thomas G. Bever. 2010. The biological nature of human language. *Biolinguistics* 4(1). 4–34.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81–138. DOI: 10.2307/416370.

Endress, Ansgar D., Marina Nespor & Jacques Mehler. 2009. Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences* 13(8). 348–353. DOI: 10.1016/j.tics.2009.05.005.

Ettinger, Urlich, Inga Meyhöfer, Maria Steffens, Michael Wagner & Nikolaos Koutsouleris. 2014. Genetics, cognition, and neurobiology of schizotypal personality: A review of the overlap with schizophrenia. *Frontiers in Psychiatry* 5(18). DOI: 10.3389/fpsyt.2014.00018.

Fábregas, Antonio, Jaume Mateu & Michael T. Putnam (eds.). 2015. *Contemporary linguistic parameters*. London: Bloomsbury Academic.

Fowler, Anne E. 1995. Linguistic variability in persons with Down syndrome: Research and implications. In Lynn Nadel & Donna Rosenthal (eds.), *Down syndrome: Living and learning in the community*, 121–131. New York: John Wiley.

Grohmann, Kleanthes K. 2003. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins. DOI: 10.1075/la.66.

Grohmann, Kleanthes K. 2011. Anti-locality: Too-close relations in grammar. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*, 260–290. Oxford: Oxford University Press.

Hagoort, Peter. 2014. Linguistics quo vadis? An outsider perspective. Paper presented at the 47th Annual Meeting of the Societas Linguistica Europaea, Adam Mickiewicz University, Poznań.

Hawkins, John A. 2010. The FOFC from a processing typology perspective. Paper presented at the Workshop on the Final-Over-Final Constraint: Processing, Typology and Acquisition, University of Cambridge.

Henry, Alison. 2005. Non-standard dialects and linguistic data. *Lingua* 115(11). 1599–1617. DOI: 10.1016/j.lingua.2004.07.006.

Hinzen, Wolfram. 2014. What is un-Cartesian linguistics? *Biolinguistics* 8. 226–257.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and Universal Grammar*. Cambridge: Cambridge University Press.

Kanwisher, Nancy G. 1987. Repetition blindness: Type recognition without token individuation. *Cognition* 27(2). 117–143. DOI: 10.1016/0010-0277(87)90016-3.

Kroch, Anthony. 1994. Morphosyntactic variation. In Katherine Beals (ed.), *Papers from the 30th regional meeting of the Chicago Linguistic Society: Parasession on variation and linguistic theory*, 180–201. Chicago: Chicago Linguistic Society.

Leivada, Evelina. 2015a. *The nature and limits of variation across languages and pathologies*. Universitat de Barcelona. (Doctoral dissertation).

Leivada, Evelina. 2015b. X-within-X structures and the nature of categories. *Biolinguistics* 9. 50–73.

Leivada, Evelina. 2017. What's in (a) label? Neural origins and behavioral manifestations of identity avoidance in language and cognition. *Biolinguistics* 11. 1–30.

Leivada, Evelina, Elena Papadopoulou, Maria Kambanaros & Kleanthes K. Grohmann. 2017. The influence of bilectalism and non-standardization on the perception of native grammatical variants. *Frontiers in Psychology*: Language Sciences 8, 205. DOI: 10.3389/fpsyg.2017.00205.

Leivada, Evelina, Elena Papadopoulou & Natalia Pavlou. 2017. Functionally equivalent variants in a non-standard variety and their implications for Universal Grammar: A spontaneous speech corpus. *Frontiers in Psychology*: Language Sciences 8, 1260. DOI: 10.3389/fpsyg.2017.01260.

Lohndal, Terje & Marit Westergaard. 2016. *Grammatical gender in american norwegian heritage language: Stability or attrition?* DOI: 10.3389/fpsyg.2016.00344.

Montrul, Silvina. 2002. Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition* 5(1). 39–68. DOI: 10.1017/S1366728902000135.

Montrul, Silvina. 2008. *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins. DOI: 10.1075/sibil.39.

Odden, David. 1986. On the role of the Obligatory Contour Principle in phonological theory. *Language* 62(2). 353–383. DOI: 10.2307/414677.

Poeppel, David. 2012. The maps problem and the mapping problem: Two challenges for a cognitive neuroscience of speech and language. *Cognitive Neuropsychology* 29(1-2). 34–55. DOI: 10.1080/02643294.2012.710600.

Poeppel, David & David Embick. 2005. Defining the relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-first century psycholinguistics: Four cornerstones*, 103–118. Mahwah, NJ: Lawrence Erlbaum.

Richards, Norvin. 2010. *Uttering trees*. Cambridge, MA: MIT Press.

Roberts, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press.

Roberts, Ian. 2012. Macroparameters and minimalism: A programme for comparative research. In Charlotte Galves, Sonia Cyrino, Ruth Lopes, Filomena Sandalo & Juanito Avelar (eds.), *Parameter theory and linguistic change*, 320–335. Oxford: Oxford University Press.

Roberts, Ian. 2016a. Introduction. In Ian Roberts (ed.), *The Oxford handbook of Universal Grammar*, 1–34. Oxford: Oxford University Press.

Roberts, Ian. 2016b. Linearization, labelling and word-order parameters. Presentation at Workshop on disharmonic word order, Chinese University of Hong Kong, October 14th.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Rowe, Charley & Kleanthes K. Grohmann. 2014. Canaries in a coal mine: Native speakerhood and other factors as predictors of moribundity, death, and diglossic shift in Cypriot Greek. *Mediterranean Language Review* 21. 121–142.

Sandler, Wendy, Irit Meir, Svetlana Dachkovsky, Carol A. Padden & Mark Aronoff. 2011. The emergence of complexity in prosody and syntax. *Lingua* 121(113). 2014–2033. DOI: 10.1016/j.lingua.2011.05.007.

Sandler, Wendy, Irit Meir, Carol A. Padden & Mark Aronoff. 2005. The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences of the United States of America* 102. 2661–2665. DOI: 10.1073/pnas.0405448102.

Treisman, Anne M. & Nancy G. Kanwisher. 1998. Perceiving visually-presented objects: Recognition, awareness, and modularity. *Current Opinion in Neurobiology* 8(2). 218–226. DOI: 10.1016/S0959-4388(98)80143-8.

Trettenbrein, Patrick C. 2015. The "grammar" in Universal Grammar: A biolinguistic clarification. *Questions and Answers in Linguistics* 2(1). 1–10. DOI: 10.1515/qal-2015-0005.

van Riemsdijk, Henk C. 2008. Identity avoidance: OCP effects in Swiss relatives. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 227–250. Cambridge, MA: MIT Press.

Vetter, Peter, Grace Edwards & Lars Muckli. 2012. *Transfer of predictive signals across saccades*. DOI: 10.3389/fpsyg.2012.00176.

Walter, Mary Ann. 2007. *Repetition avoidance in human language*. Cambridge, MA: MIT. (Doctoral dissertation).

Washabaugh, William. 1986. *Five fingers for survival: Deaf sign language in the Caribbean*. Ann Arbor, MI: Karoma Press.

Yang, Charles D. 2004. Toward a theory of language growth. In Lyle Jenkins (ed.), *Variation and universals in biolinguistics*, 37–56. Amsterdam: Elsevier.