

Balancing Evolution and Acquisition in Theoretical Linguistics: Tensions and Prospects

Gabe Dupre

Keele University

Abstract

The history of generative linguistics is a history of a tension between on the one hand incorporating all the complexity suggested by detailed descriptions of natural languages and the innate knowledge humans bring to the task of acquisition, and on the other the sparsity that seems to be forced on such innate contributions by developmental and evolutionary biology and psychology. In this paper I describe these pressures in detail, and then outline the strategy currently adopted by ‘Minimalist’ linguistic theorizing, which involves deriving the complexity of linguistic competence from an underlying simple system. I close with some remarks on the remaining difficulties and prospects for such an explanatory strategy.

Keywords

Generative linguistics, Language Evolution, Minimalist Grammar, Syntax, Syntax-Semantics interface, Syntax-Phonology interface.

1 Introduction

The acquisition of a natural language is a product of, on the one hand, features of the mind of the person acquiring the language, and on the other, features of the developmental environment. This much is trivial. But, identifying the relevant features on either side of this division is far from trivial. With respect to the former, debate has raged for decades, if not centuries, about the extent to which the mind features components *specialized* for language acquisition, which leave their mark on the linguistic systems we acquire. It has been a central commitment of generative linguistics since its inception that there are such specialized systems, whereas a variety of ‘empiricist’ approaches have argued that language acquisition may feature complex and powerful innate mechanisms, but such mechanisms are used for a wide range of learning tasks, extending beyond language. On the other hand, there are complementary debates concerning the features of the environment that the child relies on in acquiring a language. The principal reason for positing language-specific innate mechanisms is the apparent dearth of suitable linguistic information in the learner’s environment. If the learner acquires a language without encountering suitable ‘instruction’ in the environment, then the explanation for why their language has the properties it does must appeal to internal properties of the learner, rather than properties of the environment. As the developed structure is linguistic, this invites the inference that these internal properties are themselves language-specific. Consideration 1, then, for a theory of language acquisition, is that the degree to which the environment provides cues to the acquired language is inversely proportional to the amount of innate structure which must be posited: the less evidence the child has to work with, the more internal structure there must be to make up for it.

On the other hand, this internal structure does not come for free. Natural selection can, of course, explain the providence of highly complex traits, but such explanations presuppose various things: such traits must be (or at least must have been) adaptive, and so must any proposed intermediate, less complex, traits, and there must have been sufficient time in evolutionary history for the processes of selection to have found and favoured the traits in question. Consideration 2, then, is that the greater the complexity attributed to

innate factors, the greater the burden placed on developmental and evolutionary theory.¹

The development of generative linguistics is a competition between these considerations. Early work stressed both the sparsity of the environment and the complexity and idiosyncrasy of the acquired linguistic system, necessitating a highly complex and language-specific innate contribution. However, the biological implausibility of such an innate system was recognized, leading to greatly stripped-down ‘Minimalist’ linguistic theories. Such theories are simple enough that the systems they describe might plausibly have evolved, but *prima facie* fall far short of accounting for the complex linguistic observations made in the earlier period. In this paper, I shall outline these developments, identifying where and why such tensions arise, and then identify the several strategies being used in attempts to solve this problem and balance the constraints of language acquisition with language evolution.

2 Complexity From Above

Perhaps the most well-known doctrine of generative linguistics is that of Universal Grammar (UG). This term has traditionally been associated with the idea that, to a significant extent, our acquired capacity with language is internally driven, rather than extracted from our environment. Corollaries of this claim include the claim that there are significant constraints on the extent to which human language can vary (i.e. that language acquisition is canalized) and that all human languages are, despite surface variation, highly similar. Traditionally, UG has been split into two components, formal and substantive universals. Formal universals concern the forms that linguistic rules or processes can take. That all rules are structure-dependent, identifying their targets by their hierarchical relations rather than their linear position, is a purported formal universal, eliminating from any possible language an infinite range of conceivable rules such as “to form a question, replace the third word in a sentence with a question particle”. Substantive universals concern the entities and categories that such rules apply to. That all

¹This difficulty is centered by the task of going “beyond explanatory adequacy” (Chomsky (2004)). That is, the project is not merely explaining how a human language is acquired, but why the mechanisms involved in this achievement have the properties they have.

languages draw a grammatical distinction between nouns and verbs is an example of a purported substantive universal.

If ‘UG’ is used just to refer to whatever aspect of the mind is responsible for our ability to acquire a language, then the existence of UG is uncontroversial. The debate then concerns what the properties of this aspect of mind are. The controversy centers on the claim that UG incorporates lots of language-specific information, constraining human languages in ways quite unlike those found in other learning domains.² The motivation for such claims has historically come from Poverty of Stimulus (PoS) arguments.

PoS arguments begin by identifying an aspect of developed linguistic competence. For example, both observation of corpora and the linguistic judgements of native speakers can tell us that in English, adjectives tend to come in a specific pre-nominal order, according to their semantic category.³ So that we have:

1. The *big old red brick house*

With the ordering *size > age > colour > material > Noun*, but never any other order⁴:

2. *The *big red brick old house*
3. *The *brick big red old house*
4. *The *old big red brick house*

And so on.

Assuming these observations are correct, we can infer that it is a fact about English speakers that they have internalized some information about their language concerning which linguistic expressions are legitimate and which are not. This fact needs explaining, and such an explanation will involve appealing to the interactions between the speaker

²The debate is sometimes framed in terms of the existence of a ‘Faculty of Language in the Narrow Sense’ (FLN). Hauser et al. (2002) distinguish the FLN from the ‘Faculty of Language in the Broad Sense’ (FLB): FLB incorporates all cognitive systems relevant to, or used in, the acquisition and use of language, including memory, perceptual systems, belief/desire psychology, etc. while, FLN is just those elements of FLB that are unique to humans and uniquely used for language.

³See Scott (2002).

⁴Some apparent exceptions can be explained away as involving complex adjectives in a specific category, rather than categories in a non-standard order. For example, “Big old brick red house” is fine, but only as referring to a house of unspecified material which is *brick red* in colour.

and their linguistic experience. One possibility is that such linguistic facts are learned in much the same way that everyday knowledge of the world is acquired, i.e. through experience and inductive reasoning. Given that all the red wine I have tried has tasted unpleasant, I infer that all red wine tastes unpleasant. Likewise, one could imagine a language-learning system which induced the order of adjectival modifiers on the basis of the adjective orderings they have encountered. In the first place, such a learner could infer that orderings they have encountered are legitimate, and then could attempt to generate a complete ordering on this basis. For example, if they had heard “red brick house”, and “new yellow house”, they could infer that age comes before material even though they have never heard these items in construction with one another.

However, problems for such a proposal abound. For one thing, they face the notorious ‘logical problem of language acquisition’. On the widely adopted assumption that language learners are provided with positive evidence, concerning what is legitimate in their language, but not negative evidence, concerning what is illegitimate, a gulf opens up between their linguistic experience and their acquired competence. Assume that the child encounters “red brick house” and “new yellow house”, and infers from this that *colour* > *material* and *age* > *colour* are licensed. Purely on the basis of this evidence, what is there to tell the child that alternative orderings (e.g. *material* > *colour*) are *not* licensed? A child will, of necessity, be exposed to a finite subset of all possible linguistic structures, and there will be indefinitely many structures which are legitimate, and recognizably so by the learner, even though they have never been encountered. This precludes the possibility of a ‘conservative’ learner, who takes as legitimate only those structures they have encountered. Thus, the child’s evidence must be supplemented by a more substantive learning rule which allows it to identify the *impossibility* of orderings like *material* > *colour* without thereby excluding legitimate structures. Note further that a simple assumption like *if one ordering is legitimate, alternative orderings of the same constituents are illegitimate* would be too powerful in exactly this way, ruling out the possibility of alterations such as “Maita picked the pizza up”/“Maita picked up the

pizza”.⁵

The problems for such a proposal are exacerbated by the complexity of the acquired knowledge. The above fairly simple ordering, featuring just four adjectival classes, allows for 16 different options out of a possible 65 adjective-noun combinations. But the more adjectival classes are featured on the list, the greater the reduction of options provided by rigid ordering constraints. Scott (2002) proposes an ordering of 16 pre-nominal adjectives. If this ordering is inviolable, there will be 256 possible orders. Without such an ordering constraint, the possibilities are enormous. If these 16 adjectival types could go in any order, even assuming we are allowing at most one instance of each semantic class per noun-phrase, there would be over five million possible phrase-types containing exactly 6 adjectives. The total number of phrase-types, just composed out of these 16 adjectival types with at most one instance of each, is in the tens of trillions.⁶ Combinatorial explosion of this sort drastically reduces the prospects of identifying the tiny legitimate set from the massive set of conceivable orderings, as the evidence will only cover a tiny proportion of the possibility space, thus again necessitating substantial constraints on the learning process.

In addition to these problems which beset an individual language learner, a theory of *human* language acquisition must apply to all learners. Thus, even if there were cases in which particular learners were exposed to sufficient evidence to help them identify the correct linguistic facts, this will be irrelevant if other learners are able to acquire the same competence without access to this evidence. The fact that ordering constraints akin to those above appear to be cross-linguistically highly robust further undermines explanations which appeal centrally to environmental evidence. For these and related reasons, much work in generative linguistics has concluded that substantial aspects of our acquired linguistic competence is not extracted from the linguistic environment, but is instead provided as part of our innate biological endowment. That is to say that our

⁵There are proposals in the literature aimed at showing how absence of evidence can be interpreted by the learner as evidence of absence, such as that proposed by Chater et al. (2015), but these face further empirical problems, as discussed in Yang (2015).

⁶While it is true that phrases with, say, a dozen adjectives start sounding pretty bad, we do have strong intuitions for intermediate numbers. “The big old square green American brick house” sounds a bit strange, but nowhere near as bad as “the brick American green square old big house”.

knowledge of, *inter alia* adjectival ordering constraints is not *learned*, but simply develops as part of normal psychological maturation (e.g. Cinque (2010)). As this knowledge does not seem, on its face, to be derivable from any non-linguistic aspects of our psychology, this suggests that humans have innate and language-specific ‘knowledge’ constraining which languages they can and cannot acquire.

The case for substantive universals is perhaps even easier to make. To have any hope at acquiring a language, children must be able to identify aspects of their environment *as* instances of language. Since Mehler et al. (1988) it has been established that the ability to differentiate human speech from mere noise is present very early in development. Beyond this, to identify properties in the speech stimulus, children must be able to more narrowly classify the linguistic items they are hearing. For example, to determine which phonemes are legitimate in their language and which are not, they must be able to identify the phonemes of their ambient language. For example, upon hearing the word ‘waiter’, the child is able to infer that the local language includes labial approximants like /w/, which rules out languages like German. Such reasoning seems essential to explain linguistic variation. But note that without the ability to identify the stimulus as *containing a labial approximant*, such learning would be impossible.⁷ The general idea here is expressed nicely by Howard Lasnik in the introduction to Lasnik et al. (2000): “The list of behaviors of which knowledge of language purportedly consists has to rely on notions like “utterance” and “word.” But what is a word? What is an utterance? These notions are already quite abstract. Even more abstract is the notion “sentence.” Chomsky has been and continues to be criticized for positing such abstract notions as transformations and structures, but the big leap is what everyone takes for granted. It’s widely assumed that the big step is going from sentence to transformation, but this in fact isn’t a significant leap. The big step is going from “noise” to “word.”” (p.3) Here Lasnik is driving home the point that the very process of identifying environmental soundwaves as linguistic expressions requires a significant ‘head-start’ by the learner; a head-start which will presumably be accounted by attributing to the learner sophisticated representational

⁷For reasoning of this sort, see Hale & Reiss (2003).

tools, i.e. substantive universals.

These sorts of arguments have thus been applied to a huge range of cases within generative linguistics. For just about any interesting property of natural language syntax, semantics, morphology, and phonology, it has been argued that the bulk of work in the acquisition process is done by innate facts about the mind, not by extraction of relevant structures from the environment. If these arguments are good, which they often appear to be, this suggests a huge amount of innate and language-specific structure, and thus an ornate and highly complex UG. This complexity was easily accommodated by early approaches to generative grammar, namely Transformational Grammar, and the Principles and Parameters approach, to which I turn to next.

3 A Brief History of Nativism in Generative Grammar

If PoS arguments show us that aspects of natural language must be innately provided, rather than learned through perception of environmental stimuli, this raises the question of how such innate knowledge is encoded or realized. In this way, theories of language acquisition and theories of developed competence are entangled. What shape we think our mature competence takes will influence how we think of the innate contribution to the learning process. And indeed, in the course of the development of generative linguistics, these two projects have been mutually informing.

The earliest generative approach (Chomsky (1965, 1957/2002)), Transformational Grammar, viewed linguistic competence as, roughly, a collection of rules mapping linguistic representations onto other linguistic representations. Assuming a stock of primitive lexical items (morphemes and/or words), Phrase Structure Rules identify the ways that lexical items can be combined into more complex expressions. For example, a rule could state that a Verb-Phrase could be generated by combining a Verb and a Noun-Phrase. These Phrase Structure Rules are then supplemented by Transformations, which map complex expressions onto other complex expressions, as when in English a passive is generated from its corresponding active by deleting the subject, raising the object to subject

position, inserting an auxiliary verb and inflecting the main verb to form its past participle. Each rule includes a statement of which expressions it can, and cannot, apply to, and what the result of applying it is. That the Phrase Structure Rules are recursive (i.e. they can apply to expressions containing as constituents expressions of the same type) enables this approach to account for linguistic productivity, the ability of finite speakers of a language to produce and interpret indefinitely many distinct linguistic expressions. Grammaticality consisted in generability by these rules, and all expressions not so generable were thereby ungrammatical.

Acquiring a language was viewed as a process of considering which possible rules best accounted for the encountered linguistic data. While not the only way in principle to incorporate innate contributions, the standard way to do so in this era was to posit innate constraints on which rules were possible. So, for example, substantive universals would ensure that the child begins the process of language acquisition by analyzing encountered speech (or signed) stimuli as exemplifying Nouns, Verbs, etc. and complex phrases containing these categories, organized into hierarchical constituent structure. Formal universals might then preclude certain sorts of conceivable structural hypotheses (e.g. as in Stowell (1981) and Jackendoff (1977)), or transformations (e.g. Ross (1967)). If acquisition is akin to hypothesis testing, then UG could serve as a restriction on the hypothesis space under consideration.

For a variety of reasons, this approach to grammar and to acquisition fell out of favour in the 1980s (Chomsky (1981)). Language-, and construction-, specific rules were replaced by general principles, which were found across languages. Linguistic variation was accounted for by parameters, or underspecified principles which could be fully specified by fairly minimal lexical experience. For example, Principle C of Binding theory states that R-expressions (nominals which are neither pronouns nor anaphors, such as proper names) cannot be bound by (i.e. grammatically identified as referentially dependent on) any other expression. And the Head-Directionality Parameter stated that heads universally either precede or follow their complements, with experience of the local language determining which option is adopted. A grammar, on this picture, was not a collection

of rules which determined which expressions could be constructed and which could not. Instead, it was a collection of distinct ‘modules’, featuring principles governing different aspects of grammar (phrase-structure, co-reference, etc.). Grammaticality was typically viewed as generability consistent with the constraints provided by all such modules: so long as a structure didn’t violate any of these constraints it was legitimate.

Once again, such an approach was easy to combine with PoS arguments. When some aspect of developed language was not identifiable from the linguistic stimulus, then information concerning this features could be encoded into the architecture of the grammar. The main question was then whether such linguistic knowledge seemed to be universal to all languages, in which case a principle should be posited, or whether it displayed constrained variation, in which case a parameter was appropriate.⁸

Aside from the empirical (e.g. Newmeyer (2004)) problems with such an approach, the principles and parameters account of grammar and acquisition faces a deep worry, which is shared by TG. Namely, the innate structures posited seem to be biologically highly implausible. Linguistically specific assumptions concerning the distribution and interpretation of anaphors, say, just don’t seem to be the kinds of things that biological evolution and development seem liable to account for. Further, such innate traits don’t seem to be plausibly adaptive. And finally, even if they were biologically plausible and adaptive, the timescale on which they have been incorporated into the human lineage seems too brief to allow for evolutionary processes to select for them. For these sorts of reasons, recent work in generative linguistics has taken a ‘Minimalist’ turn, aiming to rid the innate endowment of such baroque linguistic information. I turn to these considerations in the next section.

⁸Of course, the usual scientific considerations of simplicity provided constraints here: ideally a small number of principles and parameters could explain a wide range of linguistic observations. And indeed much work was done to show that surprising amounts of linguistic variation could be traced to the interactions between a handful of parameters. But the substance of the theory did not place any particular constraints on positing innate structure.

4 Simplicity from Below

So far, then, we have seen motivation from PoS considerations for viewing what human children bring to the task of learning a language as quite complex and language-specific. With little in the way of instruction, children know that certain word orders are available while others are excluded, that certain linguistic categories will be exemplified in their environments while others will never even be entertained, that certain linguistic rules are possible while others are not, and so on. Even with some systematization and unification of this knowledge, as in the Principles and Parameters approach, it seems that a statement of all such knowledge will require a fair degree of complexity. We can imagine creating an artificial intelligence that brings such rules and constraints to the task of language acquisition explicitly, with something akin to an axiomatic system ‘hard-wired’ into its learning procedures. But there are powerful reasons to doubt that human linguistic development works like this.

For one thing, any purported innate trait must be explicable from the perspective both of ontogeny and heredity. There must be some story about how the developmental cycle that relates a fertilized egg to an adult is able to ensure that the latter has the traits in question. Further, if these traits are indeed innate and species-universal, it must be shown how the traits of the offspring are dependent on those of the parents. While the details remain controversial, embryology provides a reasonably well-understood story, involving the regulation and modulation of cellular growth and differentiation, about how such processes work for certain gross anatomical traits such as the structure of the chordate skeleton. And the key causal agents in this story, genomic material, its developmental environment (e.g. a uterus), and other materials (a source of energy, a microbiome, etc.), are made available for the new organism by the previous generation. But it is not clear that analogous integrations into developmental biology are possible for the highly specific innate knowledge identified by generative linguistics.

Generative linguists often refer to UG as a ‘genetic endowment’. In an authoritative textbook on language acquisition, Guasti (2017) nicely describes the standard view that “Children are born expecting that, whichever language they are going to hear, it will

have properties that their genetic equipment is prepared to cope with.” (p. 19). While not an entailment of the generative view, this does suggest a picture of the genome as a ‘blueprint’, specifying in advance the properties of the adult organism. And indeed, it is difficult to see how linguistically highly-specific knowledge (e.g. that natural languages allow for multiple-wh-questions in which a raised wh-expression begins in a higher position than a wh-expression that remains in situ, as in *What did Aike give to whom?*, but not when the raised wh-expression begins in a lower position, as in **To whom did Aike give what?*) could be inherited and develop without assuming a ‘genetic information’ model of this sort. The problem, then, is that such models are roundly taken to be incompatible with modern biology.⁹

Of course, if the human genome were indeed a blueprint, specifying the traits of the developed phenotype, then it is perfectly conceivable, assuming a suitably ‘sized’ genome, that psychological traits could be encoded to an arbitrary degree of specificity. That is, if the genome were simply a list of developed traits, there is no reason in-principle why purported linguistic universals should not feature on such a list. However, it is near unanimous in philosophy of biology that this is a deeply mistaken conception of the genome and biological development more generally. If genes can be said to ‘code’ for anything, it is amino acids. And the genome is no more a source of information about the developed phenotype than is the environment, or the various other epigenomic factors that play essential roles in determining how an organism will develop. But without the guiding hand of a genetic blueprint, it is far from clear how a biological system could ensure that such developmental processes could produce a psychological ‘database’ of linguistic knowledge.

Even worse, whatever prospects remain for substantive innate linguistic knowledge once the notion of a genetic blueprint has been rejected, highly complex innate traits seem to require long histories of natural selection. But there is compelling evidence that the history of human language is fairly brief, and that many of the apparently innate features of language are not in any clear sense adaptive. On the first count, evidence from

⁹See e.g. Griffiths (2001).

genetics (Cavalli-Sforza (2001)), paleoanthropology (Tattersall (2017, 2019)) and theoretical linguistics (Bickerton (1995)) all points to the fact that genuine human language, featuring recursive syntax and combinatorial semantics, is a trait only of anatomically modern humans who came onto the scene around 200kya. Standard proxies for linguistic capacities, such as symbolic artifacts, are very rarely found prior to this. In the other direction, relative isolation between human populations in the last 100kya does not seem to be reflected in significant linguistic variation¹⁰, which suggests that there has been minimal evolutionary change in this period. So, if language-specific constraints are to be innate, they must all have been introduced into the human lineage in the space of around 100k years.

While biological change can happen relatively quickly, this is most plausible alongside strong selection pressure. But this seems to be absent in the linguistic case. What, if any, adaptive function human language has in general is hotly disputed. Historically, it has been largely assumed that language has been selected for communication.¹¹ And there is a clear sense to this: human languages give us the ability to communicate arbitrarily complex pieces of information, thereby potentially increasing our store of world-knowledge to encompass that of all those we interact with. However, there are problems with this idea (e.g. in competitive environments the benefits of deception may undermine any benefits of pooling information), leading various theorists (e.g. Chomsky (2015), Reboul (2015)) to argue that language is adapted not for inter-personal communication, but instead as the medium for thought. Whatever one thinks about such general claims about the “function of language”, it seems clear that the specific features of language identified earlier seem unhelpful, and sometimes even positively perverse, from the perspective of biological selection.

Take the above mentioned case, of adjectival ordering. What sort of imaginable adaptive benefit could be provided to an organism on the grounds that it constructs sentences with adjectives identifying the age of an object closer to their nouns than adjectives indicating size? Or for an organism that can question multiple arguments of a verb so long

¹⁰With some possible exceptions, such as phonemic inventory. See Huybregts (2017).

¹¹Pinker & Bloom (1990)

as they are ordered one way (“Who brought what?”) but not another (*“What did who bring?”)? Other apparently innate constraints don’t merely mandate a preference for one word-ordering over another, but seem to make certain sorts of otherwise intelligible constructions impossible without significant paraphrasing. For example, one can perfectly well imagine a scenario in which a speaker knows that Mahima saw Aike and someone else at the bar, but not know who the latter person was. But to ask about this person, clunky paraphrases (“Mahima saw Aike and someone else at the bar. Who was the other person?”) are needed, as the standard procedures for forming questions are impossible (*“Who did Mahima see Aike and at the bar?”). And the same goes for myriad other examples from the literature.

For these sorts of reasons, many linguists have come to believe that the linguistic generalizations, including principles and parameters, proposed in early generative work cannot be simply incorporated into developmental psychological stories. Linguistically specific rules cannot easily be traced to features of the genome, and even if they could it is hard to see how any plausible evolutionary story could account for how they got there. However, this realization does nothing to undercut the compellingness of the arguments for such innate knowledge. Poverty of Stimulus arguments remain untouched. Thus, a deep tension arises. On the one hand, it seems clear that humans are innately prone to adopting some kinds of languages and precluded from even considering others. On the other, there seems to be no way to make sense of this fact consistent with what we know of biology and developmental psychology.

As of yet, there has been no satisfactory resolution of this tension, but the attempt is ongoing. In the remainder of the paper I will outline some standard strategies, highlighting their promises and problems.

5 The Search for the Middle

While both assumptions are controversial, I shall assume that to some extent the previous points have been established. That is, that there is some substantial innate component

to our linguistic competence, but that viewing the conclusion of each successful PoS argument as motivating an independent innate trait would significantly over-burden any plausible biological or psychological theory of development. While the specific cases of innate linguistic knowledge are up for debate¹², I believe that there are enough varied examples of this style of argument that some such innate linguistic competence must be accounted for. Likewise, while there are proposals for how a highly complex language faculty could have plausibly evolved (e.g. Pinker & Bloom (1990)), I believe the evidence cited earlier strongly suggests that something must be done to lighten the load that linguistics places on evolutionary and developmental theory.

In the last couple decades, this tension has been at the heart of one of the major developments in linguistic theory: The Minimalist Turn. The central explanatory strategy of the Minimalist program involves asking what is minimally necessary for the acquisition and use of any language, and then seeing the extent to which apparently *sui generis* features of human language can be explained by positing only this minimal machinery. Taking the conclusions of PoS arguments as identifying linguistic phenomena to be explained, showing that they stem purely from the minimal requirements on the existence of language undermines the worries that these are specific features which must be ‘built in’ to the human phenotype by evolutionary and developmental stories, undermining the worries raised in the previous section. Such traits, as it were, come for free. I will identify several categories of traits of this sort, before turning to empirical examples.

What one views as the essential components involved in language use and acquisition will determine the form and scope of minimalist explanations. Standard generative assumptions have it that the functional core of a language is a mapping between publicly observable signs (e.g. sounds or gestures) and meanings. While both anatomical and psychological specification for language exists, the behavioural and perceptual capacities to produce and recognize such public signs are to a large degree shared with non-linguistic creatures, and so the ability to interact with these perceptual and motor systems seems to be an unavoidable feature of a language. The other ‘interface’, with

¹²See Culbertson et al. (2020) for a recent argument that the above case of adjectival ordering is *not* a successful PoS argument.

meaning/conceptualization is more controversial, as the degree of modification and re-organization that language imposes on thought seems much greater than that imposed on the sensory and motor systems. However, it is clear that non-linguistic animals must share in some of our representational capacities (e.g. perceptual, navigational, associative, etc.) and our ability to talk about even these features of our minds seems to require some interactions between linguistic capacities and those of broader cognition. If certain linguistic rules or constraints can be shown to be necessitated by these interactions between language proper and perception and action, on the one hand, and conceptualization, on the other, then, it is argued, they need not be independently specified as components of language-specific innate competence, and thus the burden on biology and psychology is minimized.

Alongside these ‘external’ requirements, there may be features ‘internal’ to the grammatical system that similarly seem to come along for free but which may be reflected in observable linguistic behaviour in unexpected ways.¹³ Such internal forces, molding the shape of possible languages, could come in several varieties. One category contains features of a grammar that are conceptually required: features of a grammar that no grammar could, in principle, lack. Call these ‘conceptually minimal traits’. Another contains features that could be absent from a grammar, but their absence would make the grammar more complicated, so that the default assumption should be that they are present. Call these ‘default minimal traits’. And finally, and most controversially, some traits are, for some purposes, optimal for a grammar, such that a grammar would be dysfunctional without them. Call these ‘optimal traits’.

These external and internal forces ensure that human grammars take specific forms. Specifically, they ensure that human grammars are capable of generating structures that can link observable signs to conceptual structures. Any linguistic phenomena that can be shown to follow from just these constraints will not require specific evolutionary and developmental accounting, beyond that which is required to explain what differentiates

¹³Note that ‘external’ and ‘internal’ are here used to mark the distinction between factors external or internal to *language* (or, equivalently, to the Language Faculty in the Narrow Sense (see fn. 2)), not between factors external or internal to the mind more generally.

language-using humans from other animals.

This explanatory strategy also raises the importance of so-called ‘functionalist’ linguistic theorizing. This approach aims to derive linguistic phenomena from non-linguistic features of cognition. If we can explain some linguistic behaviour as reflecting, say, non-linguistic constraints on memory, or perceptual capacities, then we do not need to incorporate features of our *grammar* which explain it. This again ‘minimizes’ the slack to be picked up by evolutionary and developmental accounts of language. The hope, then, is that some combination of these internal, external, and extra-linguistic pressures can explain all the apparently innate features of language, without the need to posit large quantities of independent and *sui generis* linguistic competence. I turn now to some plausible applications of these strategies.

At a most general level, external pressures require that any working language must be hierarchical, but linearizable. Linguistic structures must be hierarchically structured so that they can capture structural semantic relations such as the distinction between an agent and a patient, where the patient is, in some sense, more closely related to the verbal event than the agent. Plausibly, argument structure in general requires hierarchical structure. However, given that language is expressible in speech, the structures generated must be capable of being mapped onto a linear channel.¹⁴ Early generative work suggested that ‘structure dependence’ was a substantive principle of UG. However, we can now see that this was unnecessary. The role of structural relations in semantic interpretation ensures that languages will be sensitive to structural relations, without further stipulation, and thus without creating an additional biological trait for which we need a developmental or evolutionary story. If we further assume that the mapping from these hierarchical structures onto a linear channel is used *only* for production of a public sign, then we get the stronger claim that interpretation, an internal process, will be exclusively structure dependent, again without the positing of novel language-specific constraints.

More subtle examples will require getting a bit more linguistic detail on the table. As

¹⁴This is so even though some languages allow exceptions to linearity, as when signed languages allow constructions featuring simultaneous expression using each hand. I assume that the requirement that languages be linearizable is species-universal, even though not all languages need be strictly linearized.

noted earlier, in the transition between TG and P & P approaches, the burden of explaining grammaticality shifted from the ‘generation’ side, to the ‘constraint’ side. Rather than identifying a large set of rules, which apply only to very specific linguistic constructions, the rules for constructing expressions became bleached, with fairly minimal constraints on what they could apply to and fairly transparent relationships between the expression before and after application of the rule. This culminated in the Merge of minimalist grammar, which simply says that any two linguistic expressions can be combined to form a complex linguistic expression with both as constituents. This simplified generative component must then be balanced by constraints on its products, to prevent massive over-generation.

As Merge can apply equally to items which have already been Merged into the structure as to items previously outside of the structure (so called ‘internal’ and ‘external’ Merge, respectively), this freedom must be curtailed to prevent all languages being predicted to have free word-order. For example, in English, while some movement of expressions is required to generate certain constructions, as in:

5. Artemis will watch the movie.
6. Will Artemis watch the movie?
7. What will Artemis watch ~~what~~?

it is not the case that just anything can be moved:

8. *Watch Artemis will the movie.
9. *The movie watch Artemis will.
10. *What will who watch ~~what~~?

To prevent predicting that these latter examples are acceptable, there must be grammatical restrictions on movement. One such restriction is *Procrastinate*, a principle proposed in Chomsky (1995) stating that overt movement (i.e. movement with perceptible effects,

as indicated in 5-9 above) is always less favoured than covert movement.¹⁵ While Procrastinate allows that there may be all sorts of movement going on behind the scenes, changes to perceptible word order must be strongly motivated. The way this is cashed out is in terms of features of linguistic expressions. ‘Strong’ features must be moved to appropriate locations (where they can be ‘checked’) early on in a derivation, so that the effects of this movement can be heard, while ‘weak’ features may be moved later, inaudibly. Word-order variations can then be reduced to whether a given expression has a strong or a weak feature. In English, auxiliary verbs and *wh*-expressions have strong features, which leads to movement in questions, whereas in Chinese languages like Cantonese and Mandarin, these features are weak so that verbs and their arguments are produced in their semantically natural locations, and in French lexical verbs also feature strong features, so they are also raised. In this way, word-order variation can be reduced to different distributions of features in the lexicon.

However, this still leaves open the question of what motivates Procrastinate? It seems to be exactly the kind of language-specific innate knowledge that we have seen to be problematic. Richards (1997) argues that the empirical results of Procrastinate can be shown to follow from a general constraint on the interface between grammar and production. As we said earlier, for a language to be useable, the grammatical structures it generates must be *producible*, i.e. we must be able to behave in such a way as to generate public signs which can reliably indicate to others what expression we are intending to generate. This is the source of constraints on what is typically called the ‘PF interface’, where grammar meets phonology. One such constraint is that the hierarchical structure generated by the grammar must enable the unambiguous determination of a linear order by the externalization systems.¹⁶ A standard account of the capacity of natural language

¹⁵The examples of proposed explanations discussed in the remainder of this section are selected not on the grounds that they are particularly likely to be correct. Indeed, several of them are inconsistent with now widely adopted assumptions about the nature of grammar and its interfaces. They are instead chosen on the grounds that they particularly clearly exemplify some explanatory approaches within theoretical linguistics. My project is methodological, identifying the nature of a particular style of scientific inquiry, rather than first-order scientific, proposing an accurate account of the nature of human language.

¹⁶‘Unambiguous’ here indicates that the mapping from grammatical structure to phonological structure must be one-to-one or many-to-one, i.e. it must be a function. It does not, of course, require that the mapping back from public symbol to a grammatical structure is likewise unambiguous.

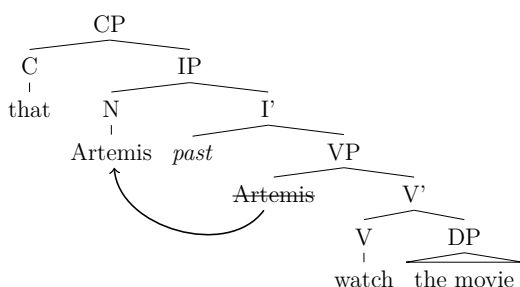
to move or displace elements, specifically the copy-theory of movement, poses problems for this requirement. According to this theory, when we find ‘displaced’ expressions, such as the question particle in “What will Artemis see?” which is semantically the agent of ‘see’ but is found sentence-initially, we should view the underlying structure as containing multiple instances of this expression, one in its phonologically-relevant position and one in its semantically relevant position, only one of which is pronounced. The question then is: how does the phonological system “know” which to pronounce. The structure seems phonologically ambiguous.

Richards’ proposal is that ‘strong’ features should be viewed precisely as phonological disambiguators. In a chain of copies, one must be marked as the one to receive pronunciation. Learning a language then consists in learning which copy is marked with which features. The predictions of Procrastinate then follow from the fact that movement of an expression with a strong feature will be unambiguous, as whichever copy retains this feature will be uniquely identifiable as pronounceable¹⁷, whereas movement of an expression without a strong feature will result in unpronounceable ambiguity. Thus, what seemed like a substantive piece of language-specific knowledge (Procrastination) can be shown to follow from requirements on any useable language.

Coming from the other side, Ramchand & Svenonius (2014) argue that what appears to be innate and language-specific knowledge of certain abstract properties concerning word order can be shown to stem from the LF-interface, where syntax meets semantics. Work in the ‘Cartographic Syntax’ program has shown that there is a cross-linguistically robust hierarchy of clausal structures, with Verb-Phrases (VP) occurring within Inflectional-Phrases (IP), which themselves occur with Complementizer-Phrases (CP).¹⁸ This can be seen in the highly simplified tree-structure for the embedded clause “that Artemis watched the movie”:

¹⁷This proposal can be extended to languages like German in which multiple copies of wh-expressions are, in certain constructions, pronounced.

¹⁸This cartographic program has identified much structure internal to these broad divisions, for example splitting VP into *vP* (‘little vP’) and VP (‘big VP’) accounting for subject/object asymmetries, but for our purposes we can focus only on these coarse distinctions. See Rizzi (2004).



The important thing for our purposes here is that the verbal domain includes both the verb specifying the action being performed, as well as all of its arguments, specifying the participants. This verbal domain is fully constructed before being embedded within the inflectional domain, which specifies morphologically required information such as tense and verb-argument agreement. The verbal subject is pronounced before any inflectional components due to movement out of the verbal domain, as argued for in Koopman & Sportiche (1991). This inflected clause is then further embedded within a CP, which adds information concerning discourse-relevant properties such as force. CP also provides a ‘landing site’ for moved elements from within the VP. And of course this structure can be further recursively embedded by larger structures, e.g. by a psychological verb such as ‘believes’.

Crucially, this hierarchical ordering, it is argued, is universal. Although surface word-order can vary, we will not find a language with underlying structures in which information about force is combined with verbal and argumental constituents before tense or agreement. The question again is: Why? Again, it seems like there are substantive structural features of our linguistic competence that are, in some sense, part of our biological endowment. But for the reasons stated above it is far from clear that positing substantive innate knowledge is viable. Ramchand and Svenonius’s argument is analogous to Richards’, but applied to the syntax-semantics interface rather than the syntax-phonology interface. Human conceptualization of the world, they argue, makes formal distinctions between events, situations, and propositions, and in order for the grammatical structures generated by the language faculty to be interpretable, analogous distinctions must be drawn by the grammar.¹⁹ The distinctions between VP, IP, and CP play these roles.

¹⁹Philosophers beware: as the following paragraph will make clear, the distinctions drawn by Ramchand and Svenonius are quite different from those typically drawn by philosophers with these terms. As

An event, for Ramchand and Svenonius, is constituted by some small number of participants and a specific property or relation holding between them. Paradigmatically, these will be dynamic, causal relations, but stative verbs require this class to be broader. *Artemis watching a movie* specifies an event. Events differ from situations in that the latter are *anchored*. That is, features of an event are connected to the broader context in ways that enable identification of a particular instantiation of the event-type. Most characteristically, this will be done with reference to a specific time. While the event-type identified by *Artemis watching a movie* can be instantiated at many times and many locations, a situation is identified only when these parameters have been set. Finally, a proposition involves relating a situation to the current discourse, particularly to the purposes of conversation. For example, while the assertion *Artemis was watching a movie* and the question *Was Artemis watching a movie?* may identify the same situation (if the time referred to is the same), they correspond to different propositions, as the speaker's attitude towards this situation differs. These distinctions thus generate a hierarchy, with higher categories constructed by expanding on lower ones.

If this is a true account of human (pre-linguistic) cognition, it provides an account of the syntactic claims noted above. If humans understand propositions as constituted by situations (plus discourse-relations), and situations as constituted by events (plus tense, agreement, etc.), then linguistic structures must, in order to package human conceptualizations, structure information in analogously nested ways. And of course this is exactly how the CP > IP > VP hierarchy does package this information. A VP specifies an event, with its participants and their actions/properties. The IP then adds 'anchoring' information, such as tense, to identify a situation. And the CP adds force markers, relating such situations to the discourse context to specify a proposition. This neat mirroring of conceptualization and grammar thus shows that the clausal hierarchy identified by the cartographic project need not be viewed as an optional piece of linguistic knowledge, which must be specifically encoded for, raising the problems noted earlier, but instead as

they note, Barwise & Perry (1981) distinguish situations from propositions by viewing the former as mereological parts of the latter, not with respect to discourse-relations. Similarly, a standard view in philosophy of language views declaratives and their corresponding interrogatives as expressing the *same* proposition.

a requirement on any possible interface between grammar and (human) cognition.²⁰

We have now seen two paradigmatic examples of ‘external’ constraints on grammar, and how such external constraints can help explain what initially seemed like arbitrary, and therefore deeply puzzling, features of UG. Turning now to ‘internal’ constraints, we can start with the most basic: what I called ‘conceptually minimal’ features of language.

Whatever one thinks about the project and prospects of conceptual analysis in general, it seems that all parties agree that any account of natural language must account for the ability to combine expressions to create novel, more complex, expressions. There are of course systems of communication which lack this feature (e.g. Paul Revere’s famous signaling system). But I am very hesitant to call such things ‘languages’. Given this, it seems that any account of human language will have to posit some mechanism, or suite of mechanisms, capable of taking multiple linguistic items and combining them. This role is played in contemporary generative theory by ‘Merge’.

Merge plays a unique role in the Minimalist program. Whereas the explanations given earlier (in terms of phonological and semantic interfaces) aimed to reduce linguistically specific and *sui generis* knowledge to that required by already existent constraints on human cognition, Merge is taken to be a genuine posit of the approach. Merge is what differentiates us from the non-linguistic animals. And it must indeed be accounted for by some positive evolutionary/developmental story. External pressures influence how Merge is reflected in observable behaviour, and how it can operate internally, but its existence is, from the perspective of linguistics, explanatory bedrock. The hope is that Merge is simple enough that the above arguments against biological accounts of substantial linguistic knowledge do not apply. This hope is strengthened by the fact that it is quite unclear what would be simpler than it: if this hunch is correct, then there could be no ‘intermediate stage’, and so Merge must simply be posited, perhaps as a catastrophic “great leap forward” as claimed in Chomsky (2015). Merge is viewed as a substantial and language-specific component of UG, and this proposal thus differentiates Minimalists

²⁰Note however, that the parenthetical ‘(pre-linguistic)’ is crucial here. If, as suggested by Hinzen & Sheehan (2013) and Dupre (2020), these features of cognition are not independent of language, but are instead made possible by human grammar, this form of reasoning would be circular.

from empiricists and anti-nativists (e.g. Ibbotson (2020)) who eschew language-specific innate endowments entirely.

This brings up the “default minimal” traits, which are not strictly required for the existence of a language, but which are simpler than any other traits that would do the job.²¹ Merge is, it is hoped, a trait of this sort. On the one hand, less powerful mechanisms could be imagined, such as the ability to combine elements but only to a finite degree of depth. As human behaviour is necessarily finite, one could even imagine that the positing of such mechanisms would be empirically adequate. However, it seems that such systems would be strictly more complex than Merge, incorporating as they do both the combinatory capacities of something like Merge, *and* the restriction on iterated re-application. For this reason, one would need strong motivation for imposing such a restriction, with unrestricted Merge as the default. On the other hand, one could posit more powerful mechanisms, such as a ‘tertiary Merge’, capable of combining three distinct expressions. Again, empirical observations could force such systems onto us, but without such pressures we are best off assuming binary Merge. The capacity to combine expressions is required for language, and Merge seems to allow for this in the minimal way, with both more and less powerful mechanisms seeming to require more substantive biological accounting.

Once again, we can leverage these minimal linguistic posits into more empirically surprising claims about language. For example, the impetus for favouring binary over ternary branching, and thus keeping the biological commitments relatively minimal, might seem disconfirmed by a variety of linguistic structures which seem to involve combining three or more linguistic items. Ditransitive verbs, such as ‘show’, which require both direct and indirect objects provide a classic example (witness: “Artemis showed Alexis the movie” vs. **“Artemis showed Alexis”*²²). These have traditionally been analyzed with a tertiary branching structure, as they were in early TG approaches. However, work by Larson

²¹See Chomsky (2021) for a recent discussion of the centrality of simplicity in evaluating grammatical theories, and for both empirical and theoretical arguments that a simple theory is more likely to be a true theory.

²²The acceptability of the latter sentence in certain circumstances, which initially seems problematic, actually further supports the claim above, as these circumstances are only those wherein the direct object is understood as elided. E.g. “Pim showed Ifeoma the movie. Artemis showed Alexis.”

(1988) and Kayne (1994) has shown that in these structures, as well as many others, the insistence on binary branching leads to significant empirical support. For example, asymmetric potential for binding is accounted for by viewing the indirect object as c-commanding the direct object:

11. Artemis showed Alexis₁ herself₁.
12. *Artemis showed herself₁ Alexis₁.

These data are accounted for perfectly by independently motivated principles of binding theory if binary branching is enforced, and the verbal domain is decomposed into nested ‘shells’, with the constituent containing the direct object within the larger constituent containing the indirect object. However, if we posit a tertiary branching structure, this explanation is no longer viable. Thus yet again we have a piece of grammatical knowledge, that ditransitive verbs decompose into a pair of binary-branching structures, which can be accounted for by appeal to apparently required or default features of a natural language.

The final, and most controversial, force for Minimalist grammatical explanation, is what I called ‘optimal traits’. These are traits that seem required not by just any linguistic system at all, nor by the simplest possible systems, but instead by systems which are, in some sense of the term, optimal. Of course, there are about as many different conceptions of optimality as there are desirable traits, and so any account of this type will require specification of exactly in what way the traits in question are optimal. Most frequently discussed in the literature are broadly computational considerations involving search space.²³

From the very beginning of generative linguistics, dependencies between linguistic constituents have taken center-stage in discussion of grammar. Agreement relations between verbs and their arguments [13], between anaphoric expressions and their antecedents [14], between functional expressions and their scope markers [15], between expressions as they are pronounced and as they are interpreted [16], and so on:

²³Such explanations are generally referred to by Chomsky as “3rd factor” explanations, appealing to general features of cognition, rather than to language-specific/UG (“1st factors”) aspects of mind, or to linguistic experience (“2nd factors”). (Chomsky (2005))

13. She sees the movie.
14. She₁ showed me her₁ favourite movie.
15. She either watched *True Romance* or *Pulp Fiction*.
16. Which movie₁ did she watch ~~Which movie₁~~?
17. Either Artemis will tell you which movie she watched or Alexis will ask you which movie you want to watch.

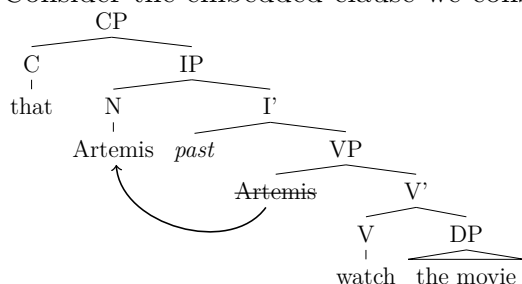
As 17 shows, such dependencies can iterate. We can easily come up with longer and longer cases, with highly intricate dependency patterns. Grammaticality generally depends on these dependencies. The grammatical system therefore needs a way of determining what depends on what, and whether these dependencies are legitimate. Some dependencies require the presence of certain elements (e.g. whenever we find an ‘either’, there had better be a suitable related ‘or’) while others require the absence of certain elements (non-reflexive anaphoric pronouns such as ‘him’ preclude local antecedents). Given that the dependent elements can be indefinitely far apart with respect to the linear order of the utterance, this can create a significant computational burden. This thus raises issues of computational optimality: are there ways of limiting the search space within a grammatical structure so as to ensure that its (un)grammaticality is relatively easily identifiable by the grammatical system?

Chomsky (2008) argues that a surprising range of linguistic facts can be derived by assuming certain features of the grammar which serve this role of minimizing search space. The central idea is that derivation of a complex linguistic expression proceeds in a series of ‘phases’, and that once a phase is complete, further operations have no access to its internal structure. Given that derivation can proceed beyond the construction of a phase, however, there must be some feature of the phase with which further grammatical processes can interact. We can thus differentiate between a ‘phase edge’, which must contain all information relevant to further derivation, and a ‘phase domain’, which is inaccessible to these further processes. Ensuring that requirements on grammatical dependencies are

met thus requires that the dependent elements can be found within the edge of a phase, rather than (merely) within the domain.

If this proposal about phases were true, it would significantly reduce the burden on the computational system tasked with ensuring that dependencies formed with linguistic expressions meet these conditions. Rather than having to check whether the relevant dependencies are consistent with every element of the structure (which can, in principle, be of any finite length), only a small subset of the structure (the phase edges) need be considered.²⁴ Thus, if we can trace observed linguistic facts to these types of computational constraints, the innate knowledge governing these facts need not be viewed as arbitrary, but is instead motivated by computational optimality. Chomsky sketched how such explanations can work.

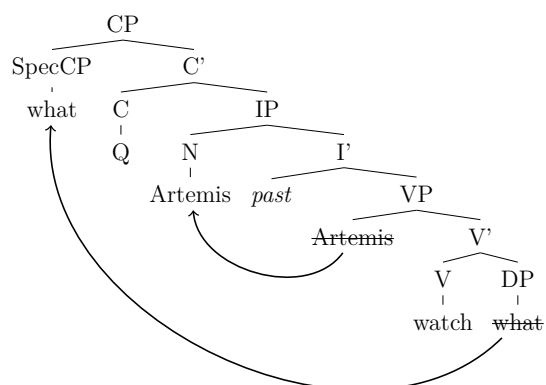
Consider the embedded clause we constructed earlier:



Assuming that CP is a phase, grammatical processes operating on this structure have access only to its edge, in this case the complementizer ‘that’. This explains why grammatical rules can specify whether a verb can take a CP-complement or not, but none which specify, say that it can only take a CP-complement featuring a transitive verb. Rules of the latter kind would require looking inside the phase.

Dependencies can be generated between phase-internal elements and those higher up in the tree, as in *wh*-questions, but only by moving these elements to the edge:

²⁴I am making the point in terms of the checking of the legitimacy of a structure, but it could be made equally well by instead discussing conditions on the generation of a structure. If only phase edges are relevant for future derivation, the system has much less work to do in determining which ‘continuations’ of a tree are possible. This latter approach is more in line with the ‘crash-proof’ approach to syntax pioneered by Frampton & Gutmann (2002).



Here the head (Q, an unpronounced complementizer indicating interrogative force) and its specifier (‘what’) form the phase edge, and so can be accessed by further operations, such as further movement (as in the question “What₁ did Alexis say *what*_T Artemis watched *what*_T?”).

However, there are some configurations in which these long-distance dependencies cannot be formed consistently with these constraints on operations on phase-internal material. For example, if we assume that DP and CP are phases, we can now explain why complex DPs generate movement islands:

18. *What did Alexis believe the rumour that Artemis watched?

18 is predicted to be impossible, on the grounds that, while the DP ‘what’ is able to move to SpecCP in the embedded clause “what₁ that Artemis watched *what*_T”, as in the tree above, from here it cannot move to the edge of the DP clause. Movement is motivated by feature-checking, and there is no motivation for such a move.²⁵ Wh-expressions have force features (+Q), but force is checked at the clausal level, not by arguments. Thus, this expression is trapped within the DP, and the dependency between the interrogative feature of the matrix clause and the question-particle originating in the embedded verbal phrase cannot be formed, leading to ungrammaticality.²⁶

²⁵This explanation would need to be modified if we follow Chomsky (2013) and view movement (“internal Merge”) as freely applicable, in line with the general thrust away from constraints on generation and towards constraints on the legitimacy of the resulting structure. As in the rest of the paper, the goal here is not to present the correct linguistic approach, but rather to identify explanatory strategies for resolving the tension between complexity and simplicity in language. The proper balance here, and the theoretical machinery underlying such a balance, will of course depend on many subtle empirical details.

²⁶Some explanations can also of course combine elements from multiple such strategies. One example

Both these ‘internal’ arguments stemming from computational optimality, and the ‘external’ arguments focusing on constraints imposed by the interfaces, have the same form: show that a well-ordered linguistic system would have such-and-such properties, and then show that the presence of such properties allows us to explain otherwise unexpected features of linguistic behaviour. This goes some distance to dissolving the tension between the apparent complexity of language and the simplicity seemingly required by the evolutionary history of language. The final strategy works slightly differently, in effect aiming to show that certain behavioural observations which have been taken to reflect linguistic competence are actually products centrally of non-linguistic cognition.²⁷

Perhaps the most famous example of strategies along these lines is the argument in Miller & Chomsky (1963) that multiply center-embedded sentences, like **“The mouse the cat the dog chased caught squeaked”*, while near-uniformly viewed as unacceptable, should be viewed as grammatical. One strand of this argument involves showing that a grammatical system which precluded such expressions would be dysfunctionally more complex than one which allowed them. The other involves showing how to explain their unacceptability without appeal to the ungrammaticality of their underlying structure. In this case, this was done by appeal to perceptual mechanisms involved in parsing. Assume that each noun-phrase encountered generates a ‘task’ for the parser: find the verb this is an argument for. The successive presentation of three noun-phrases then requires that

of this is Adger (2018)’s account of the adjective-ordering facts discussed above. Adger’s proposal is that some semantic classes of adjectives are found structurally closer to the noun because different orders of intersection can make the process of sentence verification more efficient. So, we say “four red houses”, rather than “red four houses”, because it is more efficient to first identify the perceptible red houses, and then the subsets of this set with 4 members, than it is to identify the set of sets containing four houses and then subsequently identify which of these sets have only red members. Such an explanation thus straddles the boundary between general efficiency-based explanation, and specific appeals to the interface between language and broader cognition.

²⁷For reasons of space, I will not discuss perhaps the most promising recent work in this line of inquiry, the ‘emergentist’ approach to parameters, developed in Biberauer et al. (2014) (see also Epstein et al. (2021)), which aims to show how existence of linguistic parameters stems from the way that the grammar generates structures which are, from the perspective of the perceptual/motor interfaces, under-specified. This work is akin to Richards’ proposal discussed earlier, in that it aims to derive linguistic complexity from demands made by the interfaces, but differs in that the solutions to these interface tensions are not resolved by features of the grammar, but instead by viewing different languages’ parameter settings as different possible cognitive arrangements capable of enforcing a mapping between abstract grammatical structures and linearized performances and perceptual stimuli. Parameters on this view are ways that extra-grammatical factors can ‘fill the gaps’ left by the grammar. As always, determining whether this is best viewed as an instance of elimination or reduction is subtle, but not particularly pressing.

the parser simultaneously attempt three tasks, which overwhelms the system's memory capacity, leading to an unacceptability result.

Beyond these classical cases, much work in 'functional' linguistics has suggested similar lines of argument. In opposition to the phase-based account of *wh*-islands discussed above, Hofmeister & Sag (2010) (H&S) argue that the unacceptability of expressions violating 'island-constraints' does not reflect constraints imposed by the grammar, but instead reflects the difficulty the parser faces in identifying the location at which the moved element is to be semantically interpreted. Hofmeister and Sag identify a variety of features of sentences which make them difficult to parse, even for uncontroversially grammatical expressions, and claim that *wh*-islands typically reflect nothing more than the confluence of such processing factors. For example, the complex-NP constraint, noted earlier, precludes extracting a *wh*-expression from within a larger complex Noun-phrase (or Determiner-phrase):

19. *What did Alexis believe the rumour that Artemis watched?

According to H&S, this unacceptability stems from various processing difficulties created by the nominal expression ('the rumour') intervening between the *wh*-expression and the 'gap' at which it is to be located. For one thing, definite NPs like 'the rumour' are taken to be discourse-familiar, and so when they are encountered, a search for previous mention of their referent is initiated. When no antecedent is found, as would be typical in linguistic 'experiments', this creates perceptual uncertainty. Further, complex-NPs are, as the name suggests, complex, requiring substantial structure-generation by the parser. This in itself predicts a certain degree of unacceptability. And so on. In this way, functionalists apply the strategy of Chomsky and Miller to a wide range of perhaps intuitively grammatical phenomena, arguing that upon reflection grammar is not the culprit. This again reduces the burden on evolutionary accounts of the language faculty.²⁸

²⁸Note that this is a useful strategy in the context under discussion only if the functioning of the perceptual systems can itself be explained without positing language-specific perceptual strategies. In many cases this will be the case, as in the examples described above which appeal to constraints on memory or simultaneous processing power, which would be features of any biological system. Note also that this strategy is different from that of functionalists like Hawkins (2004, 2014) who argue that these functionalist pressures on extra-linguistic systems are 'internalized' by the grammar. Such

6 Worries and Open Questions

Thus far, I have argued that the relatively recent birth of human language and the apparent complexity and arbitrariness of human linguistic competence create a stark tension for explaining what language is, where it comes from, and why it has the properties it does. In the previous section, I identified two main strategies for resolving this tension. Firstly, if a linguistic observation can be derived from the properties of any well-functioning linguistic system, this relieves some of the pressure on an evolutionary/developmental account of language. And secondly, if we can appeal to features of pre-existing non-linguistic cognition to account for some linguistic behaviour, this reduces the scope of the phenomena an account of linguistic development is supposed to account for, thus making this task significantly easier. In this final section, I will briefly identify some crucial remaining difficulties for this program.

Of course, as with any linguistic, or indeed scientific, theory, there are many observations and phenomena yet to be accounted for. Even for those adopting a broadly Minimalist approach, which phenomena remain to be explained remains a topic of controversy, depending on what the strictly linguistic phenomena are taken to be, and how plausible one finds the Minimalist accounts provided for them. Almost 20 years ago now, Newmeyer (2003) claimed that “the empirical coverage of the [Minimalist Program] is vastly reduced compared to [Government and Binding Theory].” (p. 589). In the subsequent decades, this gap has gotten smaller, but my guess is it is still true. Precisely the features of Minimalism that make it attractive from the perspective of the tension discussed in this paper make empirical coverage much harder to come by. As detailed above, in previous approaches, the flexibility of the theory made modification to cover novel phenomena relatively straightforward. However, the Minimalist insistence that it be shown how these phenomena derive from otherwise unavoidable features of the linguistic or extra-linguistic mind creates significant barriers to empirical adequacy. On the flip-side, when such derivations can be provided, these results are liable to be significantly

explanatory styles are problematic in light of the relatively recent introduction of language into the human cognitive system.

more stable, in virtue of relying on less substantive developmental and evolutionary assumptions. For these reasons, I will not focus on outstanding empirical phenomena, but instead on broader ‘conceptual’ difficulties with Minimalist explanation.

One major worry is that, while the ideology surrounding Minimalist theorizing involves analyzing linguistic complexity with appeal only to those aspects of mind that are absolutely required for a system to possess language, in actual practice, more than this is inevitably made use of. Let us assume that Merge is indeed a suitably minimal operation for grammatical construction, and further that the lexical items to which Merge applies cannot, on pain of massive overgeneration or vacuity, be simply indistinguishable, so that the existence of lexical features is minimally required as well. These assumptions alone, however, are insufficient for just about any of the Minimalist explanations provided in the literature: for these we need also to posit the particular features actually found in the human lexicon. Depending on the theory, these might include grammatical categories, phonological features, functional morphological features, semantic universals, and so on. But it is far from obvious that such features are in any sense ‘minimally required’.²⁹ For example, in Richards’ account of why some copies get pronounced and other do not, he appealed to strong and weak features of lexical items. But, as he notes (pp. 1-2) such features themselves seem in need of reduction to something more basic, and thus biologically explicable.

A related point is that in appealing to non-linguistic aspects of cognition, as Ramchand and Svenonius do in their account of the CP > IP > VP hierarchy, these aspects had better themselves be more easily explicable than the linguistic traits they are supposed to explain. But again it is far from clear that this is the case. On the one hand, conceptualizing the world in terms of propositions, which relate to the current discourse in a certain way, is not clearly found in any non-human organisms. Indeed, it is not clear that

²⁹Similar remarks could be made concerning proposals which posit additional operations, such as ‘Label’, which identifies the grammatical category of a newly Merged expression (as in Chomsky (1995)). Although various strands in recent research suggest that such additional operations are not needed. See e.g. Narita (2014) for an attempt to do syntax without labeling, Hornstein & Pietroski (2009) for an attempt to reduce labeling to more basic, and potentially genuinely minimal, operations, and Collins & Seely (2020) for an argument that the widely adopted strategy stemming from Chomsky (2013) in fact makes no appeal to such additional mechanisms, and instead reduces to 3rd factors properties.

properties of tense or aspect are found in comparative psychology either. And on the other, such an explanatory strategy appears to clash with the possibility of explaining the distinctiveness of human cognition precisely by appeal to the human capacity for language.

This is not necessarily to say that these explanations are misguided. Replacing one puzzle with another is often a significant step forward. And it may well be that replacing the question of where innate linguistic rules and constraints come from with the question of where lexical features come from is progressive in just this way. Likewise with replacing innate linguistic knowledge with innate conceptual schemas. But replacing one question with another is not the same as answering the question, and these cases both show that the Minimalist resolution to the tension discussed above is far from complete.

Another major worry is raised by some of the appeals to derivation from minimal traits in Minimalist explanation. Take the statement that Minimalist explanations involve showing that some linguistic phenomenon follows from what is minimally required for possession of a language. If this latter expression “what is minimally required for possession of a language” is simply taken to refer to Merge, then I have no problem with it. Merge is admitted to be a genuine posit of the theory, not derivable from some other source, and does seem to fit the bill for a minimal mechanism for language. However, once we go beyond mere appeal to Merge, and include in this designation also requirements on interfaces, for example, we seem to be incurring greater empirical costs. Consider again Richards’ claim that it is a requirement of the PF interface that it be able to ‘disambiguate’ a grammatical structure for the purposes of linearization. While I agree that this is a requirement on the public use of a language, it does not follow from the core posit of Minimalist theorizing, Merge. There seems to be a sort of ambiguity at work here: some linguistic traits are necessitated by the presence of Merge in the sense that they are entailed by this operation. Others, however, are necessitated by Merge in the sense that Merge could not do what it does (e.g. allow for externalized language) without them. But these latter do not ‘come for free’ in the way the former do, once we have posited Merge in the first place. In general, that a given trait is required for some other trait

to be useful/useable does not provide an account of where the former trait came from. Indeed this interdependence between traits has historically been seen as a deep worry for evolutionary biology.

If the Minimalist proposal is really to show that all the traits of human language can be traced back to one catastrophic development in human pre-history, this poses a deep worry. The existence of Merge is not on its own sufficient for interfacing with the perceptual-motor system, as innovations such as Richards' rules of copy-pronunciation are required. And indeed, Huybregts (2017) has argued that Merge predated human ability to externalize linguistic symbols. However, we can instead view proposals like Richards' as providing accounts of why these independent features of our broader linguistic competence were incorporated into the human lineage. Assuming that these are not themselves too baroque or too numerous, this strategy could still take us some ways towards resolving the tension.

Yet more problematic than these appeals to traits required for some specific application of Merge are appeals to 'optimality'. Again, it may well be that a phase-based system is computationally more efficient than one which didn't neatly package grammatical information. But showing that some organismic organisation is more functional than some other is not in general sufficient for explaining how it came to be so. Chomsky (2005), due in part to the reasons discussed earlier, is very clear that the notion of optimality in question is not the same as that found in adaptationist approaches to biology, and need not imply optimal capacities for any behavioural function. So the worry is not simply that pure adaptationism is false. Chomsky instead appeals to broader notions of optimality as posited in 'evo-devo' accounts of organismic self-organisation, and the organisation of physical systems more generally. Organisational 'laws of form' are appealed to in explaining regularities at all levels, from ensuring that soap bubbles adopt spherical shapes, to making it so that spotted organisms can develop striped tails but striped organisms cannot develop spotted tails (Murray (1988)). The hope here is that analogous forces acting on neural/psychological development and behaviour could ensure that linguistic processes and structures are optimal in this sense, but as Chomsky is very keen to note,

this is, as of yet, little more than a hope of future results. We are a long way away from deriving linguistic optimality from such basic principles.³⁰

What these worries have in common is that they point to further assumptions that apparently must be made in order for these Minimalist ‘reductions’ to work. It is not clear that these assumptions are correct. But it also is not clear that they are not. The question for working linguists (and psychologists, philosophers, etc. interested in the foundations of language) is whether these assumptions are more or less troubling than the tension which they purport to dissolve.

7 Conclusion

I hope, in this paper, to have presented as clear a case as possible for the tension that animates recent Minimalist theorizing in generative linguistics. On the one hand, decades of early generative work uncovered large amounts of apparently *sui generis* linguistic information composing and constraining human linguistic development. On the other, evolutionary and developmental considerations suggested that all this linguistic structure simply couldn’t be viewed as both innate and independent. This has led to a variety of strategies for accounting for linguistic complexity with a minimum of linguistic posits. Centrally, this involves showing either that traits that appear to be arbitrary are instead either necessitated by the core properties of human language, or are required in order for language to be usable, or showing that a linguistic system with these properties is, contrary to appearances, ideally organized, or finally showing that apparently linguistic traits are instead reflective of extra-linguistic cognition. While I believe this is the most promising approach to explaining human linguistic cognition and its evolutionary origins, I have closed by flagging some of the main remaining open questions.

³⁰Cherniak (2012, 2006) presents some suggestive work on optimal organization of neural structures, but one should be wary of inferring from these any claims about optimal processing at the *psychological* level.

References

- Adger, David. (2018). The autonomy of syntax. In *Syntactic structures after 60 years*, (pp. 153–176). De Gruyter Mouton.
- Barwise, Jon., & Perry, John. (1981). *Situations and attitudes*. Cambridge, MA: MIT Press.
- Biberauer, Theresa., Holmberg, Anders., Roberts, Ian., & Sheehan, Michelle. (2014). Complexity in comparative syntax: The view from modern parametric theory. *Measuring grammatical complexity*, (pp. 103–127).
- Bickerton, Derek. (1995). *Language and Human Behavior*. Seattle, WA: University of Washington Press.
- Cavalli-Sforza, Luigi Luca. (2001). *Genes, peoples, and languages*. Univ of California Press.
- Chater, Nick., Clark, Alexander., Goldsmith, John A., & Perfors, Amy. (2015). *Empiricism and language learnability*. OUP Oxford.
- Cherniak, Christopher. (2006). Innateness and brain-wiring optimization: Non-genomic nativism. In *evolution, rationality and cognition*, (pp. 115–124). Routledge.
- Cherniak, Christopher. (2012). Neural wiring optimization. *Progress in brain research*, 195, 361–371.
- Chomsky, Noam. (1957/2002). *Syntactic structures*. Walter de Gruyter.
- Chomsky, Noam. (1965). *Aspects of the Theory of Syntax*. The MIT Press.
- Chomsky, Noam. (1981). *Lectures on government and binding: The Pisa lectures*. Walter de Gruyter.
- Chomsky, N. (1995). *The Minimalist Program*. The MIT Press.

- Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.) *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3, (pp. 104–131). Oxford University Press New York.
- Chomsky, Noam. (2005). Three factors in language design. *Linguistic inquiry*, 36(1), 1–22.
- Chomsky, Noam. (2008). On phases. *Current Studies in Linguistics Series*, 45, 133.
- Chomsky, Noam. (2013). Problems of projection. *Lingua*, 130, 33–49.
- Chomsky, Noam. (2015). *What kind of creatures are we?*. Columbia University Press.
- Chomsky, Noam. (2021). Simplicity and the form of grammars. *Journal of Language Modelling*, 9(1), 5–15.
- Cinque, Guglielmo. (2010). *The syntax of adjectives: A comparative study*, vol. 57. MIT press.
- Collins, C., & Seely, TD. (2020). Labeling without labels. NYU/Eastern Michigan University.
- Culbertson, Jennifer., Schouwstra, Marieke., & Kirby, Simon. (2020). From the world to word order: deriving biases in noun phrase order from statistical properties of the world. *Language*, 96(3), 696–717.
- Dupre, Gabe. (2020). What would it mean for natural language to be the language of thought? *Linguistics and Philosophy*, (pp. 1–40).
- Epstein, Samuel D., Obata, Miki., & Seely, T Daniel. (2021). Is linguistic variation entirely linguistic? In *A Minimalist Theory of Simplest Merge*, (pp. 82–110). Routledge.
- Frampton, John., & Gutmann, Sam. (2002). Crash-proof syntax. In S. D. Epstein, & T. D. Seely (Eds.) *Derivation and explanation in the Minimalist Program*, (pp. 90–105). Wiley Online Library.

- Griffiths, Paul E. (2001). Genetic information: A metaphor in search of a theory. *Philosophy of Science*, 68(3), 394–412.
- Guasti, Maria Teresa. (2017). *Language acquisition: The growth of grammar*. MIT press.
- Hale, Mark., & Reiss, Charles. (2003). The subset principle in phonology: why the tabula can't be rasa. *Journal of Linguistics*, 39(2), 219–244.
- Hauser, Marc D., Chomsky, Noam., & Fitch, W Tecumseh. (2002). The faculty of language: what is it, who has it, and how did it evolve? *science*, 298(5598), 1569–1579.
- Hawkins, John A. (2004). *Efficiency and complexity in grammars*. Oxford University Press on Demand.
- Hawkins, John A. (2014). *Cross-linguistic variation and efficiency*. OUP Oxford.
- Hinzen, Wolfram., & Sheehan, Michelle. (2013). *The philosophy of universal grammar*. OUP Oxford.
- Hofmeister, Philip., & Sag, Ivan A. (2010). Cognitive constraints and island effects. *Language*, 86(2), 366.
- Hornstein, Norbert., & Pietroski, Paul. (2009). Basic operations: Minimal syntax-semantics. *Catalan journal of linguistics*, 8, 113–139.
- Huybregts, MAC Riny. (2017). Phonemic clicks and the mapping asymmetry: how language emerged and speech developed. *Neuroscience & Biobehavioral Reviews*, 81, 279–294.
- Ibbotson, Paul. (2020). *What it takes to talk: exploring developmental cognitive linguistics*, vol. 64. Walter de Gruyter GmbH & Co KG.
- Jackendoff, Ray. (1977). *X-bar Syntax: A study of phrase structure*. Linguistic Inquiry Monographs. MIT press Cambridge, MA.
- Kayne, Richard S. (1994). *The antisymmetry of syntax*. 25. MIT Press.

- Koopman, Hilda., & Sportiche, Dominique. (1991). The position of subjects. *Lingua*, 85(2-3), 211–258.
- Larson, Richard K. (1988). On the double object construction. *Linguistic inquiry*, 19(3), 335–391.
- Lasnik, Howard., Depiante, Marcela A., & Stepanov, Arthur. (2000). *Syntactic structures revisited: Contemporary lectures on classic transformational theory*, vol. 33. MIT Press.
- Mehler, Jacques., Jusczyk, Peter., Lambertz, Ghislaine., Halsted, Nilofar., Bertoncini, Josiane., & Amiel-Tison, Claudine. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143–178.
- Miller, George A., & Chomsky, Noam. (1963). Finitary models of language users. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.) *Handbook of Mathematical Psychology*. Wiley and Sons.
- Murray, James D. (1988). How the leopard gets its spots. *Scientific American*, 258(3), 80–87.
- Narita, Hiroki. (2014). *Endocentric structuring of projection-free syntax*, vol. 218. John Benjamins Publishing Company.
- Newmeyer, Frederick J.. (2003). On nature and language, and: The language organ: Linguistics as cognitive physiology, and: Language in a darwinian perspective (review). *Language*, 79(3), 583–599.
- Newmeyer, Frederick J. (2004). Against a parameter-setting approach to typological variation. *Linguistic Variation Yearbook*, 4(1), 181–234.
- Pinker, Steven., & Bloom, Paul. (1990). Natural language and natural selection. *Behavioral and brain sciences*, 13(4), 707–727.
- Ramchand, Gillian., & Svenonius, Peter. (2014). Deriving the functional hierarchy. *Language sciences*, 46, 152–174.

-
- Reboul, Anne Colette. (2015). Why language really is not a communication system: a cognitive view of language evolution. *Frontiers in Psychology*, 6, 1434.
- Richards, Norvin. (1997). In full pursuit of the unspeakable. In *Proceedings of NELS*, vol. 28.
- Rizzi, Luigi. (Ed.) (2004). *The Structure of CP and IP: The Cartography of Syntactic Structures, Volume 2*. Oxford University Press.
- Ross, John Robert. (1967). *Constraints on variables in syntax..* Ph.D. thesis, Massachusetts Institute of Technology.
- Scott, Gary-John. (2002). Stacked adjectival modification and the structure of nominal phrases. *Functional structure in DP and IP: The cartography of syntactic structures*, 1, 91–120.
- Stowell, Timothy Angus. (1981). *Origins of phrase structure*. Ph.D. thesis, Massachusetts Institute of Technology.
- Tattersall, Ian. (2017). The material record and the antiquity of language. *Neuroscience & Biobehavioral Reviews*, 81, 247–254.
- Tattersall, Ian. (2019). The minimalist program and the origin of language: A view from paleoanthropology. *Frontiers in psychology*, 10, 677.
- Yang, Charles. (2015). Negative knowledge from positive evidence. *Language*, 91(4), 938–953.