

On doing theoretical linguistics: a reply to Haspelmath

David Adger
Queen Mary University of London

17th January 2021

1 Introduction

Haspelmath 2021 (hence H2021) seeks to address the question of how linguistics should deal with what he sees as a paradox at the heart of the enterprise, his General Linguistics Paradox in (1).

- (1) We want to explore and understand the nature of Human Language, but what we can observe directly is particular languages.

H2021 sets up the problem as follows:

- (2)
 - a. (1) implies that the study of human language in general must be based on the study of language universals.
 - b. However, a number of recent studies, including Adger, Harbour, and Watkins (2009), base general claims on the study of particular languages and while this is “not as naive as Antoine de Rivarol’s claims about the universality of French,” it is problematic.
 - c. The confusion that has led to this state of affairs arises to a large extent because the clear notion of *general linguistics* has been replaced by the unclear one of *theoretical linguistics*.
 - d. An alternative to studying universals is to adopt an approach that H2021 calls a “Natural Kinds Programme,” working with hypothetical innate building blocks. This entails that there is a “rich innate grammar blueprint.”

- e. This is problematic for a number of reasons: (i) it has been abandoned by major proponents; (ii) there is “no clear methodology for progressing” and (iii) there are “no clear criteria for success.”

I want to take issue with virtually all of this. I will argue

- (3)
 - a. The study of human language in general can be based on the development of *theories* of the human language capacity (H2021’s “linguisticity”).
 - b. Studies like Adger, Harbour, and Watkins (2009) make claims about theories of the human language capacity, and it is perfectly possible, by analysing one language, to show certain theories are inadequate for all languages (a general claim).
 - c. What makes theoretical linguistics theoretical is that it is concerned with building theories. This is quite distinct from General Linguistics.
 - d. H2021’s example of a “Natural Kinds Programme,” generative grammar, does not require a “rich innate grammar blueprint” because rich categories can be built from more minimal ones with no information loss and with explanatory gain.
 - e. Generative Grammar has (i) not been abandoned by major proponents; (ii) has a clear methodology for progressing and (iii) has the same criteria for success that any scientific programme has.

H2021 makes a number of other points I will not take up here for reasons of space.

In section 2 I lay out what I take to be the relevant assumptions of generative grammar, and the implications that has for both what theoretical linguistics is and for whether one can make progress in the effort to “explore and understand” human language in general by pursuing it. In section 3, I very briefly respond to H2021’s terminological suggestion. In section 4, I clarify the nature of the “Natural Kinds” adopted in generative grammar and argue for the merits of the approach. It will be clear, I hope, throughout this response that there are actually a number of areas where H2021 and I share assumptions, which raises the question of why we come to such radically different conclusions. In the concluding section, I suggest that the core problem with the argument in H2021 arises from applying a particular epistemological perspective (Quinean behaviourism) to an explanandum for which it is not suited (a natural object).

2 On doing theoretical linguistics

I assume, with H2021, that there's some cognitive property humans have that other species don't and that this property is responsible for our capacity to acquire any extant (or historical) natural language. Generative grammar usually calls this the Faculty of Language and H2021 calls it linguisticity. I'll abbreviate it as FL/L, incorporating both. There are important issues as to the domain specificity of FL/L, but they are not relevant to the argument I make here, so I leave them aside (see Adger and Svenonius 2015 for one perspective).

I assume, also with H2021, that FL/L is part of the natural (biological) world.

I assume, and I think that this is consistent with H2021, that to understand FL/L, application of scientific methods and approaches are appropriate. A core aspect of scientific method is developing theories of natural phenomena (in this case FL/L) and evaluating them empirically. This allows us to explore the properties of FL/L and deepen our understanding of it.

What is a theory? I heartily agree with H2021 that virtually all statements in all approaches to linguistics are theory-laden. But that does not mean that all these approaches are *theoretical linguistics*. Theoretical linguistics is linguistics whose primary concern is building and exploring theories.

Building a theory involves stating a set of propositions (hypotheses) in a language which has a basic vocabulary of entities, relations, etc. (the concepts of the theory). The syntax of this language and its interpretation is whatever we need it to be, but at least some aspects of it are drawn from mathematical and logical concepts which we have a good understanding of, and, when a theory is fully formalised, all of the propositions can be stated mathematically. I do not, however, think that a theory must be formalised to be a theory. Formalization is a useful method for ensuring that you understand what your theory is doing, but it is not necessary for making progress. In linguistics, a theory can be stated as a model, via (the interpretation of) a set of propositions, or as a logical deductive system via a set of axioms and deduction rules. It doesn't matter, though different ways of thinking about it may give rise to different insights (viz HPSG vs CCG, or GB vs Minimalism). Theoretical work in the variant of theoretical linguistics I adopt generally involves building partial theories, focussing on some hypothesized aspects of FL/L (phrase structure, argument structure, binding relations, etc). This of course means that there is further theoretical work to be done in integrating different partial theories with each other, ensuring that the

deductive consequences are correct empirically, providing an overall framework, etc.

Given this approach to theories, we need to hypothesise what concepts we need, and to hypothesise the propositions that constitute the theory. Obviously, given we are *in medias res* in virtually all research, much of this is given by previous work, which has developed (partial) theories of the phenomena we want to understand.

In fact, usually multiple theories will be available, some very close to each other, some quite different in fundamental ways. To determine whether one theory is better than another you need a means of evaluating them. To evaluate a theory empirically, you need to state an analysis. You can also evaluate theories in their own terms, and I return to this below.

I take an analysis in linguistics to be a mapping between data and theory. Data will usually be “pre-processed” to a point where the analysis is viable. For example, in syntax we assume that phonetic-level variation in the pronunciation of ‘dog,’ produced on different occasions of utterance, is irrelevant, and that we can work with an abstract element. I put such pre-processing, which is of course a deeply theory-laden task, aside in what follows.

If an analysis takes a piece of data and links it to a theoretical posit, how exactly does this work? I’ll illustrate by an example. If one is interested in the phenomenon exemplified by the English sentences in (4), a number of analyses of ‘dog’ might be posited:

- (4) a. The dog is hungry.
 b. It dogged my footsteps.

One analysis, A_1 , might claim that ‘dog’ has the category N and that ‘dog’ in ‘dogged’ maintains that category. Another, A_2 , might claim that ‘dog’ may have either the category N or the category V. Yet another analysis, A_3 , might claim that ‘dog’ has no grammatical category.

Each of these distinct analyses makes use of certain theoretical terms, which are assumed to refer (if indirectly) to parts of the natural object FL/L: N, V, *has the category*, etc. This is relevant to H2021’s point about “Natural Kinds” and I return to it in section 4.

The analysis, then, maps between the data and these terms. Different theories make available different analyses. A theory, T_1 , incorporating a proposition (constraint) disallowing ambiguity in the assignment of categories will rule out

A₂, and therefore require some other proposition (rule or mechanism) allowing tense morphology to attach to ‘dog’ in 3b. A theory, T₁, which lacks the predicate *has the category* will rule out A₁ and A₂, need some other mechanism to explain why English does not allow a verb to appear in a nominal subject, **The despise frightened me*, etc.

An analysis, then, like a theory, is a set of propositions (hypotheses), but rather than relating the concepts of the theory, an analysis dictates how data maps to the theoretical concepts (and vice versa). Note that this does not, in principle, require data to be directly observable. A pattern of data may lead to an analysis where a theoretical posit corresponds to nothing (as in ellipsis, or null pronominals). One might have a theoretical constraint that rules this out, of course, but that is as much a theoretical claim as the claim that null elements are possible.

Let’s now put all this together and see how it addresses the argument made in H2021, using H2021’s example of Adger, Harbour, and Watkins (2009).

Research like that presented in Adger, Harbour, and Watkins (2009) does indeed analyse a particular language (in this case Kiowa), but its purpose is to determine which theories are better than other theories in coping with the analytical challenges the language poses. It is an exercise in using the analysis of data to evaluate theories. The same can be said for most of the other cases that H2021 lists as research on a single language being used to make general claims.

In Adger, Harbour, and Watkins (2009), we looked at four theories of the relationship between word structure, phrase structure and order, all of which were, at least at the time, important contenders. These theories had been developed on the basis of non-Kiowa data, usually through multiple monographs and papers, and by many researchers. Our purpose was to see whether these particular theories (sets of propositions) were successful in handling a puzzling set of phenomena in Kiowa involving the order between richly inflected verbs, particles signifying aspect, modality, negation, etc., and arguments of the verb. The theoretical positions we examined were:

- (5) a. a theory that treats complex verbs as single words derived via syntactic movement of heads, and that allows heads in phrase structure to either precede or follow their complements (Koopman (1984)’s Head Parameter)
- b. a theory that treats complex verbs as consisting of syntactically sep-

- arate units (Julien (2002)), and that allows heads in phrase structure to either precede or follow their complements;
- c. a theory that requires heads to precede their complements (Kayne 1994) and derives complex verbs by syntactic movement of phrases (Koopman and Szabolcsi 2000);
 - d. a theory that treats complex verbs as single words derived by concatenating morphological units that correspond to a series of syntactic heads in a complement relation (Brody 2000).

We examined how plausible analyses of the Kiowa data would look in these various theories, and argued that the first three had incorrect empirical consequences. These consequences were determined by what the analysis plus the theoretical propositions adopted in the relevant theory predicted about a range of phenomena in the language. In general the theories predicted meanings, orders, or overt syntactic dependency markings that native speakers of the language rejected.

We also examined how one would have to augment or change the various theories to overcome these consequences. We showed that in theories (a) - (c), these led to contradictions, or to such levels of extra theoretical stipulations that any insight was lost. We concluded that Kiowa forced us to adopt a version of theory (d), which treats complex verbs as single words constructed via a particular mechanism argued for by Brody (2000) incorporating an extension to that theory proposed by Brody and Szabolcsi (2003).

This is how much work in theoretical linguistics proceeds. It draws conclusions about theories on the basis of empirical data, and sometimes those conclusions will have universal import. Its ultimate purpose is not to provide an analysis of the data, but rather to improve the theories. The aim is to achieve insight and understanding of FL/L through this method.

Given that it is only through theories that we can deepen our understanding of FL/L, and that the arguments in Adger, Harbour, and Watkins (2009) suggest that theories (a) - (c) are not up to the task, I think it is reasonable to conclude that that work achieved its goal of “exploring and understanding” FL/L. This goal was reached by providing analyses of phenomena from a single language which allowed us to argue that certain approaches to complex word formation were inadequate to the task.

Of course, the analyses we developed in our book may be wrong (though the perspective we argued for has become more prominent in discussions of these

issues in recent years, see Bye and Svenonius 2012, Ramchand 2014, Merchant 2015, Dékány 2018, Arregi and Pietraszko 2018, Harizanov and Gribanova 2019, amongst many other works). In doing this kind of linguistics, many things are in play at once, and juggling complex data, analyses, and theories is a delicate and difficult business. Problems we didn't notice may emerge, and new theoretical ideas may be developed that will improve over our efforts. Science is, after all, a never-ending conversation.

There is a second way in which theory development may take place. A theory can be evaluated also for whether it is the simplest theory, whether it contains contradictions, whether it is equivalent to another theory, how well it is integrated with theories of other domains that are relevant, etc. For example, Adger (2013) spends a fair amount of effort developing a theory which is broadly equivalent in many respects to Brody (2000). The difference is that the theory in Adger (2013) does not take complex lexical items to be listed in the lexicon and then mapped to syntactic structure, but rather to be built in the syntax derivationally. This was an attempt to simplify Brody's theory in certain ways, especially by removing a lexicon of complex words in favour of a lexicon of simplex functional categories. It was also an attempt to integrate it into a more widespread set of assumptions. The theory in Adger (2013), it turns out, had some problems, in that it was more complex than it should have been to define two extended projections as equivalent in certain circumstances, and so I modified it in (still unpublished) later work, which solved this problem, simplified the theory further overall, but opened up new issues, whose implications I still haven't fully worked out. It will probably need some further modification to make it compatible with more recent proposals in Adger (2017), which seeks to provide a more restricted set of grammatical operations.

This kind of work is also a core part of theoretical linguistics, but rather than empirical evaluation, it explores how theories interact, how they can be improved, whether such improvements open up potentially new empirical consequences that were unforeseeable from the perspective of earlier versions, whether different mathematical conceptualizations are possible, etc. This kind of theoretical work is perhaps less common than work which uses empirical phenomena to evaluate theories, but it is still an important source of new proposals which may or may not deepen understanding.

Returning to the argument of H2021, I think I have shown that it is false that “the study of human language in general *must* be based on the study of language universals” (my emphasis). It can also be based on the analysis of

particular languages, and what those analyses tell us about theories of FL/L. This of course doesn't mean that it can't also be based on language universals. Multiple bases surely make things more stable.

One might make a different argument against the kind of approach outlined here, which is that it is too abstract, and that we would have been better spending our time recording more data, working with the Kiowa people to improve their linguistic situation, writing more detailed grammatical descriptions of the language, publishing analysed texts in the language, developing teaching materials, etc. In fact the project team have done all of these things, in addition to the theoretical work. Each of these ways of doing linguistics has its own goal. The goal of the theoretical work we did in Adger, Harbour, and Watkins (2009) was to understand FL/L, and, as the microbiologist Andreas Wagner wrote, "the price of understanding is always abstraction" (Wagner 2014). I don't think that this is the only way to do linguistics, I don't think that it should have any particular priority, epistemologically or sociologically, over other ways to do linguistics, but I do think that it is an important approach and that we can learn much about FL/L by pursuing it.

3 On the "theoretical" in Theoretical Linguistics

It will be clear from the discussion in the last section that I have a very different view of what theoretical linguistics is from that in H2021. I agree with that paper in taking all statements to be somewhat theory-laden. I think that theory is integral to every domain of linguistics, from phonetics to sociolinguistics. Sometimes, given the nature of the research questions, a more formal methodology like that described in the previous section is necessary, sometimes, given the complexity of the domain, or the focus of the research, it is not possible, or would not be helpful. However, for understanding an aspect of the natural world, FL/L, which is not amenable to direct observation or experimentation, an approach of the sort outlined in section 2 is, I think, indispensable, and can be very useful for other areas of linguistics.

The common term "Theoretical Linguistics" is used, I think, in just this way: the modifier *theoretical* isn't to be distinguished from *applied* as H2021 suggests; it signifies rather that the task that is being engaged in is development and investigation of theory using analyses of data (and, less commonly, investigating

the structure, consistency, and ramifications of the theories themselves, irrespective of their empirical consequences). Theoretical linguistics is not, then, just a term that has replaced General Linguistics, and I don't think there is confusion in the field about this. Theoretical linguistics is a particular approach to *doing* linguistics, roughly following the methodology laid out in the previous section. Though the way I do theoretical syntax involves attempting to build theories of certain aspects of FL/L, one can obviously pursue this same methodology with other aims and other assumption and in other linguistic domains, as the rich theoretical literature spanning linguistics as a field shows.

Anticipating an objection here, I accept that not everyone who is analysing the syntax of languages using a particular theory might be motivated by the broad aims I set out above. They may just be interested in a particular phenomenon in their language, and they might find that the theory they learned in graduate school is useful for that task. This work applies a theory to data via an analysis, and uses the theory as a fundamentally descriptive tool, leveraging the wide acceptance of the theoretical posits. This is important and necessary work which has an essentially descriptive focus. H2021 is keen for such researchers to adopt less theory-bound descriptions, and H2021's objection to this research is that it presupposes the existence of a rich innate set of universal building blocks, and that presupposition is problematic. I turn to this in the next section.

4 Theoretical Linguistics and Natural Kinds

The final plank of H2021's argument (at least the aspects I am addressing here) has two parts. First, a solution to the General Linguistics Paradox is to adopt what H2021 calls a "Natural Kinds Programme" that crucially requires a "rich innate grammar blueprint". Second, this kind of programme is problematic: it has been abandoned, it lacks a clear methodology for progress and it lacks criteria for success.

I take it that the approach sketched in section 2 instantiates what H2021 sees as a "Natural Kinds Programme". It certainly adopts the view that there is something in the natural world (FL/L) which is what distinguishes humans' grammatical abilities from those of other species that we know of. The approach develops theories of FL/L on the basis of evidence from particular languages (and on the basis of other evidence too, from language acquisition, language processing, typological patterns, etc.). Those theories consist of posits that are

assumed to refer to aspects of the natural object FL/L. So when a theorist says that “dog” has the category N, that “despise” has the category V, and that “the dog” is a phrase, that does indeed involve a claim that the natural object FL/L has the following real world properties: there are distinct structural components of FL/L which we can call N and V, and these share a property, which we call “being a grammatical category”, which is distinct from the property we call “being a phrase”. All of the theoretical posits used in a grammatical analysis, as H2021 quite correctly says, are hypothesized to refer to aspects of the natural world. The only way we can get to know those aspects of the world is through the success or failure of our theories.

I do have some quibbles about H2021’s terminological proposal, that generative grammar is a “Natural Kinds Programme,” mainly because I think the terminology is inexact. Spike (2020) argues that previous work by Haspelmath on the topic (e.g. Haspelmath 2010) places too strong requirements on what might constitute natural kinds, to the extent that biology and chemistry would lack natural kinds, never mind linguistics. He proposes an alternative, which is categories that are clusterlike and variable, but such categories are not, I think, what Haspelmath would take to be the “Natural Kinds” he has in mind. I also think it’s terminologically odd to take an operation, such as Merge, or Agree, to be a kind. Natural kinds are usually thought of as substances, not operations (Bird and Tobin 2018). H2021 points out, in footnote 18, Chomsky’s use of the term “naturalistic approach,” which is just the standard philosophical viewpoint of naturalism: there are objects of the natural world. This I certainly agree with. H2021 states that this seems to be “more or less what I mean here by Natural Kinds Programme.”

So I agree with H2021 about adopting a naturalistic approach, but what of the “rich innate grammar blueprint”? H2021 (section 6.5) argues that only a “rich set of innate building blocks” is sufficient to escape the General Linguistics Paradox, and the discussion in Section 2 would seem to back that up. That theory adopted various categories, features, operations etc. But where I think H2021 goes wrong is in its identification of the rich set of posits with the innate set of posits.

Consider the natural numbers between one and nine inclusive. We can think of these as involving a set of nine distinct building blocks (1, 2, 3, 4, 5, 6, 7, 8, 9), and lets take nine to be a rich set of building blocks for nine things. But we can also think of this issue as involving two building blocks, call them 1 and 0, plus another building block which is principle of positional interpretation extending

infinitely in one direction, with each position signifying that its content is to be multiplied by the power of the position. That approach involves three distinct building blocks, one of them admittedly fairly rich in scope, but also gives us the natural numbers between one and nine: 1, 10, 11, 100, 101, 110, 111, 1000, 1001. So we have two ‘theories’ of the numbers between one and nine: one involves a rich set of building blocks, the other involves a substantially reduced set of building blocks. Of course the second approach is far more powerful than the first, and gives us a theory of all numbers. To allow the first approach to do the same, we must augment it too with a principle of positional interpretation. If we do this, we now have approaches of equal scope, but in one there is a richer set of ‘categories’ than there is in the other.

Of course, the analogy is inexact but, I think, still informative. You can build richer sets of categories out of poorer ones. The richer set can be used in analysis, and may be much easier to use for some tasks than others (e.g. stating in the minimal number of syllables how many things there are on the table), while the poorer set may be better for other tasks (e.g. determining whether a number is even or not).

This is exactly the job of the more theoretical side of theoretical linguistics, which was discussed at the end of section 2. Returning to the “dog” example, the theoretical syntactician does not stop at N, V, *category* and *phrase*, as is evident from all of the work in generative grammar over the past seven decades. The task is to keep abstracting, trying to get down to more fundamental units and their principles of combination. It is the interactions between these that derive the richness we see in human languages. H2021 conflates “rich” and “innate”. In fact, the goal is to find the poorest set of “innate” categories (and it is not even necessary that these be domain specific—that is an empirical question, Adger and Svenonius 2015). H2021 also claims that “by eliminating ‘richness’ of UG, generative grammarians have also tended to reduce the explanatory scope of their analyses.” Again, this doesn’t follow, and for the same reasons. Richness is eliminated, but the information that the rich categories contained is not eliminated. It is reduced to a more minimal set of units and their structures. Indeed, this approach often profoundly *extends* the explanatory scope of the analyses (see Harbour 2016 for an example of how person and number features can be reduced dramatically and how such a reduction provides a more restrictive and empirically adequate typology of pronominals).

We can see this drive for abstraction in almost every area of generative grammar. A good example is in how phrase structure rules were re-theorized

over the decades. Generative grammars used to have many hundreds of phrase structure rules of different sorts, but as the theory developed it abstracted out of them shared properties of their categories (in the form of features, Chomsky 1965), imposed on them similar overall structure (X-bar theory in Chomsky 1970 and Jackendoff 1977), removed the notion of order from the rules (Stowell 1981 and GPSG's ID/LP format, Gazdar, Klein, Pullum, and Sag 1985), and, within Minimalism, removed the residue of the concept of phrase structure rule entirely, replacing it with Merge, which allowed a unification of the local syntactic combination that phrase structure rules captured with the non-local syntactic combination that transformational rules had been used for (Chomsky 1995). This theoretical move unified the phrase structure and transformational components of the previous theory, resulting in a new theory of similar scope but simpler structure. It also opened up new ways to think about phenomena such as reconstruction for anaphor binding etc. Similar comments could be made about the notions of category and feature, about reduction of islands to subjacency, etc. Not every such theoretical move is successful, but that is part of the usual to and fro of theoretical debate. H2021 is concerned about the lack of agreement in theoretical syntax, but at least from my perspective there is overwhelming agreement about a great deal once you learn to look beyond notational and terminological distinctions. Conceptually, theoretical syntax is fairly unified, with even quite distinct frameworks sharing much (e.g. the massive borrowing of ideas between Relational Grammar and GB in the 1970s, the closeness of HPSG and GB analyses in the 1990s, how broadly similar Combinatory Categorical Grammar and Minimalist Syntax are now). Of course there are many fundamental disagreements, but that's hardly surprising. I find myself in fundamental disagreement with things I myself accepted or proposed over the course of a mere few years. Science moves forward by being wrong.

This is why when H2021 suggests that influential linguists have abandoned the "Natural Kinds Programme" that is a misreading of the situation. They have not. Their work rather aims to reduce, as much as possible, the variety of theoretical posits. The idea is that the relevant properties of the theory are minimized (like 1, 0 and interpreting position as numerical powers) so as to more accurately comport with the properties of the world that they refer to. This work doesn't reject the rich set of categories; it argues that the same or better explanations can be got by driving towards more abstract elements out of which these categories are built. In that process we come to see that our rich categories were perhaps not exactly right. They were rough generalizations at

a certain level of description that a better theory improves on.

So much for the abandonment of the programme. What of H2021's other criticisms: that this approach lacks a clear methodology for progress and that it lacks criteria for success.

In section 2 I laid out a clear methodology for progress, which is what is standardly used in the kind of work I've been discussing here. I don't claim that it is the only methodology for progress, and indeed my own research has used many other methods (including descriptive work, corpus work, experimental work etc.). However, rejection of theories on the basis of evidence is clearly possible in such an approach, and that constitutes progress. Further, the method of abstraction just discussed can also lead to appreciable progress, both in terms of deepening understanding and in terms of empirical coverage.

Finally, what of the claim that such an approach lacks clear criteria for success. I think this is of a piece with the issue about a methodology for progress. Generally the criteria for success of a programme are whether it opens up new empirical phenomena for study and provides insight into the object of study. Both are clearly true.

5 Conclusion

I have argued that the study of human language need not be based on linguistic universals, but can rather involve the development of theories, and that analysis of a single language can inform which theories should be taken forward. Theoretical linguistics is the variety of linguistics that does this. This does not preclude the contribution of many other approaches, which are needed to develop a full picture. Evidence for evaluating theories will also come from cross-linguistic comparison, from typological study, from dialectal and individual variation, from how language is acquired, how it undergoes attrition, how it is processed, etc. Further, H2021's criticism of the naturalistic approach (the "Natural Kinds Programme") misses the crucial role of abstraction in explanation in theoretical linguistics, and that leads to misreadings of the current situation.

However, what I have argued here actually shares many assumptions with H2021. It is the conclusions that we draw from the assumptions that are very different. The question is why.

The reason is, I think, not unrelated to the debate between Quine and Chom-

sky over five decades ago on the indeterminacy of translation (Quine 1969, Chomsky 1969, Quine 1970). Their disagreement stemmed from different views on what linguistics was. For Chomsky, linguistics (or at least the part of it he was concerned with) was about an object of the natural world, and so theories of that object were of a piece with theories of other objects of the natural world (like planets, or gasses). For Quine, linguistics was about a collection of learned conventions, as he famously details in *Word and Object* (Quine 1960), and theories of that are quite distinct in nature from theories of natural objects. For Quine, linguistic theory in Chomsky's sense (a theory of FL/L in our terms) is too distant from observed utterances. What is accessible to the linguist, according to Quine, is observations of utterances, and it is from these that grammars are born: a grammar is, in the end, a theory of behaviour. For Chomsky, a grammar is a theory of an object of the natural world (a cognitive state), and behaviour is simply evidence for or against that theory. Clearly I have taken the Chomskyan viewpoint here, but I think that the mistake that H2021 makes is that it accepts the existence of the cognitive state FL/L, but at the same time it also takes a grammar to be ultimately a theory of behaviour. This is why H2021 takes the General Linguistics Paradox in (1) to be a paradox. It is paradoxical indeed to insist on a behaviourist interpretation of a cognitive object.

References

- Adger, David. 2013. *A Syntax of Substance*. Cambridge, MA: MIT Press.
- Adger, David. 2017. A memory architecture for Merge, <http://ling.auf.net/lingbuzz/003440>.
- Adger, David, Harbour, Daniel, and Watkins, Laurel. 2009. *Mirrors and Microparameters: Phrase Structure beyond Free Word Order*. Cambridge: Cambridge University Press.
- Adger, David and Svenonius, Peter. 2015. Linguistic explanation and domain specialization: a case study in bound variable anaphora. *Frontiers in psychology* 6:1421.
- Arregi, Karlos and Pietraszko, Asia. 2018. Generalized head movement. *Proceedings of the Linguistic Society of America* 3:1–15.
- Bird, Alexander and Tobin, Emma. 2018. Natural kinds. In Edward N.

- Zalta, ed., *The Stanford Encyclopedia of Philosophy (Spring 2018 Edition)*.
- Brody, Michael. 2000. Mirror theory: syntactic representation in perfect syntax. *Linguistic Inquiry* 31:29–56.
- Brody, Michael and Szabolcsi, Anna. 2003. Overt scope in Hungarian. *Syntax* 6:19–51.
- Bye, Patrick and Svenonius, Peter. 2012. Non-concatenative morphology as an epiphenomenon. In Jochen Trommer, ed., *The Morphology and Phonology of Exponence*, Oxford, UK: Oxford University Press.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1969. Quine’s empirical assumptions. In Donald Davidson and Jaakko Hintikka, eds., *Words and Objections*, 53–68, New York: Humanities Press.
- Chomsky, Noam. 1970. Remarks on nominalization. In R. A. Jacobs and P. S. Rosenbaum, eds., *Readings in English Transformational Grammar*, 184–221, Waltham, MA: Ginn-Blaisdell.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Dékány, Éva Katalin. 2018. Approaches to head movement. *Glossa: a journal of general linguistics* 3:10–5334.
- Gazdar, Gerald, Klein, Ewan, Pullum, Geoff, and Sag, Ivan. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwells.
- Harbour, Daniel. 2016. *Impossible persons*. Cambridge, MA: MIT Press.
- Harizanov, Boris and Gribanova, Vera. 2019. Whither head movement? *Natural Language & Linguistic Theory* 37:461–522.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86:663–687.
- Jackendoff, Ray. 1977. *\bar{X} -Syntax*. Cambridge, MA: MIT Press.
- Julien, Marit. 2002. *Syntactic Heads and Word Formation*. New York: Oxford University Press.
- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

- Koopman, Hilda. 1984. *The syntax of verbs. From verb movement rules in the Kru languages to universal grammar*. Studies in generative grammar 15, Dordrecht: Foris.
- Koopman, Hilda Judith and Szabolcsi, Anna. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.
- Merchant, Jason. 2015. How much context is enough? two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46:273–303.
- Quine, W. V. 1970. Methodological reflections on current linguistic theory. *Synthese* 21:386–398.
- Quine, Willard Van Orman. 1969. Linguistics and philosophy. In Sidney Hook, ed., *Language and Philosophy*, 95–96, New York: University Press.
- Quine, Willard v.Orman. 1960. *Word and Object*. Cambridge, MA: MIT Press.
- Ramchand, Gillian. 2014. Deriving variable linearization. *Natural Language & Linguistic Theory* 32:263–282.
- Spike, Matthew. 2020. Fifty shades of grue: Indeterminate categories and induction in and out of the language sciences. *Linguistic Typology* 24:465–488.
- Stowell, Tim. 1981. Origins of phrase structure. Ph.D. thesis, MIT, Cambridge, MA.
- Wagner, Andreas. 2014. *Arrival of the fittest: solving evolution's greatest puzzle*. London: Penguin.