22 Universal Grammar and doing comparative linguistics without an a priori assumption of a (strong) UG

The following two sections deal with the tools that I believe to be necessary to capture generalizations and the way one can derive such generalizations.

22.1 Formal tools for capturing generalizations

In Chapter 13, it was shown that all the evidence that has previously been brought forward in favor of innate linguistic knowledge is in fact controversial. In some cases, the facts are irrelevant to the discussion and in other cases, they could be explained in other ways. Sometimes, the chains of argumentation are not logically sound or its premises are not supported. In other cases, the argumentation is circular. As a result, the question of whether there is innate linguistic knowledge still remains unanswered. All theories that presuppose the existence of this kind of knowledge are making very strong assumptions. If one assumes, as Kayne (1994) for example, that all languages have the underlying structure [specifier [head complement]] and that movement is exclusively to the left, then, while it is possible to develop a very elegant system, the basic assumptions must be part of innate linguistic knowledge since there is no evidence for the assumption that utterance in all natural languages have the structure that Kayne suggests.

As we have seen, there are a number of alternative theories that are much more surface-oriented than most variants of Transformational Grammar. These alternative theories often differ with regard to particular assumptions that have been discussed in the preceding sections. For example, there are differences in the treatment of long-distance dependencies that have led to a proliferation of lexical items in Categorial Grammar (see Section 8.6). As has been shown by Jacobs (2008), Jackendoff (2008) and others, approaches such as Categorial Grammar that assume that every phrase must have a functor/head cannot explain certain constructions in a plausible way. Inheritance-based phrasal analyses that only list heads with a core meaning in the lexicon and have the constructions in which the heads occur determine the meaning of a complex expression, turn out to have difficulties with derivational morphology and with accounting for alternative ways of argument realization (see Section 21.2.2, 21.4.1, and 21.4.2). We therefore need a theory that handles argument structure changing processes in the lexicon and still has some kind of phrase structure or relevant schemata. Some variants of GB/MP as well as LFG, HPSG, TAG and variants of CxG are examples of this kind of theory. Of these theories, only HPSG and some variants of CxG make use of the same descriptive tools ((typed) feature descriptions) for roots, stems, words, lexical rules and phrases. By using a uniform description for all these objects, it is possible to formulate generalizations over the relevant objects. It is therefore possible to capture what particular words have in common with lexical rules or phrases. For example, the *-bar* derivation corresponds to a complex passive construction with a modal verb. (1) illustrates.

(1) a. Das Rätsel ist lösbar.
the puzzle is solvable

b. Das Rätsel kann gelöst werden.
the puzzle can solved be
'The puzzle can be solved.'

By using the same descriptive inventory for syntax and morphology, it is possible to capture cross-linguistic generalizations: something that is inflection/derivation in one language can be syntax in another.

It is possible to formulate principles that hold for both words and phrases and furthermore, it is possible to capture cross-linguistic generalizations or generalizations that hold for certain groups of languages. For example, languages can be divided into those with fixed constituent order and those with more free or completely free constituent order. The corresponding types can be represented with their constraints in a type hierarchy. Different languages can use a particular part of the hierarchy and also formulate different constraints for each of the types (see Ackerman & Webelhuth 1998: Section 9.2). HPSG differs from theories such LFG and TAG in that phrases are not ontologically different from words. This means that there are no special c-structures or tree structures. Descriptions of complex phrases simply have additional features that say something about their daughters. In this way, it is possible to formulate cross-linguistic generalizations about dominance schemata. In LFG, the c-structure rules are normally specified separately for each language. Another advantage of consistent description is that one can capture similarities between words and lexical rules and also between words and phrases. For example, a complementizer such as *dass* 'that' shares a number of properties with a simple verb or with coordinated verbs in initial position.

(2) a. [dass] Maria die Platte kennt und liebt
that Maria the record knows and loves
'that Maria knows and loves the record'

b. [Kennt und liebt] Maria die Platte?
knows and loves Maria the record
'Does Mary know and love the record?'

The difference between the linguistic objects mainly lies in the kind of phrase they select: the complementizer requires a sentence with a visible finite verb, whereas the verb in initial position requires a sentence without a visible finite verb.

In Section 9.1.5, a small part of an inheritance hierarchy was presented. This part contains types that probably play a role in the grammars of all natural languages: there are head-argument combinations in every language. Without this kind of combinatorial operation, it would not be possible to establish a relation between two concepts. The ability to create relations, however, is one of the basic properties of language.

As well as more general types, the type hierarchy of a particular language contains language-specific types or those specific to a particular class of languages. All languages presumably have one and two-place predicates and for most languages (if not all), it makes sense to talk about verbs. It is then possible to talk about one and two-place verbs. Depending on the language, these can then be subdivided into intransitive and transitive. Constraints are formulated for the various types that can either hold generally or be language-specific.

In English, verbs have to occur before their complements and therefore have the INITIAL value +, whereas verbs in German have the INITIAL value − and it is the lexical rule for initial position that licenses a verb with an INITIAL value +.

The differing settings of the INITIAL value for German and English is reminiscent of parameters from GB-Theory. There is one crucial difference, however: it is not assumed that a language learner sets the INITIAL value once and for all for all heads. The use of an INITIAL value is compatible with models of acquisition that assume that learners learn individual words with their positional properties. It is certainly possible for the respective words to exhibit different values for a particular feature. Generalizations about the position of entire word classes are only learned at a later point in the acquisition process.

A hierarchy analogous to the one proposed by Croft (see Section 21.4.1) is given in Figure 22.1. For inflected words, the relevant roots are in the lexicon. Examples of this

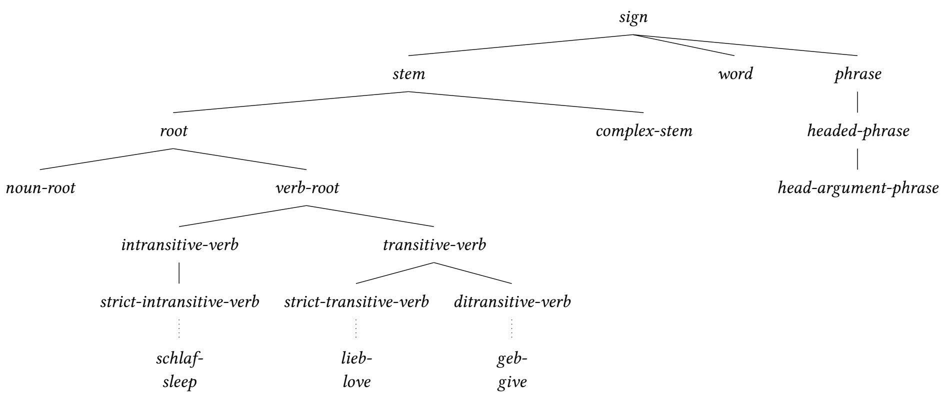

Figure 22.1: Section of an inheritance hierarchy with lexical entries and dominance schemata

are *schlaf*- 'sleep', *lieb*- 'love' and *geb*- 'give'. In Figure 22.1, there are different subtypes of *root*, the general type for roots: *intrans-verb* for intransitive verbs and *trans-verb* for transitive verbs. Transitive verbs can be further subdivided into strictly transitive verbs (those with nominative and accusative arguments) and ditransitive verbs (those with nominative and both accusative and dative arguments). The hierarchy above would of course have to be refined considerably as there are even further sub-classes for both transitive and intransitive verbs. For example, one can divide intransitive verbs into unaccusative and unergative verbs and even strictly transitive verbs would have to be divided into further sub-classes (see Welke 2009: Section 2).

In addition to a type for roots, the above figure contains types for stems and words. Complex stems are complex objects that are derived from simple roots but still have to be inflected (*lesbar-* 'readable', *besing-* 'to extol'). Words are objects that do not inflect. Examples of these are the pronouns *er* 'he', *sie* 'she' etc. as well as prepositions. An inflected form can be formed from a verbal stem (*geliebt* 'loved', *besingt* 'extols'). Relations between inflected words and (complex) stems can be formed again using derivation rules. In this way, *geliebt* 'loved' can be recategorized as an adjective stem that must then be combined with adjectival endings (*geliebt-e*). The relevant descriptions of complex stems/words are subtypes of *complex-stem* or *word*. These subtypes describe the form that complex words such as *geliebte* must have. For a technical implementation of this, see Müller (2002a: Section 3.2.7). All words can be combined to form phrases using dominance schemata. The hierarchy given here is of course by no means complete. There are a number of additional valence classes and one could also assume more general types that simply describe one, two and three-place predicates. Such types are probably plausible for the description of other languages. Here, we are only dealing with a small part of the type hierarchy in order to have a comparison to the Croftian hierarchy: in Figure 22.1, there are no types for sentence patterns with the form [Sbj IntrVerb], but rather types for lexical objects with a particular valence (V[SUBCAT ⟨ NP[*str*] ⟩]). Lexical rules can then be applied to the relevant lexical objects that license objects with another valence or introduce information about inflection. Complete words can be combined in the syntax with relatively general rules, for example in head-argument structures. The problems from which purely phrasal approaches suffer are thereby avoided. Nevertheless generalizations about lexeme classes and the utterances that can be formed can be captured in the hierarchy.

There are also principles in addition to inheritance hierarchies: the Semantics Principle presented in Section 9.1.6 holds for all languages. The Case Principle that we also saw is a constraint that only applies to a particular class of languages, namely nominative-accusative languages. Other languages have an ergative-absolutive system.

The assumption of innate linguistic knowledge is not necessary for the theory of language sketched here. As the discussion in Section 13 has shown, the question of whether this kind of knowledge exists has still not been answered conclusively. If it should turn out that this knowledge actually exists, the question arises of what exactly is innate. It would be a plausible assumption that the part of the inheritance hierarchy that is relevant for all languages is innate together with the relevant principles. It could, however, also be the case that only a part of the more generally valid types and principles is innate since something being innate does not follow from the fact that it is present in all languages (also see Section 13.1.9).

In sum, one can say that theories that describe linguistic objects using a consistent descriptive inventory and then make use of inheritance hierarchies to capture generalizations, are the ones best suited to represent similarities between languages. Furthermore, this kind of theory is compatible with both a positive and negative answer to the question of whether there is innate linguistic knowledge.

22.2 How to develop linguistic theories that capture cross-linguistic generalizations

In the previous section I argued for a uniform representation of linguistic knowledge on all descriptive levels and that type hierarchies are a good tool for representing generalizations. This section explores a way to develop grammars that are motivated by facts from several languages.

If one looks at the current practice in various linguistic schools one finds two extreme ways of approaching language. On the one hand we have the Mainstream Generative Grammar (MGG) camp and on the other hand we have the Construction Grammar/Cognitive Grammar camp. I hasten to say, that what I say here does not hold for all members of these groups, but for the extreme cases. The caricature of the MGG scientist is that he is looking for underlying structures. Since these have to be the same for all languages (poverty of the stimulus), it is sufficient to look at one language, say English. The result of this research strategy is that one ends up with models that were suggested by the most influential linguist for English and that one then tries to find ways to accommodate other languages. Since English has an NP VP structure, all languages have to have it. Since English reorders constituent in passive sentences, passive is movement and all languages have to work this way. I discussed the respective analyses of German in more detail in the Section 3.4.2 and Chapter 20 and showed that the assumption that passive is movement makes unwanted predictions for German since the subject of passives stays in the object position in German. Furthermore, this analysis requires the assumption of invisible expletives, that is, entities that cannot be seen and do not have any meaning.

On the other extreme of the spectrum we find people working in Construction Grammar or without any framework at all and who claim that all languages are so different that we not even can use the same vocabulary to analyze them and within languages we have so many different objects that it is impossible (or too early) to state any generalizations. Again, what I describe here are extreme positions and clichés.

In what follows, I sketch the procedure that we apply in the CoreGram project[1] (Müller 2013a, 2015a). In the CoreGram project we work on a set of typologically diverse languages in parallel:

- German (Müller 2007b, 2009c, 2012a; Müller & Ørsnes 2011, 2013a; Müller 2014b, 2015b)
- Danish (Ørsnes 2009b; Müller 2009c, 2012a; Müller & Ørsnes 2011, 2013a,b, 2015)

[1] http://hpsg.fu-berlin.de/Projects/CoreGram.html, January 14, 2016.

- Persian (Müller 2010b; Müller & Ghayoomi 2010)
- Maltese (Müller 2009b)
- Mandarin Chinese (Lipenkova 2009; Müller & Lipenkova 2009, 2013, 2015)
- Yiddish (Müller & Ørsnes 2011)
- English (Müller 2009c, 2012a; Müller & Ørsnes 2013a)
- Hindi
- Spanish (Machicao y Priemer 2015)
- French

These languages belong to diverse language families (Indo-European, Afro-Asiatic, Sino-Tibetan) and among the Indo-European languages the languages belong to different groups (Germanic, Romance, Indo-Iranian). Figure 22.2 provides an overview. We work

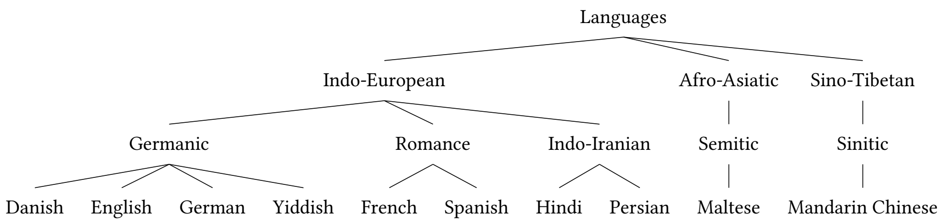

Figure 22.2: Language families and groups of the languages covered in the CoreGram project

out fully formalized, computer-processable grammar fragments in the framework of HPSG that all have a semantics component. The details will not be discussed here, but the interested reader is referred to Müller (2015a).

As was argued in previous sections, the assumption of innate language-specific knowledge should be kept to a minimum. This is also what Chomsky suggested in his Minimalist Program. It is even possible that there is no language-specific innate knowledge at all, a view taken in Construction Grammar/Cognitive Grammar. So instead of imposing constraints from one language onto other languages, a bottom-up approach seems to be more appropriate: grammars for individual languages should be motivated language-internally. Grammars that share certain properties can be grouped in classes. This makes it possible to capture generalizations about groups of languages and language as such. Let us consider some example languages: German, Dutch, Danish, English and French. If we start developing grammars for German and Dutch, we find that they share a lot of properties: for instance, both are SOV and V2 languages, both have a verbal complex. One main difference is the order of elements in the verbal complex. The situation can be

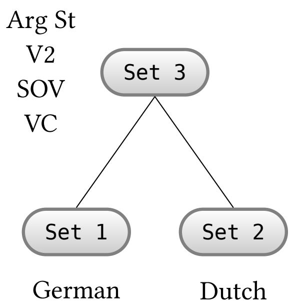

Figure 22.3: Shared properties of German and Dutch

depicted as in Figure 22.3. There are some properties that are shared between German and Dutch (Set 3). For instance, the argument structure of lexical items, a list containing descriptions of syntactic and semantic properties of arguments, and the linking of these arguments to the meaning of the lexical items is contained in Set 3. In addition to the constraints for SOV languages, the verb position in V2 clauses and the fronting of a constituent in V2 clauses are contained in Set 3. The respective constraints are shared between the two grammars. Although these sets are arranged in a hierarchy in the figures this has nothing to do with the type hierarchies that have been discussed in the previous subsection. These type hierarchies are part of our linguistic theories and parts of such hierarchies can be in different sets: those parts of the type hierarchy that concern more general aspects can be in Set 3 in Figure 22.3 and those that are specific to Dutch or German are in the respective other sets. When we add another language, say Danish, we get further differences. While German and Dutch are SOV, Danish is an SVO language. Figure 22.4 on the following page shows the resulting situation: the top-most node represents constraints that hold for all the languages considered so far (for instance the argument structure constraints, linking and V2) and the node below it (Set 4) contains constraints that hold for German and Dutch only.[2] For instance Set 4 contains constraints regarding verbal complexes and SOV order. The union of Set 4 and Set 5 is Set 3 of Figure 22.3.

If we add further languages, further constraint sets will be distinguished. Figure 22.5 on the next page shows the situation that results when we add English and French. Again, the picture is not complete since there are constraints that are shared by Danish and English but not by French, but the general idea should be clear: by systematically working this way, we should arrive at constraint sets that directly correspond to those that have been established in the typological literature.

The interesting question is what will be the top-most set if we consider enough lan-

[2] In principle, there could be constraints that hold for Dutch and Danish but not for German and for German and Danish, but not for Dutch. These constraints would be removed from Set 1 and Set 2 respectively and put into another constraint set higher up in the hierarchy. These sets are not illustrated in the figure and I keep the names Set 1 and Set 2 from Figure 22.3 for the constraint sets for German and Dutch.

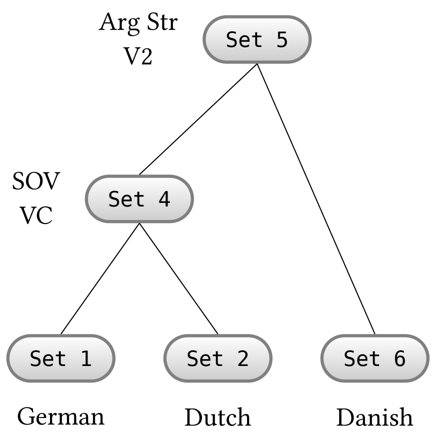

Figure 22.4: Shared Properties of German, Dutch, and Danish

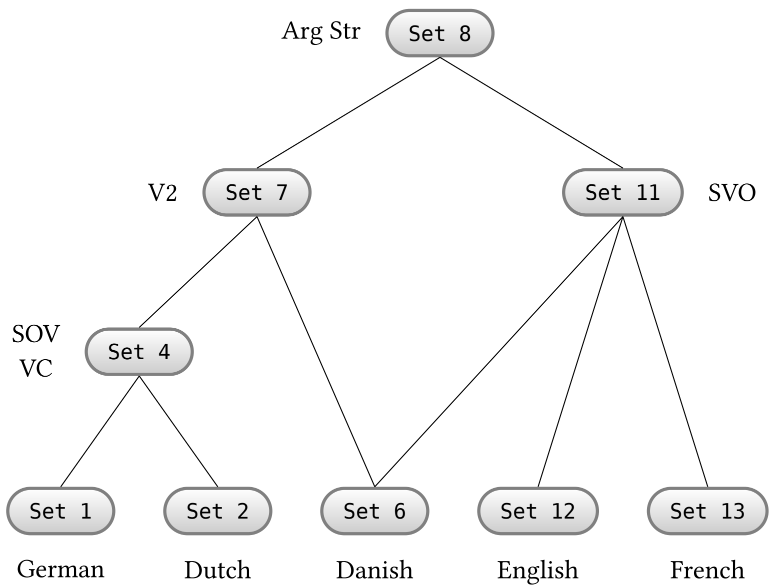

Figure 22.5: Languages and language classes

guages. At first glance, one would expect that all languages have valence representations and linkings between these and the semantics of lexical items (argument structure lists in the HPSG framework). However, Koenig & Michelson (2012) argue for an analysis of Oneida (a Northern Iroquoian language) that does not include a representation of syntactic valence. If this analysis is correct, syntactic argument structure would not be universal. It would of course be characteristic of a large number of languages, but it would not be part of the top-most set. So this leaves us with just one candidate for the top-most set from the area of syntax: the constraints that license the combination of two or more linguistic objects. This is basically Chomsky's External Merge without the binarity restriction[3]. In addition, the top-most set would of course contain the basic machinery for representing phonology and semantics.

It should be clear from what has been said so far that the goal of every scientist who works this way is to find generalizations and to describe a new language in a way that reuses theoretical constructs that have been found useful for a language that is already covered. However, as was explained above, the resulting grammars should be motivated by data of the respective languages and not by facts from other languages. In situations where more than one analysis would be compatible with a given dataset for language X, the evidence from language Y with similar constructs is most welcome and can be used as evidence in favor of one of the two analyses for language X. I call this approach the *bottom-up approach with cheating*: unless there is contradicting evidence we can reuse analyses that have been developed for other languages.

Note that this approach is compatible with the rather agnostic view advocated by Haspelmath (2010a), Dryer (1997), Croft (2001: Section 1.4.2–1.4.3), and others, who argue that descriptive categories should be language-specific, that is, the notion of *subject* for Tagalog is different from the one for English, the category *noun* in English is different from the category *noun* in Persian and so on. Even if one follows such extreme positions, one can still derive generalizations regarding constituent structure, head-argument relations and so on. However, I believe that some categories can fruitfully be used cross-linguistically, if not universally, then at least for language classes. As Newmeyer (2010: 692) notes with regard to the notion of *subject*: calling two items *subject* in one language does not entail that they have identical properties. The same is true for two linguistic items from different languages: calling a Persian linguistic item *subject* does not entail that it has exactly the same properties as an English linguistic item that is called *subject*. The same is, of course, true for all other categories and relations, for instance parts of speech: Persian nouns do not share all properties with English nouns.[4] Haspelmath

[3] Note that binarity is more restrictive than flat structures: there is an additional constraint that there have to be exactly two daughters. As was argued in Section 21.10.4 one needs phrasal constructions with more than two constituents.

[4] Note that using labels like *Persian Noun* and *English Noun* (see for instance Haspelmath (2010a: Section 2) for such a suggestion regarding case, e. g. Russian Dative, Korean Dative, …) is somehow strange since it implies that both Persian nouns and English nouns are somehow nouns. Instead of using the category *Persian Noun* one could assign objects of the respective class to the class *noun* and add a feature LANGUAGE with the value *persian*. This simple trick allows one to assign both objects of the type *Persian Noun* and objects of the type *English Noun* to the class *noun* and still maintain the fact that there are differences. Of course, no theoretical linguist would introduce the LANGUAGE feature to differentiate between Persian and

(2010c: 697) writes: *Generative linguists try to use as many crosslinguistic categories in the description of individual languages as possible, and this often leads to insurmountable problems.* If the assumption of a category results in problems, they have to be solved. If this is not possible with the given set of categories/features, new ones have to be assumed. This is not a drawback of the methodology, quite the opposite is true: if we have found something that does not integrate nicely into what we already have, this is a sign that we have discovered something new and exciting. If we stick to language-particular categories and features, it is much harder to notice that a special phenomenon is involved since all categories and features are specific for one language anyway. Note also that not all speakers of a language community have exactly the same categories. If one were to take the idea of language-particular category symbols to an extreme, one would end up with person specific category symbols like *Klaus-English-noun.*

After my talk at the MIT in 2013, members of the linguistics department objected to the approach taken in the CoreGram project and claimed that it would not make any predictions as far as possible/impossible languages are concerned. Regarding predictions two things must be said: firstly, predictions are being made on a language particular basis. As an example consider the following sentences from Netter (1991):

(3) a. [Versucht, zu lesen], hat er das Buch nicht.
tried to read has he.NOM the.ACC book not
'He did not try to read the book.'

b. [Versucht, einen Freund vorzustellen], hat er ihr noch nie.
tried a.ACC friend to.introduce has he.NOM her.DAT yet never
'He never before tried to introduce a friend to her.'

When I first read these sentences I had no idea about their structure. I switched on my computer and typed them in and within milliseconds I got an analysis of the sentences and by inspecting the result I realized that these sentences are combinations of partial verb phrase fronting and the so-called third construction (Müller 1999a: 439). I had previously implemented analyses of both phenomena but had never thought about the interaction of the two. The grammar predicted that examples like (3) are grammatical. Similarly the constraints of the grammar can interact to rule out certain structures. So predictions about ungrammaticality/impossible structures are in fact made as well.

Secondly, the top-most constraint set holds for all languages seen so far. It can be regarded as a hypothesis about properties that are shared by all languages. This constraint set contains constraints about the connection between syntax and information structure and such constraints allow for V2 languages but rule out languages with the verb in penultimate position (see Kayne (1994: 50) for the claim that such languages do not exist. Kayne develops a complicated syntactic system that predicts this). Of course if a language is found that puts the verb in penultimate position for the encoding of

English nouns, but nouns in the respective languages have other features that make them differ. So the part of speech classification as noun is a generalization over nouns in various languages and the categories *Persian Noun* and *English Noun* are feature bundles that contain further, language-specific information.

sentence types or some other communicative effect, a more general top-most set has to be defined, but this is parallel for Minimalist theories: if languages are found that are incompatible with basic assumptions, the basic assumptions have to be revised. As with the language particular constraints, the constraints from the top-most set make certain predictions about what can be and what cannot be found in languages.

Frequently discussed examples such as those languages that form questions by reversing the order of the words in a string (Musso et al. 2003) need not be ruled out in the grammar since they are ruled out by language external constraints: we simply do not have enough working memory to do such complex computations.

A variant of this argument comes from David Pesetzky and was raised in Facebook discussions of an article by Paul Ibbotson and Michael Tomasello in the Guardian[5]. Pesetzky claimed that Tomasello's theory of language acquisition could not explain why we find V2 languages but no V3 languages. First, I do not know of anything that blocks V3 languages in current Minimalist theories. So per se the fact that V3 languages may not exist cannot be used to support any of the competing approaches. Of course the question could be asked whether the V3 pattern would be useful for reaching our communicative goals and whether it can be easily acquired. Now, with V2 as a pattern it is clear that we have exactly one position that can be used for special purposes in the V2 sentence (topic or focus). For mono-valent and bi-valent verbs we have an argument that can be placed in initial position. The situation is different for the hypothetical V3 languages, though: If we have mono-valent verbs like *sleep*, there is nothing for the second position. As Pesetzky pointed out in the answer to my comment on a blog post, languages solve such problems by using expletives. For instance some languages insert an expletive to mark subject extraction in embedded interrogative sentences, since otherwise the fact that the subject is extracted would not be recognizable by the hearer. So the expletive helps to make the structure transparent. V2 languages also use expletives to fill the initial position if speakers want to avoid something in the special, designated position:

(4) Es kamen drei Männer zum Tor hinein.
EXPL came three man to.the gate in
'Three man came through the gate.'

In order to do the same in V3 languages one would have to put two expletives in front of the verb. So there seem to be many disadvantages of a V3 system that V2 systems do not have and hence one would expect that V3 systems are less likely to come into existence and if they do exist that they will be subject to change in the course of time, e. g. omission of the expletive with intransitives, optional V2 with transitives and finally V2 in general. With the new modeling techniques for language acquisition and agent-based community simulation one can actually simulate such processes and I guess in the years to come, we will see exciting work in this area.

Cinque (1999: 106) suggested a cascade of functional projections to account for reoccurring orderings in the languages of the world. He assumes elaborate tree structures to

[5] *The roots of language: what makes us different from other animals?*. published 05.11.2015. http://www.theguardian.com/science/head-quarters/2015/nov/05/roots-language-what-makes-us-different-animals

play a role in the analysis of all sentences in all languages even if there is no evidence for respective morphosyntactic distinctions in a particular language (see also Cinque & Rizzi 2010: 55). In the latter case, Cinque assumes that the respective tree nodes are empty. Cinque's results could be incorporated in the model advocated here. We would define part of speech categories and morpho-syntactic features in the top-most set and state linearization constraints that enforce the order that Cinque encoded directly in his tree structure. In languages in which such categories are not manifested by lexical material, the constraints would never apply. Neither empty elements nor elaborate tree structures would be needed. So, we have shown that Cinque's data could be covered in a better way in an HPSG with a rich UG but we, nevertheless, refrain from introducing 400 categories (or features) into the theories of all languages and, again, we point out that this is implausible from a genetic point of view and wait for other, probably functional, explanations of the Cinque data.

Note also that implicational universals can be derived from hierarchically organized constraint sets as the ones proposed here. For instance, one can derive from Figure 22.5 the implicational statement that all SVO languages are V2 languages since there is no language that has constraints from Set 4 that does not also have the constraints of Set 7. Of course this implicational statement is wrong, since there are lots and lots of SOV languages and just exceptionally few V2 languages. So as soon as we add other languages as for instance Persian or Japanese, the picture will change.

The methodology suggested here differs from what is done in MGG in that the general constraints that are supposed to hold for all languages are just stipulated on the basis of general speculations about language. In the best case, these general assumptions are fed by a lot of experience with different languages and grammars, in the worst case they are derived from insights gathered from one or more Indo-European languages. Quite often impressionistic data is used to motivate rather far-reaching fundamental design decisions (Fanselow 2009; Sternefeld & Richter 2012; Haider 2014). It is interesting to note that this is exactly what members of the MGG camp accuse typologists of. Evans & Levinson (2009a) pointed out that counter examples can be found for many alleged universals. A frequently uttered reply is that unanalyzed data cannot refute grammatical hypotheses (see for instance Freidin 2009: 454). In the very same way it has to be said that unanalyzed data should not be used to build theories on (Fanselow 2009). In the CoreGram project, we aim to develop broad-coverage grammars of several languages, so those constraints that make it to the top node are motivated and not stipulated on the basis of intuitive implicit knowledge about language.

Since it is data-oriented and does not presuppose innate language-specific knowledge, this research strategy is compatible with work carried out in Construction Grammar (see Goldberg (2013b: 481) for an explicit statement to this end) and in any case it should also be compatible with the Minimalist world.