

Evelina Leivada*

Mid-level generalizations of generative linguistics: Experimental robustness, cognitive underpinnings and the interdisciplinarity paradox

<https://doi.org/10.1515/zfs-2020-2021>

Abstract: This work examines the nature of the so-called “mid-level generalizations of generative linguistics” (MLGs). In 2015, *Generative Syntax in the 21st Century: The Road Ahead* was organized. One of the consensus points that emerged related to the need for establishing a canon, the absence of which was argued to be a major challenge for the field, raising issues of interdisciplinarity and interaction. Addressing this challenge, one of the outcomes of this conference was a list of MLGs. These refer to results that are well established and uncontroversially accepted. The aim of the present work is to embed some MLGs into a broader perspective. I take the Cinque hierarchies for adverbs and adjectives and the Final-over-Final Constraint as case studies in order to determine their experimental robustness. It is showed that at least some MLGs face problems of inadequacy when tapped into through rigorous testing, because they rule out data that are actually attested. I then discuss the nature of some MLGs and show that in their watered-down versions, they do hold and can be derived from general cognitive/computational biases. This voids the need to cast them as language-specific principles, in line with the Chomskyan urge to approach Universal Grammar from below.

Keywords: Universal Grammar, adjectives, adverbs, Final-over-Final Constraint, linguistic theory

1 Introduction

It has been argued that linguistics today is characterized by a certain degree of fragmentation and lack of terminological clarity. As a consequence, field-internal coherence, disagreements over unanimous identification of primitives, and decreased field-external visibility are some of the challenges the field faces at present (Hagoort 2014; Déchaine 2015; Legate 2015; Varlokosta 2015). Most of these opinions on the strengths, weaknesses, and challenges of present-day

*Corresponding author: **Evelina Leivada**, Department of English and German Studies, Universitat Rovira i Virgili, Av. Catalunya 35, 43002 Tarragona, Spain, e-mail: evelina.leivada@urv.cat, ORCID: <https://orcid.org/0000-0003-3181-1917>

generative linguistics were formulated on the occasion of an academic event that took place in 2015: *Generative Syntax in the 21st Century: The Road Ahead*. According to the call for papers, the absence of a clear, uncontroversial canon posits an important challenge for the field, raising difficulties that pertain to interaction and external visibility.¹ Addressing this challenge, one of the outcomes of that conference is a list that involves the most significant “mid-level generalizations of generative linguistics” (MLGs) (Svenonius 2016; D’Alessandro 2019). These refer to results that are (i) concrete (Ramchand 2015), (ii) established reasonably well and (iii) usable as background for the pursuit of other research, such that “the negation of any of them would be a proposal that would have to be argued for” (Svenonius 2016). Given the need to address challenges that carry implications for interaction with other disciplines, visibility, and the future of linguistics (Hornstein 2015a), the aim of this work is to examine these MLGs from a broader perspective.

2 Defining MLGs

To the best of my knowledge, there is no definition of the term “MLG” in any of the sources that discuss them. Definitions of specific MLGs are offered (Svenonius 2016; D’Alessandro 2019), but there is no list of criteria according to which something is classified as a MLG. Although there is consensus that these are mid-coverage results that should be treated as observations made possible thanks to the tools and approaches of generative grammar, the precise nature of *these* specific generalizations remains somewhat unclear. For example, what does it mean to be a *mid*-level generalization, and what kind of generalizations do the other levels involve? Svenonius (2016) suggests that it means that these observations are general enough to apply to a large sample of languages as well as fairly clear to test. Following the types of universals proposed in de Vries (2005), it is unclear whether these MLGs should be understood as (i) absolute universals (i. e. for all checked languages: p), (ii) implicational universals (i. e. for all checked languages: p \rightarrow q), (iii) general tendencies (i. e. for many checked languages: p), (iv) implicational tendencies (i. e. for many checked languages: p \rightarrow q), or (v) a mix of all the above, such that the list of MLGs groups together different types of observations. With respect to the claim that they are fairly clear to test, again it seems that the list groups together observations of different granularity. For instance, some of these generalizations are marked as “robust”, “very ro-

¹ The call for papers and the description of the goals is available at <http://site.uit.no/castl/events/road-ahead/>

bust” or “robustness not known” in Svenonius’ list, while others are not marked at all. Scarcely any explanation is offered about these different levels of robustness and how they were established. Similarly, in D’Alessandro’s (2019) presentation of these generalizations, experimental robustness is not demonstrated in the sense that the entries in that list do not come with a set of references that provides experimental evidence for each of the mentioned generalizations. Of course, the theoretical work that gave rise to these generalizations is data-driven, and this evidence should not be ignored. However, the total absence of references *outside theoretical linguistics* is somewhat surprising in a list that was developed as a response to the need to discuss the road ahead, while maximizing *field-external visibility*. Put differently, if this list is meant to showcase our most solid findings, its entries should probably count as our best candidates for interdisciplinarity, and yet there is nothing in their presentation that suggests so. Of course, it is possible that this list of MLGs is not intended to make sense to non-linguists and that if the task was to produce a list that would make sense to scholars from other disciplines, the wording would have been different.² If that is indeed the case, Hornstein’s challenge (i. e. think of interaction with other disciplines, visibility, and the future of linguistics in the greater scheme of things in cognitive science) has not really been addressed through this list. Given that five years after the conference, this list still hasn’t reached the stage of “translation”, one would think that it still has not been addressed.

The second factor that lends some ambiguity to the nature of these generalizations is the absence of any explanation as to what sets them apart from other well-established generalizations in linguistics. More specifically, the question is whether this is really a well-established canon or a reflection of the research interests of some of the scholars that attended the conference. As D’Alessandro (2019) puts it, the list is subjective and had other linguists been invited to the meeting, this list would probably look different. In this sense, these MLGs are important observations, but it is unclear whether they are uncontroversially accepted as some of the most robust results of generative linguistics. It seems that they rather reflect a discussion at the conference, that may not be representative of the opinion of the entire list of invitees in that event.³

Recognizing the importance of Hornstein’s challenge, this work aims to take a crucial step towards the “translation” stage for some MLGs, approaching them from a broader perspective that combines insights from neurocognition and experimental psycholinguistics. Taking the Cinque hierarchies for adverbs and

2 I thank Terje Lohndal for discussion on this point.

3 Thanks to Terje Lohndal and Volker Struckmeier for their comments on this point.

adjectives and the Final-over-Final Constraint as case-studies, it is argued that while the phrasing of certain MLGs may rule out data that are attested, the essence of these generalizations holds across languages, albeit not in a way that always agrees with the way they are presented at the linguistic level. Reaching the “translation” stage is an important and non-trivial move, given the gulf between the generative linguistic tradition and research in neurocognition (Marantz 2005; Grimaldi 2012; Hagoort 2014). To reconcile these mid-level observations made in the study of language data with neuroscientific studies of human language and cognition requires focusing on the right type of primitives as well as creating a shared, interdisciplinary understanding of them, something that is still not attained (cf. the Granularity Mismatch Problem and the Ontological Incommensurability Problem in Poeppel and Embick 2005). In other words, the fact that the list of MLGs did not reach the “translation” stage, either at the conference or five years later, may mean that this translation is not possible for all the MLGs in that list. The question then is the following: Can granularity considerations be addressed for all MLGs, in the sense that they are all of the same ilk when it comes to granularity mismatches? It seems that the only way to an answer goes through addressing Hornstein’s challenge.

The forthcoming exploration of the experimental robustness and the cognitive underpinnings of some MLGs leads to a discussion of the *interdisciplinarity paradox* (Weingart 2000): there is a significant mismatch between the proclaimed interest for maximizing interaction with other disciplines and the actual presentation of MLGs, which almost exclusively involves discipline-based terminology and lacks an outlook that would either discuss the possible connections of this list with other disciplines or make some suggestion about the interaction envisioned from the linguist’s side.⁴ In order to remedy this paradox, I discuss the nature of some MLGs and show that in their watered-down versions, they can be derived from general cognitive and/or computational biases. This voids the need to cast them as linguistic principles of Universal Grammar (UG), something that is in line with the urge to approach the latter from below (Chomsky 2007a).

⁴ This thesis presupposes that MLGs are understood as well-established, empirically robust results of minimalism. The other view is that MLGs are targets for future minimalist explanation (Norbert Hornstein, pers. comm.). These two views differ in how much they ascribe to these MLGs. Under the first view, these are some of the *achievements* of generative grammar; “the concrete results of bread and butter generative syntax” (Ramchand 2015). Under the second view, these are claims whose empirical robustness and overarching theory are to be determined in future research within minimalism. If one supports this second view, an interesting question would be why the Athens event, which according to the call for papers aimed to “reaffirm the theoretical core of the discipline”, did not produce a list of achievements whose robustness has been already demonstrated, instead of or in parallel to this list of targets for future research.

3 The Cinque hierarchies for adverbs and adjectives

The list of MLGs involves two entries that make reference to two different Cinque hierarchies (D'Alessandro 2019):

- (1) Cinque hierarchy: There are semantically defined classes of Tense-Aspect-Mood functors that appear in the same hierarchical order in all languages in which they exist overtly. (Cinque 1999)
- (2) Cinque hierarchy for adverbs: There are semantically defined classes of adverbs that appear in the same hierarchical order in all languages in which they exist overtly. (Cinque 1999)

There is a third Cinque hierarchy that is not mentioned in the list, but as the latter is subjective, and given that the Cinque hierarchies are interconnected, it seems that its absence is more of an accident, than a deliberate choice.

- (3) Cinque hierarchy for adjectives: Different types of adjectives appear in the same hierarchical order in all languages in which they exist overtly. (Cinque 2010)

(2) entails the syntactic hierarchy in (4), while (3) the one in (5). There are slight variations of these hierarchies in the literature in terms of their level of detail, but the bottom line is the same: An innate, syntactic hierarchy encoded in Universal Grammar (UG) specifies the relevant orderings.

- (4) Mood_{speech act} (*frankly*) > Mood_{evaluative} (*fortunately*) > Mood_{evidential} (*allegedly*) > Modality_{epistemic} (*probably*) > Tense_{past} (*once*) > Tense_{future} (*then*) > Modality_{irrealis} (*perhaps*) > Modality_{necessity} (*necessarily*) > Modality_{possibility} (*possibly*) > Aspect_{habitual} (*usually*) > Aspect_{repetitive I} (*again*) > Aspect_{frequentative I} (*often*) > Modality_{volitional} (*intentionally*) > Aspect_{celerative I} (*quickly*) > Tense_{anterior} (*already*) > Aspect_{terminative} (*no longer*) > Aspect_{continuative} (*still*) > Aspect_{perfect} (*always*) > Aspect_{retrospective} (*just*) > Aspect_{proximative} (*soon*) > Aspect_{durative} (*briefly*) > Aspect_{generic} (*characteristically*) > Aspect_{prospective} (*almost*) > Aspect_{completive I} (*completely*) > Voice (*well*) > Aspect_{celerative II} (*fast*) > Aspect_{repetitive II} (*again*) > Aspect_{frequentative II} (*often*) > Aspect_{completive II} (*completely*)
(Cinque 1999: 106)

- (5) Subjective Comment > Evidential > Size > Length > Height > Speed > Depth > Width > Temperature > Wetness > Age > Shape > Color > Nationality/Origin > Material

(adapted from Scott 2002: 114)

Focusing on the experimental robustness of (2) and (4), both corpus data of naturalistic speech and data that come from judgment tasks suggest that orderings that are not deemed possible, according to the Cinque hierarchy for adverbs, are actually attested or accepted as well formed respectively. Cinque (1999) discusses potential explanations for apparent counterexamples to his adverb hierarchy, but as Payne (2018) shows, even after taking these explanations into account, some examples of unexpected orderings still remain. More specifically, Payne (2018) provides data that do not comply with (4), but are nevertheless attested in the Corpus of Contemporary American English (e. g., *already quickly*, *almost just*). Of course, a series of adverbs is a rare phenomenon in naturalistic speech, hence rigorous testing is needed to properly adjudicate on the robustness of certain orderings and on whether the accepted orderings always comply with the Cinque hierarchy for adverbs. Comparing the hierarchy-deviating patterns in the Corpus of Contemporary American English with their hierarchy-compliant counterparts (i. e. *almost just* vs. *just almost*) in an acceptability judgment task, Payne's results showed that out of the 18 comparisons, only 5 showed a statistically significant higher judgment of the hierarchy-compliant ordering, at $p=0.001$, after adjusting alpha for multiple comparisons. Grouping together items in the two conditions (hierarchy-compliant and hierarchy-deviating), the average rating for compliant orderings was 4.39 on an 1–7 Likert scale, while the rating for the deviating orderings was 3.67. Payne (2018) suggests that this difference is statistically significant at $p<.045$, however it is not clear that this difference would survive the process of adjusting for multiple comparisons. More importantly, not all differences in the comparison of the two conditions favored the hierarchy-compliant orders. Overall, these results raise some concerns about the experimental robustness of the MLG stated in (2), which seems to rule out data that are actually attested in naturalistic speech.

Even if one interpreted the findings from the acceptability judgment task as showing a statistically strong preference for the hierarchy-compliant orders, this result would not be informative about the (un)acceptability of the hierarchy-deviating orders. The fact that one sentence is deemed more acceptable or more preferred/natural than another sentence, does not necessarily speak about the status of the latter, unless many variables are controlled. For example, the simultaneous, forced-choice juxtaposition of two orders, one hierarchy-complying and the other one hierarchy-deviating, may not tap into the acceptability of each order, but into the strength of the preference that may arise over the explicit comparison

of the deviating order with the *prescriptively correct* variant of it (Leivada and Westergaard 2019). Especially speakers that are literate and belong into groups that have specific sociolinguistic characteristics (i. e. college students, Amazon Mechanical Turk workers) are likely to recognize the prescriptively correct sentence when this is juxtaposed with its deviating counterpart. As the literature on acceptability judgment tasks has shown, prescriptive correctness interferes with acceptability, given that acceptable sentences may be rejected either on a prescriptive basis or because they are formulated in a way that is not the most preferred one for the informant (Tremblay 2005). The bottom line is that as long as more than one order exist in the repertoire of a speaker, the less preferred one does form part of this repertoire, even if it is less preferred and given lower ratings in acceptability judgment tasks. The fact that two (or more, if one considers instances of three adjectives) orderings seem acceptable is a fact that a theory that postulates innate (and as such fixed) hierarchies cannot accommodate.

This *fixity problem* of UG has been identified and discussed in the context of parametric hierarchies (Boeckx and Leivada 2013; Leivada 2015): This problem concerns the essence of UG as – what Chomsky has called – an *innate fixed nucleus* (Piattelli-Palmarini 1980). In 1974, a debate took place between Jean Piaget and Noam Chomsky, and in this debate, the nature of this innate fixed nucleus was discussed. As Piattelli-Palmarini (1980) documents, the positions of Chomsky and Piaget are indicative of their different persuasions (empiricism vs. nativism):

In this sense, the final position of Piaget at Royaumont represents a manifestation of the “empiricist” position. Once the existence of a fixed nucleus is acknowledged, the contrast between the paradigms is even more remarkable. For Piaget, accounting for the stability of the fixed nucleus in terms of self-regulating mechanisms becomes the first goal of epistemology, whereas for Chomsky, the fundamental issue is precisely the specificity of the fixed nucleus and not the manner in which its fixity is attained. (Piattelli-Palmarini 1980: 353)

Chomsky and Piaget disagree about certain things, but they agree on accepting the fixed character of UG. If one accepts their view that fixity is indisputable, one cannot argue that innate hierarchies that were thought to be rigid – and as such giving rise to a single ordering – can at the same time be viewed as adjustable primitives that allow for flexibility in the orderings they permit.

Returning to the adverb hierarchy, perhaps the most important observation about the experimental robustness of (2) comes from Wilson and Saygın (2001). Upon observing some orderings in Turkish that cannot be properly accounted for by the Cinque hierarchy for adverbs, they make a crucial link between actual experimental robustness and the explanatory power of (4). In their words:

Cinque does postulate some lower heads which duplicate the functionality of certain higher heads; these are marked with ‘(II)’ in (4) above. *It is always going to be possible to accom-*

modate any observed ordering facts simply by duplicating heads. However, if heads can be duplicated as required, the motivation for having a hierarchy in the first place is called into question. (Wilson and Saygin 2001: 5; emphasis mine)

In other words, patterns that fall outside the domain of predictions of (2) entail that we must seek different functional positions on the hierarchy for identical elements that do *not* have different functions. This also relates to the fixity issue, because one could have argued that fixity is not a problem for as long the innate hierarchy has different positions for the same category of adverbs such that any flexibility in the attested orderings can be accommodated, without challenging the rigidity of the innate hierarchy. Wilson and Saygin's (2001) observation is accurate: If the hierarchy was created on the basis of the data, it seems that nothing is preventing us from amending it, by duplicating portions of it in order to be able to (re)accommodate the data. However, this is a problematic move, because we neither *explain* the distribution of the data this way, nor do we *uncover* an innate primitive. Taking into account that the Cinque hierarchies that form part of the list of MLGs are taken to be encoded in UG according to standard cartographic assumptions (see Cinque and Rizzi 2008: 48), one understands that some explanation is necessary as to why machinery that seems to be adjustable to fit the data is cast as innate. This explanation is conspicuously absent from the literature. As Cheng (2015) noted in the context of the *Generative Syntax in the 21st Century: The Road Ahead* conference, "one of the main weaknesses of the field is that there is no criterion concerning what counts as an explanation or an analysis. Take cartography as an example. Do cartographic analyses count as explanations?" This *problem of explanation*, as Larson (2018) calls it, comes together with what he identifies as the problems of *plenitude* and *rigidity*. Plenitude refers to the fact that we must assume that the entire hierarchy is covertly present even if only two elements are overtly realized, while rigidity boils down to the fact that functional selection is not a gradable notion, and yet the predicted rigidity of the hierarchy is not reflected in the speaker's judgments, which eventually show flexibility (Larson 2018). This matters because it attests to the mismatch between theory and results: at the empirical level the hierarchy is not as rigid as suggested at the theoretical level.

Although the Cinque hierarchy for adverbs suffers from three major problems (i. e. it rules out data that are attested, it lacks an explanation about why the data appear the way they do, and it inserts unnecessary complexity in UG by allocating to it a hierarchy that seems an ad hoc invention that is customizable on the basis of the data), it should also be acknowledged that most scholars find that the majority of the adverb patterns that have been examined across different languages do conform with (4) (see Payne 2018 and references therein). This probably

suggests that the essence of the hierarchy is correct and possibly rooted in some general cognitive/computational constraint, although we have not yet discovered which one. Instrumental in understanding this point is the Cinque hierarchy for adjectives [(3)–(5)].

Similar to the idea that adverbs are ordered in a specific way cross-linguistically, it has been suggested that different types of adjectives occupy fixed positions via rigid hierarchies that permit specific orderings. Once again, this allegedly universal order for adjective placement has been argued to be the result of a syntactic hierarchy that is encoded in UG (Scott 2002; Cinque and Rizzi 2008; Panayidou 2013). Although certain preferences about the hierarchy-compliant adjective orderings have been demonstrated (Scontras et al. 2017), once again it has been shown that the hierarchies are not rigid, given that hierarchy-deviating orderings are judged as highly acceptable by participants in an experimental setting (Leivada and Westergaard 2019). More importantly, evidence from reaction times has suggested that the hierarchy-deviating orders do not take longer to process than their hierarchy-compliant counterparts (Leivada and Westergaard 2019). This result suggests that not only aren't the deviating orderings unacceptable, but they are not even more marked, because if they were, the reaction times should have reflected this, given that deviations from an unmarked order induce an extra processing cost (Erdocia et al. 2009). Still, exactly as happened in the case of the hierarchy for adverbs, in the case of adjectives too, it seems that the predictions of the hierarchy are largely borne out, in the sense that the hierarchy-compliant orderings are the most preferred ones and they are also the ones eliciting slightly higher acceptability ratings (Scontras et al. 2017; Leivada and Westergaard 2019). Importantly, unlike the hierarchy for adverbs, the hierarchy of adjectives has been successfully explained in terms of cognitive notions, voiding any need to appeal to a UG hierarchy. Several studies have suggested that the preference for the ordering that complies with the hierarchy is due to cognitive factors such as inherentness (Whorf 1945), absoluteness (Sproat and Shih 1991) and subjectivity (Hetzron 1978; Stavrou 1999; Scontras et al. 2017): more objective/absolute adjectives (e. g., *red* is more objective than *pretty*) that denote noun-inherent properties appear closer to the noun. These cognitive notions facilitate certain orderings in order to serve *communicative purposes*: subserving the distinction between speaker-oriented and object-oriented information (Stavrou 1999), adjectives are often ordered in a way that reflects this difference in subjectivity in order to aid successful reference resolution (Franke et al. 2019).

Although the Cinque hierarchy for adverbs has not been explained yet along the lines of cognitive notions that facilitate certain orderings, this has happened for the Cinque hierarchy in (1), which was also included in the list of MLGs. It has been argued that (1) is rooted in general cognitive constraints and the hierarchical

ordering of Tense-Aspect-Mood functors reflects how our perception organizes experiences in terms of events, situations, and propositions (Ramchand and Svenonius 2014). It is likely that extralinguistic accounts such as those developed for Cinque hierarchy (1) and the Cinque hierarchy for adjectives (3) will also be developed for the Cinque hierarchy for adverbs (2) and other MLGs.

4 The Final-over-Final Constraint

Another MLG in the list that emerged from the *Generative Syntax in the 21st Century: The Road Ahead* conference is the Final-over-Final Constraint (FOFC):

- (6) FOFC: A head-final phrase cannot (immediately) dominate a head-initial phrase in the same extended projection.⁵
(Biberauer et al. 2013)

FOFC was originally argued to be a *linguistic universal*, determining that three of the four logically possible combinations in the ordering of two phrases are licit options, while also noting that the third possibility is less common due to the preference for harmonic orders (Biberauer et al. 2014): consistent head-initial, consistent head-final, and inverse FOFC (i. e. head-initial dominating head-final).

However, in V2 languages like German, head-final VPs immediately dominate head-initial PPs. This rendered necessary a modification of FOFC that narrowed down the empirical coverage of the universal by introducing the factor of categorial agreement, predicting (7).

- (7) *_{[Aux'} [_{VP} V O] Aux]
(Biberauer et al. 2014)

Yet, the new formulation of FOFC is not immune to counterexamples either.⁶ Hindi, for instance, involves V-O-Aux structures, where a head-initial VP is dom-

⁵ This is not how FOFC is presented in the actual list of MLGs. The phrasing in Svenonius (2016) – and in D'Alessandro (2019: 11) who uses Svenonius' presentation – is the following: "It is relatively difficult to embed head-final projections in head-initial ones, compared to the opposite (132 but not *231, where 1 takes 2 as a complement and 2 takes 3)." Although this definition does not correspond to FOFC (but to inverse FOFC), its gist is correct. As the forthcoming discussion will show, it is indeed relatively difficult to embed head-final phrases in head-initial ones due to the preference for harmonic patterns.

⁶ The discussion of the forthcoming examples from Latin is a shortened version of the analysis of FOFC in Leivada (2015).

ulations are not explanations, but ad hoc amendments that modify the definition of the universal to make it immune to disconfirming data. For instance, consider the following claim by Sheehan about the apparently FOFC-violating final discourse particles in Mandarin Chinese: “As a result, it is impossible even to add a stipulation to FOFC stating that it does not apply to particles, as at present the only diagnostic for particlehood is insensitivity to FOFC.” (Sheehan 2013: 440) The nature of this argument again brings forward Larson’s *problem of explanation*: structure α is not sensitive to FOFC on the basis of an added part to the definition of FOFC that posits that structure α is not sensitive to FOFC. One can add stipulations to a definition ad infinitum, but has one captured and explained an innate universal this way or has one created it by tweaking its definition accordingly?

With respect to the experimental robustness of FOFC, the original proposal that this is an absolute universal (Sheehan et al. 2017) has been replaced in recent work with the weaker claim that FOFC is a strong tendency, again in light of the growing number of the attested counterexamples (Clem 2018). Also, despite the fact that MLGs are taken to be results that are well established and uncontroversially accepted, it seems that some controversy does exist not only about whether FOFC is a hard constraint, but also about whether it is a syntactic constraint or a cognitive/extralinguistic one. Although Sheehan et al. (2017) argue that FOFC is neither a tendency nor a general cognitive constraint in processing, it seems that for some scholars a cognitive explanation is possible. Hawkins (2013), for example, proposed that the human parser has a preference for minimizing domains of processing, hence the preference for harmonic patterns. Grohmann and Leivada (2020) suggest that this preference is due to a combination of computational conservatism and the way human memory works: Putting together Roberts’ (2016) *Input Generalization* – a 3rd factor principle that suggests that there is a preference for a given feature of a functional head to generalize to other functional heads – with the results of experiments on statistical learning in infants, which have shown that sequence edges are exceptionally salient positions and facilitate learning in a way that gives rise to harmonic patterns (Endress et al. 2009), the infrequency of disharmonic patterns can be explained.

In sum, this discussion of FOFC aimed to show three things: first, it is not clear that FOFC-violating patterns are inexistent across languages. If they exist, the nature of this MLG as an absolute universal vs. statistical tendency is unclear, and this has implications about the expected robustness of this phenomenon. Second, even if one accepts that FOFC-violating patterns are not attested, it is not clear that their inexistence is due to a syntactic universal. The arguments given in favor of viewing it as such are not backed up by explanations that address the “why” question. This is crucial if MLGs are proposed to be the most robust findings of theo-

retical linguistics and, as such, our best candidates for meeting Hornstein's challenge. Third, as in the case of the Cinque hierarchies of adjectives and adverbs, the essence of FOFC in its watered-down version (i. e. a cross-linguistic tendency, not a hard, syntactic constraint) is correct, because disharmonic patterns are indeed less frequent compared to consistent head-initial or head-final patterns. It is possible that syntax has something to do with this infrequency, but no definitive claims can be made until a cognitive account of FOFC has been formulated in detail, tested, and (dis)confirmed.

5 Outlook: The interdisciplinarity paradox

Weingart (2000) used the term *paradox of interdisciplinarity* in order to talk about a phenomenon that has been apparent in the science policy arena ever since the 1970s. This paradox has been defined as “a strange simultaneity of the proliferation and continued persistence of a programmatic discourse about interdisciplinarity on the one hand, and an ever increasing disciplinary and sub-disciplinary differentiation and specialisation taking place in academic research on the other” (Woelert and Millar 2013: 756). In Weingart's (2000) discussion, the tension is between a proclaimed interest for interdisciplinarity and an increase in disciplinary rigidity, manifested through boundary demarcating, niche seeking, and the exclusive use of discipline-specific concepts, terminologies, and technologies.

The *Generative Syntax in the 21st Century: The Road Ahead* conference aimed to address the challenge of coherence and visibility. This aim was presented in the call for papers in the following way:

A major challenge concerns the coherence of the field. Given the large number of different analytic approaches, it has resulted in small groups working on x, y, or z. From a scientific point of view, this is not problematic, but it raises difficulties when it comes to interaction, funding, recruitment and external visibility. We want to discuss ways of improving this situation. We believe that this is especially important given that linguistics and generative syntax are not major fields compared to e. g., psychology or physics. In addition to being problematic in its own right, the proliferation of approaches further exacerbates the problem of teaching and supervision. (<http://site.uit.no/castl/events/road-ahead/>)

Sub-disciplinary boundary demarcation is evident in the above juxtaposition of linguistics and psychology, for this view of the two as different fields is not an uncontroversial one. In fact, the most prominent figure of generative syntax, Noam Chomsky, has repeatedly suggested that there is no other way to conceive linguistics but as part of psychology. Actually, he has made a much stronger claim: “In

my opinion one should not speak of a ‘relationship’ between linguistics and psychology, because linguistics is *part of* psychology; I cannot conceive of it in any other way.” (Chomsky 2007b: 43) If one agrees with Chomsky, the issue is not that linguistics and psychology are not conceptually connected, but that in practice over the last years they have somewhat drifted apart in their tools and methodologies. Undermining this conceptual connection by talking about two different fields, while at the same time recognizing the need to improve visibility in neighboring fields, does not help in increasing interaction. It rather seems to be an instantiation of subdisciplinary identity shielding that is typical in the context of the interdisciplinarity paradox.

Another exemplification of the latter comes from some of the statements offered by various prominent linguists in the context of the *Generative Syntax in the 21st Century: The Road Ahead* conference. Consider the following view in Merchant’s statement:

We can stop tying our analytical proposals to old debates about Universal Grammar, innateness, and learnability, and stop even paying lip service to positions in these debates. These are independent issues, orthogonal to the central theoretical issues we face, and a wonderful red herring for those who would seek to ignore or dismiss all generative syntax work. *One can argue for or against UG as a theory of the language faculty, but it makes no difference to whether our proposals about selection, agreement, movement, phrase structure, etc. are right.* (Merchant 2015; emphasis mine)

Again, this is not an uncontroversial view. Chomsky (2007b) has argued that no discipline can productively concern itself with the utilization of a form of knowledge, unless it deals with crucial “why” questions about the nature of the system that underlies this form of knowledge. In other words, our theories of UG, the type of assumptions we make about it, and the kind of primitives we allocate to it are *highly relevant* to whether our proposals about movement and phrase structure are right. If one endorses Chomsky’s view, drawing a sharp line between proposals about, say, phrase structure and (i) the reason and the way the overall system implements them, (ii) the path through which the individual learns them and (iii) the way the learning process reshapes them, is a dangerous move, not because it leaves the crucial “why” questions unanswered, but because it suggests that leaving these questions unanswered does *not* carry implications about whether proposals on phrase structure are right. Without answering these questions, the proposals are bound to have important gaps (along the lines of Larson’s problem of explanation), which are usually filled in with stipulative, circular arguments. Put another way, proposals about selection, agreement, movement, and phrase structure evoke certain primitives and operations. For example, certain features have been argued to drive movement operations and the most common

answer to the question “where does that feature come from?” is UG. This is true also of many MLGs, and certainly of the ones discussed in this work.⁸ In this context, proposing a disconnection between the theory that evokes a primitive (be it feature, parameter, constraint, etc.) and the very component from which the primitive is argued to come from is not only a wonderful red herring (to use Merchant’s words), but also another example of disciplinary identity shielding, when this proposal occurs in a context that has identified field-internal interaction and field-external visibility as two of the major challenges it aimed to address.

Hornstein’s (2015b) statement in the same event makes a valuable point that relates to the above discussion. He argues that there is a danger in treating our descriptions about grammars (and by extension about how movement, selection, and agreement occur in these grammars) as ends in themselves instead of way stations to the deeper “why” questions of how the system is organized. Of course, not being interested in answering these questions is fine, and one is free to pursue their research aims until the point that concerns them. The problem arises when one suggests that the two sets of inquiries (i. e. grammars and the system that implements them) can be disconnected, because they are independent issues. They are not, because the answer to certain questions about the former goes through the developing knowledge about latter. One can ignore these questions, but this does not entail that the gaps in one’s theory about the former will disappear. Last, if we embrace the view that the research questions that link the field of linguistics with other closely connected disciplines are irrelevant to the field’s core aims and concerns, Hornstein’s challenge is bound to remain unaddressed for longer, with all the consequences this may have for disciplinary visibility and interdisciplinary progress.

Acknowledgements: I thank the editors of this special issue for their help throughout the process. For helpful discussions that helped me sharpen some of the ideas presented here, I thank Kleanthes Grohmann, Norbert Hornstein, Terje Lohndal, Richard Larson, Volker Struckmeier, and two anonymous review-

⁸ The Cinque hierarchies are explicitly allocated to UG (cf. Cinque and Rizzi (2008: 53): “UG expresses the possible items of the functional lexicon and the way in which they are organized into hierarchies”), while FOFC has been presented by its proponents as a syntactic universal (Biberauer et al. 2014), therefore it should fall in the 1st factor: principles specific to language. Sheehan et al. (2017: 31) leave open the possibility that FOFC “is derivable from UG, or from UG in combination with other aspects of language design.” In recent work Biberauer (2019) has put forth a more detailed neo-emergentist approach of FOFC that agrees with the idea presented in this work: FOFC seems to derive by 3rd factor, computational biases.

ers. This work received support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement n° 746652 and from the Spanish Ministry of Science, Innovation and Universities under the Ramón y Cajal grant agreement n° RYC2018-025456-I. The funders had no role in the writing of the study and in the decision to submit the article for publication.

References

- Biberauer, Theresa. 2019. Factors 2 and 3: Towards a principled approach. *Catalan Journal of Linguistics* [Special Issue]. 45–88. DOI: 10.5565/rev/catjl.219.
- Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2). 169–225.
- Biberauer, Theresa, Ian Roberts & Michelle Sheehan. 2013. No-choice parameters and the limits of syntactic variation. In Robert E. Santana-LaBarge (ed.), *Proceedings of WCCFL 31*, 46–55. Somerville, MA: Cascadilla Proceedings Project.
- Boeckx, Cedric & Evelina Leivada. 2013. Entangled parametric hierarchies: Problems for an overspecified Universal Grammar. *PLOS ONE* 8(9). e72357. DOI: 10.1371/journal.pone.0072357.
- Cheng, Lisa Lai-Shen. 2015. Statement at *Generative Syntax in the Twenty-First Century: The Road Ahead*. <http://site.uit.no/castl/files/2017/03/cheng.pdf> (24 September 2020).
- Chomsky, Noam. 2007a. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language?*, 1–29. Berlin & New York: Mouton de Gruyter.
- Chomsky, Noam. 2007b. *On language*. New York: The New Press.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads*. Oxford: Oxford University Press.
- Cinque, Guglielmo. 2010. *The syntax of adjectives: A comparative study*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo & Luigi Rizzi. 2008. The cartography of syntactic structures. *CISCL Working Papers on Language and Cognition* 2. 43–59.
- Clem, Emily. 2018. Disharmony and the Final-Over-Final Condition in Amahuaca. Talk at *GLOW 41*, Hungarian Academy of Sciences.
- Danckaert, Lieven. 2011. *On the left periphery of Latin embedded clauses*. Ghent: University Ghent dissertation.
- Danckaert, Lieven. 2012. Word order variation in the Latin clause: O's, V's, Aux's, and their whereabouts. Paper presented at *SyntaxLab*, University of Cambridge.
- D'Alessandro, Roberta. 2019. The achievements of Generative Syntax: A time chart and some reflections. *Catalan Journal of Linguistics* [special issue]. 7–26.
- Déchainé, Rose-Marie. 2015. Statement at *Generative Syntax in the Twenty-First Century: The Road Ahead*. <http://site.uit.no/castl/files/2017/03/dechainé.pdf> (24 September 2020).
- de Vries, Mark. 2005. The fall and rise of universals on relativization. *Journal of Universal Language* 6. 1–33.
- Endress, Ansgar D., Marina Nespors & Jacques Mehler. 2009. Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences* 13(8). 348–353.

- Erdocia, Kepa, Itziar Laka, Anna Mestres-Missé & Antoni Rodriguez-Fornells. 2009. Syntactic complexity and ambiguity resolution in a free word order language: Behavioral and electrophysiological evidences from Basque. *Brain & Language* 109. 1–17. DOI: 10.1016/j.bandl.2008.12.003.
- Franke, Michael, Gregory Scontras & Mihael Simonič. 2019. Subjectivity-based adjective ordering maximizes communicative success. In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*, 344–350.
- Grimaldi, M. 2012. Toward a neural theory of language: Old issues and new perspectives. *Journal of Neurolinguistics* 25(5), 304–327.
- Grohmann, Kleanthes K. & 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: Synchronic and diachronic perspectives, volume 1: Syntax inside the grammar*, 561–578. Berlin: Language Science Press.
- Hagoort, Peter. 2014. Linguistics quo vadis? An outsider perspective. Talk presented at *SLE 2014: 47th Annual Meeting of the Societas Linguistica Europaea*, Adam Mickiewicz University, Poznań, Poland.
- Hawkins, John A. 2013. Disharmonic word orders from a processing-efficiency perspective. In Theresa Biberauer & Michelle Sheehan (eds.), *Theoretical approaches to disharmonic word order*, 391–406. Oxford: Oxford University Press.
- Hetzron, Robert. 1978. On the relative order of adjectives. In Hansjakob Seiler (ed.), *Language universals. Papers from the conference held at Gummersbach/Cologne, Germany, October 3-8, 1976*, 165–184. Tübingen: Narr.
- Hornstein, Norbert. 2015a. The road ahead from Athens; some scattered initial remarks. *Faculty of Language* (blog), May 31, 2015. <http://facultyoflanguage.blogspot.com/2015/05/the-road-ahead-from-athens-some.html> (accessed 24 September 2020).
- Hornstein, Norbert. 2015b. Statement at *Generative Syntax in the Twenty-First Century: The Road Ahead*. <http://site.uit.no/castl/files/2017/03/hornstein.pdf> (24 September 2020).
- Larson, Richard. 2018. *Rethinking cartography*. Ms., Stony Brook University.
- Legate, Julie Anne. 2015. Statement at *Generative Syntax in the Twenty-First Century: The Road Ahead*. <http://site.uit.no/castl/files/2017/03/legate.pdf> (24 September 2020).
- Leivada, Evelina. 2015. *The nature and limits of variation across languages and pathologies*. Barcelona: Universitat de Barcelona dissertation.
- Leivada, Evelina & Marit Westergaard. 2019. Universal linguistic hierarchies are not innately wired. Evidence from multiple adjectives. *PeerJ* 7: e7438. DOI: 10.7717/peerj.7438.
- Mahajan, Anoop Kumar. 1990. LF conditions on negative polarity licensing. *Lingua* 80(4). 333–348.
- Marantz, Alec. 2005. Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review* 22. 429–445.
- Merchant, Jason. 2015. Statement at *Generative Syntax in the Twenty-First Century: The Road Ahead*. <http://site.uit.no/castl/files/2017/03/merchant.pdf> (24 September 2020).
- Panayidou, Fryni. 2013. *(In)flexibility in adjective ordering*. London: Queen Mary University of London dissertation.
- Payne, Amanda. 2018. *Adverb typology: A computational characterization*. Newark, DE: University of Delaware dissertation.
- Piattelli-Palmarini, Massimo (ed.). 1980. *Language and learning: The debate between Jean Piaget and Noam Chomsky*. Cambridge, MA: Harvard University Press.

- 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.
- Ramchand, Gillian. 2015. Athens Day 1. *Language* (blog), May 29, 2015. <http://generativelinguist.blogspot.com/2015/05/athens-day-1.html> (24 September 2020).
- Ramchand, Gillian & Peter Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46B. 152–174.
- Roberts, Ian. 2016. Linearization, labelling and word-order parameters. Presentation at *Workshop on Disharmonic Word Order*, Chinese University of Hong Kong.
- Scontras, Gregory, Judith Degen & Noah D. Goodman. 2017. Subjectivity predicts adjective ordering preferences. *Open Mind. Discoveries in Cognitive Science* 1. 53–65.
- Scott, Gary-John. 2002. Stacked adjectival modification and the structure of nominal phrases. In Guglielmo Cinque (ed.), *Functional structure in DP and IP. The cartography of syntactic structures*, 91–120. New York: Oxford University Press.
- Sheehan, Michelle. 2013. Explaining the Final-over-Final constraint: Formal and functional approaches. In Theresa Biberauer & Michelle Sheehan (eds.), *Theoretical Approaches to Disharmonic Word Order*, 407–444. Oxford: Oxford University Press.
- Sheehan, Michelle, Theresa Biberauer, Ian Roberts & Anders Holmberg. 2017. *The Final-over-Final Condition. A Syntactic Universal*. Cambridge, MA: MIT Press.
- Sproat, Richard & Chilin Shih. 1991. The cross-linguistic distribution of adjective ordering restrictions. In Carol Georgopoulos & Roberta Ishihara (eds.), *Interdisciplinary approaches to language*, 565–593. Dordrecht: Kluwer.
- Stavrou, Melita. 1999. The position and serialization of APs in the DP. In Artemis Alexiadou, Geoffrey Horrocks & Melita Stavrou (eds.), *Studies in Greek syntax*, 201–226. Dordrecht: Kluwer.
- Svenonius, Peter. 2016. Significant mid-level results of generative linguistics. (blog), August 30, 2016. <https://blogg.uit.no/psv000/2016/08/30/significant-mid-level-results-of-generative-linguistics/> (accessed 24 September 2020).
- Tremblay, Annie. 2005. Theoretical and methodological perspectives on the use of grammaticality judgment tasks in linguistic theory. *Second Language Studies* 24. 129–167.
- Varlokosta, Spyridoula. 2015. Statement at *Generative Syntax in the Twenty-First Century: The Road Ahead*. <http://site.uit.no/castl/files/2017/03/varlokosta.pdf> (accessed 24 September 2020).
- Weingart, Peter. 2000. Interdisciplinarity: The paradoxical discourse. In Peter Weingart & Nico Stehr (eds.), *Practising Interdisciplinarity*, 25–41. Toronto: University of Toronto.
- Whorf, Benjamin Lee. 1945. Grammatical categories. *Language* 21(1). 1–11.
- Wilson, Stephen & Ayşe Pinar Saygın. 2001. Adverbs and functional heads in Turkish: Linear order and scope. In Leslie Carmichael, Chia-Hui Huang & Vida Samiian (eds.), *WECOL 2001: Proceedings of the 13th Western Conference on Linguistics*, 410–421. Fresno, CA: Department of Linguistics, California State University, Fresno.
- Wolert, Peter & Victoria Millar. 2013. The ‘paradox of interdisciplinarity’ in Australian research governance. *Higher Education* 66. 755–767.