Explanation in syntax: generative syntax from a functional perspective and the incommensurability of syntactic theories (concept januari 2021)[1]

Egbert Fortuin

e.fortuin@hum.leidenuniv.nl

Leiden University

Abstract

Within linguistics there are several theories which deal with syntax. A general division can be made between Chomskyan generative syntactic theories and various functional syntactic theories. This paper addresses the question (i) to what extent these theories and their explanations deal with the same type of phenomena, (ii) to what extent they use the same concept of a successful explanation and (iii) to what extent they are incommensurable in the sense of Kuhn (1962). This topic is addressed by a discussion of a number of generative analyses (case studies), which are analyzed and evaluated from a functional perspective. This analysis shows how a number of phenomena that are analyzed and explained within the generative paradigm can be explained differently from a functional theoretical perspective and framework. The paper concludes that generative and functional approaches to syntax are incommensurable in the sense of Kuhn, even though there are ways to engage in a common discussion about syntax.

---

[1] To appear in: Kosta, Peter & Schlund, Katrin, eds. (forthcoming): Proceedings from the International Conference on Explanation and Prediction in Linguistics (CEP): Formalist and Functionalist Approaches. Heidelberg, February 13th and 14th, 2019. Potsdam Linguistic Investigations. Berli ; Bern ; Bruxelles: Peter Lang.

1. Introduction

Explanation essentially deals with providing an answer to a why-question (Skow 2016). In the study of syntactic phenomena, not only what is seen as a good "why-question", but also what is seen as a good answer to such a question differs considerably. From an abstract perspective two types of theories can be distinguished in the study of syntax. One the one hand various functional theories of syntax exist, for example various descriptive theories (see for example Beedham 2001 for this term), Dependency Grammar (Ágel and Fischer 2015 for a discussion), (European) Structuralist theories such as Universal Semantic Syntax (see Fortuin & Geerdink-Verkoren 2019 for an overview), Cognitive Linguistics (e.g. Cognitive Grammar by Langacker 1987, 1990, 1991), Construction Grammar (e.g. Goldberg 1995), Functional Discourse Grammar (Hengeveld & Mackenzie 2008), Simpler Syntax by Culicover and Jackendoff (2005), Head Driven Phrase Structure Grammar (Pollard and Sag 1994), Lexical Functional Grammar (Bresnan 2001), etc. One the other side of the spectrum, the dominant syntactic theory is the Chomskyan generative theory, which has been developed over time from the sixties, resulting for example in the Minimalist Program (Chomsky 1995; reprinted 1996]). Whereas functional theories of syntax are interested in questions that deal with form and meaning (linguistic signs) and with the communicative function of language, this is not the case, or at least less central, in generative Chomskyan syntactic theories (henceforth also 'generative syntax' or 'generative grammar'). The difference between the theoretical starting points of these theories also means that what counts as a *unified* explanation (for example an explanation which covers various domains of grammar) or as a *deep* explanation (for example an explanation that can account for details of specific

languages) for similar phenomena differs greatly between the two types of theories or frameworks.

In this paper, I will address the topic of explanation in syntax by discussing the theoretical starting points of the two competing theories (Section two), and by discussing a number of syntactic phenomena that have been the focus of attention of generative syntactic theory (Section three). I will provide an analysis of these phenomena from a different, viz. functional, perspective. In all cases I will argue that the same empirical phenomena can be analyzed and explained differently within a functional theory. In Section four, I conclude that generative theories and functional theories and their explanations are incommensurable in the sense of Kuhn, which explains why there is generally little discussion between the different syntactic paradigms, or discussions which lead to a change of the theoretical positions. Nevertheless, in the conclusion in Section five, I will also provide a number of ways in which people from different frameworks can engage in meaningful discussions about language and syntax.

2. Theoretical starting points to syntax

2.1 Functional approaches to syntax

According to functional approaches to language and grammar, language is a communicative, functional and cultural system, and linguistic data and phenomena are often explained with reference to this communicative function. The functional view of language is elegantly expressed in the following quote from Sapir's *Language: An Introduction to the Study of Speech*:

> ‘Language is a purely human and non-instinctive method of communicating ideas, emotions and desires by means of a system of voluntarily produced symbols.’ (Sapir 1921: 7)

The functional basis of language and grammar is reflected in the different domains or building blocks of language: phonemes (‘meaningful’ sounds), linguistic signs (form-meaning pairings), and (meaningful) combinations of linguistic signs (constructions). The term ‘syntax’ can refer to the rules, regularities, principles, etc. that govern how constructions are formed, or to the discipline that deals with this domain of language. Within functional theories, explanation in syntax may deal with questions like: how is the meaning of a specific construction in a specific language derived from the meanings (and syntactic functions) of the component parts? How is the construction used in actual language? In doing so, it may involve theoretical concepts like: subject-predicate, valency, quantification, compounding, modifier-head, etc. etc. Some of these notions can be applied to various or even all languages, but others will apply to a subset of languages only. As such, functional syntax may focus on individual languages, and explain the syntactic rules and regularities of that language, or focus on syntactic patterns that reoccur in different languages, and try to account for such crosslinguistic patterns.

Within the various functional approaches there are also different points of view or different domains on which the theory focusses. To give one example, Construction Grammar focuses on the fact that a grammar is a network of constructions, and that constructions can be more or less idiomatic, for example due to various diachronic processes. Other functional approaches to syntax may focus more on the regular (compositional) patterns in syntax. Some functional syntactic theories try to link linguistic data to general cognitive capacities, for example Cognitive Linguistics, which uses concepts from psychology like ‘Figure’,

'Ground'. Other functional theories refrain from using insights from what is known about cognition, and focus purely on the linguistic system (for example European branches of Structuralism such as Prague structuralism).

Within functional approaches to syntax, explanation may make a difference between explanation in terms of syntactic rules ("norms") of a specific language and an explanation of the rules themselves. This can be illustrated with two examples that have to do with word order. In Dutch the basic word order of the main clause is "finite verb second", whereas in the subordinate clause the finite verb comes as the last element in the sentence:

(1) Ik **weet** dat hij thuis **is**. (Dutch)
I **know** that he home **is**
'I know that he is home.'

We can explain the difference in word order in Dutch by pointing at the two rules or norms that govern Dutch word order. One can, however, also try to explain why these rules or norms are as they are. The rules of a particular language may be explained with reference to its function, non-linguistic factors (e.g. cognitive), diachronic factors, etc. and may use typological evidence to sustain the explanation.[2] This is formulated by Sandra Thompson with respect to grammar:

> 'Since grammar arises from competing motivations involved in the cognition and pragmatics of human interaction, the most sensible approach to explaining grammar would seem to be to seek to understand

[2] I am not aware of any (diachronic) explanations of the V2 rule in Germanic languages. An explanation of V2 in Germanic might answer questions like: How did the V2 word order in Germanic come about? Are these factors similar to the development in other V2 languages? Or: Why do we find V2 languages in the World's languages, but not V3 languages? Perhaps explanations of V2 might have to refer to the specific semantic, syntactic and prososodic properties of the finite verb in older stages of Germanic.

> the cognitive and pragmatic principles, as well as the 'routinization' principles which provide the complex of forces shaping the design of grammars.' (Thomposon 1991: 96)

Note that this differs from the Chomskyan generative approach to the data, which defines a basic word order (for example SOV for Dutch), from which other orders are derived. Bartsch (1987) discusses the difference between norms (rules) of language, and an explication of these norms in terms of abstract Chomskyan rules, which are not rules that can be learned:

> 'According to Chomsky's theory of grammar, the rules of a grammar are not knowledge in the normal sense of 'knowledge'; they are rather a biologically based mechanism that is automatically filled in and completed in language learning to become the grammar of the language that is learned. These rules are by no means norms. In my view, the rules of these grammars can only be (biologically restricted) systematizations of set of norms. These systematizations are not norms themselves but they make it possible to store the systematic aspects of norm contents in such a way that a fairly regular part of a system of linguistic norms can be derived from it in an effective manner.' (Bartsch 1987: 187)

In the same vein, from a functional perspective a difference must be made between properties of specific languages, and an explanation of these properties, for example in terms of general communicative or cognitive factors (see also Haspelmath 2020 for a discussion, and the difference between "particular linguistics" and "general linguistics").

2.2 Generative approaches to syntax

Within Chomskyan-generative approaches to language as laid down in Transformational Grammar, Government and Binding, Principles and Parameters, Minimalist Program, etc. language is seen as a biological "faculty" (cf. Haspelmath 2020 for the term "linguisticality" for the biological capacity). Since the beginning of generative grammar, an important

argument has been the argument of the *Poverty of stimulus,* and *autonomy of syntax*. In the book *The Minimalist Program* this argument is formulated in the following way:

> 'In general, it is not the case that language is readily usable or "designed for use". (…) Similarly (…), there is no a priori reason to expect that the languages permitted by UG be learnable – that is, attainable under normal circumstances.' (Chomsky 1996: 18; chapter co-authored with Howard Lasnik)

Another argument that has played in important part in generative grammar is the argument of explanatory adequacy as formulated by Chomsky (1964). Rizzi presents this as follows:

> 'A particular adequate analysis meets explanatory adequacy when UG provides general principled reasons for choosing it over imaginable alternatives.' (Rizzi 2017: 100)

As such, an inherent part of explanation in syntax deals with Universal Grammar (UG). Within generative grammar, several ontological claims have been made about the nature of UG, for example its presumed simplicity:

> '…the minimalist program has a particularly strong commitment to the Galilean vision of natural phenomena and theory construction. That is, minimalists endorse the belief (held by all major proponents of modern science, from Kepler to Einstein) that nature is the realization *of the simplest conceivable mathematical ideas* …. The minimalist program calls attention to the underlying *simplicity of this Universal Grammar*…' (Boeckx 2006: 8, 150)

Similarly, Chomsky suggests the Ockham's razor principle in syntactic explanation within the minimalist program:

> ‘In pursuing a minimalist program, we want to make sure that we are not inadvertently sneaking in improper concepts, entities, relations, and conventions. (…) The more sparse the assumptions the more intricate the argument is likely to be.’ (Chomsky 1996: 225)

> ‘To what degree would the simplest imaginable procedures for generating unboundedly many meaningful expressions be procedures that give rise to the observed phenomena of constrained ambiguity?’ (Berwick, Pietroski, Yankama, Chomsky 2011)

The starting points of generative grammar have resulted in various theoretical concepts that are absent in functional approaches to grammar such as movement, traces, etc. They are the result of the idea that one has to explain the data (e.g. constraints on configuration of words) with the *simplest imaginable procedures*. Even though movement, traces, blocking, deleting, etc. are non-directly observable phenomena, they are treated as actual empirical phenomena and as a *mental reality* underlying actual linguistic behavior within generative syntax.

Within a functional syntactic theory an explanation which links syntax to semantics (i.e. explain syntax in terms of semantics) provides a deeper and therefore better explanation. This is not necessarily the case for generative syntax. Within generative grammar, an explanation which links visible (surface structure) and reconstructed (deep structure) syntactic phenomena to one another without reference to semantics, and which refers to the simplest imaginable procedures that generate unboundedly many meaningful expressions considering the relevant observed phenomena, is the best explanation (e.g. Berwick et al 2011). This is, however, not necessarily the case for functional syntax, which may provide a different explanation for syntactic phenomena that can be explained in terms of semantics or information structure within the synchronic structure, and for syntactic phenomena which have no semantic explanation at all, and which may for example be explained in terms of

diachronic terms and/or by referring to grammar external factors such as general cognitive capacities such as processing effort (see for example Yngve 1960; Gibson 1998 and van Trijp 2014 for an analysis of linguistic complexity and processing efforts).

3. Case studies: generative analyses from a functional perspective

In this section, I will discuss three case studies, to show how data analyzed within generative syntactic theory, can be explained differently from a functional theoretical perspective:

1. Elements used in syntactic description not directly related to linguistic signs
2. Movement (raising) and traces
3. Structure dependency and the Poverty of Stimulus (POS) argument

I discuss these case-studies in quite some detail since I think this is the best way to show the different theoretical approaches to syntax.

3.1 Case Study one: elements used in syntactic description not directly related to linguistic signs

An important feature of generative analyses is that the syntax of particular constructions is analyzed in terms of phrase structures, with a hierarchical structure, which are usually modelled as syntactic trees. These syntactic trees are rather complex, and contain model theoretic information about properties that are not necessarily directly reflected in the linguistic data they describe. This can be illustrated with the following Russian sentence taken from the introduction to Russian syntax by Bailyn (2012):

(2)	Ja	ščitaju	Sašu	vdoxnovljajuščim. (Russian)
	I.NOM	consider.1sg.pr	Sasha.ACC	inspirational.INSTR

'I consider Sasha inspirational.'

The sentence is an example of a construction with three valency slots:

x_{NOM} *ščitaet* y_{ACC} z_{INSTR} = 'after careful consideration x is of the opinion that y has the property z'

In Russian, the first valence x (i.e. the subject), is expressed by a nominative form, the verb is marked for tense and aspect, the second valence y, the object, occurs in the accusative, and the third valence y, occurs in the instrumental case. In Bailyn (2012), who works within the Minimalist Program, this sentence is analyzed as in Figure 1.

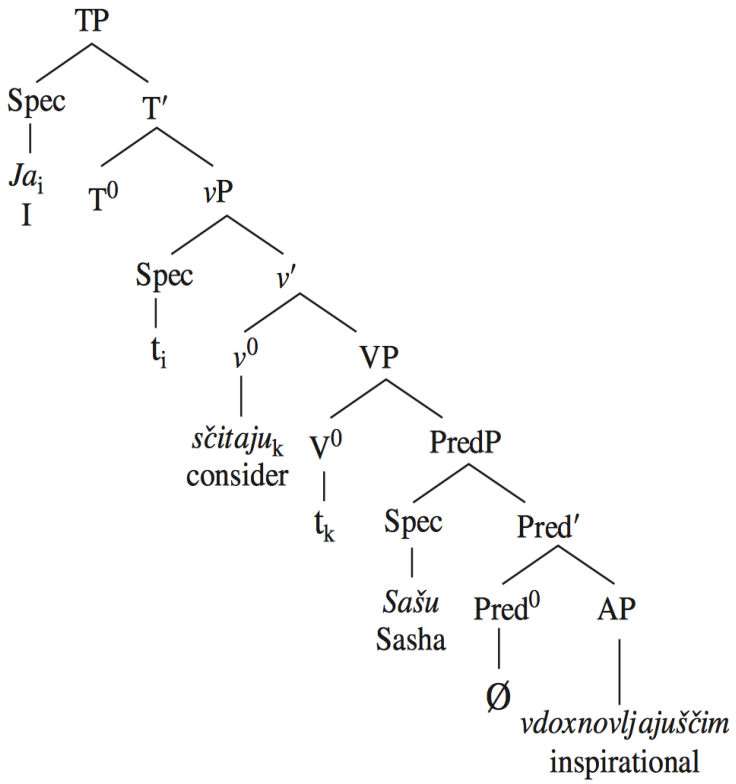

Figure 1: Analysis of example (2) by Bailyn (2012: 110)

Note that we see a clear phrase structure since the sentence is broken up in constantly smaller pieces starting with the whole structure (TP "tense phrase"), which is split into a Spec ("specifier"), in this case filled with the subject *ja* 'I', and a T' (Tense), which is further divided into two categories, etc. What is striking in this analysis is that even though some linguistic elements have a clear function in the constituency tree (for example the nominative form *ja* 'I' occupies the position SPEC), in many cases there are nodes in the tree that do not clearly correlate with linguistic forms (e.g. Pred^{o}: ø; V^{o}: t_k, SPEC: ti; T': T^{o} and *v*P). The reason to postulate these syntactic nodes is not based on the behavior of this specific construction, but on theory internal reasons, which have to do with other constructions from Russian or even other languages. This the result of the fact that grammar is essentially conceived of as system which generates sentences with the simplest imaginable procedures from particular building blocks and the application of particular rules or principles. This can

be illustrated with the tense node (T'). In Russian, verbs are marked for tense. As such, tense is a relevant form-meaning category in Russian. Semantically, tense has to do with the temporal location of the situation (state of affairs) relative to some vantage point, often the moment of utterance. Within the generative framework, however, tense is seen as something which actually occupies a position in a constituency tree, in the same vein as an actual word can occupy a position in a sentence, described in a tree-structure. In this case, the Tense position is located immediately after the SPEC-position filled by the sentential subject. This is in fact based on properties of other Russian constructions. Bailyn (2012: 32) notes for example that tense can be marked on an element other than the primary lexical verb, for example in the case of sentences where the primary lexical verb is an infinitive, and the tense is expressed by the auxiliary verb *budet* 'will':

(3) Ja ne budu prodavat' mašinu. (Russian)
I.NOM not will.AUX sell.INF car.ACC
'I won't sell the car.'

He also notes that ellipsis can eliminate the second appearance of the VP, probably to illustrate that tense is not the same as the lexical verb:

(4) Sonja budet zanimat'sja jogoj, (Russian)
Sonja.NOM will.AUX do.INF yogha.INSTR
i Aleksandra ne budet.
and Alexandra not will.AUX
'Sonya will do yogha but Alexandra won't [do yogha].'

Having established that tense can be expressed by an auxiliary verb, Bailyn argues that this auxiliary is not part of the VP, but outside of the VP, and then further states: 'Thus we have reason to believe that there is a tense position outside of VP (…): in fact, because all complete sentences are tensed, it is generally assumed that the sentence is a projection of the T head, hence a TP. (…) Elements that sit in the Tense position, such as *budet*, carry the feature [T].' (Bailyn 2012: 32). Bailyn provides the following minimal sentence structure in Russian:

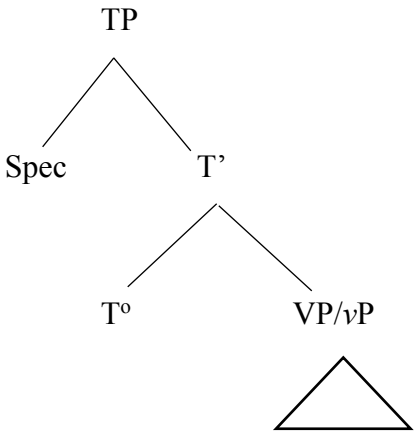

Figure 2: Minimal sentence structure in Russian according to Bailyn (2012: 32)

From a functional perspective, however, the status of the formalization and the analysis is not clear. Why would tense be a position in a phrase structure? And why would the fact that tense can be expressed on the auxiliary *budet* have implications for the analysis of the constituency structure of sentences with just a finite (tensed) verb? More in general, the generative way of analyzing syntax raises many questions, perhaps the most important one being: What is the status of a generative constituency tree? Does it have a specific ontological status? If that is the case, the hierarchical structure used in the generative theory should be clearly reflected in the linguistic data. There is, however, no real evidence for this. Especially because Russian

has a fairly free word order, there is no directly observable proof for the existence of a (strictly) hierarchical syntactic structure. There is even less proof for the existence of "Spec", "T'" or "T°" as actual empirical phenomena, or that language users have tacit knowledge of them. In addition, it is not clear what data would falsify the claim that there is such a thing as a tense position. In my view, all this points to the fact that these notions must be seen as theoretical notions, which have an explanatory status within the theory. One could therefore argue that the analyses provided in generative syntax is supposed to be the most convenient, elegant or successful way to model the linguistic data. But if that is the case, one can question whether the generative theory is in fact successful: why would it be necessary to have such a complicated theory to account for relatively simple data? Even though it is claimed within the generative framework that "nature is the realization of the simplest conceivable mathematical ideas" and that the "minimalist program calls attention to the underlying simplicity of this Universal Grammar" (Boeckx 2006: 8, 150), this simplicity is not part of the actual linguistic analysis. This is the result of the fact that the theory tries to explain all human languages (i.e. linguistic systems) on the basis of a sparse set of assumptions (Chomsky 1996: 225), or universal or innate "building blocks", and has to use these and 'the simplest imaginable procedures for generating unboundedly many meaningful expressions' (Berwick, Pietroski, Yankama, Chomsky 2011). As such, one could speak about a "simplicity-paradox". On the one hand the theory starts out with a "minimal" number of given building blocks, and has a generative character, such that one structure is derived from the other, using these building blocks. On the other hand, this starting point makes the actual linguistic analysis of a particular language extremely complicated. From a functional perspective, the type of analysis that is used in generative syntax raises more questions than it answers. From this perspective, a sentence like (2) does not need to be represented in an overly complicated way, and reference to rules of Russian, of the kind that one can find in a descriptive ("normative")

grammar, suffices to describe and analyze the sentence. Also note that the use of the auxiliary *budet* 'will' and its syntactic behavior can be described and explained independently from sentence (2).

3.2 Case study 2: movement and raising

Another striking property of generative syntactic analyses is the use of notions such as movement, raising, transformation, traces, blocking of movement, etc. Notions like these are used in various stages of the theory (under different labels) to account for the fact that words or groups of words (constituents) that belong together syntactically and semantically, for example in terms of valency or in terms of an element which modifies another, are not placed in each other's vicinity. As such movement has to do with the relation between syntax and meaning, and with the idea that linguistic structures must somehow be derived from other structures where meaning and syntax coincide. An early example of a debated notion of "movement" is "negative transportation" or "negative raising" as introduced by Fillmore (1963). This can be exemplified with (5a), where the negation is placed before the verb *think*, even though semantically the negation pertains not to the action 'think' but to the event expressed in the complement clause. According to the generative analysis (at least of that time), (5a) is therefore derived from (5b) where the negation is in the right place:

(5) a. I don't think it is going to rain.
b. I think it is not going to rain.

The notion of negative transportation was criticized by Bartsch (1973) in her paper *"Negative transportation" gibt es nicht.* According to her, the reason why (5a) normally does not

express ‘I have no opinion about p’ is due to the rule of the “excluded middle inference”: you either think that something is the case or not the case. Whether or not this rule applies depends on the lexical meaning of the verb. To give an example, in the case of the verb *hope*, we do not put the negation before the verb as in (6a), since the rule of the excluded middle inference does not apply; instead, the negation is placed in the complement clause as in (6b):

(6) a. ?I don’t hope it will rain.

b. I hope it won’t rain.

The conclusion drawn by Bartsch is that we do not need rules of transformation, movement, or syntactic derivation, and that the phenomena that have to be explained are of a semantic and pragmatic nature.

The use of various notions of movement have, however, been central to generative grammar, including in the version of the Minimalist Program developed by Chomsky in his 1996 [first edition 1995] book. This can be illustrated with two small case studies taken from Chomsky (1996). In both instances, the style of reasoning by Chomsky is the following: (i) a particular pairing of form and meaning is not acceptable in English, (ii) this is strange because on the basis of the grammar one would not expect this, (iii) the restriction is explained in terms of (innate) syntactic rules, which involve movement (raising, traces, blocking of raising, etc.).

In the first example, Chomsky observes that the copular verb *seem* does not occur with a *that*-clause which expresses the property of the subject as in (7), where the ‘t’ refers to a trace of a moved element that is raised:

(7) *John seems [(that) t is intelligent] (Chomsky 1996: 261)

The absence of this structure is perhaps surprising, since raising does in fact occur in regular sentences with *seem* as in (8):

(8) John seems to be intelligent.

Note that the theoretical concept of "raising" is used for sentences with the verb *seem* in order to account for the fact that the grammatical subject of the verb *seem* is, according to the generative framework, not semantically its subject:

(9) John seems (to be) intelligent.
Semantically: 'It seems to be the case that John is intelligent.'

In order to account for the perceived "mismatch" between the actual form of the sentence and its semantic structure, it is argued that the subject is raised (moved) from the subordinate clause to the main clause. This is therefore the original structure before raising (where 'e' refers to an empty element):

(10) [e seems [John to be intelligent]]

Put differently, *seem* is a so-called "raising-to-subject" verb. Chomsky asks himself why *John* could not raise from the embedded subject position to the matrix subject position in (7). In order to account for this, Chomsky (1996: 261) introduces a rule called "Greed". He argues that the raising is blocked by "Greed", "since at the time of the raising, the Case

feature of *John* is already checked." It is not clear to me what the status of this type of rule is. But since it is not a rule that can actually be learned, or stated in a descriptive grammar, it seems to me that it must somehow be innate, or at least triggered automatically by learnable and language specific rules. As such, the explanation of the data refers to a principle, which presupposes various theory-internal notions such as movement and "checking". In turn, these notions are directly related to the axiomatic theoretical starting points of the generative theory.

Let's have another look at the data from a semantic perspective, and forget about the generative theoretical starting points and theoretical apparatus. First consider copular constructions of the type 'x cop y_{ADJ}' and their meanings:

(11) John is intelligent.
'x has the property y'

(12) John seems intelligent.
'on the basis of preliminary observation one can conclude that x has the property y'

The copular verb *seem* can also occur with an infinitival construction filling the y-slot:

(13) John seems to be a leader.
'on the basis of preliminary observation one can conclude x is a participant in situation y'.

In the case of the infinitive, which expresses a situation type (i.e., a situation abstracted from a specific subject or time), the identity of the first participant associated with the infinitive is

inferred from the (surrounding syntactic) context, and is established by coreference. In the case of these sentences this means that 'John' is interpreted as the agent (subject) of the event expressed by *to be a leader*. The same construction can in fact also be used in the case of properties such as 'nice', in which case we find the infinitive of *be*:

(14) John seems to be intelligent.

The question is now: why do copular verbs such as *seem* or *be* not combine with a *that*-complement clause? For this we also need to look at the semantics and syntax of the *that*-construction:

(15) I know that John is intelligent.

(16) He said that he was ill.

In both cases we find a main-clause with a verb which expresses a particular (mental) activity which presupposes a content, for example:

x say y = 'x utters words conveying information y'

The information (y) is expressed by the accompanying *that*-clause in (15) and (16). Why do copular verbs (*seem*, *be*) not occur with a *that*-clause?:

(17) *John seems that (John) is intelligent.

(18) *John is that (John) is intelligent.

The main answer is: ‘intelligent’ does not express the content of ‘seem’ or ‘is’. As such, they are different from verbs that occur with *that*-clauses. Interestingly, there is an exception. *That* can occur with copular forms with a so-called ‘preliminary subject’, in which *it* refers to a non-further identified situation, which is further identified by the *that*-clause:

(19) It seems that John is intelligent.
Cf. John seems (to be) intelligent.

Note that in this case ‘it’ is not coreferential with the subject of the subordinate clause as in (17) or (18), but with the whole *that*-clause. Because of this, this construction cannot occur with an infinitival clause, where the subject of the infinitive is coreferential with the subject of the main-clause:

(20) *It seems John to be intelligent.

Now, let’s move on to the second example given by Chomsky (1996: 158). He notices that we find structures like in (21), but we do not find the same structure without *to be* as in (22), even though this is in fact possible in the case of (23):

(21) I consider [there to be a solution].
(22) *I consider [there a solution];
(23) I consider John intelligent.

Chomsky states:

> 'In [our example 21] it must be that *be* assigns Case directly to *a solution*; *there* also receives Case (from *consider*), so that the Chain Condition is satisfied after LF raising.' (Chomsky 1996: 158).

Again, we find various theoretical notions in the explanation (raising, Chain Condition, satisfaction, LFG, etc.). And again, it is also possible to look at the data from a functional-semantic perspective, not using these theoretical notions. Consider the meaning of the construction with *consider* in (23):

'x considers y z': 'after careful thought x is of the opinion that y has the property z'.

In this construction, 'z' is either expressed by an adjective, or by an infinitival clause with the verb 'be':[3]

(24) I consider John intelligent.

(25) I consider John to be intelligent.

If we take the definition of this construction into account it makes sense that (22) is ungrammatical: 'there' is not a referential object that can have a property on the par with John in (24). It is not a valency (with case) of *consider*. Instead, sentences with *there* like (21) can better be analyzed as consisting of two valency slots, someone who considers ('I') and the content of the consideration ('there is a solution'):

(26) I consider [there to be a solution].

'x consider y ': y 'there to be a solution'

[3] There are also other versions of this construction, for example as in the following example with a gerund, where consider means something like 'thinking about', 'anticipating': *I consider coming out to my parents*.

In this construction *there* opens up a space for the existence of the (infinitive) situation as a whole, like in the construction *there is a solution*. In the same vein, in the following sentence, 'it' is not the object valence of 'consider', but part of the content of 'consider':

(27) I consider [it (to be) important to be tidy].

'x consider y ': y 'it is important to be tidy'

Such sentences are in fact closely related to instances where the pronoun is the y-valence; cf.:

(28) I consider that (to be) important.

To conclude, if we proceed from the generative axiom that syntax is separate from semantics, and that one has to somehow derive structures from other underlying structures in order to account for the correct relation between syntax and semantics, the approach taken by Chomsky makes sense. But if we leave aside the Chomskyan axioms, and proceed from a functional and semantic perspective, where semantics is at the core of many phenomena that are called purely "syntactic" in generative grammar, one can explain the data (as discussed in this section) in a different way and provide an explanation that does not make use of notions denoting various types of movement, such as "raising", etc.[4] Of course, whether one finds this type of explanation, which focuses primarily on semantics, successful or not, depends on the theoretical framework one works in (as I will discuss in more depth in Section four).[5]

[4] Generative syntax has been incorporating semantics more and more from the nineties on, and semantics plays a part in many syntactic analyses, but semantics is still seen as fundamentally separate from syntax, and generative linguists generally do not actively look for a semantic explanation of presumed syntactic phenomena.
[5] One can of course criticize my specific analysis here, but my main goal is to provide an insight into the *type* of analysis. The same is true for the analysis provided by Chomsky. One can argue that this specific analysis is also incorrect from a generative point of view, but my main point is to analyze the *type* of reasoning.

3.3 Case Study 3: structure dependency and Poverty of Stimulus (POS)

Within generative Chomskyan syntax the phenomenon of auxiliary fronting has played a crucial role in the argumentation in favor of innate syntactic rules (so called "Universal Grammar") and poverty of stimulus, and therefore I will devote relatively much space to this topic. In various writings Chomsky (e.g. 1965, 1968, 1975) has illustrated how a simple linguistic analysis of a sentence with auxiliary fronting leads to the conclusion that some syntactic rules are not learned but must be innate. Chomsky (1975: 32–33) does so by providing an analysis of a so-called polar interrogative with a relative clause (also called "PIRC" in the literature), namely *The man who is tall is in the room*. The specific examples used in the literature have changed over the years, but the in essence rather marginal phenomenon of auxiliary fronting has remained central in the theory of generative (Chomskyan) grammar. In this paper I will focus on a somewhat newer version of the argument which is provided by Berwick, Pietroski, Beracah & Chomsky (2011), Chomsky (2011a, 2011b) and Berwick, Chomsky & Piattelli-Palmarini (2013). They provide a semantic version of the previous argument focusing on the question of what interpretations interrogatives can and cannot have. Berwick *et al.* (2011) use the following example to illustrate the argument of Poverty of Stimulus and a-priori structure-dependent constraints:

(29) Can eagles that fly eat?

According to the generative approach (Berwick *et al.*'s view (2011), Chomsky (2011a, 2011b), syntax is a computational system, which should generate all possible grammatical sentences. Linguistic theory should contain an axiomatic system and a rule system that is as

economical as possible. In the explanation of (29), the syntactic model has to contain rules that somehow link the declarative structure to the interrogative structure.[6] There are two possible ways in which this can be done. In regular English, *can* is derived from a position next to the verb *eat*. Note that "_" indicates the original position of *can*, and the line indicates to which position *can* is moved. The brackets in (30) and (31) indicate the constituent structure of the clause as indicated by Berwick *et al.* (2011: 1214)[7]:

(30) English: [Can [[eagles that fly] [__ eat]]]

|_____|

The meaning is therefore the same as in English 'Can eagles that fly eat'? It would also be possible, however, that the corresponding declarative structure is *Eagles that can fly eat*, in which case *can* is somehow derived from a position in the relative clause (that _ fly), and moved to a position outside of the subject (eagles that _ fly); let's call this English-2:

(31) English-2: [Can [[eagles that __fly] eat]]

|_____|

[6] Berwick et al. (2011) are not very explicit whether one has to assume a derivational system or not. They refer to various terms from different stages of generative grammar such as 'transformations' (1211), 'rules' (p. 1216), 'paired⁄raised' (1217) and an element as 'somehow rooted in' (1214) to relate the interrogative sentence to another structure. Also note that they use symbols that suggest moved elements ("_"). Berwick, Chomsky & Piattelli-Palmarini (2013: 23) are more explicit about this and argue that their model deliberately takes a neutral stance as to how pairings are determined: whether by derivational rules or by relational constraints and lexical redundancy rules. Berwick et al. (2011) also do not explicitly discuss whether the underlying structure has declarative illocution or whether it is 'illocution-less'. Here as well, Berwick, Chomsky & Piattelli-Palmarini (2013: 25) are more explicit and state that "from the earliest works in generative grammar in the 1950s, both declaratives and corresponding interrogatives were assumed, for good reasons, from common underlying forms that yield the basis shared semantic interpretations of the paired constructions." This seems to imply that the underlying form is "illocution-less".

[7] More specifically, we can see that both in English-1 (regular English) and hypothetical English-2 the sentence consists of three chunks: (i) eat, (ii) eagles that fly eat, (iii) can eagles that fly eat.

In that case the meaning of the interrogative structure would be: 'Is it the case that eagles that can fly do eat?'. This is because *can* is derived in some way or another from the relative clause (cf. the term "somehow rooted in the relative clause" used by Berwick *et al.* (2011)) where it belongs to *fly.* As Berwick *et al.* (2011) and Chomsky (2011a, 2011b) argue, one would in fact expect that this is the right interpretation since for the language user it would be computationally easiest to choose the closest verb as the place of origin for *can*. And that is *fly* and not *eat*. But, as Berwick *et al.* (2011: 1216) note, "[s]imply ''counting''—say, until reaching the first occurrence of *can*, and ignoring sentential structure— yields the wrong results." It is not linear distance ("counting rules") but structural distance, which is the governing principle here. As such, the linguistic theory must refer to innate structures such as the division of the sentence into hierarchical constituents (i.e. bracketing) and labels like 'subject noun phrase'. Berwick at al. (2011) also point at other innate structure-dependent constraints, namely island constraints. They mention that:

> 'subject relative clauses act as ''islands'' (see Ross, 1967), barring the pairing of an element inside the relative clause with an element outside it (whether an auxiliary verb, a verb, a do-auxiliary, an adjective, or a wh-word).' (Berwick at al. 2011: 1217)

Put differently, in the case of (31) *can* belongs to the subject relative clause *eagles that _fly*. However, since subject relative clauses function as islands the pairing of *can* to *eagles that fly* as in (31) is barred. In conclusion, Berwick at al. (2011) argue: "[c]ompetent speakers of English know this, upon reflection, raising the question of how speakers come to know that [*Can eagles that fly eat?*] is unambiguous in this way." The answer to this rhetorical question

is that competent speakers of English know this because of their language specific endowment:[8]

> 'This leads us to believe that some 50-plus years after examples like (5a) [*Can eagles that fly eat?*; *author*] were initially offered, the basic points still hold—unchallenged by considerable efforts, in the interim and earlier, to show that real progress lies with the study of how contingent experience shapes cognition. In our view, the way forward begins with the recognition that environmental stimuli radically underdetermine developmental outcomes, and that grammar acquisition is a case in point.' (Berwick at al. 2011: 1239)

It is clear that the linguistic analysis of auxiliary fronting plays a vital role in conclusion that some parts of grammar must be innate. As is remarked by Pullum & Scholz (2002: 39) with respect to auxiliary fronting, "[t]his instance of the APS [the argument from the poverty of the stimulus; *author*] has an extraordinarily high citation index. (…) It is not an exaggeration to say that this instance is being passed around and repeated over and over again; APS enthusiasts never tire of citing it." Chomsky (2011a, 2011b) also uses this example to compare generative Chomskyan linguistics to the Galilean scientific revolution, and it keeps being cited as a key example (see e.g. Smith & Allott 2016).

Whether or not the arguments in favor of innate syntactic rules are valid depends to a large extent on the question whether one shares the general assumptions and starting points of generative syntax. More specifically, the argument put forward by generative Chomskyan grammar hinges very much on the idea that syntax is a computational system with an

[8] From a functional linguistic perspective, one might as well ask the question of how it is possible that humans are capable of learning *language specific* complex syntactic, semantic and pragmatic rules or regularities, even though these rules and regularities are not explicitly taught. To account for this, reference to innate linguistic rules or regularities is of course of no use. Instead, in order to account for this one has to refer to the general capacities of humans for pattern recognition and finding analogies, and the fact that they are tuned into all things which have to do with interaction and communication (see for example Tomasello 2003, 2005).

autonomous ontological status. If we proceed from the point of view that language is primarily a communicative system, then the argument loses most of its power.[9] In such a view, the syntactic structure has to account for the information or meaning that is expressed in the construction, referring to form-meaning elements and semantic-syntactic relations between these form-meaning elements.[10] In the case of *Can eagles that fly eat?* the relevant syntactic information consists of the meanings of the words (*can*, *eagles*, *that*, *fly*, *eat*), and the semantic-syntactic relations between these meanings such as the division of the sentence into a subject ('eagles that fly'), and predicate ('can eat'), and the division of the subject into a head noun indicating 'eagles' and a relative clause *that fly*, indicating a property of 'eagles'. Berwick *et al.* (2011) and Chomsky (2011a, 2011b) argue that children are not able to learn these syntactic relations because from a perspective of simplicity it is easier to relate *can* to the nearest verb. But there is no indication that this is correct. One can imagine very well that children first learn the relevant rule by confrontation with linguistic data to associate fronting of *can* with a 'can'-question:

(32) Eagles can eat.

(33) Can eagles eat?

English children must learn at an early stage that words that belong together semantically are not always adjacent. A clear case are simple interrogatives as in (33), where the infinitive is separated from the auxiliary even though the infinitive can be seen as a valence of the

[9] In such a view one structure does not need to be derived from the other or even linked to the other in terms of phenomena such as "movement", since the linguistic system is not a computational system that needs to be as economical as possible. It should be noted, however, that there are also non-derivational models that do not place much emphasis on the communicative role of language either (e.g. LFG or HPSG).

[10] This is not to say, of course, that the model of grammar used by Berwick et al. (2011) does not deal with meaning at all. In minimalist syntax the assignment of meaning to syntactic structures is regulated by the Conceptual-Intentional Interface, and Berwick et al. are in fact very concerned with the issue of assigning meaning to structures. The main point I want to make is that meaning and syntax are seen as very different modules.

auxiliary, but there are of course much more constructions where this is the case. When a child learns that interrogatives involve an auxiliary followed by its subject and complement as in (33), there is no reason to assume that they will have problems figuring out that the same rule also applies in cases where the subject noun occurs with a relative clause or with another modifying element:

(34) **Flying/big** eagles can eat. > Can **flying/big** eagles eat?

(35) Eagles **that fly** can eat. > Can eagles **that fly** eat? (repeated (29))

Learners know that sentences are not just strings of individual words, and that subjects and complements (for example modifying elements) can consist of phrases. As such, if we proceed from the idea that syntax is also about meaning, there is plenty of evidence for the learner how to proceed while learning the language (cf. the discussion by Pullum and Scholz 2002: 40-45).[11]

The example that Berwick *et al.* (2011) provide – *Can eagles that fly eat?* – is also rather idiosyncratic because it involves a sentence with very little inflection or morphological clues which indicate what word syntactically belongs to another word.[12] As such, Berwick *et al.*'s example is less helpful than many because both verbs could be either a non-finite form or a present tense form. This is not the case with an aspectual auxiliary or a singular subject:

(36) a. Are eagles that fly eating? ('are' cannot belong to 'fly')

[11] Berwick et al. (2011) are aware of this paper, since they mention in footnote 1: "On alternatives to the classical POS argument and its extensions, see Lewis and Elman (2001) and Pullum and Scholz (2002)." They do, however, not actually evaluate the arguments and data presented here.

[12] The verb *can* is not inflected for person and the infinitive (*eat*) does not occur with *to* (because of the rule that in English infinitive complements of modal auxiliaries do not occur with *to*). The third person plural verb (*fly*) shows no inflection (according to the rule that in the present tense only the third person singular shows agreement with the subject).

b. Have eagles that fly eaten? (‘have’ cannot belong to ‘fly’)

c. Can an eagle that flies eat? (‘can’ cannot belong to ‘fly’)

In general, there are several morphological clues in English that interrogatives involve an auxiliary followed by its subject and complement. By focusing on a very atypical example, Berwick *et al.* neglect the ample evidence there is in English for learners how to construe and interpret an interrogative sentence with auxiliary fronting (see also Tomasello 2005).

Even if we set aside that meaning and the function of morphological clues is fully neglected in Berwick *et al.*’s approach to syntax, the assumption they put forward about “counting rules” is still far from evident. This is shown by Bod (2009) who argues that auxiliary fronting can be accounted for in terms of an innate structure learning algorithm that is based on the idea of simplicity rather than on innate syntactic rules or language learning. The basic idea presented by Bod is that linguistic phenomena such as sentences with auxiliary fronting can be learned from simpler sentences by means of structural analogy, and that the rule of auxiliary fronting is an emergent property of a structure learning algorithm. The basic idea can be explained as follows. Take the following declarative sentence with an auxiliary:

(37) The man is hungry.

This sentence can be turned into an interrogative by fronting *is*:

(38) Is the man hungry?

The language learner might learn the rule that in order to construe an interrogative the first auxiliary has be fronted, but this is not the case in declarative sentences with a relative clause containing *is* such as (39), where this rule would result in incorrect (40); instead (41) is correct:

(39) The man who is eating is hungry.

(40) Is the man who is eating hungry? (cf. (30)) (second is *is* fronted)

(41) *Is the man who eating is hungry? (cf. (31)) (first *is* is fronted)

Bod argues that one can correctly predict that (40) should be derived instead of (41). For this, the model needs input sentences on which the analogy can be based. The necessary sentences are (39) (a declarative sentence containing a relative clause), and a sentence without relative clause and auxiliary fronting such the following:

(42) Is the boy hungry?

Note that these two sentences do not contain an example of auxiliary fronting with a relative clause. Within the model, these structures should then be assigned the following constituency structure:

(43) [[[The man [who is eating]] [is hungry]]].

(44) [Is [[the boy] hungry]]?

The simplest way to derive (40) is simply by substituting the simple NP *the boy* as in (44) for the complex NP *the man who is eating* as in (43), leading to the correct fronting. This differs

from sentence (41), which is much more complex to derive because a larger number of subtrees or subconstructions is necessary that need to be combined in order to arrive at (41).[13] In fact, in order to derive this sentence one needs the following four subtrees or subconstructions:

- [is [X X]];
- [[the man] [who X]]
- [eating];
- [is hungry]

The actual construction would then have to derived by:

- substituting X in [[the man] [who X]] for *eating*;
- substituting the first X in [is [X X] for *the man who eating*
- substituting the X in [is [the man who eating X]] for *is hungry*

In the same vein (31) is just more complex to relate to the existing structures than (30).[14] The analysis provided by Bod suggests that it is possible to explain language production and

[13] In his model (U)-DOP, the smallest subtrees correspond to the traditional notion of phrase-structure rule, while the largest subtrees correspond to full phrase-structure trees. But DOP also takes into account the middle ground between these two extremes, which consists of all intermediate subtrees that are larger than phrase-structure rules and smaller than full sentence-structures (Bod 2009: 754).

[14] In the example given here the constituency structure is assigned to the sentences by Bod himself, but Bod (2009: 779-786) also tests his model with respect to auxiliary fronting on actual data, namely in the Eve part of the Childes corpus. In this cas,e U-DOP computes the constituency structure (tree) of a sentence based on probability (see 2009: 761-763). The results are somewhat mixed in that U-DOP still generates relatively many ungrammatical sentences, even though the grammatical sentences are the most frequent. In my view, this may have to do with the fact that Bod does not take semantics into account in his model. In actual reality even the division into a sentence into a NP and a VP is in the end also a matter of semantics (namely the division into subject and predicate), and the same is true for the idea of a complex NP such as the man who is eating, where 'who is eating' modifies 'the man'. Furthermore, it may be that Bod contributes too much importance to the idea of simplicity. In actual language, simplicity may not be the sole driving force, and languages may exhibit several irregular features, which may sometimes be explained from a diachronic perspective.

understanding without reference to innate syntactic rules, but instead with reference to an innate tendency for simplicity. Furthermore, Bod also points at the relation between language learning and the suitability of the language structure. Within a theory of grammar that proceeds from the communicative function of language one can also ask the question of *why* the system is as it is. Why is it the case that according to the rules of grammar in *Can eagles that fly eat*, *can* belongs to the verb in the main clause? Why are the rules of grammar not such that *can* belongs to the relative clause? One can approach this question as follows from a cultural-evolutionary (dynamic) perspective. Language needs a form to signal interrogative meaning, in other words, it needs a linguistic sign for that. In the hypothetical language Berwick *et al.* (2011) refer to, fronting of *can* is not always associated with signalling a *can*-question, which makes it from the perspective of economy of the communicative system very complex. In Berwick *et al.*'s non-existing version of English, one signals questions by fronting the modal verb in the subordinate clause. Fronting of *can* to signal a "*can*-question" therefore depends on the presence or absence of a subordinate clause:

(45) Eagles can eat. → Can eagles eat? (meaning: 'Do eagles have the capacity to eat?')

(46) Eagles that can fly eat. → Can eagles that fly eat? (meaning: 'Is it the case that eagles that can fly eat?')

Since fronting (inversion) of *can* is *not* available for sentences consisting of a main-clause and a subordinate clause, another configuration must be chosen to convey the message 'Can eagles that fly eat?', or 'Do eagles that fly have the capacity to eat?' As such, in Berwick *et al.*'s hypothetical English we find a complex system only to facilitate a communicatively peripheral function, namely to mark the interrogative meaning of sentences with *can* (or other

modal auxiliaries) in a subordinate clause. It is not very likely that such a system would ever emerge in the history of language.

The argument of structure dependency and poverty of stimulus (POS) is already from the sixties, but still used in generative grammar to argue for the presence of UG, and turns up in different shapes now and again. A different version can be found in Adger (2013). In his paper, in which he responds to a paper by Adele Goldberg, he asks why a *wh*-element can question the positions Z and X in English, but not position Y (note that in this example, I have indicated the different valences directly with Z, X and Y):[15]

(47) a. Z teased X before she devoured Y.

b. [Which cat]$_Z$ teased [the mouse]$_X$ before she devoured [her food]$_Y$?

c. What$_X$ did [the cat]$_Z$ tease before she devoured [her food]$_Y$?

d. What$_Y$ did [the cat]$_Z$ tease [the mouse]$_X$ before she devoured?

Intended meaning: 'The cat teased the mouse before she devoured what?'

The data point at a relevant difference between the main-clause (with valence elements Z and X), and the subordinate clause (with valence element Y). In this example from English, it is not possible to have a long-distance dependency between a *wh*-element in the main-clause, and the verb in the subordinate clause of which it is a valence. This explains why in the acceptable sentences (47) we find pieces of sentence that are all grammatical, but not in ungrammatical ((47)d):

[15] Note that Botha (2016: 178) remarks that "It is interesting in this regard that in a lengthy response to Adger's (2013) criticisms of Construction Grammar, Goldberg (2013b) refrains from addressing his point that this approach is unable to provide a theory of how simple syntactic rules such as wh-Question Formation are learned." Clearly Goldberg and Botha do not speak the same scientific 'language'.

(48) a. Which cat teased the mouse / she devoured her food.

b. What did the cat tease / she devoured her food.

c. *What did the cat tease the mouse.

In (47d) (*What did* [*the cat*]$_Z$ *tease* [*the mouse*]$_X$) the language user will first interpret *what* as the object of tease, only to find that the expression of *the mouse* blocks the interpretation. These data show that the main clause functions as a separate semantic-syntactic unit from the temporal subordinate clause, which is introduced by the bivalent conjunction *before*, which expresses that the situation as depicted by the main clause occurs prior to the situation in the subordinate clause. As such, the phenomenon of the syntactic "independent" status of main-clauses and subordinate clauses, is a given fact of English and other languages. It is known from the literature, however, that in some contexts long-distance dependencies crossing the "boundary" of the two clauses are in fact possible, for example in the case of *that*-clauses:

(49) I believe that she devoured X.

(50) What$_X$ did I believe that she devoured?

More generally, long distance dependencies with main-clause/subordinate clause are possible if triggered by special instances of topicalization/focalization and in limited contexts (see for example Fortuin and Davids 2013 for data from Russian, and Reinöhl 2020, for a typological study). Put differently, if there is a functional or communicative trigger for a construction to exist, it will exist (even though there are language specific possibilities and restrictions). From a functional perspective a question could therefore be: would a hypothetical linguistic structure as in (47d) be effective to express two situations with a temporal relation ('before') and a what-question:

(51) The cat teased the mouse before she devoured what?

The answer is probably not. First of all, as I indicated, the structure in (47d) provides the language processor with parsing problems, such that if it were used, it would be an instance of a garden-path sentence. Second, there are other more efficient ways of conveying the intended message, for example by using a conjunction of posteriority *after*:

(52) What did the cat devour after she teased the mouse?

Also note that the sentence (51) without long-distance dependency, which corresponds to (47d) with long-distance dependency, is already rather communicatively peripheral. This sentence is only used to quote or repeat someone's words when the speaker wants to make sure that (s)he heard the identity of the object of 'devour' correctly, or when it was not clear what the identity was. As such, the limitation on particular constructions, or restrictive rules does not *necessarily* point at innate *linguistic* capacities and restrictions, or even innate *cognitive* capacities and restrictions, but may also very well point at *communicative* restrictions.[16]Adger (2013) states, however, that he finds the Construction Grammar approach to phenomena like the one given in (47) of no help:

> 'The CxG [Construction Grammar] response is to say that no such theory is needed beyond a language specific statement of the regularities that obtain between form and meaning. No theory relating form

[16]Adger (2013) argues that sentences like (47 b, c) are very infrequent in the linguistic input, which for him is extra evidence for POS. One can, however, also argue that if language users try to make sense of constructions incrementally (cf. Gibson 1998, who also looks at other processing factors), and if they consider main clauses and temporal subordinate clauses as separate conceptual units, it does in fact make sense that they will produce and understand (47b,c) and not (47d). As such we can also understand the data without reference to POS and innate syntactic rules.

> and meaning is required beyond a taxonomy of form-meaning pairs which is organized so as to capture the pattern. How that taxonomy is learned is never made explicit. The fact that, in language after language, the same descriptive regularities arise has nothing to do with the linguistic capacity of human beings (because there is none) but is a result of accidents of history, language independent processing strategies, or the exigencies of communicative function. I have yet to see any half-way successful analysis of this kind of pattern which does not assume the existence of G [i.e., innate cognition only used for language, a term used by Adger 2013] (see, for example, the Sign Based CxG approach to wh-dependencies in Sag (2012)). A usage based variant of CxG of the sort Goldberg proposes is of no help in understanding these fundamental and general properties of human linguistic cognition.' (Adger 2013: 475)

Adger does not make explicit in this fragment what kind of functional explanation, which does not assume the existence of what he calls "G", he would find convincing, and exactly why he finds the explanations formulated within (Goldbergian) Construction Grammar unconvincing. But his discussion suggests that he finds the functional explanations known to him inadequate for two reasons. First, they are not unified, and refer to various different factors (accidents of history, language-independent processing strategies, exigencies of the communicative function of language). Second, in his opinion, it not made explicit how phenomena like the one discussed in (47) could ever be learned at all, suggesting that any successful theory would have to resort to innate cognition used only for language. As such, this fragment shows that even when functional and generative theories try to discuss how each theory deals with the same data, they are not able to convince one another even in the smallest possible way.[17] In the next section, I will say more about that.

[17] In the course of time, there have been various authors from functional frameworks that have argued against the idea of UG and POS (for a recent discussion and criticism on UG see for example Dąbrowska 2015), and discussions between proponents of competing theories such as the one by Goldberg and Adger referred to here. Even though various linguists have switched sides, as far as I know, criticism has never led to a real and fundamental change in the "theoretical core" of the theory.

4. Which theory is better?

I have discussed a number of case studies to illustrate how generative analyses can be evaluated from a functional theoretical perspective, and how the same data can be explained differently. I have chosen the three case studies exactly because they are very typical of the type of analysis, argumentation and explanation that can be found in generative linguistics.

From a functional perspective, generative explanations, as illustrated by the case studies, are not unified or deep since they leave out important dimensions (semantics, pragmatics) in the explanation of syntax. The omission of semantics/pragmatics may lead to the postulation of model-specific notions (movement, traces, blocking, a Tense position, etc.) which are inelegant and difficult to falsify. Furthermore, generative explanations do not differentiate between actual norms/rules of a specific language, and an explanation of the norms/rules of language. In the same vein, they make no principled distinction between the analysis of a specific language (structure), and the analysis of general properties of language.

Of course, it is also possible to evaluate functional analyses of syntactic phenomena from a generative perspective. From such a perspective, one might argue, functional explanations are not unified since syntax is explained in different ways – some properties are semantic, some cognitive, and some diachronic, etc. Furthermore, functional analyses use semantic notions which are fuzzy (not precise) and difficult to formalize. They also, one might argue, do not inspire the linguist to look for the boundaries of the syntactic system. So even though typologists or general linguists have tried to explain syntactic patterns that can be found in all or various languages with reference to principles such as economy, etc., it were generative

linguists who found empirical patterns that are described within the theory by notions such as binding, island-constraints, etc.

Which approach one prefers depends on the theoretical starting points (axioms, "hard core") of the theory used by the linguist. These, in turn, are often defended with reference to the data they are meant explain within the theory, which may lead to a particular circular style of reasoning.

One could of course pose the question of which theory is better? To some extent the theories are difficult to compare because they deal with different "why-questions", but in some cases, the different theories deal with exactly the same data and phenomena; an example was the issue of syntactic "islands" discussed in the previous section. On a deep theoretical level, however, the a priori theoretical starting points (the so-called "hard core" of the theory) are so different that there is no meaningful way in which linguists from different theories can interact. Put differently, the theories are *incommensurable* in the sense of Kuhn. In the same vein, the two theories (generative versus functional) have different technical scientific "languages", with their own theoretical concepts, which makes it impossible to engage in a meaningful discussion about syntax. To put it differently, the theoretical concepts used in generative syntactic theory cannot be translated in a meaningful manner into a functional syntactic theory. As such, the perceived empirical data that need to be explained are strikingly different within each theory. This can be illustrated very well with the following fragment from a discussion on the Internet about different theories in syntax, in which a generative syntactician responds to a remark by linguist Martin Haspelmath. I provide this fragment since it is typical of the way generative linguists evaluate their own scientific enterprise:

> 'We study the patterns as a window into cognitive architecture and computational capacity, not simply to look for patterns of patterns. This is why universal claims like "all syntactic dependencies involve c-command" or "movement dependencies are cyclic" have real content – they reveal something about the basic computations that underlie the human capacity for language. And they make quite clear predictions for particular languages. If you want to know whether a given dependency is syntactic, check to see if the two elements necessarily stand in a c-command relation. If they don't, then either (i) it's not a syntactic dependency, (ii) the premise was false or (iii) there is some other language particular property that is masking the c-command relation. All three of those possibilities suggest avenues for further research and can lead to further insight about the computational system that underlies (the syntactic component of) human language.'[18]

The thing I would like to point out here is that a notion like "c-command" is a theory internal notion, and not a theory neutral empirical fact. Because of this, propositions like "all syntactic dependencies involve c-command" or "movement dependencies are cyclic" do *not* have real content outside of the theory in which they are used.

In his later work, Kuhn also made a difference between "taxonomic incommensurability" and "methodological incommensurability" (see for example Oberheim and Hoyningen 2018). The term "taxonomical incommensurability" is used to indicate that two competing theories do not share the same scientific language or "lexical taxonomy". The term "methodological incommensurability" is used to indicate that there are no objective criteria to determine what is a good theory. Depending on the dominant paradigm, different criteria are used (such as scope, simplicity, fruitfulness, accuracy). In my view, one can apply the concept of "taxonomic incommensurability" to the competing theories of functional linguistics and Chomskyan (generative) linguistics. Because of the different starting points of the two

[18] https://dlc.hypotheses.org/2282 (march 31 2020)

theories or approaches, both theories use different theoretical concepts and speak a different scientific language. This explains why there is little communication between the competing theories. At the same time, of course, one can also speak of "methodological incommensurability" since the ideas about how one can measure a successful theory also differ considerably, in many cases exactly because of the different starting points of the theory. This can be illustrated with the claim by Chomsky given earlier that "in pursuing a minimalist program, we want to make sure that we are not inadvertently sneaking in improper concepts, entities, relations, and conventions. (…) The more sparse the assumptions the more intricate the argument is likely to be." (Chomsky 1996: 225) In other words, because of the starting point that the concepts, entities, relations, and conventions of theory have to be economical or "minimalistic", but still account for all language structures, the theory itself becomes highly complex. This complexity is valued as something positive. This differs, of course, from how the complexity of the generative theory is evaluated from a functional perspective. From such a perspective, there is no need to start out with the idea that the concepts, entities, relations, and conventions of the theory should be minimal in the same sense, and that one should use the same abstract syntactic rules to account for all languages. This explains why functional theories evaluate the complexity of generative theory as something negative. In the same vein, functionalist theories of syntax which do not presuppose universal grammar, are judged to be inadequate according to Adger discussed above, since they are not unified. But again, it is an a priori starting point of generative syntax that it is possible to provide one (abstract) grammar for language as such, whereas according to functional theories of syntax it is an a priori starting point of the theory that this is not possible, and that one should differentiate between the type of rules of individual grammars and factors that may play a part in language as such. This difference between generative and functional grammars can also be illustrated with the following fragment from

a response by D' Alesssandro to a paper by Martin Haspelmath. In this paper she defends the idea of laws as formulated within generative linguistics, which apply to all languages. In her view, exactly because of its unifying character, generative linguistics is superior to theories that only try to explain facts from individual languages:

> 'A formal theory is built through the convergence of a number of proven hypotheses. These hypotheses apply at different levels of the analysis: we observe an empirical phenomenon and we describe it in abstract-structural terms. We hypothesize that this phenomenon is related to another phenomenon. In other words, we reduce 2 phenomena to a more general phenomenon, and then we proceed by reducing these issues as much as we can, until we hit the general "law". In this sense, we can make predictions all along: we know that if one phenomenon exists, the corresponding one will also exist, we abstract away from the single element and try to find the general law governing these phenomena. (…) These generalizations are always based on data and checked on data, but they also bring about implications and predictions. This is what a theory is: not just using technical metalanguage, but the search of the general laws governing language based on converging proven hypotheses.' (D' Alessandro 2021)

One might argue, however, that in the case of explaining surprising language-specific properties with reference to universal rules, something is explained with reference to phenomena that are just as unlikely and unexplained, or in fact even more unlikely and unexplained. As such, unification does not necessarily amount to a good explanation (see also Gijsbers 2011 for unification in explanation in general).

One could also ask whether it is not possible to devise a crucial experiment in the sense of Popper, which will determine for each hypothesis which theory is correct. However, as is well known from philosophy of science, such a crucial experiment is difficult if not impossible to conduct. According to the famous Quine-Duhem thesis (based to a large extent

on Quine 1951) one cannot isolate one hypothesis, because it is embedded within a bundle of interconnected hypotheses. Therefore, each single "falsification" is likely to be rejected by an adherent of the theory, and in order to save the theory new *ad hoc* auxiliary hypotheses will be introduced. Furthermore, prediction in generative syntax is not based on "laws" that can be tested by crucial experiments since the laws themselves are often framed in generative theoretical terms. Because of this, testing the hypothesis mentioned in the fragment above that "all syntactic dependencies involve c-command" is useless as a way to settle which theory (generative or functional) is correct, not only because functional theories do not accept the existence of this notion, but more importantly because the notion of c-command is related to various other theoretical notions within the generative theory, which forms a web of interrelated hypotheses. In the same vein, all the examples of "Significant mid-level results of generative linguistics" drawn up by Peter Svenonius after the Athens 2015 meeting, [19] or the list drawn up by D'Alessandro (2019) about the achievements of generative syntax, are useless as a means to settle a dispute between generative and functional syntax, since most of these results are purely theory-internal claims that have no meaning outside of the generative theory, even if they are extremely meaningful within generative syntactic theory.

This can also be illustrated by D' Alesssandro (2021). She emphasizes the unifying character of linguistic laws with predicable power as formulated within generative linguistics. A scientific law describes natural phenomena which always apply to a physical system under repeated conditions. According to some philosophers of science, such laws also imply a causal relationship involving the elements of the system (see for example Woodward 2003, and Gijsbers 2011 for a discussion). The type of laws D' Alessandro is talking about deal with correlations between seemingly unrelated phenomena in one language, which can be found, in one form or another, crosslinguistically. D'Alessandro llustrates her point of

[19] See: https://blogg.uit.no/psv000/page/3/

prediction by discussing by, what she calls, surprising similarities between English and Chinese, which, again, involve island constraints: [20]

> '[W]h-movement (…) exists in languages, and therefore it is a possible grammatical phenomenon in language. Therefore, since it is possible, it can be found in languages other than English because we are all human, and we all make use of the same restricted set of grammatical tools. Variation is not unlimited. This means that we can use what we see in English to search for the same phenomenon (and the correlations it has, see the reductionism above) in other languages. In fact, thinking about movement in this sense turns out to be very valuable when we comparing English (which evidently has it [wh-movement]) and Chinese (which seems not to) because closer consideration of the predictions made for overtly wh-moving languages with what we see in Chinese reveals that this language also, surprisingly, exhibits a subset of the English movement-restrictions, like the fact that adverbial wh-phrases in Chinese may not take scope across islands (see Huang 1982, Xu 1990, Lin 1992, Aoun and Li 1993, Cheng 1997, Cole and Hermon 1998 and many others).'

In the case of explanation in science based on laws with a predictive power (see for example Hempel's 1965 Deductive-Nomological model of explanation) an explanation is an argument in which the thing to be explained (the explanandum) can be subsumed under one or more general laws of science. In this case, however, it is not clear, at least not in any theory-neutral sense, what the explanandum is. The phenomena that are mentioned here are all framed within the theoretical framework of generative linguistics (involving notions such as "wh-movement", "wh-moving languages", "scope across islands", etc.). The element of surprise D'Alessandro is talking about in this fragment is a purely theory-internal surprise, which has

[20] D'Alessandro also gives an example of a law in the sense of UG, namely the *ABA constraint on adjectival suppletion discovered by Bobalijk. This law is, however, of a different kind. It describes that a particular pattern is not found in any known natural language. Of course, the absence of such patterns needs an explanation, no matter in what theory one works.

to do with theory-internal predictions.[21]As such, the surprising facts that are presented in generative analyses are not likely to convince linguists working in competing theorerical frameworks of the existence of universal grammar, theory-internal constructs like "wh-moving languages", or the relevance of English data to understand Chinese data as referred to by D'Alessandro (2021). This clearly shows that, in order to engage in a meaningful discussion across theories, the *a priori* starting points or axioms of the theory have to be discussed first. But even if we do this, we become part of a circular discussion I referred to above. This can be illustrated with the example of case study three, which evolved around structure dependency and the argument of poverty of stimulus. If one engages in a discussion about the existence of UG, which is one of the starting points of generative grammar, one also has to discuss the various phenomena such as the one listed in section three that are presented as evidence for the existence of UG. But as soon as one provides alternative explanations for the same phenomena, they are considered not even "half-way successful", because they are not framed in familiar generative syntactic theoretical concepts.

Perhaps, then, it is possible to show objectively which theory is more successful? Lakatos (1970) provides some ways to measure successfulness in science. According to him, a research program (a term which was later borrowed by Chomsky for his Minimalist Program) is *degenerative* when it only changes auxiliary hypotheses to deal with troublesome evidence, and *progressive* if it achieves greater explanatory/predictive power. In the same vein, one could try to determine which theory (generative or functional) is more progressive than the other. This criterion is, however, not applicable very well to syntax: what a functionalist finds a good answer to a why question (or troublesome evidence), is not in accordance with what

[21] Every explanation is always about surprising facts, since, as is shown by Van Fraassen (1980), an explanation always presupposes a *contrast-class*, a set of alternatives. As such, in this case, the contrast is why Chinese behaves like English does, and not, as one might expect on the basis of the theory, differently from English.

generative linguists consider a good answer (and vice versa). The field of syntax is not comparable to the empirical sciences (e.g., physics), where there is agreement that Einstein provides a better theory than Newton does, since it provides a more unified explanation and predicts more facts. More generally, "laws of science" are not identical to "generalizations of syntax", even if generative syntacticians perhaps like to compare their theory to scientific theories, and argue, like in the fragment above from an internet discussion, that their theories "make quite clear predictions for particular languages". And, as I mentioned above, even if a generative theory predicts a particular "fact", or a functional theory predicts a particular "fact", these "facts" are often only "facts" within the theory, and not theory neutral empirical observations that everyone agrees on.

One could, of course, hypothetically speaking, imagine that generative syntax would predict much more theory neutral empirically observable facts than functional linguistics (or vice versa). Or alternatively, one could imagine that generative linguistics would produce much more successful technology, for example translation machines, than functional linguistics (or vice versa). But in actual reality, this is not the case. This explains why syntacticians are not able to settle the dispute which theory (functional or generative) is superior, and, with a few exceptions, keep on working within separated competing theories.

5. Conclusion

In this paper I have given an overview of how within the field of syntax, the same phenomena are analyzed and explained differently. I have done so by discussing a number of case-studies, which are typical for the generative theory, from a functional theoretical perspective. In each case I have argued that the same data can be explained differently from a

functional theoretical perspective. In many cases, the question what the data are that need to be explained, also differs greatly. Because of the very different a priori theoretical starting points ("hard core") of the theories, the idea of what is an explanation differs considerably between generative and functionalist linguists. As such, in my view, there is *incommensurability* on the theoretical level in the sense of Kuhn between functional and generative syntactic theories, and it is not possible to design a crucial experiment in the sense of Popper to determine which theory or hypothesis of a particular phenomenon is correct. As such, the conclusion of this paper is rather negative.

To offer some hope, I think that generativist and functionalist linguists can find each other in a number of ways. To name a few:

- By discussing interesting data (patterns) that everyone agrees on;
- By attempting to describe such data in an as theory-neutral way as possible;
- By acknowledging how theoretical frameworks point at interesting empirical phenomena. To give an example, it is exactly because of the idea of UG that generative linguists have tried to look for the boundaries of syntactic rules. The phenomena that have been found and described are interesting and important and need an explanation, irrespective of the theory one works in;
- By agreeing with one another that there are different points of view, and be explicit about this. Syntax is not synonymous with generative or functional syntax;
- By acknowledging that theoretical notions such "c-command", "raising" (to name a few from generative grammar), are theory internal notions, which have an important explanatory status within the theory. They are, however, not theory-neutral empirical or ontological facts.

Finally, it is important to keep on discussing important issues such as "What is an explanation?", or "What is a good explanation?". By doing so, we can at least keep informed about each other's theories and theoretical starting points.

References

Adger, David. 2013. Constructions and Grammatical Explanation: Comments on Goldberg. *Mind & Language* 28(4), 466–478.

Ágel, V., & Fischer, K. 2015. Dependency Grammar and Valency Theory. In B. Hein, H. Narrog (Eds), *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press.

Bailyn, J. 2012. *The syntax of Russian.* Cambridge: Cambridge University Press.

Bartsch, R. 1987. *Norms of Language: Theoretical and Practical Aspects*. Addison-Wesley Longman Limited.

Bartsch, R. 1973. "Negative Transportation" gibt es nicht. *Linguistische Berichte* 27, 1-7.

Beedham, Christopher. 2011. A critique of the model-building approach. *Word* 52(1), 79-94.

Berwick, R. C., Chomsky, N., & Piattelli-Palmarini, M. 2013. Poverty of the Stimulus Stands: Why Recent Challenges Fail. In M. Piattelli-Palmarini & R. C. Berwick (Eds.), *Rich Languages from Poor Inputs* (19-42). Oxford, UK: Oxford University Press.

Berwick, R., P. Pietroski, B. Yankama, N. Chomsky & M. Piattelli-Palmarini. 2011. Poverty of the Stimulus Revisited. *Cognitive Science* 35, 1207–1242.

Boeckx, C. 2006. *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford: OUP.

Bod, R. 2009. From Exemplar to Grammar: A Probabilistic Analogy-Based Model of

Language Learning. *Cognitive Science* 33, 752–793.

Botha, R. 2016. Language Evolution: The Windows Approach. Cambridge: Cambridge University Press.

Bresnan, J. 2001. *Lexical-Functional Syntax*. Blackwell.

Chomsky, N. 2004. Beyond explanatory adequacy. In A. Belletti (ed.), *Structures and Beyond*. Oxford: Oxford University Press, 104–31

Chomsky, Noam. 1996. *The Minimalist Program*. Cambridge, MA: The MIT Press. [second print]

Chomsky, N. 1968. *Language and Mind. New York*: Harcourt, Brace and World.

Chomsky, N. 1975. *Reflections on language*. New York: Pantheon books.

Culicover, P. W. and R. Jackendoff. 2005. *Simpler Syntax*. Oxford (etc.): Oxford University Press.

Dąbrowska, Ewa. 2015. What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology* 6, 852, https://doi.org/10.3389/fpsyg.2015.00852

D'Alessandro, Roberta. 2021. Not everything is a theory. A reply to Martin Haspelmath. To appear in *Theoretical Linguistics*. Published online (lingbuzz/005659).

D'Alessandro, Roberta. 2019. The achievements of Generative Grammar: a time chart and some reflections. *Catalan Journal of Linguistics, Special Issue*, 7-26.

Evans N, Levinson SC. 2009. The myth of language universals: language diversity and its importance for cognitive science. *Behav Brain Sci*. 32(5):429-494.

Fillmore, Ch. 1963. The position of embedding transformations in grammar. *Word* 19, 208-231.

Fortuin, Egbert & Hetty Geerdink-Verkoren. 2019. *Universal Semantic Syntax*. Cambridge: Cambridge University Press.

Fortuin, E., Davids, I. 2013. Subordinate clause prolepsis in Russian. *Russian Linguistics* 37,

125–155

Fraassen, van, Bas. 1980. *The Scientific Image*. Oxford University Press.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.

Gibson, Edward. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition* 68, 1–76.

Gijsbers, Victor. 2011. *Explanation and determination.* Dissertation Philosophy Department, Faculty of the Humanities, Leiden University

Haspelmath, Martin. 2020. Human Linguisticality and the Building Blocks of Languages. Front. Psychol. (https://www.frontiersin.org/articles/10.3389/fpsyg.2019.03056/full)

Hempel, Carl G. 1965. *Aspects of Scientific Explanation*. Free Press.

Human linguisticality and the building blocks of languages. Frontiers in Psychology 10(3056). 1-10

Hengeveld, Kees & Lachlan Mackenzie. 2008. *Functional Discourse Grammar: A Typologically-Based Theory of Language Structure*. Oxford: Oxford University Press

Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lakatos, Imre. 1970. Falsification and the methodology of scientific research programmes. In: Lakatos, Musgrave eds, pp. 91–195.

Langacker, R. W. 1987. *Foundations of Cognitive Grammar* Vol. 1: Theoretical Prerequisites. Stanford: Stanford University Press.

Langacker, R. W. 1990. *Concept, Image, and Symbol. The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter.

Langacker, R. 1991. *Foundations of Cognitive Grammar* Vol. 2: Descriptive Application. Stanford: Stanford University Press.

Oberheim, E. & Hoyningen-Huene, P. The Incommensurability of Scientific Theories. *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.) (retrieved online December 16 2020: https://plato.stanford.edu/archives/fall2018/entries/incommensurability/)

Pollard, C. I. A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Pullum, G. & Scholz, B. 2002. Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19, 9–50.

Quine, W. V. O. 1951. Two Dogmas of Empiricism. *The Philosophical Review*. 60 (1): 20–43.

Reinöhl, U. 2020. Continuous and discontinuous nominal expressions in flexible (or "free") word order languages – Patterns and correlates. *Linguistic Typology* (published online ahead of print)

Rizzi, L. 2017. The Concept of Explanatory Adequacy. In: Roberts, I. (Ed.), *The Oxford Handbook of Universal Grammar*. New York: Oxford University Press.

Thompson, S. 1991. On addressing functional explanation in linguistics. *Language & Communication* 11 (1/2), 93-96

Tomasello Michael. 2005. Beyond formalities: the case of language acquisition. *Linguistic Review* 22, 183–197.

Tomasello, Michael. 2003. *Constructing a language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press

Sapir. 1921. Language: An Introduction to the Study of Speech. Language. New York: Harcourt Brace, 1921

Skow, Bradford. 2016. Scientific Explanation. In Paul Humphreys (ed.), *The Oxford Handbook of Philosophy of Science,* 524–543. Oxford: Oxford University Press.

Van Trijp, Remi. 2014. Long-Distance Dependencies without Filler-Gaps: A Cognitive-Functional Approach in Fluid Construction Grammar. *Language and Cognition*, 6(2), 242–270.

Woodward, James. 2003. *Making Things Happen*, Oxford University Press.

Yngve, Victor. 1960. A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society* 104(5), 444–466.