June 2021, revised May 2022

Formalism and, not vs, Functionalism

Željko Bošković

Baker & McCloskey (2007) examine the relationship between traditional typology and generative theoretical syntax. Since then typology has started to play an increasingly more important role within the latter camp to the point that we can actually talk about generative typology. Given that traditional typology is generally associated with functional approaches (see e.g. Nichols 2007), this then gives us two approaches to typology, which I will refer to as traditional (TT) and generative typology (GT). [1] In a way, then, typology is setting grounds for a potential rapprochement of the functional and the formal approach to language more generally. This paper will provide a number of remarks to this effect, within a unificational view where both of these approaches have a place, i.e. where they are in principle not in competition with each other. [2]

To this end, the paper will show that many of what are assumed to be clear demarcation lines between the traditional generativist and the traditional functionalist/typological camp (and the two approaches to typology) are not clear, in fact may not be there at all—the two approaches are not in opposition as much as they used to be, and as the practitioners of the two camps still seem to think they are. In this respect, it will be shown that many of the perceived irreconcilable differences and antagonism between the two fields are there because, to put it a bit more abstractly, there are differences between the actual state of affairs in field X and the way field X is perceived by field Y, where the negative reaction of Y to X is based on Y's perception of X.

The paper will also discuss points of convergence between the two traditions. One recent point of convergence in fact concerns the emergence of generative typology. Several other points of convergence will be discussed, including the minimalist assumption that language is characterized by efficient design, which opens the door for bringing in functional considerations into formalist approaches like minimalism. More generally, the paper argues for an overall view of the field where the functional and formalist approaches are seen not as being in competition but as complementary to each other (much of which will be based on a re-evaluation of some of the fundamental issues regarding the field where the formalist and functionalist approaches have been assumed to be in opposition (more precisely, where the practitioners of the two camps have been actively antagonistic to one another). Concrete examples of complementarity will also be provided.

1. Formalism vs functionalism: Not that much of a divide

This section is intended to show that the divide between the generativist and the functionalist/typological camp is bigger in the slogans that are used as characterizations of the respective camps than in actual research practice. Much of it is due to misunderstandings of the slogans in question, as well as taking them too literally, at their face value, without actually trying to see what is behind them.

---

[1] Below, I will use TT and GT to refer either to traditional typology and generative typology or the practitioners of these approaches, a distinction which should be clear from the context. Additionally, because of the TT/GT distinction, the term typology will often be used neutrally below.

[2] The paper will thus also contain a more general discussion of these approaches (see Thomas 2020 for a recent overview). The two approaches have been in opposition for so long (and rarely communicating with each other, the only communication often involving one-sided attacks and dismissals), that it is really impossible to find anyone who is completely neutral between these two approaches. This paper is written from the perspective of a formalist who is looking for a rapprochement between the two approaches. The perspective will inevitably in some places give the discussion a somewhat subjective (but also activist) flavor (for another work from a formalist perspective which is still quite different from the current one, see Newmeyer 1998). An important note: there are various formal approaches which considerably differ from each other (functional approaches are also far from being monolithic); when the differences are important enough to affect the discussion, what is assumed by the formal approach will be the Chomskian tradition, broadly characterized by what is referred to as the Principles and Parameters or the Minimalist approach.

1.1. Every language in its own terms

A position that is often associated with the traditional functionalist/typological camp by the generativists, in fact often looked at as an insurmountable and fundamental difference, is that every language should be described in its own terms. This is often perceived by generativists as a there-is-no-universal-grammar attitude.[3] The position in question is actually not universally adopted in the TT camp (more on that below). But there is a bigger issue here. What modern typologists mean by this is not what the generativists think they do. There are actually two misconceptions at work here. The generativists assume that what is meant by this stance is what American structuralists, who were the originators of the position in question (see Boas 1911), meant by it, which is not true. One of the reasons why it is not true is rather simple: for TTs, the stance in question arose at least in part as a reaction to some of the generativists' views discussed below, which American structuralists obviously could not have reacted to. Furthermore, there are misconceptions in the TT camp regarding the generativists' views in question. In other words, the negative reaction of generativists regarding the stance in question is (at least in part) based on misconceptions regarding what TTs mean by this view, and the TT view in question is in turn (at least in part) based on misconceptions regarding certain generativists' views.

To start untangling the cobweb of misconceptions concerning the generativist's reaction to the view in question, it does not seem that the practitioners of the TT camp truly believe it. The typologists from that camp have made incredibly important contributions to the field at large in terms of Greenberg-style generalizations, which the practitioners of the generative camp are increasingly relying on. Reaching such generalizations would not have been possible if they truly believed the slogan in question. American structuralists did believe it, but as a result, they also did not engage in typological work (see Greenberg 1974). A number of typologists have actually attempted to demonstrate that typological work is still possible while adhering to the slogan in question. For brevity, I will focus on one such work, Haspelmath (2010), one of the reasons being that the generativists who do look into the issue seem to take it to be a TT cannon, i.e. to reflect the general state of affairs in the TT camp (which actually is not true, as we will see). Haspelmath (2010) attempts to demonstrate that typological work is possible while adhering to the slogan in question regarding grammatical categories.[4] However, he also provides universal definitions of

[3]The perception among generativists that the stand is a reflection of a no-universal grammar (UG) attitude is somewhat misguided due to a difference in the phenomena that are investigated. As noted in Baker (2015), the kind of phenomena typological works typically explore are not considered by generativists to be the prime source of UG universals; those concern more abstract properties involving phrase structure, structural dependencies, locality relations involved in movement, coreference (im)possibilities…, which traditional typologists in turn generally do not deal with. To illustrate, here is one such generalization from Bošković (2012) (based on earlier work regarding only Slavic by Uriagereka 1988, Corver 1992; (i) is restated in the standard implicational universal way in fn 18, which also discusses other conditions on the possibility of (ii). Note that when checking (i), it is necessary to ensure that (ii) in the language considered does not involve a base-generated topic (something like "as for expensive (things), John likes expensive cars") and/or NP ellipsis in the sentence-initial constituent (something like 'as for expensive ~~cars~~, John likes expensive cars", where 'likes' or 'John' would likely be focalized). The most straightforward test to control for this would involve introducing an island between *expensive* and *cars*.)

(i) Left-branch extraction of adjectives (and adjectival-like elements), as in (ii), may be allowed only in languages without definite articles.

(ii) *Expensive$_i$, Mary sells [t_i houses]

[4]The following quote from Haspelmath (2020) indicates, however, that he does not take the slogan in question as literally as generativists assume that traditional functionalists/typologists in general take it (in fact, generativists would likely label the underlined part below UG—as discussed in section 4, there is actually much less disagreement between the two camps regarding the notion of "UG" than what is widely

the relevant categories, which are applicable to all languages and which make typological work possible.

He treats them as artificial linguistic constructs and not real (in fact not part of the grammar of individual languages), but there is really no deep reason (and, more importantly, nothing in the general TT worldview) why they should not be considered real. Haspelmath makes a distinction between a language particular descriptive category, call it X, and its crosslinguistically applicable comparative concept, call it Y (which is used in typological generalizations). But there does not seem to be a real issue here—it is possible that in some language, X is exactly like Y, while in another language, where this is not the case, we have a more complex situation where X=Y+Z; so there is still Y in that language as well. As an illustration, consider the typological generalization in (1a) and the definition of the relevant element in (1b), an example of Y.

(1) a. GENERALIZATION: In all languages, markers of future tense are less bound than markers of present tense or past tense, or equally bound, but never more so.

b. DEFINITION: A future tense is a grammatical marker associated with the verb that has future time reference as one prominent meaning. (Haspelmath 2010:671)

To make his point regarding X and Y, Haspelmath observes crosslinguistic differences regarding future tense, e.g. in Spanish it is also used to express probability (but not habituality), while in Lezgian it is also used to express habituality (but not probability). This shows future tenses are not synonymous crosslinguistically, which then necessitates making a distinction between a language particular descriptive category (X) and its crosslinguistically applicable comparative concept (Y). But what we really have here is the more complex X=Y+Z situation, where Y is still always present. Haspelmath gives similar definitions of other concepts (e.g. question words and ergative case), with similar crosslinguistic differences, all of which instantiate the X=Y+Z situation. Haspelmath considers Ys to be concepts created by linguists for the purpose of formulating typological generalizations. While they are applicable to all languages, they are supposed to be artificial, i.e. not psychologically real and not part of particular language systems. But there is no real reason why those Ys could not be real (and in fact part of UG from the perspective of a formalist; note I am putting aside here the question of what the real primitives of UG in the relevant domain are, which is irrelevant to the general point made here). In fact, a number of TT works have expressed this view, see e.g. Gill (2016), Lander and Arkadiev (2016), Round and Corbett (2020) (for a criticism of Haspelmath's position in question, see especially Spike 2020). In some languages those abstract categories would happen to map straightforwardly to surface categories, and in others that would not be the case: in such a case we could have the X=Y+Z situation (with Y applicable to all languages). This kind situation would become more obvious if it is accepted that the grammar of each language that is studied in its own terms is, as Baker (2015:936) puts it, "abstract to some non-trivial degree" (which is what generativists generally accept)—this would result in more Y=X situations and more generally make the Y-X relationship more transparent.

At any rate, the relevant concepts can be defined differently for each language, or in a way that would at the same time make them universally applicable (which would be more abstract; abstract does not need to mean not real and artificial[5]). The latter is anyway needed for typological work, which makes Occam's razor ('use what must be there as much as possible so as to avoid positing additional things') pertinent here. As Haspelmath (2010) observes, a number of non-

assumed): "First, language description is true to the categories of each language, but is inspired by the accumulated knowledge of comparative linguistics" (Haspelmath 2020:14).

[5]There is nothing strange in what is more abstract being psychologically real. Consider e.g. the concepts of allophones and phonemes. In a typical case of allophonic variation, what is psychologically real is the abstract phonemic level, which actually does not correspond to anything that is physically real, since the phoneme will always be physically realized as one of its allophones.

generative typologists refute the view that every language must be described in its own terms (e.g. Dahl 1985, Bybee and Dahl 1989, Lehmann 1989). In this respect, Lehmann 1989:142 says: "Describe your language in such a way that the maxim of your description could serve, at the same time, as the principle of general comparative grammar—and thus, the maxim of description of any other language." This will give us fewer mechanisms, which reflects Occam's razor as a general scientific principle (in addition to being a prerequisite for doing crosslinguistic typological work). Haspelmath's position seems to be a result of accepting a certain level of abstractness in doing typological work but not in doing analyses of individual languages, which essentially leads to separating the two into different fields (as Haspelmath 2010:682 puts it, "the analysis of particular languages and the comparison of languages are thus independent of each other as theoretical enterprises"). Allowing the same level of abstractness for both, which would also be in the spirit of Occam's razor, would, however, dissolve this distinction (see also Round and Corbett 2020).

There are other typologists with positions similar to Haspelmath's (see especially Dyer 1997, Croft 2001, who antecede Haspelmath's work). The above discussion would extend to them. In fact, generative typology does not really differ from Haspelmath's position in that works in this tradition also essentially assume what I have referred to as X and Y above (so there is really no disagreement here), the only difference being that Y, which Haspelmath considers a linguist's construct, is treated as real and in fact part of UG. Thus, the same point that was illustrated with (1) can be illustrated with any of Bošković's (2008, 2012) generalizations regarding definite articles (see e.g. (i) in fn 3, which is the counterpart of (1a)), and Bošković's definition of *definite article* (which superficially shows similar variation across languages as future tense) in Bošković (2016b), which is stated in semantic terms (as is the case with many of Haspelmath's Ys).

Haspelmath's explicit distinction between what was referred to as X and Y above is nevertheless a welcome and useful warning that should be heeded; those Ys that both traditional typologists and generative typologists are using are very often not quite the same as Xs used in individual languages, a difference which does get overlooked, especially by the latter.

At any rate, this is a case where there is less disagreement between traditional and formal typologists in practice than what is generally assumed; note that the main point is actually methodological—assuming the X/Y distinction is necessary methodologically to be able to do typological work, whether those Xs are real or not is a separate issue (where in fact there is no full field vs field disagreement; as Round and Corbett 2020 (see also Spike 2020) put it, this is an issue of a more general philosophical understanding of science, which is independent of the two approaches to language discussed here.

There is another aspect of the describe-every-language-in-its-own-right view which should be taken by the generativists as a methodological warning, be careful before jumping to conclusions that something is in UG—it's a warning not to follow without further checking a detailed investigation of a single language with a proclamation that it is all UG. Unfortunately, this tendency is still there among the generativists to some extent—I am not talking here about very abstract properties like investigations of e.g. c-command and domination where a single language can be used as an illustration basically for ease of exposition (see Epstein 1999)—I am not aware of any language where the notions of c-command/domination (which essentially means structure) do not hold, but detailed investigations of the structure of a single language which is immediately followed by a proclamation that all languages are like that without even checking the language next door (or simply by forcing other languages into the mold set up by that detailed investigation of one language without seriously looking at what does not fit).[6] From this perspective, as a generativist, I understand and am sympathetic to the mantra look-at-languages-in-their-own-right.

[6]There has been strong emphasis on investigations of understudied languages in the generativist camp in recent years. However, Germanic and Romance to some extent still hold a privileged place when it comes to UG proclamations of the kind noted above, with understudied languages often being used to confirm

At any rate, the view endorsed here is that there are a lot of similarities across languages, but also a lot of differences—the quest for the former should not ignore the latter (and the other way round). While there are extremes in both camps (those who in an American-structuralist style over-emphasize and truly believe the mantra look at languages in their own right, which leads to missing crosslinguistic similarities, and those who overdo it in the opposite direction by ignoring cross-linguistic differences, forcing all languages into one of those 'privileged' languages from fn 6), the majority of both TTs and GTs seem to hold the view expressed above, regarding TTs, see e.g. Gill 2016:458-459, which underscores the lack of a fundamental difference between them in this respect.

The look-at-languages-in-their-own-right mantra is used by TTs as a reaction to what they seem to take to be an assumption held by everyone in the generative camp, in fact one of the defining beliefs of that camp, which is that there is a universal sentence structure holding for all languages, the underlying assumption being that if we were to look at languages in their own right it would become clear that there is no such thing.[7] But there is a misconception here regarding the actual state of affairs within the GT camp (in fact, the situation here is similar to the misconception that the generativists have regarding the look-at-languages-in-their-own-right position). While it is true that the universal structure claim is often made in isolation by the generativists, in actual research practice it is often given up. In fact, the most radical departures come from generative typology works, where the claim in question is argued against on typological grounds (which can actually be interpreted as a point of convergence between the two approaches to typology). E.g., Todorović (2016), a typological study within the formalist tradition, argues against the universal presence of TensePhrase—in particular, she argues for a broad typological distinction between languages with and without Tense (following a suggestion from Bošković 2012), correlating this distinction with a number of properties).[8] Similarly, Bošković (2008, 2012) gives a number of crosslinguistic generalizations where languages with and without definite articles (henceforth with/without articles) are shown to differ regarding numerous syntactic and semantic properties, which cannot be accounted for if the distinction between languages with and without articles is simply a matter of phonology, namely whether articles are overt or null.[9] Based on this, Bošković argues that languages without definite articles do not simply have a null definite article, they lack

---

those proclamations (as a result of which they are often used as mold fitters)—there are likely political and sociological reasons behind this but discussing them is beyond the scope of this paper (see Bošković 2021b).

[7]See in this respect Gill (2016:458): "Many of us have developed our views of language at least in part in reaction to a dominant and sometimes domineering universalist approach that attempts to impose an aprioristic set of universal categories on languages with respect to which the categories in question are completely irrelevant. Our response was to reject such universal categories, while returning to the old American structuralists' ideal of describing each language on its own terms (Boas 1911)." Much of what is referred to as (universal) grammatical categories in discussions of this kind in TT is framed in terms of (universal) clausal structure in generative literature. The discussion below will be stated in these terms.

[8]Todorović also shows that the labels that traditional grammars use, and which TTs often rely on, can be very misleading; thus, she shows that what is traditionally called Aorist and Imperfectum tenses in Bulgarian and Serbo-Croatian (SC) are actually very different things, in fact not even the same categories—in Bulgarian these are indeed tenses, while in SC they actually represent aspect.

[9]See fn 3 for one such generalization; three more are given in (i).

(i) a. Second-position clitic systems are found only in languages without articles.
b. Only languages without articles may allow scrambling (section 1.3).
c. Only languages with articles may allow clitic doubling.

The works in question give a number of other generalizations, concerning phenomena as varied as interpretation of superlatives, negative raising (see section 3), sequence-of-tense, pro-drop in the absence of agreement, head-internal relatives, polysynthesis, multiple wh-fronting, possessives, numeral classifiers, subject reflexives, number morphology, scope, negative constituents, adjunct extraction, and focalization (see also fn 12 for additional semantic arguments).

it altogether, they in fact lack the DP projection (which is the only possibility for the structural placement of the definite article; other D/DP-related elements from a language like English can be located in other projections, in which case they show different behavior from the corresponding elements in English, see Bošković 2012[10]). This position argues against the universal structure hypothesis, where we would expect all languages without articles to have phonologically null articles (it would simply be a phonological accident that articles are unpronounced in some languages). Now, variation regarding the presence of DP is not a universally adopted assumption within the generativist camp. This is actually tied to the issue noted in fn 6: fundamental proposals of this sort are generally made on the basis of Romance and Germanic. Consequently, most of the time those who adopt the universal DP assumption (which comes with a "phonological accident") adopt it with no further discussion (after all, Romance and Germanic have DP, so all languages must have it), or attempt to fit other languages into the Romance/Germanic DP mold while ignoring relevant differences, i.e. ignoring what does not fit.[11] Even worse, they do it at a rather significant cost. As discussed in Bošković (2012), extraction patterns out of the nominal domain are completely different in languages with and without articles (for a partial illustration, see (i) in fn 3 and section 3.2). Locality restrictions on movement are currently stated within the phase theory. There are two mechanisms that can be used to capture crosslinguistic variation regarding extraction of the kind that is found in the nominal domain: assuming structural differences (as in the NP/DP proposal) or variation in the locality system, i.e. phases. The latter concerns variation within the computational system itself; the former, on the other hand, can be easily stated in lexical terms (in terms of a particular feature, +definiteness). Most generativists assume that there is no variation in the computational system itself, which means that there should be no variation regarding phases—all variation should be tied to lexical properties (e.g. Borer 1984, Boeckx 2008). The issue here is that those who assume universal nominal structure would generally also assume that there should be no variation regarding phases (i.e. the computational system). But it is simply not possible to assume both (unless we ignore differences in extraction patterns). Either we have variation in the structure (structural variation) or in the locality system (which means the computational system)—something has to give. (The universal DP analysis is often seen as having an appealing universalist character, but that is actually not true: it leads to a non-universal locality, i.e. phasal, system.) Adopting universal structure thus has a rather significant consequence, which those who adopt it don't seem to be aware of (since they generally do not discuss the issue).

Another point is worth noting. It's often assumed that there is a universal structural hierarchy (referred to as functional sequence). Take the abstract structure in (2) to represent it.

(2) [XP [YP [ZP [KP

The standard universalist approach to the structural sequence is not that the whole sequence is always projected (CP e.g. is not projected in raising infinitives like *she seems <u>to like me</u>*; even a more drastic case of non-projection of full clausal structure concerns restructuring infinitives, which are even smaller, see e.g. Wurmbrand 2001, Cinque 2004). Rather, the standard universalist approach (even this is actually not really widely accepted) is that the structural hierarchy in (2) is

[10]See also Fukui (1988), Corver (1992), Zlatić (1997), Chierchia (1998), Cheng & Sybesma (1999), Lyons (1999), Willim (2000), Baker (1996), Trenkić (2004), Despić (2013b), Marelj (2011), Takahashi (2011), Jiang (2012), Talić (2013), Cheng (2013), Runić (2014), Kang (2014), Bošković & Şener (2014), Zanon (2015), among others, for no-DP analyses of individual languages without articles.

[11]The universal DP literature often cites Progovac (1995), who argued for a DP in SC based on certain alleged parallelisms in word order between SC and Italian, completely ignoring the fact that Despić (2011, 2013a) has subsequently quite conclusively shown that these parallelisms do not hold at all (for similar situations, see Franks 2019 regarding DP claims in LaTerza 2016, and Bošković 2009 regarding Pereltsvaig 2007).

respected: (a) there cannot be a language where YP is higher than XP and (b) it is not possible to project KP, ZP, and XP without projecting YP. There is nothing in the NP/DP approach that goes against either of these. In fact, there is nothing in the NP/DP variation approach that goes against anything in the standard assumptions regarding sentence structure: requiring that every nominal domain must project to DP would not be any different from requiring that every clause be a CP, including, e.g., restructuring infinitives, which are standardly assumed not to be CPs. The works which argue against the NP/DP approach on universalist grounds thus seem to be based on very non-standard assumptions regarding these universalist grounds.[12]

The main point of the above discussion is that in this particular case (universal structure) the difference between the two camps is smaller than it is assumed to be. TTs are concerned with the structural issue in question (although they don't state it in these terms); they are generally not concerned with theoretical issues pertaining to the locality of movement. They generally assume that there is no universal structure (the look-at-languages-in-their-own-right mantra being a reflection of that stand) and that they differ in this respect from GTs, but they are actually not aware of the full range of views among the generativists. Many generativists who espouse the universal structure view are in turn unaware of the full range of consequences of that view (most of them would not want the computational/locality system parameterized, but they seem to be unaware that this is a consequence of the universal structure view). In this respect, it's worth noting that Cinque (1999) provides evidence that different adverbials are located in different projections, with more than 30 such projections in a fixed hierarchy partially given in (3).

(3) [*frankly* $\text{Mood}_{\text{speech act}}$ [*fortunately* $\text{Mood}_{\text{evaluative}}$ [*allegedly* $\text{Mood}_{\text{evidential}}$ [*probably* $\text{Mod}_{\text{epistemic}}$ [*once* T(Past) [*then* T(Future) [*perhaps* $\text{Mood}_{\text{irrealis}}$ [*necessarily* $\text{Mod}_{\text{necessity}}$ [*possibly* $\text{Mod}_{\text{possibility}}$ [*usually* $\text{Asp}_{\text{habitual}}$ [*again* $\text{Asp}_{\text{repetitive(I)}}$ [*often* $\text{Asp}_{\text{frequentative(I)}}$ [*intentionally* $\text{Mod}_{\text{volitional}}$ [*quickly* $\text{Asp}_{\text{acelerative(1)}}$ [*already* T(Anterior) [*no longer* $\text{Asp}_{\text{terminative}}$…

The full universal structure view, which is what Cinque (1999) espouses (and which the functional sequence hypothesis discussed above would require) is that all these projections are present even in both clauses of *I said that he left*, a sentence with no adverbials, in any language. While Cinque does assume they are all present, this is not a widely held assumption (see also (4)). It is thus fair

[12]Syntacticians occasionally assume DP is necessary for semantic reasons, to be able to interpret a nominal as an argument. Curiously, this is not what a semanticist would assume, see e.g. Chierchia (1998) and Dayal (2004), who crucially adopt the NP/DP distinction. In fact, there are rather strong semantic arguments against the universal-DP Hypothesis. The hypothesis assumes that the only difference between a language like English and a language like SC regarding articles is phonological: SC simply has a null *the*. A number of Bošković's generalization actually concern semantic phenomena, which shows that this simply cannot be right. Consider also Jenks's (2018) and Despić's (2019) arguments regarding the anaphoric use of nouns. Jenks observes that a bare noun cannot be used in Mandarin in donkey anaphora contexts like (i), which is surprising if Mandarin has a definite article just like English, which just happens to be phonologically null (Mandarin requires a demonstrative on the anaphoric/bound reading of "donkey").
(i) Every farmer that has a donkey_i beats <u>the donkey</u>$_i$.
Consider also Despić (2019) on the anaphoric use of mass nouns, illustrated by (ii). SC (iii) cannot have the meaning English (ii) has, with *fruit* anteceded by *grapes* (that reading requires a demonstrative). Despić shows this is a more general difference between languages with and without articles, also noting that if the latter had a definite article, which would just happen to be phonologically null (so the only difference would be in phonology), this would be totally unexpected. The conclusion is that the difference between the two language types is deeper—it's not a matter of phonology but syntax and semantics—there is no null *the*/DP in languages without articles in the counterparts of constructions where languages like English use it.
(ii) We have been growing grapes for generations – and you know, we have made millions on the fruit.
(iii) Naše mesto već generacijama proizvodi belo grožde. Sve dugujemo <u>voću</u>.
our town already generations produces white grape everything owe fruit-dat.
'Our town has been producing white grape for generations. We owe everything to that fruit.'

to say that the assumption that all projections from (3) are present in every clause, which is what the all-structure-is-universal hypothesis would entail, is not uniformly accepted in the GT camp.

The universal structure proclamation is still often made by generativists, but not really adhered to (which means not truly believed), which is especially clear regarding works on the left periphery of the sentence.[13] To give some illustrations, most accounts of the voiding of the *that*-trace effect in (4b) (e.g. Bošković 2016a, Erlewine 2020) would actually also void it in (4a) if the projection where the adverb is located in (4b) is also present in (4a).

(4) a. *Who do you think that would leave Mary?
b. Who do you think that under no circumstances would leave Mary?

Erlewine (2016) shows that in Kaqchikel, where wh-phrases and indefinites have the same form (a common crosslinguistic pattern, see Haspelmath 1997), the relevant elements are fronted to the specific projections in the left periphery on both functions, with the first element interpreted as a wh-phrase and the second as an indefinite pronoun when two of them are present. It is pretty clear that indefinites cannot be undergoing this kind of fronting universally (either overtly or covertly), assuming this (even covertly) would create havoc regarding e.g. scopal interpretation of indefinites.

A similar non-universality conclusion follows from the works on languages that front all wh-phrases. Thus, Rudin (1988) and Bošković (2002) show that such multiple wh-fronting languages (MWF) differ regarding the landing site of MWF; the highest clausal projection (CP) or lower than that. Thus, despite superficial similarity, there are numerous differences between (5) and (6), e.g. regarding the penetrability of fronted wh-phrases, their ordering, the availability of single-pair answers, inversion, the possibility of fronted wh-phrases following subjects, which can all be accounted if SC MWF lands lower than Bulgarian MWF (see Bošković 2002).

(5) Koj kakvo kupuva?
who what buys (Bulgarian)
(6) Ko šta kupuje?
who what buys (SC)

If the lower and the higher wh-fronting projections were always universally present, given the standardly assumed shortest move requirement, wh-fronting would always have to go to the lower wh-fronting projection (furthermore, due to the well-known and standardly assumed freezing effect (see e.g. Rizzi 2006), further movement from this projection would not be possible). It follows then that Bulgarian, in fact any language that fronts wh-phrases to the same projection as Bulgarian (this is also what English does, the only difference being that English fronts only one wh-phrase) cannot have the lower wh-fronting projection that SC has.

Or consider Russian (7a). Russian has been argued to have a high NegPhrase, where negation in this high NegP has semantic effects but does not have the true meaning of negation, as the translation shows. This negation also does not license negative concord (Brown and Franks 1995). It is clear that we don't want this NegP to be present in English (where negation always means negation) or other negative-concord languages, e.g. Serbo-Croatian, where negation always licenses negative concord. (Zanon 2020 notes that (7a) would be used in a context where John promised to stop by at some point this week but did not specify the day; *Didn't John stop by today* is not possible in this context in English; instead, *Did John stop by today* would be used.)

[13]There is a fallacy of universality syndrome among generativists, where the word is sometimes jumped at and adopted without really thinking about it (or with clear arguments to the contrary ignored, see e.g. fn 11). Obviously, not everything is universal. But the fallacy of universality syndrome sometimes leads to pulling out the "universal" card too quickly, without thinking about its consequences in the particular context in which it is used.

(7) a. Ne zaxodil li Sergey segodnja? (Zanon 2020)
neg stop.by Q Sergey today
'Did Sergey stop by today by any chance?'
b. Zaxodil li Sergey segodnja?
'Did Sergey stop by today?'

Rizzi (1997) argues that the traditional CP should be split into a number of projections, shown in (8), primarily based on Italian.
(8) [ForceP [TopP* [IntP [FocP [TopP* [FinP [IP]]]]]]]

While (8) is often cited as a universal structure holding for all languages, there is evidence even internal to Italian that all the structure from (8) cannot always be present, see in this respect the anaphor binding data from Petrossino (2018). Furthermore, Abels (2003) provides an account of the general immobility of IPs dominated by CP (which holds crosslinguistically), illustrated by (9), where it is crucial that the CP here is not split at all.

(9) *[IP His mother left]$_i$ everyone believes [CP that t_i]

It thus seems clear that CP cannot be always or uniformly split in either Kaqchikel, Russian, English, or MWF languages—there is no uniform split CP field that is present either crosslinguistically or in all constructions of a single language (Rizzi 1997:314-315 actually acknowledges this possibility; in fact, the facts discussed above indicate that even the version of the universal structural hierarchy where it's not possible to project KP, ZP, and XP in (2) without projecting YP cannot be right). The above illustrations are really just the tip of the iceberg. There is a great deal of crosslinguistic variation regarding left periphery which pretty strongly argues against structural universality of the left periphery.

At any rate, the point here is that there is a plurality of views within the generative camp regarding the notion of universal structure, a proclamation that is often made but not really adhered to, the look-at-languages-in-their-own-right TT stand being (at least in part) a reaction to that notion, as a result of which there is actually less disagreement here than what is believed within the TT camp. It should also be noted that many grammatical categories whose universality is questioned in TT works in what is taken to be disagreement with generativism are not taken to be universal, or even real at all, in generative works (this e.g. includes the notion of *subject*).[14]

In fact, just like there is a plurality of views regarding universality in the case discussed above within the generative camp, there actually is also a plurality of views regarding non-universality in the TT camp. The quote from Gill (2016) from fn 7 continues as follows: "However, some of us have gone beyond rejecting specific proposals for universal categories, such as subject, adjective, or whatever, and allowed our prejudices against such categories to lead us to deny the very possibility of universal categories. It is this latter move that seems to me to be an unwarranted overreaction…I have been outspoken against the Eurocentrically-motivated imposition of universal categories such as noun, verb, and their various phrasal projections on languages that offer no evidence for their presence. However, it does not follow from this that universal syntactic

[14]The issue here is that what TTs have been reacting to with the stance in question may have been true (to some extent) of the past research in the generative paradigm, but this is no longer the case (the development of GT did contribute to this). In fact, the current state of the field within generativism encourages investigation of crosslinguistic differences (contrary to the perception among TTs that it "actively discourages the investigation of such differences" (Gill 2016:459)) as well as investigation of understudied languages to the point that it is almost a must on the job market (field methods classes are also becoming a must in the curricula of generative departments; in my department they are regularly offered, the last one was on Mandinka). So the reaction is still there, but what is being reacted to is actually no longer there (except in the slogans, as discussed above).

categories do not exist; it's just a question of choosing the right ones… while pursuing linguistic diversity, it is important not to lose sight also of the ways in which languages may resemble each other, and of the possibility that all languages may embody a fundamental unity." Gill (2016) is certainly not unique in this view within the TT camp, see e.g. Round and Corbett (2020) and Lander and Akadiev (2020 (the two issues of *Linguistic Typology,* 20.2 and 24.3, are actually very useful starting points for generativists who would like to familiarize themselves with the broader range of views regarding the issues under consideration within the TT camp).

There is a perceived all-or-nothing/either-or difference between the two camps—it's all universal or nothing is universal. Note we are talking here about perception of X by Y, not the actual state of affairs in X (where X/Y stand for the two camps). In reality, what we are dealing with here is a matter of degree, i.e. how much is universal. This is very different from what's perceived. Different TTs and different GTs differ regarding the exact degree, but this is a very different situation from an all-or-nothing difference that would hold across the two fields, with everyone in completely opposite corners, which would not leave any room for common ground or an opening for a dialogue, since one side would have to be 100% wrong. The degree difference in fact opens the door for what should be a productive dialogue regarding how much, and what exactly, is universal.

At any rate, the upshot of the above discussion is that the two camps are really unaware of the full range of views within the respective camps regarding the issues/slogans discussed in this section, and what is really behind them, which leads to the impression that there is more disagreement between the two camps than there really is.

There is a difference here between what is proclaimed and what is truly believed (as shown by the actual research practice): the generativists react negatively to the every-language-in-its-own-right mantra since they take it at face value (taken as such, it does make comparative work impossible and reflects a no-UG attitude (an issue I return to below)); while it was intended to be taken at face value by American structuralists, this is not the case with the practitioners of TT who adopt it nowadays; they in turn use this mantra partly in reaction to a particular bad practice of the generativists (pulling the UG card too easily when examining details of the structure of a particular language) and in reaction to a universal structure proclamation that the generativists make (they in fact also always make it in reaction to the every-language-in-its-own-right stand), though they do not really believe it, as revealed by the actual research practice.

In other words, much of perceived disagreement comes from misinterpretations of pragmatically motivated slogans, where the two sides react negatively to what the other side is saying because they don't realize that what is said does not straightforwardly reflect what is really believed. Just like the TTs don't really believe in the mantra under discussion in this section (otherwise they would not be engaging in typological work), the generativists don't really believe in the universal structure slogan (which is easy to show by looking at the actual research practice).

1.2. Syntax as a tool

Another widely assumed difference between functionalists and formalists concerns their stands regarding the role of syntax. The perception of the difference is so significant here that there are functionalist works where simply showing that something is a semantic or pragmatic (rather than syntactic) phenomenon is taken to argue against the generative approach in general. Functionalists generally rely on much more impoverished syntax than generativists. The reason for this is mostly methodological, which in turn concerns their primary goal: for them pragmatics (and semantics) is more basic than syntax; they often look at syntax simply as a tool for expressing pragmatic functions and semantic roles—as a result, they generally do not consider syntactic relations that go beyond the tool role of syntax (they also generally do not consider what is not possible, since their goal is to determine how to capture what is possible, i.e. how to express the needed pragmatic and semantic notions; this is in fact something they have in common with the semantic approaches that generative syntacticians rely on, a point I return to). A generativist pursues a different methodology

here, which is again connected to their primary goal that in a way gives primacy to syntax: A generativist is interested in examining the full complexity of syntactic relations, unbounded by the tool role of syntax; they are interested in uncovering syntactic principles that determine well-formed and ill-formed sentences—pragmatics and semantics take the former and assign them interpretation/pragmatic use (for some relevant discussion of formalism vs functionalism in this context, see also Baker 2015:21). Importantly, there is really no deep worldview difference here. The approaches to semantics and pragmatics generative syntacticians rely on also assume much poorer syntax than generative syntacticians do and are also not concerned with what is not possible syntactically. But the reason for this difference is simple: they investigate semantics and pragmatics, not syntax—they go into syntax only to the extent it is relevant to their concerns. Due to the nature of their inquiry, syntax is just a tool for them (and the same holds for syntacticians when it comes to semantics and pragmatics). There is no fundamental difference in worldview here, they just do different things. The same in fact holds for functionalists and generative syntacticians in this respect. What is taken to be a deep-seated difference in the worldview is actually just a byproduct of them doing different things—the difference here is very similar to the difference between generative syntacticians and the approaches to semantics/pragmatics that generative syntacticians rely on. As a result, there is really no deep reason why many of the results reached in functionalist works could not be incorporated into generative works. This is not happening in practice due to the pervasive perception that the two approaches are so fundamentally incompatible that the practitioners of the two approaches generally do not read each other's works, even when examining the same topic. They are in fact not incompatible, to a large extent they are complementary (just like generative syntax and the approaches to semantics/pragmatics that generative syntacticians rely on).[15]

Such complemenarity can be easily illustrated. Consider the phenomenon of ergativity. There are numerous syntactic differences between a verb like *work* and a verb like *arrive* crosslinguistically, which in the generative approach (in the Principles and Parameters tradition) are accounted for by having *Mary* start the derivation in different positions in (10) and (11).

(10) Mary$_i$ [$_{VP}$ t_i works]
(11) Mary$_i$ [$_{VP}$ arrived t_i]

A functionalist (e.g. DeLancey 2001) would complain that these structures do not explain why *work* and *arrive* differ in the relevant respect. This is certainly a valid complaint, and the ultimate explanation will likely not be syntactic—it may very well turn out to involve cognitive or communicative factors (for some discussion relevant to these issues, see Levin & Rappaport Hovav 1995, Kuno & Takami 2004). But that would not invalidate all the structural/syntactic reflexes of the *work*/*arrive* distinction that hold across a variety of different phenomena crosslinguistically,[16] and which the derivations in (10) and (11) unify. Providing a non-syntactic explanation for why *work* and *arrive* differ in the relevant respect can be complementary to the syntactic differences that generative syntax has uncovered in this respect.

We are dealing with a broader issue here: functionalists often raise very valid "why" questions which, when taken seriously, indicate we need more than just syntax (even in the broad sense the generativists understand it) for particular phenomena, but the non-syntactic answers to those why questions very often can be added to the syntactic accounts, which would then give us better, more comprehensive accounts of the relevant phenomena. (Instead, the functionalists often interpret we-

[15]Of course, sometimes it is not clear whether a particular phenomenon should receive a functional or a formal explanation, just like sometimes it is not clear whether a particular phenomenon should receive a syntactic or a semantic explanation—there is no deep incompatibility here.
[16]To mention a few, *ne*-cliticization in Italian and genitive of negation and *po*-phrases in Russian (for an overview, see Alexiadou, Anagnostopoulou, and Everaert 2004).

need-more-than-syntax as we don't need syntax at all, and then ignore the syntactic part; the generativists, on the other hand, should be faulted for not raising, and missing, those why questions (which includes ignoring possible functional answers to those questions)).

In the next section I will discuss the status of Greenberg-style typological generalizations regarding the concept of Universal Grammar. From a formalist perspective, at the right level of abstractness that also dissolves exceptions to them (see below) Greenberg-style generalizations reflect UG at work—this has in fact prompted the development of generative typology; from this perspective the practitioners of the traditional functionalist/typological camp have contributed a great deal to the formalist's understanding of UG.

2. Greenberg-style generalizations and Universal Grammar

2.1. On the status of typological generalizations

Above I have discussed Haspelmath's distinction between language particular descriptive category and its crosslinguistically applicable comparative concept, observing that the distinction is also adopted in generative typology, though with a difference regarding the status of the latter, which concerns the notion of UG. The notion is supposed to represent a significant difference between traditional and generative typology. However, we will see in this section (and section 4) that the difference regarding the notion of UG may also be smaller than what is generally assumed (i.e. there may not be real fundamental disagreement even here).

In fact, in many respects, again in actual practice, not the slogans associated with the respective approaches, the practitioners of the functionalist/typological camp seem to be bigger believers in universal grammar (see also sec. 4), and have contributed more to the notion (although they may deny it for reasons discussed below) than many generativists (I will refer to the two camps below as α for the former and β for the latter, strictly for expository reasons). The goal of many practitioners of the former is to use detailed investigations of individual languages to reach broad Greenberg-style crosslinguistic generalizations, while many practitioners of the latter use them (generally an investigation of an understudied language in this case)[17] to argue against proposed crosslinguistic generalizations. In doing so, the former, who are generally anti-Chomskian, do what they often accuse Chomsky of doing, and the latter, who are broadly classified as Chomskians, do what Chomsky himself would never do: in order to make sense out of what seem to be chaotic data, to be able to see patterns, parts of the chaotic data, sometimes even good chunks of it, have to be put aside. Chomsky's work is full of such examples, but this is simply the way science works, this is what is done in any mature scientific discipline. To reach those Greenberg-style generalizations, the α practitioner does exactly that, those generativists who attempt to knock off proposed crosslinguistic generalizations based on a single counterexample from an understudied language, which increases the likelihood that something has been misanalysed, do exactly the opposite.

Now, there is a reason why α practitioners would deny the label I have given them, 'believers in universal grammar' (see also sec. 4). α and β practitioners read Greenberg's generalizations very differently: here are some examples of Greenberg's generalizations, with the relevant parts bolded:

(12) a. When the descriptive adjective precedes the noun, the demonstrative and the numeral, **with overwhelmingly more than chance frequency**, do likewise.(#18)
b. Where morphemes of both number and case are present and both follow or both precede the noun base, the expression of number **almost always** comes between the noun base and

[17]While there has been a surge in the work on understudied languages within the generative approach, when theoretical issues are discussed, such work is mostly done either to confirm broader theoretical proposals made with respect to more widely studied languages like Germanic and Romance (see fn 6) or to argue against proposed crosslinguistic generalizations and/or theoretical proposals (so they are often used either to confirm or disconfirm proposals made based on Germanic, Romance…), they are seldom used to make such proposals (there are of course exceptions, like Baker's work).

the expression of case.(#39)

To α, the bolded part means "exceptions". Universal Grammar is supposed to mean no exceptions, so this is not UG, in fact it argues against UG—even when we come close to it, it is not that.[18] β, a Chomskian UG practitioner, ignores the bolded part; we then have universal generalizations here—examples of UG. (It's an interesting switch regarding the normal scientific methodology of putting exceptions aside until they can be better understood, α-s who pursue it to reach generalizations like (12) now drop it, those β-s who knock off proposed generalizations by ignoring it (i.e. by pointing out exceptions to them) now endorse it).

But they are really both right and wrong. (12) is and is not UG.[19] To β, what is supposed to be in UG is not generalizations like (12) but mechanisms that deduce them. In other words, generalizations like (12) would be theorems, not principles of UG. Very often, deductions of principles leave room for exceptions. The right deduction of (12), based on the mechanisms present in UG, should then leave room for the exceptions. Moreover, it should explain why they are rare.[20] This is what I mean by (12) is and is not UG. (12) is in UG, including the bolded part, but as a theorem. **But**: this is the case if (12) is deducible from the formal mechanisms of UG. A priori, we don't know: there could be formal explanations for (12a-b), or functional, or formal-as-a-reflection-of-functional-considerations explanations of the kind discussed below. The way to tease them apart is to try them all and see which one deduces (12), including the exceptions behind the bolded parts. It may in fact turn out that a formal/functional explanation combination is needed.

As an example of such combination, consider the generalization in (3) (see e.g. Sapir 1921, Alexander 1990, Bošković 2005).

(13) If a language has scrambling (informally, free word order), it has overt Case-marking.

---

[18]While many of Greenberg's original generalizations are stated as if they have no exceptions, the current stand on language universals within TT seems to be that none are exceptionless–they all have something like the bolded part from (12); see e.g. Bickel 2007, Nichols 2007; this, however, has not been actually shown for all of them. The reason for at least some of the exceptions may be the implicational "If-X-then-Y" form in which they are stated–exceptions may disappear if additional conditions in the *if*-clause are added (as noted by Baker & McCloskey 2007:288), as in if X and Z then Y, or if they are stated as clear one-way correlations of the form if X then no Y. Consider in this respect the LBE generalization from fn 3. If stated as in (i), it has exceptions, e.g. Japanese and Chinese, which lack definite articles but disallow LBE. However, it turns out agreement is also necessary for LBE (see Bošković 2012; thus, Serbo-Croatian has both agreeing and non-agreeing adjectives, only the former allow LBE). The statement in (ii), of the form if X and Z then Y, then takes care of Japanese and Chinese. They can also be taken care of with the weaker statement (which emphasizes what is disallowed, not what is allowed) in (iii), which is of the form if X then no Y (I am actually not aware of any exceptions to (iii)).

(i) If a language lacks definite articles it allows adjectival left-branch extraction (LBE).

(ii) If a language lacks definite articles and has agreeing adjectives, it allows LBE of such adjectives.

(iii) If a language has definite articles then it does not allow adjectival LBE.

[19]Baker (2011) raises the question of the relationship between the β notion of UG and Greenberg-style language universals. The following discussion provides a partial answer to this question. Note also that I will not be concerned here with actual deductions of (12) (though see Cinque 2005 for relevant discussion of (12a) and Harley & Ritter 2002 for (12b)).

[20]Formal explanations for why something is rare but possible are in principle possible—see e.g. Baker & McCloskey (2007) on the rarity of VSO languages (the reason being that a constellation of factors is needed for that type; for a different way of capturing rare patterns based on exceptional mechanisms from the point of view of UG, which concerns person restrictions, see Stegovec 2019).

One should be careful with exceptions though; there is an inherent noise when working with descriptive works that can arise due to the errors/misreading of the original sources (see Baker & McCloskey 2007: 290); what appear to be exceptions may then turn out not to be exceptions, which in turn means that what appear to be only strong tendencies due to such exceptions may actually be exceptionless generalizations.

The explanation for (13) quite clearly should be functional. To put it in informal terms, we need to know who does what to whom. If word order won't help in this respect, then we need Case.

But there is more to scrambling/freedom of word order than Case. One of Bošković's article generalizations in fact concerns scrambling, where only languages without (definite) articles may have scrambling (see (ib) in fn 9). Informally, this means that languages with free word order tend to lack articles (there is more to scrambling than this; at any rate, to mention some typical scrambling languages: Russian, Serbo-Croatian, Latin, Japanese, Turkish, Hindi, Chukchi, Warlpiri—they all lack articles). It looks like, then, that both the presence of Case and the lack of articles matter for the freedom of word order. While the former has a straightforward functional explanation, it is difficult to see how the latter could be explained in functional terms, it seems to call for a formal account (cf. the account in Bošković 2008). What this means is that both functional and formal factors may be at work here, i.e. a combination of the two may be in order.

At any rate, exceptions to a descriptive typological generalization do not mean that we do not have UG at work. There clearly are universal properties of language, e.g. c-command, domination, the notion of structure—they hold for any language. However, even research within the generative paradigm has reached the point where broad typological generalizations of the kind the traditional typologists have been concerned with have become crucial for understanding the nature of language and what the generativists refer to as the faculty of language. There is, however, a significant difference in focus between TT and GT here. When the latter are looking for typological generalizations they are generally looking for typological gaps (see below for illustrations[21]). If typological gaps are mentioned at all in traditional typological works they are most of the time brushed off as accidental gaps, the focus in the works in this tradition generally being on explaining what is possible, not what is not possible. There are certainly many aspects of language that are best explained by looking at the function of language. Functional explanations are certainly appropriate, but it is harder for them to rule out cases. They most of the time provide a rationale for what is found, it is harder for them to completely rule out cases, which is what accounting for typological gaps, which concern what is not possible, would take. (The rationale they provide is generally based on functional pressures, but pressures are most of the time a matter of degree, they are not either-or characterizations.)

Traditional typologists actually often assume that nothing is truly impossible, but they don't really look for what is impossible—the position is there simply by assumption; as Nichols 2007 notes, they really have no interest in what is not possible[22]—the generativists do, this is in fact one of their central interests (for reasons discussed below). The difference is even reflected in the way implicational universals are stated. While functionalists generally state them in the if-X-then-Y way, formalists often state them in the if-X-then-no-Y way (see fn 18). The former emphasizes what is found, and the latter what is not found—it directly reveals typological gaps, while the former implies them—this difference reflects the primary interests of the two approaches: for TTs typological gaps are implied, one has to look for them since the statements don't make them obvious. This is not the case with GTs, since they are looking for them.

[21]The article/adjectival LBE correlation from fn 3 actually reveals a typological gap. There are four possible language types here, (d) represents a typological gap: (a) languages with articles and no LBE; (b) languages without articles and LBE; (c) languages without articles and no LBE; *(d) languages with articles and LBE.

[22]There are exceptions though: explorations of the range of possible variation within particular domains in *Canonical Typology* (see e.g. Brown, Chumakina, and Corbett 2013, Bond 2019, Round and Corbett 2020) come close to that given that determining what is possible implies knowing what is not possible (in fact, with its stance toward universality (see section 1.1) and determining what is possible, Round and Corbett (2020) come close to the larger picture view of generativists, which underscores that the one-camp-vs-another distinction is fluid, and becoming more and more a matter of a degree.

One revolutionary aspect of the early steps in the generative tradition that is often overlooked was caring about ungrammatical sentences. What the difference between traditional and generative typology regarding typological gaps, more precisely, the obsession of the latter with typological gaps, boils down to is in fact the Chomsky revolution in caring about ungrammatical sentences. The non-existing language type issue is just a higher-level instantiation of the same concern: caring about ungrammatical sentences (i.e. what is not possible), explaining why they are ungrammatical. The driving force of research in the generative tradition has in fact been to rule out sentences, not rule them in. After all, we don't have to do anything to rule in a sentence. Suppose there is no such thing as grammar, every principle, every mechanism, every condition, none of them exist. Every good sentence of every language is still "accounted for", they are all ruled in. You don't have to do anything to 'rule in' a sentence. But all bad sentences then become a problem.

Chomsky's *Syntactic Structures* (SS) in fact set the research agenda in generative tradition in this respect. The *SS* system was extremely powerful; I'm not aware of any sentence in any language that could not be captured within that system (the system even allowed movement of non-constituents)—that was in fact the problem with the system. It could do anything. The system naturally allowed for too many things, constraining it has all been about blocking things that are not found in natural languages, in other words, ruling out ungrammatical sentences.

The same reasoning, only at a level higher than sentence, extends to generative typology. Traditional typology is all about finding broad scale generalizations regarding what is found in languages, and then providing a rationale for them. Generative typology also looks for broad scale generalizations regarding what is found in human languages. But the goal is different, it is looking for what is found in order to discover what is not found. To illustrate, say property X has 5 logical possibilities to be realized in human languages, but only four are found. A traditional typology paper will typically give the generalizations regarding what is found in this domain (and if an account is provided, a traditional typologist would typically focus on explaining why we have those existing four types—if the non-existing type is mentioned, it is typically put aside as an accidental gap). A generative typology paper will, on the other hand, give that, but it will then zero down on the gap in the paradigm, on the type that is logically possible but not actually found, and ask why that is the case.

To be more concrete, consider Greenberg's generalizations in (14) (note that they are stated in the if-X-then-Y way).

(14) a. If the relative expression precedes the noun either as the only construction or as an alternate construction, either the language is postpositional, or the adjective precedes the noun or both. (#24)

b. If either the subject or object noun agrees with the verb in gender, then the adjective always agrees with the noun in gender. (#31)

(14a) tells us which adposition/N and adjective/N order is found when a relative precedes N in a language. What a generativist is now thinking is that what (14a) really says is that we don't find a prepositional language where a relative precedes the noun and the adjective follows the noun. The next step is to block the possibility of such a language. Similarly, upon seeing (14b), a generativist is thinking how do we force A-N agreement in gender in the presence of N-V agreement in gender?

To give a concrete example of such generative typology works, consider Messick (2016, 2017) regarding attitude reports like *Bill$_i$ said that he$_i$ is smart*, which reports on *Bill said: "I'm smart"*.[23] Some languages (Amharic, Zazaki) use 1st person pronoun to refer to the attitude holder—so we get "Bill$_i$ said that I$_i$ am smart". Messick observes a new type, represented with

[23]I am putting aside *de-re* readings, where the attitude holder is unaware that the attitude is about them. On that reading, Bill could be drunk, watching himself on TV saying something smart, but being too drunk to recognize the person is himself Bill says: "he is smart" (instead of "I am smart"; some languages, e.g. Golan, mark such readings, which I ignore below, differently).

Telugu and Nuer, which uses a 3rd person pronoun controlling 1st person agreement—so we get "Bill$_i$ said that he$_i$ am smart". Some languages (Ewe) use a special logophoric pronoun here (logophors are found in embedded attitude reports and cannot be the matrix subject of out-of-the-blue sentences). Donno Sɔ and Tamil represent an additional twist in using a logophor with 1st person agreement.

By examining additional languages, Messick establishes a typology of attitude reports, and crucially observes a typological gap: there is no language that uses 1st person pronoun and 3rd person agreement, as in *Bill$_i$ said that I$_i$ is smart*. Messick then provides a comprehensive syntax/semantics/morphology account of attitude reports that accounts not only for the existing patterns but also for this typological gap. The details are not important, though it is worth noting that the crucial ingredient concerning a distinction between morphological and semantic features of agreement-controlling nominals comes from a traditional typology work, Corbett (1983). But what is particularly important is the illustration of the empirical scope of generative typology. It is about discovering what is not possible, it's about looking for typological gaps. Typological gaps are like ungrammatical sentences. This is where the soul of a generativists most happily resides.

The best way to understand why this is the case is to go back to *SS*. As noted above, there was nothing that the *SS* system couldn't do, which was taken to be a problem. Ironically, the generative revolution that started with *SS* set the broader agenda for the field (within this particular perspective), really defined its soul, in a way that appears to be quite contrary to *SS* though the appearances are somewhat deceiving. Already at the *SS* stage the field was defined as going after "all and only" the sentences of the grammar, it's just that the practical implementation of this motto with respect to the corner of the grammar of English discussed in *SS* was such that it emphasized "all", not "only".

Returning to typology, to illustrate very abstractly the difference between a traditional and a generative typologist in terms of the history of generative grammar, traditional typology would be something like the *SS* system, and generative typology something like what has happened after that; the former emphasizes what is found in languages, and the latter emphasizes, in fact is really obsessed with, what is not found. This is just a higher-level instantiation of the same obsession that is found on the level of a sentence, where the emphasis in the generative approach is on ruling out ungrammatical sentences. But generative typology is obviously enormously indebted to traditional typology.

The above is not intended to imply that functional approaches cannot capture gaps in typological paradigms. The difference here is more about the nature of inquiry, it's more methodological: formalists always try to capture such gaps (for the reasons noted above), functionalists generally don't try to capture them because they are not concerned with them—the focus being on explaining what is possible.

There is a parallel to be made here. Interface considerations constantly pop up while doing syntactic investigations in the generative framework; as a result, no generative syntactician can afford to ignore semantics; they would be doing it to the detriment of their own research. While there has been a great deal of productive interaction between syntax and semantics at the syntax-semantics border (much of that research in fact involves collaborations between generative syntacticians and formal semanticists), the difference noted above regarding generative syntacticians and functionalists also arises with generative syntacticians and formal semanticists. As noted above, generative syntax is obsessed with ungrammatical sentences. The field is almost all about blocking what is not possible, ruling out ungrammatical sentences. Principles, constraints…almost everything that has been proposed is there to rule out sentences. Formal semantics is very different in this respect, and similar to functionalist approaches. The principal goal of a formal semanticist is to write a semantics for an acceptable sentence; in fact, one seldom finds cases where a semanticist would write a semantics for an unacceptable sentence with a goal to rule it out. Most proclamations of the sort this-sentence-is-ruled-out-for-semantic-reasons,

because it cannot be interpreted, are actually made by syntacticians. All of this also has a counterpart on a higher, more abstract level of the theory itself—namely, with constraining the theory, which has in fact been the driving force of generative research since *SS*. Not surprisingly, semanticists are also much less obsessed with constraining the theory (in fact, syntacticians are quite uncomfortable with one of the most standard formal semantics tools, *type shifting*, the reason being that it is a very powerful mechanism that can do almost anything one wants it to do); in other words, the obsession with ungrammatical sentences and with constraining the theory seem to again go hand-in-hand here. Generative syntax differs from both formal semantics and functionalist approaches in these respects.

At any rate, typological generalizations should be treated the same way as all descriptive generalizations, which means they raise an immediate question why they hold. This in turn means that they need to be deduced, and the deduction should also explain why some patterns are rare (this also concerns any potential exceptions to the generalizations) and provide explanations for actual typological gaps (where they exist). A proper deduction of a typological generalization (whether formal or functional, or a combination of the two) should address all of these—this is how the success of a deduction of a typological generalization should be judged: the closer it gets to this goal, the better.

At any rate, typological generalizations need to be paid serious attention to—generativists sometimes ignore them to their own detriment. Take e.g. agreement. There are properties of agreement systems crosslinguistically which show that certain widely held theoretical assumptions regarding agreement are seriously on the wrong track. In the minimalist system of Chomsky (2000), functional heads v and T(ense) are the locus of object and subject agreement respectively (vP is right above VP and TP is higher than vP), with v and T undergoing agreement with the relevant nominals. There is, however, a serious problem with these assumptions revealed by Julien's (2002) typological study of verb morphology, based on over 500 languages. Agreement is commonly marked at the edge of the verb: while there are exceptions, it generally follows everything else on the verb.[24] Crucially, this includes object agreement, as in Itelmen (15).

(15) n-ə̲łčqu-z-um
3.pl.sub-see-pres-1.sg.obj
'They see me' (Bobaljik and Wurmbrand 2001)

This is quite surprising from the current minimalist theoretical perspective, where v, the head of a phrase right above VP, is the locus of object agreement: since v is lower in the structure than T, object agreement should occur inside of tense (i.e. be closer to the verb than tense). As Julien (2002) shows, (15) is in fact not a rare pattern, hence not something that should be easily dismissed.

The above can be taken as an illustration of the importance of typological work for generativists (given that paying attention to it is fairly recent; many generativists in fact still ignore it). Typological work needs to be paid serious attention to, sometimes it may help prevent the generativists from going down a wrong theoretical path, as the above discussion indicates regarding the current minimalist assumptions concerning agreement.

I'll now briefly discuss another example of generative typology, whose driving force is again a quest for typological gaps. The account will be compared to a functional approach. However, the two will not be opposed—while superficially they look very different, more abstractly, they share interesting points of convergence, which will be taken to indicate that convergence that comes from such radically different perspectives may provide real and significant insights. Going beyond

[24]Bantu languages may provide an exception, with tense occurring between subject and object agreement markers; however, these markers have been argued to have a different status regarding the clitic/affix distinction (Bresnan & Mchombo 1987 argue object markers are actually clitics), which may be an interfering factor here.

convergence (and in the general spirit of this paper), we will in fact see that there is room for both approaches in the comprehensive account of the relevant phenomenon.

2.2. On person restrictions

In many languages, person specification regulates co-occurrence of weak pronominal objects, a restriction referred to as the Person-Case Constraint (PCC) and assumed to come in two versions: STRONG PCC, where in the presence of indirect object (IO) direct object (DO) cannot be 1P or 2P (see Greek (16)), and WEAK PCC (found e.g. in Arabic), where if IO is 3P DO must be 3P.

(16) Tha **me**/**se**/**tu** **ton**/***me**/***se** stilune. *1P/2P/3P.IO≫1P/2P.DO
fut 1_{GEN}/2_{GEN}/$3_{M.GEN}$ $3_{M.ACC}$/1_{ACC}/2_{ACC} $send_{3PL}$
'They will send **him**/***me**/***you** to **me**/**you**/**him**.'

While less-known, PCC-like restrictions also exist for external/internal argument combinations. Stegovec (2019) argues for unifying the two restrictions, observing that external (EA)-internal (IA) argument PCC also comes in strong and weak version. STRONG EA-IA PCC is found in e.g. Christian Barwar and WEAK EA-IA PCC in e.g. Southern Tiwa. Stegovec observes that both strong and weak restriction can be unified for EA-IA and IA-IA pairs, e.g. for the weak one: when pronominal markers co-occur, if the S/IO is 3P the O/DO must be 3P.

Almost all formalist approaches to the PCC focus on IO/DO pairs, analyzing it in such a way that it is crucial that the first argument, IO, bears an inherent Case (the typical situation is that IO is dative and DO accusative). This obviously cannot be extended to EA-IA pairs. Furthermore, Stegovec (2019) shows that even IO-DO PCC is insensitive to the Case type of the first argument based on languages that allow both orders of IO/DO clitics. Thus, Slovenian shows PCC effects with DAT-ACC clitic order (17). If the clitic order is reversed, a reverse PCC arises: the IO's, not DO's, person is restricted (18).

(17) Mama {**ti** **ga**}/ *{**mu** **te**} bo predstavila. ✓2P.IO≫3P.DO;*3P.IO≫2P.DO
mom $2P_{DAT}$ $3P_{M.ACC}$ $3P_{M.DAT}$ $2P_{ACC}$ will introduce
'Mom will introduce **him** to **you**/* **you** to **him**.'

(18) Mama {**te** **mu**}/*{**ga** **ti**} bo predstavila. ✓2P.DO≫3P.IO;*3P.IO≫2P.DO
mom $2P_{ACC}$ $3P_{M.DAT}$ $3P_{M.ACC}$ $2P_{DAT}$ will introduce
'Mom will introduce **you** to **him**/***him** to **you**.'

This shows that person restrictions (PR) are not limited to inherent-Case-on-the-first/higher-argument contexts and that the position in the syntax, not grammatical function, is what matters.

Stegovec also conducts a survey, which builds on Haspelmath (2004) and Albizu (1997) but significantly broadens their scope by including more than 100 languages from more than 20 families. While the survey reveals more patterns than have been previously reported, it's not the case that anything goes. PRs always have these properties:

(19) a. The restriction always applies to the structurally lower marker.
b. The restriction either forces the lower marker to be 3P or bans it from being either 1P or 2P.
c. A language can only have a reverse PR if it also has the standard one ((18) is found only in the presence of (17)).
d. If a language has both EA-IA and IA-IA PR, the latter cannot be weaker than the former; cf. (20), where the first pattern refers to EA-IA PR and the second to IA-IA PR—(20d) is a typological gap

(20) a. WEAK+STRONG (Southern Tiwa, *Kiowa-Tanoan*)
b. WEAK+WEAK (Alutor*, Chukotkan*)
c. STRONG+STRONG (Telkepe*, Semitic*)
d. STRONG+WEAK unattested

Stegovec provides a formal account that deduces the generalizations in (19)—it allows attested but not unattested patterns and also explains why some patterns are rare due to them using exceptional theoretical mechanisms (e.g. the pattern where only 3P≫2P is banned is found only in Salish). The gist of the account is that there is no crosslinguistic variation in argument structure, locality domains and the way agreement works. Relevant crosslinguistic variation comes from independently motivated variation in the internal structure of pronouns, certain movement possibilities (DO-IO order arises through movement) and the presence/properties of certain functional projections (so there are no PCC-specific mechanisms). What's crucial in the account is the structural placement of a particular head, *v* (which introduces external arguments); in particular, EA is higher and IAs are lower than *v* (so this is not simply an issue of argument hierarchy). Regarding the issue of whether particular pronominal elements are involved in a strong or a weak pattern, what matters is whether they are weak pronouns or clitics.

It is instructive to compare Stegovec's approach with a predecessor he relied on, Haspelmath (2004), which provides a functionalist perspective on the phenomenon (for ditransitives). For Haspelmath, PR restrictions are about frequency and grammaticalization. The basic idea is that there are prominency scales for person, where 1/2P is ranked higher than 3P, and semantic roles, where recipient is ranked higher than theme. Person-role association is more natural when high (1/2) persons are associated with high (Recipient) role and low (3rd) person with low role (Theme). The gist is that the requirement for alignment of semantic-role prominence with person prominence disfavors ranking DOs higher than IOs in person. PCC then reflects alignment of grammatical function prominence with person prominence (the approach is most straightforward if *Weak PCC* is the core of PCC). At any rate, the gist of the approach is that the higher a pronominal element is on the thematic scale, the more likely it is to be 1/2P. Based on corpus studies, Haspelmath claims this is the case even in non-PCC languages as revealed by the frequency of usage (German is not a PCC language, allowing all pronoun combinations, but the most frequent one is 3P DO and 1/2P IO)—what is different with PCC languages is that this preference is grammaticalized (though it's not really clear what is meant by 'grammaticalized', which is an important issue here; the following discussion will in fact address it). It seems difficult to capture all relevant variation (see Stegovec 2019) under this approach. Also, it would need to be shown that all types of PCC effects in PCC languages have their type counterparts in non-PCC languages, which doesn't seem likely. However, syntactic accounts of PCC effects like Stegovec's do not have anything to say about non-PCC languages. There may then be an opening for both approaches, where something along the lines of Haspelmath (2004) would be correct for non-PCC languages, where the effect is not syntactically encoded. Syntactic encoding of the effect (cf. the above remark about grammaticalization) would then lead to a Stegovec-style approach. This can be looked at more broadly as formalization of functional factors.[25] In fact, it would not be surprising to see such a situation arising quite often, where a formal issue arises as a reflex of a functional (or a broader cognitive) concern, getting a life of its own upon this kind of syntactic encoding, to the point that it can go quite beyond the original functional factor.[26] The reason why this kind of situations are not seen more often may be the formalist/functionalist division in the field, which minimizes the interaction that would be necessary to locate such situations.[27] (Another case where a combination

[25]When functional considerations give rise to a formal mechanism, they may end up being incorporated in a rather abstract and formal way once the formal properties of the computational system of language are brought into the picture.

[26]The lack of a complete correspondence between PCC and non-PCC languages hinted at above would then not be a problem.

[27]Another relevant case may concern the requirement that subjects in Chinese be interpreted as specific/definite. While English doesn't have this requirement, specific/definite subjects are much more frequent than indefinite ones in English, which could also be looked at as a syntactization/formalization (in Chinese)

of functional and formal factors may be involved, concerning free word order, was noted above; see also the discussion of ergativity in section 1.2.; For different suggestions regarding potential ways of combining the two, see Baker & McCloskey 2007 and Polinsky & Kluender 2007).

From this perspective, while there are differences between Haspelmath's and Stegovec's account there are also important similarities. Haspelmath crucially relies on semantic roles and their prominence, which corresponds to argument structure in Stegovec's account, where the relative prominence is reflected in the structural height of arguments (in other words, generativists also have a counterpart of Haspelmath's semantic-role prominency scale; in fact, they also have something similar for the person scale, which has been used in accounts of the PCC, e.g. Bejar & Rezac 2009, Anagnostopoulou 2003). The reverse PCC pattern, however, raises a problem for Haspelmath's account. Semantic-role prominence doesn't matter here, it in fact cannot be appealed to: what matters is the syntactic height (the first object can be shown to be structurally higher than the second). However, putting this issue aside, both the semantic-role hierarchy and the person hierarchy have their counterparts in generative mechanisms, and both are in fact used in Stegovec's account. Furthermore, in both Hasplemath's and Stegovec's account Case simply doesn't matter. Previous accounts of the PCC crucially relied on Case (but see Baker 2008). Neither Stegovec's nor Haspelmath's does. So, what we have here is a convergence on a particular issue from two very different perspectives. And the facts regarding reverse PCC quite conclusively show Case indeed doesn't matter. So, yes, we could dwell on the differences and emphasize them, but abstractly there are clear similarities between these accounts: when convergence comes from two rather different perspectives, that can be interpreted as a rather strong sign that those points of convergence are indeed on the right track. And in fact, as noted above, there may actually be room for both perspectives in a comprehensive account of the phenomenon, which could be an instantiation of a more general situation where functional factors get encoded syntactically; the relevant syntactic mechanisms then arise as a reflex of a functional factor, but their "syntactization" can take them quite beyond the original functional factor (see also section 5 for a more general discussion).

In this respect, functionalists often question whether a typological generalization requires a formal analysis of the data, but they also often use the term 'grammaticalization' in a way that in fact indicates such a need. More generally, what the functionalists often mean by grammaticalization (and what Haspelmath also means by it in his approach to person restrictions), is that X started out as a tendency due to functional/pragmatic reasons, but has "rigidified" into a hard rule, without really explaining the rigidification. This is a place where a formalist can step in to provide an actual account of the rigidification, including the details of exactly what X has hardened into, and why it has hardened into a particular format rather than a different logically possible format. This would lead to a productive and complementary investigation of a phenomenon from both perspectives that would in turn provide a more comprehensive understanding of the phenomenon. Paying closer attention to what is meant by something being "grammaticalized" can then not only be progress toward rapprochement between the two perspectives, but also make the complementarity of the two perspectives more obvious.

3. Efficient design and the nature of language

Above, I have discussed what the goals of a deduction of a typological generalization should be. There is a related point to be made. One of the tenets of the Minimalist Program is that language faculty is characterized by efficient design, which appears to be a contentious issue in functionalist circles. But there should be nothing controversial about this. Occam's razor as a research

---

of a broader cognitive tendency that English reflects (for a more general discussion along these lines, see Hawkins 2010).

methodology in fact leads to the efficient design hypothesis—it is a dictum that your subject of inquiry should be only as complex as it needs to be, hence the efficient design hypothesis. But what do we know about efficient design? Binary computer language systems and Markovian finite state devices are pretty efficient and simple, why not them? The reason why not them is due to an obvious question that arises here: efficient design for what? The *what* question has to be answered, and this is where the two candidates in question fail. So the crucial question is efficient design for what? Is it the nature of interaction with non-linguistic interfaces (namely, conceptual-intentional and articulatory-perceptual), as is often assumed by generativists? Or the function of language? The two actually may not be very different (though generativists generally talk only about the former). The point made here is that to look for efficient design one actually needs to pay attention to functional considerations, efficient design should include what language is used for. Efficient design concerns the nature of language—interaction with the interfaces, what language is used for—they are fundamental to the nature of language.[28] The efficient design hypothesis, which is often seen as a point of divergence between the generativist and the functionalist camp (though see Golumbia 2010), is then actually another potential point of convergence between the two.

Now, one of the tenets of minimalism is that as much as possible, and as directly as possible, should follow from bare output conditions (see Chomsky 1995 on this notion), i.e. the nature of language, which means that formal reasons that have nothing to do with the nature of language should be minimized. The latter often goes unacknowledged by generativists (though see Chomsky 2004). There is a related aspect of all this. As noted above, the nature of language inevitably includes functional considerations. While functional explanations for typological gaps might be more difficult to achieve than strictly formal explanations for reasons discussed above, a priori they should actually be preferred, on the grounds that they would be more likely to tie the explanations to something else, i.e. to explain the relevant properties of language by appealing to the nature of language. The above statement obviously overgeneralizes—each specific case of a typological gap and its explanation need to be looked at separately and evaluated with respect to the overall goal regarding what an explanation should achieve. But the point is that everything else being equal, functional explanations should be preferred on conceptual grounds, in fact based on one of the core minimalist tenets. Note also that what I have referred to above as the minimalist tenet, tie as much as possible, and as directly as possible, to the very nature of language, is essentially the Occam-razor strategy—simplify to what is unavoidable. So, Occam's razor, which is to say the nature of language, favors functional explanations. More generally, Occam's razor then disfavors, and requires re-evaluation of, strictly formal mechanisms, like e.g. the EPP, which in its generalized form from Chomsky (2000) is not simply a property of a single head (Infl) but drives all movement (see Bošković's 2021a discussion of the EPP from this perspective). There is

[28]It should be noted in this respect that Chomsky (2020) suggests that I-language (which is what the generativists mean by *language faculty/FL*) doesn't care about use (i.e. parsing) or communication, but it cares about expression of thought; the design of language should then capture the fundamental aspects of thinking. It is not, however, easy to make a clear distinction between communication and expression of thought (in fact, I am not doing it here); in a sense thinking involves communication with oneself; further, communication with oneself or with others, it involves expression of thought. On the other hand, use, i.e. parsing, is a different issue, I-language may indeed not care much about that—I refer the reader to Chomsky (2020) for relevant discussion, adding only one point: everything we know about syntax indicates structure building proceeds bottom-up (there have been occasional attempts at top-down syntax (Phillips 1996), but doing it that way raises numerous non-trivial issues). Parsing of course proceeds left-to-right. There is an obvious conflict here between the fundamental specification of FL and parsing, which indicates that FL indeed does not particularly care about the latter, as Chomsky suggests (Chomsky notes in this context issues that filler-gap dependencies raise for parsing).

actually an implicit acknowledgment of the point that strictly formal mechanisms should be disfavored in Chomsky's (1995) position on Agr, where he argues for elimination of Agr (i.e. agreement phrases) on the grounds that this is a strictly formal element which does not contribute to interpretation (Chomsky 2000 also attempts to minimize the role of Case, which for him is also a strictly formal property). Chomsky's (1995) stand on Agr can be interpreted as a minimize-purely-formal-considerations strategy. The other side of that coin, generally ignored by the generativists, should lead to adopting a maximize-functional-considerations-strategy (the strategy in this context really follows from Occam's razor).

Consider now a more concrete example of what it means to minimize purely formal considerations. Consider the operation of movement, which immediately raises a question: why do we have movement in the first place? Chomsky's (2000:120-121) position is that this has to do with "externally imposed legibility conditions", i.e. it is due to "conditions imposed by the external systems". What this means is that the reason for it is essentially functional, or more broadly non-syntactic: to be able to express notions that go beyond the basic argument structure (for which Merge alone suffices), i.e. more complex semantic notions involving issues like scope/scopal ambiguities, pragmatic notions like topic/focus interpretation, specificity... In other words, movement is there to express various semantic and pragmatic relations (in essence, everything that goes beyond the basic argument structure). A question that arises here is to which extent these notions have led to the development of formal requirements which then drive movement (see fn 25)? In other words, do the functional reasons in question directly motivate movement, where movement would be directly interpretation driven, or are there formal requirements that essentially serve as intermediaries, leaving syntax free of semantic considerations? To make the question a little more concrete: when α moves to SpecTopicP and receives topic interpretation, does α move there in order to receive such interpretation or is there a formal reason behind the movement, with α moving to satisfy this formal reason, as a result of which it is interpreted as a topic?[29]

While the question may seem innocent in this particular case (in that there doesn't seem to be much of a difference between the two positions), it isn't in other cases. Consider e.g. the different behavior of Bulgarian and Japanese in multiple questions, where in Bulgarian all wh-phrases move to the interrogative SpecCP while in Japanese they all remain in situ. There are various ways of implementing this formally. However, without appealing to formal properties that would cause this difference, i.e. on the direct syntax-interpretation mapping approach, we are led to the conclusion that Bulgarian and Japanese wh-phrases are interpreted differently, i.e. they differ semantically (the gist would be that Bulgarian wh-phrases must function as operators binding a variable, which is not the case in Japanese, where wh-phrases are interpretable in-situ). This means that on the direct syntax-interpretation mapping approach, Bulgarian and Japanese questions cannot have the same LF (interestingly from this perspective, in French, which has optional wh-movement, wh-movement and wh-in-situ constructions have different semantics, see Boeckx 1999, Zubizarreta 2003). Note that positing LF/covert wh-movement in wh-in-situ languages like Japanese (e.g. Huang 1982), as a result of which Japanese and Bulgarian would have the same LF, would not work—we would still need a formal difference between Japanese and Bulgarian that would be responsible for the overt/covert movement difference, which would go against the spirit of the direct syntax-semantics mapping approach.

The two cases noted above can thus be handled through either directly interpretation-driven

[29]There is a cartology/mapping debate regarding discourse notions like topic and focus within formalist approaches (see e.g. Rizzi 1997 and Neeleman & van de Koot 2008). I will not go into it here, apart from noting that the former involves more formal factors than the latter, which means that the minimize-formal-considerations strategy (on its own) would favor the latter, see Lacerda (2020).

movement[30] or formal requirements as intermediaries (which would essentially be an indirect syntax-semantics mapping approach), but with non-trivial implications under the latter regarding the semantics of questions/wh-phrases. Those non-trivial consequences, however, set the scene for a typological investigation; in particular, what under a purely formal approach of the kind argued for in e.g. Huang (1982) was considered to be a crosslinguistic distinction regarding whether wh-movement takes place overtly (Bulgarian) or covertly (Japanese) should now become a matter of typology of wh-phrases: there are languages where wh-phrases are only interpretable as operators/in moved position (Bulgarian), languages where they are interpretable as variables/in-situ (Japanese), and languages where they are interpretable as variables/in-situ only if another wh-phrase is interpreted in the moved position/as an operator (English). This essentially semantic typology in turn raises a question: to what extent is all this reflected in the morphological make up of wh-phrases and/or other functions that they may perform. Regarding the latter, it is well-known that in many languages wh-phrases can have a range of non-question functions (see especially Haspelmath 1997). There are several language types in this respect. One is represented by Serbo-Croatian, a multiple wh-fronting (MWF) language, which has a fully productive system where addition of an inseparable affix to a wh-phrase results in a series of meanings shown below. I will refer to this pattern as the *sub-wh system* (intended to indicate the morphological subset-superset relationship between the wh/question usage and other usages).

(21) ko 'who'; *i*ko 'anyone'; *ni*ko 'no one'; *ne*ko 'someone'; *sva*ko 'everyone'; bilo *ko* 'whoever'

Note that the sub-wh system is different from the situation found in Chinese, where the exact same form can have a series of usages (including wh), depending on the context it finds itself in, as well as the situation found in Japanese, where (except in the case of the counterpart of 'someone'), the element added to the stem can be detached from it (there is a special element present even on the wh-usage; in standard Japanese it is obligatorily detached, but in Okinawan Japanese it can be attached to the stem itself, which indicates we are not dealing here with a sub-wh system). Finally, English has a somewhat similar situation to SC, cf. *where, somewhere, everywhere, nowhere*. However, English is still quite different from SC; SC has a fully productive sub-wh system where the process in question is fully productive for all wh-phrases. This is not the case in English, as indicated by e.g. **somewho/*everywho/*nowho* or *somewhat/*everywhat/*nowhat.*

Consider now the list of MWF languages given in Bošković (2012): Polish, Czech, Russian, Slovenian, Ukrainian, Bulgarian, Macedonian, Romanian, Hungarian, Basque, Yiddish, and Mohawk. It turns out all these languages have a fully productive sub-wh system. Particularly interesting is Romance. Latin was a MWF language (Ledgeway 2012, Dadan 2019). Importantly, it had a fully productive sub-wh system, which got lost in all modern Romance languages except one: Romanian, the only modern Romance language that still has MWF.

(22)

Latin	**interrogative**	**existential, etc…**
person	quis	ali-quis
thing	quid	ali-quid
place	ubi	ali-cubi
time	quando	ali-quando

(23)

Italian		
person	chi	qualcuno
thing	che	qualche cosa/qualcosa
place	dove	in qualche luogo
time	quando	qualque volta

[30]Quantifier raising, which is assumed to take place because quantifiers cannot be interpreted in situ, might be the best candidate for such movement.

(24) **Romanian**

person	cine	cine-va
thing	ce	ce-va
place	unde	unde-va
time	cînd	cînd-va

We then have a correlation between MWF and the sub-wh system, the former requiring the latter.

(25) If a language has multiple wh-fronting, it has a sub wh-system.

Recall that on the direct syntax-semantics mapping, MWF indicates that wh-phrases in the language in question must function as operators binding a variable. What this then indicates is that this semantic property has a reflex in morphology: it requires a sub-wh system.

In summary, the minimalist tenet that as much as possible, and as directly as possible, should follow from bare output conditions, i.e. the nature of language, leads to preferring, in principle, functional over strictly formal explanations and more generally to a minimize-strictly-formal mechanisms strategy (where formal reasons that have nothing to do with the nature of language should be minimized). This in turn prefers (again in principle) direct syntax-semantics mapping, where movement is directly interpretation-driven, over indirect syntax-semantics mapping, where formal requirements would serve as intermediaries (and would drive movement). We have seen one area where the strategy has non-trivial consequences: it sets the scene for a semantic typology of wh-phrases (where the formal distinction overt vs covert wh-movement is recast as a difference in the semantics of wh-phrases), which in turn raises the question to which extent the relevant semantic differences are correlated with the morphological make up of wh-phrases and/or other (non-question) functions that the relevant elements may perform. We have seen that there may in fact be a correlation in the case of one particular type.

Before closing this section, I will briefly note a new argument for the direct syntax-semantics mapping approach (given its importance in the context of the current discussion). Consider the so-called Superiority effect, illustrated by (26a-b) ((27) gives their structure prior to wh-movement).

(26) a. Who did you tell that she should buy what?
b. ?*What did you tell who that she should buy?
(27) You tell who that she should buy what

In Minimalism, the effect from (26) follows from economy of derivation, which requires every condition to be satisfied through the shortest movement possible. Wh-movement has to take place in English. In (27), the requirement can be satisfied by moving either *who* or *what*. Since the former results in shorter movement (cf. the pre-movement structure in (27)), (26a) is preferred to (26b).

There is, however, evidence that Superiority is also semantically conditioned. Consider (28).

(28) a. I wonder who bought what?
b. *I wonder what who bought?
c. Who wonders who bought what?
d. Who wonders what who bought?

What cannot undergo wh-movement in (28a-b), a familiar superiority effect. However, (28d) is grammatical in spite of *what* undergoing wh-movement. Importantly, (28d) is not ambiguous: it is acceptable only if the second *who* takes matrix scope (i.e. as a matrix multiple question on the two *who*-s); it is unacceptable on the multiple indirect question reading, where *what* and second *who* are interpreted in the embedded SpecCP. Apparently, superiority effects are conditioned on the relevant wh-phrases being interpreted in the same interrogative SpecCP (*what* and second *who* are not interpreted in the same interrogative SpecCP on the allowed reading of (28d)). Now,

superiority effects are a result of a formal derivational requirement of the computational system. A question then arises how to bring in semantic considerations into a requirement of the computational system itself. At the point of movement of *what* to the embedded SpecCP, which occurs before the matrix clause is built in (28d) (given that structure building proceeds bottom-up, the embedded clause is built before the matrix clause), (28b) and (28d) are in fact identical syntactically. Movement to the embedded SpecCP, then, needs to be made sensitive to the final semantics of these examples. In the indirect syntax-semantics mapping approach, where syntax is free of semantic considerations, it is very difficult to achieve this. On the basis of such examples, Golan (1993) suggests only constructions with the same interpretation can be compared regarding economy of derivation. On the direct syntax-semantics mapping approach, where movement is directly driven by the relevant semantic considerations, those semantic considerations are directly involved in the relevant movements, hence the movements can be made sensitive to them. The effect in question may then provide an argument for the direct syntax-semantics mapping approach.

3.1. Language acquisition

As discussed above, Occam's razor leads to the efficient-design hypothesis. There is an important point in this respect that is often overlooked nowadays in generative circles: efficient design should make language easily learnable. Generative syntax in fact essentially started with language acquisition, i.e. what is often referred to as the logical problem of language acquisition: how come children can learn something as complex as language so easily. In principle, then, explicit concern for language acquisition should be a significant difference between GT and TT. Recent research in generative syntax has, however, simplified the syntactic design of language a great deal (at the risk of some exaggeration, a simple merge operation and a locality domain). From the perspective of a generativist, if we simplify the design too much, and say that only that is innate, a question arises how can the children learn the rest. The Occam-razor driven research in syntax has essentially raised an Occam-razor problem in another area, namely regarding language acquisition (which may have become a bit of a back-burner for a generative syntactician. In this respect, GT should lead to re-raising the issue of language acquisition—after all, the child needs to learn all the differences typologists are concerned with.) From this perspective, the above discussion regarding differences between languages with and without articles gains wider significance. There have been proposals regarding so-called octopus parameters, where one parameter has consequences for a number of phenomena. The DP/NP parameter (the word parameter is used for ease of exposition) is by far the biggest octopus ever, it is a veritable monster (more than 20 morphological, syntactic, and semantic properties crucially depend on it, see fn 9). What is important in the context of the current discussion is that it greatly eases the burden of language acquisition. Given all the generalizations where the NP/DP difference is involved, all the phenomena where the NP/DP difference reaches into, the difference gives the child so much, more than anything else. But how can the child learn the NP/DP difference? The NP/DP generalizations all involve potential triggers but most of them (maybe even all of them) are not plausible candidates. They either involve rather sophisticated phenomena, like the interpretation of superlatives, or phenomena that are simply not widely available crosslinguistically, like multiple wh-fronting. How about the definite article itself? Bošković (2010, 2016c) suggests the definite article is indeed the trigger for learning the whole monster.[31] In this respect it should be noted that *the* is in fact the most frequent word in English. *The* then seems to be a perfect trigger for learning the monster. Importantly, Koulidobrova (in press) shows that definite article is actually the trigger for proper emergence of D-items in

[31]This implies that there cannot be a null definite article language, a language with null *the* would essentially be unlearnable. (In fact, Bošković's NP/DP generalizations presuppose that there is no null *the*.) Strong evidence that there is indeed no null *the* in languages without articles is provided by Despić (2019) and Jenks (2018), see fn 12.

language acquisition. Early on, children make a lot of mistakes in the usage of D-related items in English, which also includes omitting *the*. Once *the* is properly learned, mistakes regarding other D-related items stop, which indicates that *the* is indeed the trigger for proper acquisition of D-related items, and more generally DP. Koulidobrova in fact interprets her data as supporting the DP/NP analysis.[32] Everything thus starts falling into place once language acquisition is taken into consideration. Most importantly, we are addressing here the issue of efficient design when it comes language acquisition: learning *the* helps the child with so many things.

In light of this, consider the semantic role of the definite article. In formal semantics terms, iota-operator, which is the primary semantic job of the definite article, turns NP, which is of type <e,t>, into type e; what this essentially means is that it turns a predicate into an argument. This way, it makes it possible to integrate the NP into the clause/VP. On a more informal level, taking verbs and nouns to be the primary categories, we want to be able to integrate the two, which will give us a clause. This is in fact the job of an iota-operator, definite article being an instantiation of an iota-operator.[33] This certainly looks like a pretty important and extremely basic job; see in fact Progovac 2010 for an idea that this was all there was in one stage of evolution of syntax; according to Progovac, at one point there was only a single argument-predicate combination. Definiteness is then the easiest way of making a distinction between predicates and arguments, and of integrating the NP into the VP/clause. From this perspective (which is essentially functional), it does not seem surprising that definiteness is so important, that its tentacles reach into so many areas—it can in fact be considered part of efficient design.

3.2. On the methodology of investigating crosslinguistic variation

Returning to efficient design, when studying it one should use whatever one can, syntactic phenomena, semantic phenomena, PF aspects, language acquisition, language change, different languages…the use-whatever-you-can methodology is in fact the standard stand in the generative approach, often reiterated by Chomsky, and a normal approach in natural sciences. There is a related point: it is often assumed in functionalist circles that the generative approach is so syntax-centric that just showing that some phenomenon is not syntactic, or has non-syntactic aspects, argues against the generative approach in general (see sec. 1.2). This relies on a view of the generative approach which may have been true long time ago, but certainly not any more.[34] The generative approach has reached the point where linguistic phenomena are looked at in their totality, not compartmentalized by subfields. A generative syntactician must constantly pay attention to semantics, morphology, phonology, language change, language acquisition, and look seriously into understudied languages and typology. Typology is particularly important here. We have reached the point where typology should be at the center of investigating what generativists refer to as UG. At this point of our understanding, broad crosslinguistic comparisons, Greenberg-style typological generalizations, are more enlightening regarding the nature of language and UG than detailed investigations of particular languages (the latter are a prerequisite for the former though the practitioners of the latter are often a reviewing stop sign for generative typology works for reasons discussed below).

[32]There have in fact been proposals that children go through the NP stage even in English, which would then be a default (Guasti et al 2008; Mathewson et al 2001). See also Petroj (2020) for an interesting parallelism between language acquisition and code switching (a process which involves alternation between different languages within a single utterance): she provides evidence for the NP/DP hypothesis based on code switching between a DP language (Romanian) and an NP language (SC).

[33]In a language without definite articles this would then be a purely semantic operation.

[34]Independently of this, even if focused on generative syntax the attitude in question would be rather strange: generative syntax cannot account for palatalization in Slavic or bring peace to the Middle East, but that does not show that it is wrong.

Now, what has dominated the generative camp regarding crosslinguistic studies is the so-called microparametric approach, which compares languages that are otherwise very similar. (Kayne's work is one of its cornerstones; the typological approach (the term is used neutrally here, not to be equated with TT) is sometimes referred to as macro(parametric) approach, and contrasted with the micro(parametric) approach. For consistency with the literature I will use that term in discussion of the opposition.) The micro approach seems to be considered 'easier', but it really isn't. Consider languages A and B within the micro/macro approach distinction. In the micro approach, A and B, say Florentino and Trentino, varieties of Italian, share a lot of things; as a result, it may be easier to spot and focus on the differences. Within the macro approach, languages A and B, say Russian and Chinese, differ in many respects; it may then be easier to spot and focus on the similarities. There is really no sense in which one approach is 'easier' than the other; but they do differ methodologically, with respect to what we focus on: differences or similarities, something that has often escaped attention. In many respects, the macro approach can be more fruitful, and can help us locate the points of crosslinguistic variation that would otherwise be difficult to notice. The reason for this is that those similarities from the micro approach, which are often taken to represent invariant FL properties and considered principles of UG, can turn out to be points of variation from a broader point of view. There is, however, one situation where the macroparameteric approach, taken as a search for similarities among very different/unrelated languages, is particularly useful which is generally neglected: when very different languages are compared it is difficult to pinpoint the reason for the differences unless we also find some similarities. Consider, e.g. the comparison of Mohawk and English in Baker (1996).[35] English allows extraction of nominal complements:

(29) Who did John see friends of?

Mohawk disallows such extraction. Baker points out that Mohawk is a polysynthetic language, while English is not, and provides an account of this contrast that ties it to this difference. The issue here is that there are a lot of differences between Mohawk and English, it is difficult to put the finger on the ultimate factor behind the contrast in question. A macroparametric/typological approach (taken as a search for similarities among very different/unrelated languages) can help us here. It turns out that Serbo-Croatian behaves just like Mohawk regarding (29). In fact, SC, which is in the respect Baker discusses different from Mohawk but like English (it's not a polysynthetic language, like English and unlike Mohawk), behaves remarkably like Mohawk and unlike English regarding extraction out of nominals. While SC disallows extraction of nominal complements, like Mohawk and unlike English, it allows left-branch extraction of possessors (i.e. extraction of the possessor in cases like *Mary's house*), which is actually also allowed in Mohawk but not English. If we are going to have a uniform account of these similarities/differences, what's responsible cannot be polysynthesis. There is, however, a similarity between SC and Mohawk which, as noted above, Bošković (2008, 2012) argues is a point of a broad typological difference that affects numerous syntactic, semantic, and morphological phenomena: SC and Mohawk lack definite articles, while English has them. In fact, Bošković (2012) provides an account of the relevant extraction differences between English and SC which would extend to Mohawk, since what crucially matters there is the presence vs lack of articles. In fact, Mohawk turns out to fit a number of Bošković's article generalizations: one of them, regarding left-branch extraction, was noted above; it also fits Bošković's superiority generalization (multiple wh-fronting languages without definite articles do not have fixed order of fronted wh-phrases), head-internal relatives island generalization (head-internal relatives display island-sensitivity in languages without definite articles, but not in languages with articles), and the number morphology generalization (number

[35]The comparison scale of Baker (1996) is much broader than what is illustrated by the discussion below, whose point is strictly methodological.

morphology may not be obligatory only in languages without articles).[36] The point here is that it is difficult to pinpoint the responsible factor when comparing languages which are very different, with no point of similarity that can be used to illuminate the issue, as was the case here with English and Mohawk. But once SC entered the picture, we got that point of similarity (between SC and Mohawk), which was in fact provided by a typological comparison. If the above discussion is on the right track, the reason for the English-Mohawk differences in question is not polysynthesis, but the presence vs lack of definite articles. In this respect it should be noted that Baker (1996) actually observes that polysynthetic languages in general lack articles (this is again a one-way correlation, it does not mean languages without articles will all be polysynthetic). The polysynthesis parameter is then also tied to articles. (Baker in fact also suggests that the lack of articles implies the lack of DP.) But it was difficult to see what is responsible for the English-Mohawk differences in question, polysynthesis or definite articles, until languages without articles that are not polysynthetic were brought into the picture. I emphasize that the relevant crosslinguistic differences are not micro but macro-parametric (i.e. typological) in nature: they illustrate the benefit of looking for similarities between (unrelated) languages that are other- wise very different. (In fact, it's worth noting that Bošković's NP/DP project started with a combination of the micro and micro (i.e. typological) approach: the first point of comparison were differences among Slavic languages (only Bulgarian and Macedonian have definite articles within Slavic) and in the history of Romance (Latin, which didn't have articles, vs modern Romance), and similarities between Slavic languages without articles and East Asian languages without articles).

At any rate, the gist of the difference between the microparametric approach, which dominates the generative program, and the macroparametric/typological approach (which has only recently started to gain traction within the generative paradigm hence has often been taken to characterize the functional approach in broad comparisons of generative and functional approaches), can be summarized as follows: (a) the former: X and Y are very similar, let's see where they differ (this is what is informative); (b) the latter: X and Y are very different, let's see where they are similar (this is what is informative).

There is obviously a place for both strategies. While this used to be a point of divergence in crosslinguistic studies between generativists and functionalists (with the former not taken to be particularly interested in employing both strategies), it is increasingly (with the emergence of generative typology) becoming another point of convergence.

There may be another reason for the (initial) research strategy difference (where the difference concerns what is emphasized) regarding the kind of languages that are looked at (similar or unrelated) between the two camps which has to do with the questions the two camps are primarily asking. The generativists generally focus on similarities across languages, asking the question why is it that there are so many similarities (because their goal is to investigate UG), while the functionalists often focus on differences, asking why there is so much crosslinguistic variation. This could actually be the reason for the difference in the methodology of crosslinguistic research, where the generativists emphasize looking at similar, related languages (which have a lot in common) while TTs emphasize comparisons of unrelated languages. But this is only an issue of focus, these two lines of inquiry are not incompatible. Both of these are right questions to ask (why there are so many similarities across languages and why there are so many differences across languages)—there are both a lot of similarities and a lot of differences across languages. Both need to be addressed.

In the Principles & Parameters approach, the differences fall within the domain of parameters (see Roberts 2019 for a recent discussion of parameters). The issue of focus is, however, clear here: there is much more research on the principles of UG (i.e. invariant properties) than on the

[36]See Bošković (2012) for what is meant by the phenomena referred to in these generalizations. Most of them are one-way correlations; SC is not relevant to the last two but e.g. Japanese, another language without articles that is not a polysynthetic language, patterns with Mohawk in these respects.

parameters (i.e. crosslinguistic variation); the distinction may not have been that sharp initially, but has been increasing over time, as with the Minimalist Program. What may have contributed to this was the hypothesis that crosslinguistic variation should be confined to lexical properties, not the computational system (see sec. 1.1). The computational system itself has always been more of a focus of research than lexical properties, being regarded as an inherently more interesting subject of inquiry. Once the principles vs parameters distinction got translated into a computational system vs lexical properties distinction, the decline of research interest into the latter was inevitable. The emergence of generative typology, which by its nature looks at crosslinguistic variation, should, however, revive research interest into the latter (all this also ties with language acquisition, where the researchers have always been concerned with acquisition of parametric variation).

4. Picky middle way

The most convenient way of doing typological work (TT or GT) is to base it on descriptive works dealing with particular languages. An issue that generative typology faces is that generative works typically examine issues that are not addressed in such descriptive grammars (like islands, see e.g. (i) in fn 3), which furthermore also focus on what is possible in languages. To deal with this practical problem, Baker & McCloskey (2007:294) suggest what they call "Middle Way", "a style of research that would look at fewer languages than a typical typological study does, but at more languages than a typical generative study does". What makes this particularly feasible is that generative research is generally theoretically guided (which in fact differentiates it from traditional typology). To illustrate, one of Bošković's article generalizations concerns negative raising, where negative raising (in particular, licensing of strict clause-mate negative polarity items (SCNPIs) under negative raising) is possible only in languages with articles.[37] This is a rather surprising interaction, the kind of interaction that cannot be stumbled upon by accident. Indeed, there is a theoretical proposal that ties definiteness and negative raising. Thus, Gajewski (2005) treats negative raising predicates as plural definite descriptions—they essentially combine a modal base (set of accessible worlds) with a definite article (see Gajewski 2005, Bošković & Gajewski 2011 for details, which need not concern us here). Without a theoretical proposal that ties definiteness and negative raising, the potential correlation between articles and negative raising (which from the point of view of descriptive grammars is rather surprising) would not have been checked.[38]

In light of GT being theory-driven, using theory as its guide (which differentiates it from TT), I would like to slightly amend the research strategy suggested by B&M, renaming it *Picky Middle Way*. Picky Middle Way is guided by investigation of a particular theoretical mechanism, and picks examples from different, unrelated languages to illuminate that theoretical mechanism, without need to offer a detailed investigation of each language considered that would go beyond the mechanism in question; even when particular constructions are used, they are discussed only to the extent that they bear on the theoretical mechanism in question. Investigation of extraction out of moved elements in Bošković (2018) and out of conjuncts in Bošković (2020), whose goal was to show that, contrary to what is standardly assumed, such extractions are possible as well as to

[37]SCNPIs require clause-mate negation (cf. *he didn't leave until yesterday* and **he left until yesterday*, the underlined element being a SCNPI), except when embedded under a negative raising predicate (compare *he didn't believe that Ann would leave until tomorrow* and **he didn't claim that Ann would leave until tomorrow*, *believe*, but not *claim*, being a negative raising predicate). SCNPI licensing under negative raising is disallowed in e.g. SC, Czech, Slovenian, Polish, Russian, Ukrainian, Turkish, Korean, Japanese, and Chinese, which lack definite articles, but allowed in e.g. English, German, Spanish, French, Portuguese, Romanian, and Bulgarian, which have articles.

[38]Needless to say, such theory-driven investigation of potential typological generalizations does not always work out, but when it does it also provides evidence for the relevant theoretical proposals (in this case it provides evidence for a semantic approach to negative raising (a contentious issue, see e.g. Collins and Postal 2014 for a syntactic account), in fact a particular semantic account.

determine the exact contexts under which they are possible, can be taken as examples of Picky Middle Way since these works pick particular (crucially different) constructions from a variety of languages (Serbo-Croatian, Dutch, Janitzio P'urhepecha…) to examine particular topics which in turn illuminate the relevant theoretical issues, islandhood in this particular case.

A practical issue arises with the reviewing of Picky Middle Way works though. Many generativists spend their lifetime research doing detailed investigations of constructions of a single language. When faced with a "Picky Middle Way" work, where certain aspects of construction X from language Y are used to illuminate theoretical mechanism Z, they often require detailed discussion of all aspects of construction X (since this is how they write their papers), although those aspects have nothing to do with Z. This is a serious practical problem that the development of generative typology is facing. Investigations of UG in the generative framework used to rely on detailed investigations of particular languages. In the old-habits-are-hard-to-change spirit, those who are used to such investigations are often the reviewing stop sign for typological works since they often require the same kind of detailed investigation for every individual language considered in a typological work that would be found in a work devoted to just one language.[39]

5. Final differences and remarks

In this section I will discuss three remaining issues, which concern what many consider to be the biggest differences between the two camps, the issues in question being left for the end of the paper in the hope that the preceding discussion has taken the edge of some of the perceived irreconcilability and antagonism between the two camps. Another reason why these issues (some of which were touched upon above) are discussed together here is that they are often lumped together under "faculty of language" (with the attitude you believe in all of it, i.e. the whole package, or not)—it's, however, important to separate them to see where the real disagreement lies. In that vein, I will "decompose" the generativist's faculty of language into three distinct notions: universal grammar, innatism, and domain specificity. I will first discuss the notion of universal grammar, which should be the least controversial (the issue was also discussed in section 1.3).

In spite of the pronouncements that are associated with the respective camps, it is not clear that there is a real fundamental disagreement when it comes to "believing in universal grammar". UG is a rather loaded term, let us put it aside for a moment because of that, and use a less loaded term "general properties of language". Traditional functional typologists are concerned with uncovering general properties of language—this is in fact what Greenberg's generalizations are all about. The same holds for generativists—they just call those general properties of language UG. But what is important is that both camps seem to believe that there are general properties of language. There are differences regarding how those general properties of language look like, but that is not the point here—the point is that both camps do hold that there are such things (otherwise they would not be trying to uncover them). We can call them general properties of language or UG, or α for that matter, the terminology itself does not affect the point.

Innatism is another widely held difference between the two camps, but it is not clear that that difference is real either. To see this, let us move to another difference between the two camps, which does seem more significant, namely what is behind those general properties of language: is it (whatever it is) domain specific (the generativists often use the term faculty of language—FL)

[39]As an example, I used the Japanese numeral floating construction in a typological investigation of extraction from conjuncts. A reviewer, clearly a Japanese syntax specialist used to writing papers involving detailed studies of particular constructions of Japanese, kept on insisting the paper should have a detailed discussion of the construction in question that would discuss in detail the issue of whether floating numerals should be analyzed the way Sportiche (1988) treats similar elements in English or as adjuncts although this was totally irrelevant to the issue under discussion—nothing would have changed regarding that issue regardless of which approach to the Japanese construction in question was taken. But that didn't matter to the reviewer. The attitude was you cannot use construction X unless you discuss all aspects of construction X, period.

or are we dealing with broader cognitive mechanisms, which are not domain specific. In principle, innatism can be attached to either position (many of those broader cognitive mechanisms that the functionalists appeal to are obviously innate), so at least in principle, innatism is not necessarily a fundamental difference between the two camps (again, the point is not how those things that would be held to be innate would look like, but that in principle at least there can be such things).

So it seems that the major, fundamental difference boils down to domain specificity vs broader cognitive mechanisms. But even there the difference is not as black-and-white as it is generally held to be. As discussed above, the minimalist tenet that as much as possible should follow from the nature of language, which includes functional considerations and paying serious attention to the systems language is interfacing with (the articulatory-perceptual and conceptual-intentional, which themselves are not part of FL), leads to a maximize functional/bare-output conditions considerations and minimize formal considerations strategy (formal reasons that have nothing to do with the nature of language should be minimized), which in the bigger picture means that broader cognitive mechanisms (henceforth BCM) need to be paid attention to—generativists should not wave an FL flag as an excuse to ignore them. But the interaction between the relevant mechanisms is such that it actually blurs the FL/BCM line, to the extent that even in this case we cannot be talking about a clear line-in-the-sand demarcation between the two camps.[40]

To illustrate, consider Kayne's (1994) Linear Correspondence Axiom (LCA), a proposal where a structural relation, asymmetric c-command (where X c-commands Y but Y does not c-command X) determines word order (i.e if X asymmetrically c-commands Y, X precedes Y and everything X dominates precedes everything Y dominates). Follow up works like Kayne (2010), Di Sciullo (2015), Bošković (2021a) have considerably expanded the domain of the LCA. In fact, Bošković argues FL quite generally favors asymmetric relations, positing a general Maximize Asymmetric Relations (MAR) preference, which he shows unifies a number of phenomena,[41] namely the diachronic loss of specifiers,[42] their avoidance in language acquisition, the LCA, the Phase Impene- trability Condition (which, among other things, is crucially involved in determining locality of movement), the no-Spec-without-complement aspect of bare phrase structure (where a head cannot take a specifier unless it also takes a complement), the crosslinguistic rarity of multiple wh-fronting (which may be expected to be quite common since it has been argued to have the most transparent and simplest syntax-semantics mapping), and the mysterious *who left* effect, where subject wh-movement cannot proceed through SpecIP, i.e. the subject position, although this position otherwise must be filled in English (see Bošković 2016a, Messick 2020 for relevant arguments). Bošković (2021a) also shows that MAR has broad theoretical consequences, e.g. regarding structure building, Case-licensing and the EPP. While Kayne's original LCA looked FL-specific, the broadening raised the question of whether MAR is an irreducible formal property of

[40]More generally, what we may be dealing with here, and which would blur the line in question, is a situation involving evolutionary developments in the organization of the human brain which would represent adaptations specifically for language, where pre-existing more general cognitive structures/mechanisms got adapted (or exapted) to be used specifically for language (for a recent general discussion, see Hauser, Chomsky, and Fitch 2002).

[41]For reasons of space, the following discussion is rather packed; at any rate its goal is merely to give an impression how wide the scope of the broadening of Kayne's original LCA is (see the cited work for details).

[42]To illustrate with questions/wh-movement, Dadan (2019) shows that the general direction of diachronic change is from wh-movement to wh-in-situ (we don't find changes from wh-in-situ to wh-movement); e.g. Old Japanese, Archaic Chinese, and Sanskrit were wh-movement languages, while Japanese, Chinese, and modern Indic languages are wh-in-situ languages. Bošković (2021a) shows that the MAR-based account of this uni-directionality of diachronic change also explains why certain cases of wh-fronting are more resistant to loss than others (this has to do wh-fronting not targeting the same position in all languages).

FL or it can be traced back to FL-external factors, even related to broader cognitive mechanisms.[43] The preference nature of MAR may in fact suggest the latter, if, as often assumed (e.g. Baker & McCloskey 2007), tendencies and preferences, as opposed to absolute properties, are more likely to have functional/broader cognitive than formal/FL-specific origins. Furthermore, Kayne (2010) observes that both parsing and language production are asymmetric, in that they show a beginning vs end asymmetry. If there is indeed a connection here, it would suggest that we may be dealing with a broader property of language, which goes beyond linguistic competence, i.e. knowledge of language, that the generativists are generally concerned with—the property would extend to performance, i.e. usage of that knowledge.

Now, Minimalist Program explores the possibility that UG, taken as a property of FL, is an optimal way of satisfying requirements that are imposed on FL by the external systems FL interfaces with. From this perspective, asymmetricity can be encoded in UG even if its ultimate source is those external systems. On this view, asymmetricity would essentially be imposed on UG as the optimal way of satisfying external system requirements. What is important for the general point made here is that this kind of situation blurs the line between FL (i.e. language-specific) and BCM (i.e. broader cognitive mechanism), which is often taken to be a clear line in the sand (in fact a differentiating line between the generative (more precisely, Chomskian) and the traditional functional camp): something can in a way be both (if it is language-specific as a reflection of broader cognitive mechanisms). In fact, this is a pretty common situation.[44] Consider the very basic question, why do we have movement?[45] It is there essentially for functional reasons—due to broader cognitive needs to express things (as noted above, movement enables us to express more complex semantic and pragmatic relations that go beyond basic argument structure; as often pointed out, movement is language-specific—nothing similar to movement is found in other symbolic systems, where the relevant cognitive needs are not present). They got encoded in UG due to the mechanism of movement, with a number of constraints on movement that do appear to be language specific. But the existence of movement is a reflection of broader cognitive mechanisms/needs (see also the discussion in the end of sec. 2.2).

As another illustration, suppose there is something like (30), which is a simplified version of Rizzi's (1997) split CP, where ForceP indicates the force of a sentence, TopP houses topicalized and FocP focalized elements. Bianchi & Frascarelli (2010) further argue for splitting TopP into projections for three different types of topics, aboutness, contrastive, and given topics, with the hierarchy in (31) ((32) shows aboutness topics precede contrastive topics)).

(30) [$_{\text{ForceP}}$ [$_{\text{TopP}}$ [$_{\text{FocP}}$ [$_{\text{IP}}$]]]]
(31) [$_{\text{TopATP}}$ [$_{\text{TopCTP}}$ [$_{\text{TopGTP}}$….
(32) a. (As for) Rosa$_{\text{AT}}$, my next book$_{\text{CT}}$ I will dedicate to her.
b. *My next book, Rosa, I will dedicate to her.

A question that arises here is what is responsible for these hierarchies? If the answer is FL, a further question arises: why did FL pick these particular hierarchies? A functionalist would complain that saying that the fixed order of topics is due to a structural hierarchy does not explain anything, why do we have those structural hierarchies? That is a very legitimate complaint. We should try to understand this better, not simply stipulate a structural hierarchy. There are in fact proposals even within generative approaches that the ultimate reasons are semantic/pragmatic, see e.g. Neeleman

[43]If we are indeed dealing with a broader property here, Bošković (2021a) can be taken as an illustration of how looking seriously into formal properties of language can help us elucidate those broader mechanisms.
[44]For a smaller-scale situation along these lines, see the discussion of person restrictions in section 2.2.
[45]Or whatever is used to encode the fact that elements often don't occur in positions where they are interpreted, as in *Him, Sue likes*, where *him* is interpreted as the object of *like* but does not occur in that position.

& de Koot (2008) and Lacerda (2020).[46] If those structural hierarchies are there for reasons that are ultimately semantic/pragmatic, what do we mean by semantic/pragmatic? The natural answer is that what is meant by that is what Chomsky (1995) refers to as the conceptual-intentional system. But the conceptual-intentional system is FL-external, which means that we would then be dealing here with what I have referred to above as broader cognitive mechanisms. This is, then, another illustration of the give-and-take between FL and BCM which blurs the line between the two.

More generally, while it is not hard to draw a line between what the generativists refer to as FL and language-external mechanisms on the sound side (i.e. form side), it is much harder to do that on the meaning side. While not often explicitly articulated, it is generally implied that most of what formal semantics does is outside of overt and covert syntax, which means that it would belong to the conceptual-intentional system, i.e. it is outside of FL. Chomsky (2020), however, suggests that it is actually part of FL (for relevant discussion, see also Hauser, Chomsky, and Fitch 2002). The issue is in fact far from trivial,[47] but almost impossible to address in our current state of understanding—it has to do with what should be the demarcation line between language and our thoughts. This underscores how difficult it is to draw a clear line between the boundary of FL and the external systems it interfaces with (i.e. broader cognitive mechanisms) on the meaning side.

At any rate, the more general point made here is that generative and traditional functional approaches may not be as fundamentally opposed to each other as is generally widely believed or as the slogans standardly associated with these approaches suggest. What are standardly assumed to be clear lines in the sand demarcating these two approaches in many cases may not be there, or they are rather blurred. A wider realization that this is indeed the case should provide an opening for a rapprochement between the two camps. The first step in that direction should be talking to each other, something that is now sorely lacking. The goal of this paper is to make a small contribution to this effect.

6. Conclusion

The overarching point of this paper is that there should be less disagreement between the generative and the functional camp, as well as generative typology and traditional typology (which is often functionally based), than there actually is. Much of the disagreement comes from the two camps not interacting with each other, and misunderstandings of the slogans which are used to characterize the two camps, where the practitioners of the two camps are not really trying to see what is behind those slogans (the former is partly responsible for the latter); what also does not help in this respect is that with both camps, there are differences between what is said and what is really believed (as reflected in the actual research practice). There is a warzone out there, which is really unwarranted—there is actually little true opposition, what the two camps do is largely complementary, which is not easy to see because of the slogans that are constantly repeated but which should not be taken at face value. It's almost like a situation where a war lasts for so long that people forget what it is all about or don't see that the reasons for the war no longer hold. In this particular case, there is no real scientific reason for the war (that of course does not include human nature and sociological factors[48]). Additionally, with the minimalist tenet that as much as possible should follow from the nature of language, which by necessity includes functional considerations (and in fact naturally leads to the maximize-functional-explanations strategy), and

[46]These works adopt a different, mapping approach to discourse relations, but the details are not important to the general point made here, which is why e.g. topics precede focus, or why different types of topics are subject to a hierarchy (Neeleman & de Koot discuss the former, while Lacerda also discusses the latter).

[47]To put it more bluntly, assuming there is FL, and that this is what linguists investigate, the issue is whether a formal semanticist like Partee is a linguist or a psychologist (it would be the former if formal semantics is part of FL, and the latter if it investigates the conceptual-intentional system).

[48]Longing for good old times with fixed demarcation lines between good and bad guys (with clear good and bad guys) may also take time to overcome.

especially the emergence of generative typology, which also attempts to find Greenberg-style typological generalizations (and which is increasingly seen as essential in generative circles), we are actually reaching an exciting point of potential convergence, where the two camps should be talking more to each other, and interacting productively with each other. There will still be competition between the two approaches regarding analyses of particular phenomena: linguistic phenomena do not come with labels regarding whether a formal or a functional explanation is more appropriate for them.This is healthy competition of the kind we are all used to.[49] Importantly, there should still be competition but no incompatibility, i.e. antagonism that comes from opposing worldviews regarding the subject matter that would deny to the opposing camp the scientific right to exist.[50] In other words, the two camps should be increasingly seen as complementary, much like syntax and semantics or syntax and morpho-phonology are seen as complementary. Typology is really crucial here; it is in fact what is getting the formalists to engage with the functionalists—if you are a generative typologist you have to read functionalist works, you cannot ignore them. The problem before, which is to a great extent responsible for antagonism between the two camps, was precisely that they were ignoring each other. The emergence of generative typology is getting formalists not to ignore functionalists, to read the relevant functional literature; hopefully, the more traction and more importance typology gets in formal circles, which will inevitably lead to more typological work being done within the formalist camp, traditional functional-based typologists will find themselves in the same position. The surest way to antagonism is ignorance, i.e. ignoring, not interacting, not reading each other's works (or reading the works from the other camp only to talk to your own camp about them in order to make fun of the other "side" based on a cartoonish version of the other side). With the emergence of generative typology, which is rapidly gaining traction within the generative camp, typology in general will increasingly make this kind of ignoring more difficult, leading to hopefully more productive interaction between the two camps.

References

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral dissertation, University of Connecticut.

Albizu, Pablo. 1997. The syntax of person agreement. Doctoral dissertation, University of Southern California.

[49]E.g., it is often not clear whether a particular phenomenon should receive a syntactic or a semantic treatment within the same overall approach, a particularly common situation nowadays in the generative circles, where semantics has been increasingly gaining in importance (with a training in semantics essentially becoming a must for a generative syntactician), which has turned out to be very productive. Typical theoretical syntax oriented department will house a semanticist, due to the correct belief that theoretical syntax and semantics are not fundamentally incompatible and that a syntactician should be familiar with semantics (at least to be able to recognize a semantic issue when it arises). Such a department does not house a functionalist (the same holds in the other direction), but there is no real (non-sociological) reason why it shouldn't—the situation here is, or should be, the same as with theoretical syntax and semantics. We should be moving from the current, "our" vs "their" department situation to a predominantly "our" vs predominantly "their" department situation. The field (and the training the students get) will be better for it. The students need to be exposed to both views, at least to be familiar with them (so that they can recognize when a formal/functional issue arises, i.e. when a particular investigation starts crossing into the other domain), just like a syntactician needs to be exposed to semantics, and a semanticist needs to be exposed to syntax. In all these cases we are dealing with complementarity, not mutually incompatible worldviews. The latter would rightly require complete department separations (essentially to avoid a warzone). We are still in the separation mode, but we shouldn't be.

[50]Upon some reflection, no sane generativist would claim that there is no aspect of language that can be explained by looking at the function of language. The anatagonism between the two camps, however, can be so strong to override the "upon-some-reflection" requirement.

Alexander, Mathew. 1990. Agreement configurations: Grammatical relations in modular grammar. Doctoral dissertation, MIT.
Anagnostopoulou, Elena, Artemis Alexiadou and Martin Everaert. 2004. Introduction. In *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*, ed. by Elena Anagnostopoulou, Artemis Alexiadou and Martin Everaert, 1-21. Oxford University Press.
Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: Evidence from clitics*. Berlin: Gruyter.
Baker, Mark, and Jim McCloskey. 2007. On the relationship of typology to theoretical syntax. *Linguistic Typology* 11: 285-296.
Baker, Mark. 1996. *The polysynthesis parameter*. New York: Oxford University Press.
Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge University Press.
Baker, Mark. 2011. The interplay between Universal Grammar, universals, and lineage specificity. *Linguistic Typology* 15:473-482.
Baker, Mark. 2015. Formal generative typology. In *The Oxford handbook of linguistic analysis 2nd edition*, ed. by Bernd Heine and Heiko Narrog, 925-951. Oxford: Oxford University Press.
Béjar, Susana and Milan Řezáč. 2009. Cyclic Agree. *Linguistic Inquiry* 40:35–73.
Bianchi, Valentina and Mara Frascarelli. 2010. Is Topic a root phenomenon*? An International Journal of Theoretical Linguistics* 2:43-88.
Bickel, Balthasar. 2007. Typology in the 21st century: Major current developments. *Linguistic Typology* 11:239–251.
Boas, Franz. 1911. Introduction. In *Handbook of American Indian languages*. 1–83. Washington DC: Bureau of American Ethnology.
Bobaljik, Jonathan and Susi Wurmbrand. 2001. Seven prefix-suffix asymmetries in Itelmen. *Proceedings of CLS* 37:205-219.
Boeckx, Cédric. 1999. Properties of interrogatives in French. Ms., University of Connecticut.
Boeckx, Cédric. 2008. *Bare Syntax.* Oxford: Oxford University Press.
Bond, Oliver. 2019. Canonical Typology. In *The Oxford Handbook of Morphological Theory*, ed. by Jenny Audring and Francesca Masini, 409-431. Oxford: Oxford University Press.
Borer, Hagit. 1984. *Parametric syntax*. Dordrecht: Foris.
Bošković, Željko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33:351-383.
Bošković, Željko. 2005. Left branch extraction, structure of NP, and scrambling. In *The free word order phenomenon: Its syntactic sources and diversity*, ed. by Joachim Sabel and Mamoru Saito, 13-73. Berlin: de Gruyter.
Bošković, Željko. 2008. What will you have, DP or NP? *Proceedings of NELS* 37:101-114.
Bošković, Željko. 2009. More on the no-DP analysis of article-less languages. *Studia Linguistica* 63: 187-203.
Bošković, Željko. 2010. Another monster. Presented at MayFest 2010: Bridging typology and acquisition, University of Maryland.
Bošković, Željko. 2012. On NPs and clauses. In *Discourse and grammar: From sentence types to lexical categories,* ed. by Günther Grewendorf and Thomas Ede Zimmermann, 179-242. Berlin: de Gruyter.
Bošković, Željko. 2016a. On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. *Linguistic Review* 33:17–66.
Bošković, Željko. 2016b. On second position clitics crosslinguistically. In *Formal Studies in Slovenian Syntax. In honor of Janez Orešnik*, ed. by Franc Lanko Marušič and Rok Žaucer, 23-44. Amsterdam: John Benjamins.
Bošković, Željko. 2016c. On the typology of the nominal domain: Its syntax, semantics, and learnability. Presented at Workshop on the Syntax and Semantics of the Nominal Domain, Goethe-Universität Frankfurt.
Bošković, Željko. 2018. On movement out of moved elements, labels, and phases. *Linguistic Inquiry* 49:247-282.

Bošković, Željko. 2020. On the Coordinate Structure Constraint, across-the-board movement, phases, and Labeling. In *Recent developments in Phase Theory,* ed. by Jeroen van Craenenbroeck, Cora Pots, and Tanja Temmerman, 133-182. Berlin: de Gruyter.
Bošković, Željko. 2021a. Generalized Asymmetry. In Keynotes from the International Conference on Explanation and Prediction in Linguistics: Formalist and Functionalist Approaches, ed. by Peter Kosta and Katrin Schlund, 15-77. Berlin: Peter Lang.
Bošković, Željko. 2021b. On the (im)practicality of the state: Why do I have to have a country? In *Chomsky on State and Democracy,* ed. by Günther Grewendorf, 77-94. Nomos Verlag.
Bošković, Željko and Jon Gajewski. 2011. Semantic correlates of the NP/DP parameter. *Proceedings of NELS* 39:121-134.
Bošković, Željko and Serkan Şener. 2014. The Turkish NP. In *Crosslinguistic studies on nominal reference: With and without articles*, ed. by Patricia Cabredo Hofherr and Anne Zribi-Hertz, 102-140. Leiden: Brill.
Bresnan, Joan and Sam Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language* 63:741–782.
Brown, Dunstan, Marina Chumakina and Greville G. Corbett 2013. *Canonical morphology and syntax*. Oxford: Oxford University Press.
Brown, Sue, and Steven Franks. 1995. Asymmetries in the scope of Russian negation. *Journal of Slavic Linguistics* 3:239-287.
Bybee, Joan L. and Östen Dahl. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13:51-103.
Cheng, Hsu-Te. 2013. Argument ellipsis, classifier phrases, and the DP parameter. Doctoral dissertation, University of Connecticut.
Cheng, Lisa, and Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30:509-542.
Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6:339-405.
Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge: MIT Press.
Chomsky, Noam. 2004. Beyond explanatory adequacy. In *The cartography of syntactic structures. Vol. 3*, ed. by Adriana Belletti, 104-131. Oxford: Oxford University Press.
Chomsky, Noam. 2020. The UCLA lectures. lingbuzz/005485
Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.
Cinque, Guglielmo. 2004. "Restructuring" and functional structure. In *Structures and Beyond: The Cartography of Syntactic Structures,* ed. by Adrianna Belletti, 132-191. Oxford: Oxford University Press.
Cinque, Guglielmo. 2005. Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36: 315-332.
Collins, Chris and Paul Postal. 2014. *Classical NEG raising: An essay on the syntax of negation.* Cambridge: MIT Press.
Corbett, Greville G. 1983. *Hierarchies, targets and controllers: Agreement patterns in Slavic*. University Park: Pennsylvania University Press.
Corbett, Greville G. 2007. Canonical typology, suppletion, and possible words. *Language* 83:8-42.
Corver, Norbert. 1992. On deriving left branch extraction asymmetries: a case study in parametric syntax. *Proceedings of NELS* 22:67-84.
Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological*

Perspective. Oxford: Oxford University Press.
Dadan, Marcin. 2019. Head labeling preference and language change. Doctoral dissertation, University of Connecticut.
Dayal, Vaneeta. 2004. Number marking and (In)definiteness in kind terms. *Linguistics and Philosophy* 27:393-450.
DeLancey, Scott. 2001. *Lectures on functional syntax*. LSA Summer Institute, UC Santa Barbara.
Despić, Miloje. 2011. Syntax in the absence of determiner phrase. Doctoral dissertation, University of Connecticut.
Despić, Miloje. 2013a. Intensifiers, focus, and clitics: is pronoun position truly an argument for D in Serbo-Croatian? In *Nominal Constructions in Slavic and Beyond*, ed. by Lilia Schürcks, Anastasia Giannakidou, Urtzi Etxeberria, Peter Kosta, 39-74. Berlin: Mouton de Gruyter.
Despić, Miloje. 2013b. Binding and the structure of NP in Serbo-Croatian. *Linguistic Inquiry* 44:239–270.
Despić, Miloje. 2019. On kinds and anaphoricity in languages without definite articles. In *Definiteness across languages*, ed. by Ana Aguilar-Guevara, Julia Pozas Loyo and Violeta Vázquez-Rojas Maldonado, 259–291. Berlin: Language Science Press.
Di Sciullo, Anna Maria. 2015. On the domain specificity of the human language faculty and the effects of principles of computational efficiency: contrasting language and mathematics. *Revista Linguística* 11:57-70.
Dryer, Matthew S. 1997. Are grammatical relations universal? In *Essays on language function and language type: Dedicated to T. Givón*, ed. by Joan L. Bybee, John Haiman, and Sandra A. Thompson, 115–143. Amsterdam: John Benjamins.
Epstein, Samuel. 1999. Un-principled syntax: The derivation of syntactic relations. In *Working Minimalism*, ed. by Samuel Epstein and Norbert Hornstein, 317-345. Cambridge: MIT Press.
Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language and Linguistic Theory* 34:429–479.
Erlewine, Michael Yoshitaka. 2020. Anti-locality and subject extraction. *Glossa* 5:1-38.
Franks, Steven. 2019. Binding and phasehood in south Slavic revisited. *Studies in Polish Linguistics* 14:61-80.
Fukui, Naoki. 1988. Deriving the differences between English and Japanese. *English Linguistics* 5:249-270.
Gajewski, Jon. 2005. Neg-raising: Polarity and presupposition. Doctoral dissertation, MIT.
Gil, David. 2020. Describing languoids: When incommensurability meets the language-dialect continuum. *Linguistic Typology* 20:439-462.
Golan, Yael. 1993. Node Crossing Economy, Superiority and D-linking. Ms, Tel Aviv University.
Golumbia , David. 2010. Minimalism is functionalism. *Language Sciences* 32:28-42.
Greenberg, Joseph. 1974. *Language typology: A historical and analytic overview*. Berlin: Gruyter.
Guasti, Teresa, Anna Gavarro, Joke de Lange and Claudia Caprin. 2008. Article Omission across Child Languages. *Language Acquisition* 15:89-119.
Harley, Heidi and Elizabeth Ritter. 2002. Person and number in pronouns. *Language* 78:482–526.
Haspelmath, Martin. 1997. *Indefinite pronouns.* Oxford: Oxford University Press.
Haspelmath, Martin. 2004. Explaining the ditransitive person-role constraint: A usage-based approach. *Constructions 2*, 49.
Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86:663-687.
Haspelmath, Martin. 2020. General linguistics must be based on universals (or nonconventional aspects of language). lingbuzz/005158.
Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298: 1569-1579.

Hawkins, John. 2010. Processing efficiency and complexity in typological patterns. In *Oxford handbook of linguistic typology,* ed. by Jae Jung Song, 1-16. Oxford: Oxford University Press.
Huang, C-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.
Jenks, Peter. 2018. Articulated definiteness without articles. *Linguistic Inquiry* 49:501-536.
Jiang, Li. 2012. Nominal arguments and language variation. Doctoral dissertation, Harvard.
Julien, Marit. 2002. *Syntactic Heads and Word Formation*. Oxford: Oxford University Press.
Kang, Jungmin. 2014. On the Absence of TP and its Consequences: Evidence from Korean. Doctoral dissertation, University of Connecticut.
Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge: MIT Press.
Kayne, Richard. 2010. Why are there no directionality parameters? *Proceedings of WCCFL* 28:1-23.
Koulidobrova, Elena. In press. DP or not DP: Acquisitional evidence from English. *Language Acquisition.*
Kuno, Susumu and Ken-ichi Takami. 2004. *Functional constraints in grammar: On the unergative-unaccusative distinction.* Amsterdam: John Benjamins.
Lacerda, Renato. 2020. Middle-field syntax and information structure in Brazilian Portuguese. Doctoral dissertation, University of Connecticut.
Lander, Yury and Peter Arkadiev. 2020. On the right of being a comparative concept. *Linguistic Typology* 20:403-416.
LaTerza, Ivana. 2016. Binding in English and South Slavic and the parameterized DP hypothesis *Linguistic Inquiry* 47:741-753.
Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change*. Oxford: Oxford University Press.
Lehmann, Christian. 1989. Language description and general comparative grammar. In *Reference grammars and modern linguistic theory*, ed. by Gottfried Graustein and Gerhard Leitner, 133-62. Tübingen: Niemeyer.
Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity at the syntax-lexical semantics interface.* Cambridge: MIT Press.
Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
Marelj, Marijana. 2011. Bound-Variable Anaphora and Left Branch Condition. *Syntax* 14:205-229.
Matthewson, Lisa, Tim Bryant and Thomas Roeper. 2001. A Salish stage in the acquisition of English determiners: Unfamiliar ‘‘definites’’. *Proceedings of SULA* 25:63–71.
McCloskey, James. 2000. Quantifier float and wh-movement in an Irish English. *Linguistic Inquiry* 31:57-84.
Messick, Troy. 2016. Pronouns and agreement in Telugu embedded contexts. *Proceedings of WCCFL* 33:309-319.
Messick, Troy. 2017. The morphosyntax of self-ascription: A Cross-linguistic Study. Doctoral dissertation, University of Connecticut.
Messick, Troy. 2020. The derivation of highest subject questions and the nature of the EPP. *Glossa* 5:1-12.
Neeleman, Ad and Hans van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *Journal of Comparative Germanic Linguistics* 11:137–189.
Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge: MIT Press.
Nichols, Johanna. 2007. What, if anything, is typology? *Linguistic Typology* 11:231–238.
Pereltsvaig, Asya. 2007. On the universality of DP. *Studia Linguistica* 61: 59–94.
Petroj, Vanessa. 2020. Ignoring Language Barriers: Romanian-Serbian Code-Switching. Doctoral dissertation, University of Connecticut.
Petrosino, Roberto. 2018. The left periphery fragmented: Evidence from Italian. *Proceedings of NELS* 48:221-230.

Phillips, Colin. 1996. Order and structure. Doctoral dissertation, MIT.
Polinsky, Maria and Robert Kluender. 2007. Linguistic typology and theory construction: Common challenges ahead. *Linguistic Typology* 11:273-283.
Progovac, Ljiljana. 1998. Determiner Phrase in a language without determiners. *Journal of Linguistics* 34:165-179.
Progovac, Ljiljana. 2010. Syntax: Its evolution and its representation in the brain. *Biolinguistics* 4:234-254.
Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. by Liliane Haegeman, 281-337. Dordrecht: Kluwer.
Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *Wh-movement: Moving on*, ed. by Lisa Cheng and Norbert Corver, 97-133. Cambridge: MIT Press.
Roberts, Ian. 2019. *Parameter hierarchies and Universal Grammar.* Oxford University Press.
Round, Erich and Greville Corbett. 2020. Comparability and measurement in typological science: The bright future for linguistics. *Linguistic Typology* 24:489-525.
Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6:445-501.
Runić, Jelena. 2014. A new look at clitics, clitic doubling, and argument ellipsis. Doctoral dissertation, University of Connecticut.
Sapir, Edward. 1921. *Language: An introduction to the study of speech*. New York: Harcourt.
Spike, Matthew. 2020. Fifty shades of grue: Indeterminate categories and induction in and out of the language sciences. *Linguistic Typology* 24:465-488.
Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425-449.
Stegovec, Adrian. 2019. Person on the Edge: Lessons from Crosslinguistic Variation in Syntactic Person Restrictions. Doctoral dissertation, University of Connecticut.
Takahashi, Masahiko. 2011. Some consequences of Case-marking in Japanese. Doctoral dissertation, University of Connecticut.
Talić, Aida. 2013. Adverb extraction, specificity, and structural parallelism. *Canadian Journal of Linguistics* 60:417-445.
Thomas, Margaret. 2020. *Formalism and functionalism in linguistics: The engineer and the collector*. New York: Routledge.
Todorović, Neda. 2016. On the Presence/Absence of TP: Syntactic Properties and Temporal Interpretation. Doctoral dissertation, University of Connecticut.
Trenkić, Danijela. 2004. Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase. *Lingua* 114:1401-1427.
Uriagereka, Juan. 1988. On government. Doctoral dissertation, University of Connecticut.
Willim, Ewa. 2000. On the grammar of Polish nominals. In *Step by step*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 319-346. Cambridge: MIT Press.
Wurmbrand, Susi. 2001. *Infinitives: Restructuring and Clause Structure*. Berlin: de Gruyter.
Zanon, Ksenia. 2015. On hybrid coordination and quantifier raising in Russian. Doctoral dissertation, University of Indiana, Bloomington.
Zanon, Ksenia. 2020. Negated polar interrogatives in Russian. Presented at *Slavic Linguistics Society* 15. Indiana University.
Zlatić, Larisa. 1997. The structure of the Serbian Noun Phrase. Doctoral dissertation, University of Texas at Austin.
Zubizarreta, Maria Luisa. 2003. Intervention effects in the French wh-in-situ construction: Syntax or interpretation? In *A Romance perspective in language knowledge and use- Selected papers from the 31st Linguistic Symposium on Romance Language,* ed. by Rafael Nuñes Cedeño, Luis Lopez, and Richard Camero, 359-379. Amsterdam/Philadelphia: John Benjamins.