The Place of Universal Grammar in the Study of Language and Mind: A Response to Dabrowska (2015)

DOI 10.1515/opli-2016-0017
Received September 27, 2015; accepted August 2, 2016

Abstract: Generative Linguistics proposes that the human ability to produce and comprehend language is fundamentally underwritten by a uniquely linguistic innate system called Universal Grammar (UG). In her recent paper What is Universal Grammar, and has anyone seen it? Ewa Dabrowska reviews a range of evidence and argues against the idea of UG from a Cognitive Linguistics perspective. In the current paper, I take each of Dabrowska’s arguments in turn and attempt to show why they are not well founded, either because of flaws in her argumentation or because of a careful consideration of the available empirical evidence. I also attempt to demonstrate how evidence from the fields Dabrowska reviews actually supports the notion of UG. However, arguments are additionally presented in favor of integrating an understanding of domain-specific UG with an understanding of domain-general cognitive capacities in order to understand the language faculty completely.

Keywords: Universal Grammar, I-language, psycholinguistics, neurolinguistics, evolutionary linguistics, language acquisition

1 Introduction

Universal Grammar (UG) is the idea that the human ability to use language is derived from an underlying innate system in the mind for specifically formulating linguistic structure. It is an idea that has generated a great deal of controversy. The main opposition to this idea of ‘linguistic nativism’, or Generative Linguistics, comes from Cognitive Linguistics. Cognitive Linguistics suggests language is grounded in domain-general representational and/or processing capacities, rather than language-specific ones, though there is huge variation in ideas about how this might work in practice. Some Cognitive Linguists suggest exposure to statistical language usage patterns in the environmental stimulus is sufficient to formulate the possible outputs of the language system (e.g. Barlow & Kemmer, 2000; Bybee & Hopper, 2001), arguing, for example, that usage patterns determine the strength of the neural networks in the brain that are responsible for language, and that language is the product of a purely real-time set of weights-and-balances that are determined by usage statistics (a ‘stochastic grammar’). We will see in the course of the current paper, however, that usage patterns alone cannot account for the nature of human grammar. Meanwhile, other Cognitive Linguists suggest there are no ‘grammatical’ neural networks, and language is merely an artifact of a combination of conceptual structure and real-time processing resources such as working-memory, prediction, and control over the activation and inhibition of information. And still other Cognitive Linguists,
the ‘general nativists’, argue that there is an innate underlying representational system that is not itself grammatical but from which a representational system for grammar can be formulated.

Dabrowska (2015), What is Universal Grammar, and has anyone seen it? (hereafter referred to simply as ‘D’) falls into the latter category of Cognitive Linguistics. D reviews a range of different kinds of evidence, arguing that the human mind possesses a biologically determined ‘language making capacity’ that arises out of domain-general capabilities. In the current response to D, her cited evidence will be reviewed again -- alongside additional facts -- and it will be argued that, taken together, it does support the existence of a language-specific genetic endowment that underpins language, or UG, after all. Secondly, it will be argued that conceptual structure and domain-general cognitive faculties are indeed critical components of language production and comprehension, and therefore need to be considered alongside domain-specific UG in order to gain a complete understanding of human language. In short, whilst the current paper argues in favor of UG, it also argues Generative Linguistics needs to be integrated into Cognitive Psychology more broadly, especially as far as conceptual and real-time processing mechanisms.

The remainder of the current paper will adopt the following structure: a summary of D’s main arguments, replies to D’s arguments, a general discussion, and a concluding statement.

2 Dabrowska’s arguments against UG

2.1 Varying definitions of UG

According to D’s opening statement, one key issue with UG is that there does not appear to be one good definition of what the term actually means. She cites various quotations about how different linguists have attempted to define UG, explaining that UG is mostly regarded as the system of formal universals (and their variants) that are found throughout all human languages, i.e. linguistic principles and their parameters. In particular, she cites Chomsky (1986) specifying a list of such universals:

‘X-bar theory, binding theory, Case theory, theta theory, bounding theory ... and so forth -- each containing certain principles with a limited degree of parametric variation. In addition there are certain overriding principles such as the projection principle, FI (full interpretation), and the principles of licensing. . . [UG also contains] certain concepts, such as the concept of domain ... and the related notions of c-command and government,’ (p.2).

According to D, the problem with Chomsky’s 1986 list is that ‘every major development in the theory since then was accompanied by very substantial revisions... [that are] radically different in Minimalism’ (p.2). Her statement is situated with various citations of different linguists’ hopelessly variable lists of possible universals (e.g. Baker, 2001; Fodor & Sakas, 2004) and quotes from others about how futile the task of enumerating universals is (e.g. Newmeyer, 2008).

2.2 Universals

If UG underlies all human language, then one would expect to see some properties that occur in all human languages -- universals. D cites functionalists such as Evans and Levinson (2009), who:

‘give counterexamples to virtually all proposed universals, including major lexical categories, major phrasal categories, phrase structure rules, grammaticalized means of distinguishing between subjects and objects, use of verb affixes to signal tense and aspect, auxiliaries, anaphora, and WH-movement’ (p.2).

Under the functionalist view, the presence of these counterexamples turns “universals” into mere trends or ‘inductive generalizations about observable features of language’ rather than the ‘cognitive or “deep” universals’ of UG. The other issue D raises is that the UG theory is apparently unfalsifiable: universals may be selected by certain languages, but need not be (Newmeyer, 2008). If some typologist finds a language...
without a given “universal” Generative Linguists can simply claim the language did not select it; meanwhile, the fact that many other languages do exhibit the feature is treated as evidence for UG.

2.3 The Poverty of the Stimulus

The statistical usage patterns in the linguistic stimulus that one hears when one is acquiring language as a child is not enough by itself to build a mental grammar to derive all the possible constructions of a language, which is why we must also postulate UG. Human language is infinitely productive. That is, it has an unbounded capacity for creativity – i.e., producing novel output – that outflanks what can be heard in one’s language-learning environment by an infinitely large order of magnitude. Chomsky illustrated this with his famous sentence Colorless green ideas sleep furiously (Chomsky, 1957). His point about this sentence is that one can always form novel combinations of linguistic units to create entirely unique sentences. D attempts to fight the Poverty of the Stimulus idea by citing work by Pullum & Scholtz (2002) that analyzes a selection of four supposedly under-represented phenomena in the linguistic stimulus. The study shows that children have in fact heard tokens of the correct structures in their input; or in some cases that linguists were simply wrong about the information children need to abstract from the stimulus. D says that ‘most expositions of the poverty of the stimulus argument in the literature do not take the trouble to establish the truth of the premises’ (p.9), the relevant premise being that ‘the relevant data are not available in the input, or not frequent enough in the input to guarantee learning’. In other words, D’s argument against the Poverty of the Stimulus argument for UG is simply to suggest that it does not exist.

D also talks about the related issue of children not being told which grammatical structures do not work (i.e. they lack negative evidence). This is similar to the Poverty of the Stimulus argument in that it is assumed that since children develop certain grammatical phenomena without being told that alternatives do not work, their language development must be guided by UG principles. D responds to this argument by pointing out that children may get some ‘indirect negative evidence in the form of requests for clarification and adult reformulations of their erroneous utterances’. She also cites studies showing that corrections from adults can actually improve children’s grammars (e.g. Chouinard & Clark, 2003; Demetras et al., 1986; Saxton, 2000; Saxton et al. 1998).

2.4 Neuroscience of language

D identifies two issues with neuroscientific evidence that seems to support the existence of UG. The first concerns neural interactivity; namely, additional neural substrates to those within the Perisylvian cortex (broadly speaking, the brain’s ‘language centers’) are sometimes activated during language processing, and sometimes the language centers are activated during ostensibly non-language related tasks (e.g. Anderson, 2010; Stowe et al., 2005).

The second issue D raises with the neuroscientific arguments for UG is that ‘the local connectivity in the language areas (as well as other areas of the brain) [is] not genetically specified but emerge[s] as a result of activity and [its] position in the larger functional networks in the brain (Elman et al., 1996; Müller, 2009; Anderson et al., 2011; Kolb & Gibb, 2011)’ (p.6).

2.5 Genetics of language

In her paper, D talks about the “double dissociation” between the language and general cognitive functions of S(pecific) L(anguage) I(mpairment) populations versus W(illiams) S(yndrome) patients. The former group has good general cognition and poor language function, whilst the latter group is in the reverse situation, hence the term “double dissociation”. Both conditions are genetically determined. Double dissociations across cognitive faculties and different populations like this are thought to point to (genetic and/or neural)
autonomy between the cognitive faculties in question. In this case, it points to language being separate from the rest of cognition – an indication of UG. After reviewing a variety of literature, D concludes that ‘the dissociation is, at best, partial: older WS children and adolescents have relatively good language in spite of a severe cognitive deficit; SLI is a primarily linguistic impairment,’ (p.5). However, D goes on to question the double dissociation on the grounds that Stojanovik et al. (2004) showed that SLI and WS participants both performed poorly on linguistic tasks (although SLI participants did outperform the WS patients on non-verbal tasks).

2.6 The ease and speed of language development

Children develop complete inventories of the sound systems of their language within the first couple of years of life, much being completed within the first year, and they move through the major stages of developing their grammar by about five years old (e.g. Guasti, 2002). Furthermore, they learn their language subconsciously; that is, simply by exposure to it and without any formal instruction (e.g. Chomsky, 1962). The ease and speed of the language acquisition process has often been pointed to as evidence for the guiding hand of an underlying UG in the structure of the developing brain.

D attacks the assertions that linguistic exposure automatically permits language development in children without any formal instruction. She first calculates a conservative estimate of children’s linguistic exposure using the Manchester Corpus (Theakston et al., 2001), which she puts at around 42 million words between ages one and five, using this to argue against the quick and subconscious automaticity of language development as such development is, in fact, contingent on substantial input. Furthermore, D argues the input must be of a certain quality. She cites studies of healthy children who grew up not being able to interact with deaf parents and consequently developed articulation and comprehension problems (Todd & Aitchison, 1980; Sachs et al., 1981). The situation for these children improved when they were able to interact with functional speakers.

2.7 The maturational timetable

D briefly discusses work that shows language develops within a biologically determined timetable at the end of which a critical or sensitive period is reached. Thereafter, children’s development capacity tapers off as the neural plasticity reduces in many parts of the cortex – that is, the ability to form new neural connections reduces (e.g. Lenneberg, 1967; Curtiss, 1977; Gleitman, 1981; Crain & Lillo-Martin, 1999). Whilst D acknowledges that maturation is clearly important, she suggests that instead of being the maturation of language-specific capacities, it may be the maturation of other necessary cognitive abilities that are needed for language use. For instance, she cites ‘developmental changes in working memory’ (Newport, 1990) or ‘older learners’ greater reliance on declarative memory (Ullman, 2006),’ (p.5).

2.8 Uniformity

Related to the maturation timetable is the fact that children seem to develop language by going through similar steps as each other in roughly the same time frame and order (e.g. Guasti, 2002; Radford, 1990). For instance, children typically begin producing their first words at around one year, and have been shown to comprehend full sentences as young as 17 months (e.g. Hirsh-Pasek & Golinkoff, 1996). They then graduate onto producing two-word utterances between 18-24 months, and sentences that use mostly lexical word classes with few function words between two and three years. Between three and four years they begin producing grammatically complete sentences. Similarly, there are commonplace trends that track all manner of specific linguistic phenomena including wh-questions, relative clauses, the development of finite main clauses, and many other phenomena (see Radford, 1990 or Guasti, 2002 for discussion). This
uniformity has often been used to allude to UG, the maturation of which predetermines the development process.

D points out, though, that there is cross-linguistic variation to be factored into language development. She gives the example of passive sentences. They are typically produced quite late by English learners (around four or five years), not until eight years by Hebrew learners (Berman, 1985), but much earlier by learners of languages where the passive is particularly frequent or easy like Sesotho, around 2;8 years (Demuth, 1989), or Inuktitut, around 2 years (Allen & Crago, 1996). D also points to the individual differences in the learning curves of children even when they are acquiring the same language. For instance, children’s Mean Length of Utterance (MLU) can differ between 1.2 and 5.0 at 30 months (Wells, 1985). Finally, D observes that children have different learning mechanisms (e.g. Peters, 1977; Nelson, 1981; Peters & Menn, 1993). They may learn “top-down” (holistically), where whole units of information are stored as wholes and children have to learn how to decompose them, or “bottom-up” (analytically), where children learn individual words and build up structure from there. Taken together, D argues that uniformity is not as much a part of language development as has been argued by Generative Linguists.

2.9 Convergence

Despite the impoverished and fragmentary input of natural speech (e.g. Wiley, 2014) children end up speaking the language of their environmental stimulus. They all converge on their target language. It is thought that UG enables this convergence as it fills in the gaps left by the stimulus, helping the children on their way to a complete steady state mental grammar. However, D points out a number of studies that demonstrate individual differences in the grammar that different native speakers end up with, such as inflectional morphology, subordinate clause structures, and passives (see Dabrowska, 2012 for review). In light of these differences in linguistic knowledge amongst native speakers, she suggests people ‘do not converge on the same grammar’ and thus convergence does not support UG in the way that is often described.

3 Responses to Dabrowska’s arguments against UG

3.1 Universals and the varying definitions of UG

As discussed in §2.1 and 2.2, one of the key issues D identifies is with defining exactly what UG is, and, in particular, which properties of language are actually universal. In-line with D’s point, Newmeyer (2008) correctly says ‘even more than for categories, features tend to be proposed ad hoc in the analysis of a particular language,’ (p.53). However, D overlooks is the reason for why the analysis of language universals is posited in this ad hoc fashion. When linguists formalize the languages they happen to be working on, they are formalizing what Chomsky calls I-language (or Internal-Individual Language) (Chomsky, 1986b). We should make a crucial, yet subtle, distinction between UG and I-language. UG is something quite fundamental; it is the genetically endowed brain structure that permits the encoding of language structure and consequently constrains linguistic representations in particular ways. I-languages, on the other hand, are the formal (i.e. discrete mathematical) descriptions of how UG actually gets applied to the languages of the world in the mind. Put another way, UG is the system that is used for making linguistic representations in the mind, and I-language consists of the representations themselves. Not only does UG make available a core set of absolute universal principles that occur in all language, we must also assume it sets out the framework space for a hypothetical superset of potential principles, or iterations thereof (i.e. parameters), that is selected from within the specific linguistic environment in which UG is operating. Indeed, the basic core set of mental grammar rules that comes from UG is what allows the brain to gain traction over the impoverished linguistic stimulus available to it when acquiring and using I-language. One could conceive of the critical distinction between UG and I-language as being the similar to the difference between the
hardware of a computer (like the circuits, motherboards, cables and so on) and some of its software (like a particular program or app that makes use of the hardware in particular specified ways). UG is part of the brain’s computational hardware that evolved specifically for language - it is the system that can calculate the combined value of the constituent parts of a structure, and is unbounded in principle in its capacity to continue adding new constituents to the structure. I-language, on other hand, is the software that is laid over the top of this general capacity that adapts it to work in a particular way such that the organic computer inside the human head runs the English program or the Japanese program or whatever it “uploaded” from its environment when the speaker was a child.

Any one I-language is merely a possible iteration of UG, and therefore does not capture the whole scope of the ways in which it might be molded. If linguists conflate UG and I-language by basing their lists of universals on a subset of I-languages they have analyzed (as with the linguists D cites, §2.1), it is unsurprising that it will not hold when a larger set of them is considered. This situation has emerged precisely because many linguists have not understood that most (though not all) work on formal grammar primarily describes I-languages rather than UG directly, notwithstanding the fact that UG does determine the principles that are or are not possible for the I-languages a human brain can derive. Other researchers have similarly pointed out the need to avoid conflating UG and I-language when analyzing cross-linguistic differences. In Kosta & Krivochen (2012), for example, I-language, rather than UG directly, is specifically singled out as the source of the linguistic diversity they investigate in a synchronic comparison of modern Czech and German scrambling. Joseph (2000:192) likewise attributes diachronic (historical) variation to such a source, clearly capturing the same distinction in the following: ‘UG is ahistorical, pure nature… [whereas] core grammar is nature having left some of her facets open to historical determination’ (P.192). The crucial point is that even many of the grammatical principles (let alone their parametric variations) actually turn out to be part of I-language instead of UG itself. As such, there may be some language α that does not make use of exactly the same principles as some language β. An anonymous reviewer points out that a similar defense can be made in terms of the Borer-Chomsky Conjecture (Borer, 1984; Chomsky, 2007). It specifies that language variation amounts to parameters of features that form learned lexical items and are inserted into -- but are separate from -- the underlying UG system that determines phrase structure. In this way, variable features are divided from UG and whatever universals it might give rise to.

The reason for this conflation of UG and I-language can probably be attributed to the way modern Generative Linguistics was born. The scientist who popularized UG, Noam Chomsky, happened also to be responsible for developing the first -- and most influential -- systems for formalizing I-language, beginning with his formidable work in *Syntactic Structures* (Chomsky, 1957), and continuing with numerous refinements of it throughout his career (e.g., Chomsky, 1964; Chomsky, 1981; Chomsky, 1986a,b; Chomsky, 1995; Chomsky, 2013). However, the distinction between UG and I-language was not developed until Chomsky (1986b), meaning that the two were simply not distinguished for the first two and a half decades of the Generative Program, in much of the foundational work.

That said, the fact that proposed universals turn out to be good trends cross-linguistically, if not absolute universals, is in fact reassuring for a UG account. It is surely expected that many I-languages would feature the same or similar properties if they are underwritten using the same neural hardware as each other, since that hardware would set out the space for what is, and is not, a possible I-language. The existence of commonplace trends in intricate linguistic structure across languages may itself imply that the underlying neural hardware contains information that is rather specific to language, in fact. The point here is that the principles used by one I-language need not be identical to those used by another in order for them both to be underpinned by UG. That would be like saying the components of a computer’s central processing unit do not exist because one program does not use the same set of operations as another one. As is the case with computer programs, many -- though not all -- operations will indeed be used across a wide range of applications since they share an underlying computational mechanism.

This relates also to the falsifiability issue D raises -- that Generative Linguists take evidence of universals in favor of UG, and any exceptions to universals as not being relevant for testing UG. However, it is clearly not a fair test of UG to expect the features of a subset of its possible manifestations (I-languages) to extend across the whole set of its possible manifestations. UG does predict that human language will have some
core identical property in all its iterations, consisting of some simple structure-building mechanism that can be extended over the full scope of possible languages. Crucially, though, UG does not exclude the possibility (in fact, on the contrary, it makes it more likely) that different languages will share I-language commonalities that do not go so far as to generalize across the set of absolutely all languages.

Generative Linguists have been attempting to retrench the list of universals as much as possible to find the core property all languages share. Moreover, they have actually been correcting it in the way good theoretical scientists should. Namely, they have formalized explanations for which they have data available; they have tested the predictions the theory makes; and they have revised the theory in light of the results with the intent of simplifying it. In Generative Syntax, this has led to the Minimalist Program (Chomsky, 1995), which Chomsky (2013) describes as ‘a continuation of the origins of the generative enterprise to reduce the postulated richness of UG, to discover its actual nature’ (p.38). This is why the framework looks radically different from earlier formal syntactic frameworks, especially with regards to the list of formal universals, which, as D notes, ‘may comprise just the structure building operation Merge’ (although certain other properties like the mechanisms for assigning certain grammatical features to constituents may also be included) (e.g. Berwick et al., 2011; Chomsky, 2012). Merge itself is simply the rule that permits new linguistic constituents to be integrated into syntactic structure (e.g., for every linguistic phrase XP there is a ‘word’ category X projecting it, and X may optionally attach to another linguistic phrase YP that will in turn have the same internal structure as XP, something like XP → X (YP)). The rule can be reapplied recursively to continue the addition of new structural elements over and over. The reduction in postulated universals in Generative Syntax followed directly from the consideration of larger sets of language data on the one hand, and the further development of formal theoretical frameworks for describing the observable (I-language) facts on the other. It is curious that D proposes this redefining of UG as a reason to question its validity. Chomsky (2013) defends this himself: ‘in normal rational inquiry, the simpler the assumptions, the deeper the explanatory force; where possible, stipulation should be overcome’ (p.37). In sum, although UG has been defined and redefined in different ways throughout the years of the Generative Program, this may be considered nothing more than part of the development of our understanding of the notion. The fact that the definition has been increasingly focused on quite a small subset of properties, especially the Merge operation, strikes me as progress rather than a basis for criticism.

3.2 Poverty of the Stimulus

D argues against the Poverty of the Stimulus argument for UG by essentially denying the environmental stimulus is impoverished at all. Moreover, she does so on the basis of only a handful of phenomena, namely the four of them worked on by Pullum & Scholtz (2002). This line of reasoning can be defeated by finding examples of other phenomena that are both rare in the stimulus and yet acceptable, or that are unique or different from any tokens in the stimulus. Such phenomena would clearly not have been determined by the input. I now turn to such examples of the Poverty of the Stimulus.

If language output were based entirely on the usage statistics in the environmental stimulus it would not be possible to assemble entirely unique examples of language, such as Chomsky’s Colorless green ideas sleep furiously (Chomsky, 1957). Our output would only be able to resemble the tokens of input we get. We can therefore rule out the possibility that the environmental stimulus by itself is sufficient for language. As another illustration of this, take the elliptical dots (…) in (1). Each represents a point in the sentence where one could unboundedly continue adding more of the kind of structure that precedes it. This would create entirely novel output each and every time the rule is reapplied and a new constituent is integrated.

(1) Jim said that Jack saw that Phil mentioned that Sarah wondered if ... the huge super-enormous rather fragile broken red…
  vase and the iron and the mug and the hot-dog... were under the table in the kitchen behind the trash can next to the dog...

We can safely conclude, then, that a system that can calculate the combined value of an unbounded range of possible combinations of constituent structural units is doing a kind of mental computation that relies on
a finite inventory of rules to be present in the mind that can be applied and reapplied to new material so as to create an infinitely large array of possible outputs. In short, it requires an internalized mental grammar or I-language (internal-individual language) (Chomsky, 1986).

The existence of I-language also predicts that we should find examples of linguistic phenomena that are both rare in the stimulus and yet are grammatically acceptable. Indeed, such phenomena do exist. For instance, so-called Heavy NP Shifts fall into this category. Compare (2) and (3):

(2) James wanted desperately the black TV with all the cable channels.

(3) *James wanted desperately the TV.

Adjunctive information (e.g. desperately) intercepts a subcategorizing verb (e.g. wanted) and its complement (e.g. the black TV with all the cable channels or the TV). This interception is tolerated in (2) since the complement is “heavy” (i.e. contains a lot of information) compared to the complement in (3) which is comparatively “light”. The reasons for why heaviness should affect the acceptability of adjunctive interception need not concern us in the present paper. The point is that these interceptions are rare in English. According to Wasow (1997) only around 5-10% of English sentences with direct objects are Heavy NP shifts, and yet (2) is perfectly well-formed. This is precisely not what we predict if statistical usage patterns from the environmental stimulus are thought to be responsible for grammar -- something that is proportionately very rare would simply not remain part of the grammar. Taken together, it does appear there are many examples of a poverty in the stimulus, which points to an underlying representational I-language system in the mind. Whilst these facts defeat D’s denial of the Poverty of the Stimulus, the Poverty of the Stimulus does not, it turns out, support UG in any case. Instead, it merely supports the presence of a mental grammar that can generate linguistic output that can outstrip available linguistic input, but does not directly imply that mental grammar is underwritten by a uniquely linguistic genetic or biological base. When D argues against the Poverty of the Stimulus in order to attack UG, she is in fact arguing against a straw-man.

D also argued that the lack-of-negative-evidence argument does not work as support for UG by reporting evidence of children receiving negative evidence. If children do benefit from corrections, this does not detract from the possibility of UG in general. Indeed, there is no reason why any metalinguistic guidance that children receive should not help them to set up their I-language within the confines of UG. However, if negative evidence is indeed absent from the input, as has sometimes been illustrated (e.g. Guasti, 2002; Pinker, 1994), then it does lend support to the notion that there is a genetic underwriting that could pick up the slack of the impoverished linguistic guidance we are given as children. As with the main Poverty of the Stimulus argument, language acquisition without negative evidence does not specifically tell us whether or not that genetic endowment is unique to language or if it is part of some domain-general capacity, and so it does not follow that criticizing it helps defeat the UG theory. It also means that the existence of a Poverty of the Stimulus, and the lack of negative evidence in children’s input, are both as consistent with D’s own argument that language is part of a domain-general biological hardwiring as they are with the domain-specific notion of UG. It is thus, at the very least, puzzling why she should argue against them whilst concluding there is an innate “language making” capacity.

Now, if we can be confident about the existence of the mental grammar, as the Poverty of the Stimulus suggests we can be, the question reduces to how did the I-language get into the mind? Clearly, I-language corresponds to the language one is exposed to, so the environmental stimulus is used to set up the I-language system (i.e., environmental stimulus is necessary, but not sufficient, for language). That is, the stimulus is used to shape the specific rules one is going to apply over new variables to make one’s language outputs. This is why people who grow up in Japan speak Japanese and people who grow up in the US speak English. However, since we have established that the stimulus cannot account for those outputs by itself, there must be a role for genetics. In other words, since we have established the limitations of what nurture can do, we know there must be a role for nature to play. Put more precisely, there must be a genetic endowment that humans are born with that enables them to set up I-languages based on the stimulus to which they are exposed.
If we can now be confident that there is an innate endowment that underpins how I-languages are set up in the mind, the question now reduces even further to the following: *is the genetic endowment specific to language or is it domain-general?* In other words, we need to know if the endowment is related to language specifically, or if it is part of our broader capacities for representing information in the mind. Recall that a key part of the notion of UG is that it is an innate system for representing *language* in particular. Indeed, D’s main argument is that there is a biologically-based “language-making” mechanism that is domain-general and is therefore not UG.

### 3.3 Neuroscience of language

One of D’s complaints about the neuroscientific evidence that points to the autonomy of language in the brain is that the language centers do not work in isolation. That is, other parts of the brain are active during language tasks, and the language centers are active during non-linguistic tasks. However, it has long been understood that the brain is highly interactive (e.g. Pulvermüller, 2002). For instance, if one is processing a sentence whose content is about motion, the motor strip of the brain is active (e.g. Buccino et al., 2005; Hauk & Pulvermüller, 2004; Hauk et al., 2004; Tettamanti et al., 2005). Language is a grammatical encoding of a conceptual message and is also rendered in sound systems, and in some cases may contain descriptions relating to emotions or senses like smell, sound, taste and so on. It is surely unsurprising that the brain needs to call upon a range of different substrates that not only relate to the grammatical processing itself, but also to the message the language is attempting to convey, and the communication thereof. Likewise, language is often needed to complete non-linguistic tasks. Language centers are activated for the purpose of “inner-speech” to make the thoughts that are required to complete all manner of tasks more tangible. Stout et al. (2008), for instance, demonstrated that the language centers were active when participants were instructed to make complex tools. D’s expectation that the language centers should be activated entirely independently from the rest of the brain seems unrealistic. Not only does the brain have to do other work (such as autonomic nervous functions like breathing and organ control), the very job of language is to encode non-linguistic information, and so of course other substrates will be active during language processing and language may be used to help resolve tasks that are not themselves linguistic. Indeed, no area of the brain is able to work in the vacuum D seems to be expecting. To flip the argument around, when the language areas are active during a motor task it is likely that the brain is conceiving of the task using language so as to understand what is required and to plan what is to be done. One would presumably not say that this in any way diminishes the role of the motor strip, and any genetic underpinning thereof, in completing the motor task.

The second neuroscientific point D raises is that ‘the local connectivity in the language areas (as well as other areas of the brain) [is] not genetically specified but emerge[s] as a result of activity and [its] position in the larger functional networks in the brain’ (p.6). However, the larger functional networks of the brain, including their cognitive functions, are known to be underwritten by the genome (Marcus, 2004), although a lot of the mappings between genetics and cognitive functions are not yet understood. If genetics is involved in providing the framework for the ‘larger functional networks’ that D speaks about, and she agrees that the language centers arise out of their position within these networks, then (at the very least indirectly) the language centers must arise out of the genetics of the brain. D also focuses specifically on neural connections in her remarks, which she maintains are determined by ‘activity’. And indeed, neural networks are the result of usage; connections that are useful remain whilst the others are weakened or die back. As D notes, this gives rise to plasticity, which enables the correspondence between neural substrates and cognitive functions to be more fluid (e.g. Trauner et al., 2013). However, neurons do not form their connections in a vacuum, whether inside or outside the language centers; they do so with the genetic code they inherited that forms the template for the overall structure of the brain (e.g. Pulvermüller, 2002). The ‘activity’ that D speaks of, then, does not absolve the formation of neural networks from their fidelity to genetics. This is as true for the language centers as for any other part of the brain.
3.4 Genetics of language

I agree with D that Stojanovik et al. (2004) undermines the double dissociation between the SLI and WS populations, in that the study found poor language function for the WS as well as the SLI group. It suggests that the situation is not as simple as SLI being a language deficient group with intact cognition more broadly, whilst WS is the reverse. If the double dissociation is questionable, it cannot be used to point towards the genetic autonomy of language (notwithstanding that we should exercise caution about drawing wide ranging conclusions from a single study). Indeed, there are additional studies that have found some linguistic impairments in WS patients that D did not mention, compounding the double dissociation still further (e.g. Clahsen & Almazan, 2001). There are, however, two points of defense to make here. First, just because the double dissociation between SLI and WS patients may not hold, it does not mean other double dissociations will not hold. For instance, amusics and dyscalculics -- different groups with genetically determined non-linguistic cognitive impairments -- may make for good contrasts with SLI patients in this regard, although that is a question for future research. Second, the most important point is the genetic link to SLI (see discussion section, §4.2), which D concedes is at least ‘primarily linguistic’ in nature. Given the extent of the mystery that surrounds the links between genes and cognition, it is surely an encouraging pointer for UG that certain genetic mutations are associated with a very specific set of linguistic symptoms.

3.5 The ease and speed of language development

The first of D’s attacks against the quick, automatic nature of language development is based on the fact that it is actually contingent on an apparently substantial input during the first years of life (her Manchester Corpus analysis estimates humans receive about 42 million words between the ages of one and five). It is puzzling, though, that she treats this as if it were a counterargument for UG. The argument is about the ease and speed with which children make use of the stimulus to which they are exposed in building their I-language quickly and subconsciously. Nobody ever argued that the stimulus was small (as we saw earlier, it is impoverished compared to the creative outputs that language is capable of -- but that is something quite different). In fact, proponents of UG agree that the development of language requires some kind of substantial input or exposure to a linguistic stimulus to build a specific I-language.

D’s second line of attack centered on the claim that language development occurs without formal instruction, as a result of the influence of an innate UG. She cites studies showing children who grow up in an environment that isolates them from interactive speech are stunted in their language development -- a problem that is remedied by the introduction of interactive speech. Such findings are, however, no smoking gun for D’s argument that children need formal instruction to learn their first language. It is well known amongst generativists that exposure to language during development needs to be interactive since children develop their I-language using trial and error techniques (e.g., Trueswell et al., 2013). That is, they make hypotheses about how the system ought to be configured on the basis of the stimulus and the range of possible principles that UG permits. They try out such hypotheses and see if they work. If, on the basis of the responses from those around them a hypothesis seems wrong they will start the process again; but if it seems right, it will be codified into their I-language using the available mechanisms of UG. Nobody is surprised, then, when a lack of exposure to interactive speech results in language development problems, or that the best solution to such problems is to introduce interactive speech to the children concerned. To put it another way, even within the context of a UG framework, the linguistic exposure a child must have needs to be good quality (i.e., interactive and not just passive), given the UG mechanisms the brain has for making use of that exposure to build an I-language.

Indeed, quality interactive communication is so vital for language development that the stimulus can be extremely impoverished -- even zero -- and UG will still be able to guide the development of language, if interaction stimulates it. For instance, deaf children (who cannot hear their speaking parents) develop their own private gestured languages together through their mutual interactions, with similar developmental stages as children learning to talk with a pre-established language stimulus (Goldin-Meadow & Feldman,
The fact that the human brain can gain traction over such extremely impoverished environmental stimuli is strong support for proposing a mechanism like UG that uses an innately endowed framework to build a language system in particular. The fact that interaction is a prerequisite for language development takes nothing away from the notion of UG.

### 3.6 The maturational timetable

In reference to the well-established trend for language development capacities to cut-out as the brain matures, D suggests that ‘developmental changes in working memory’ (Newport, 1990) or ‘older learners’ greater reliance on declarative memory (Ullman, 2006)’ may be responsible, rather than a maturation of some language-specific faculty. Changes in working memory are unlikely to be the cause, however. Alloway et al. (2006) show that the working memory capacity of healthy children expands at a regular rate until around age 15 at which point it is adult-like. Thus, children are experiencing an increase in working memory capacity year-on-year as they grow, whilst the critical or sensitive period(s) for language acquisition are thought to be pre-pubescent (e.g., Guasti, 2002). In other words, working-memory capacities continue expanding after the language development capacities have already tapered off. More critically, it means that when most of the developmental work is being carried out by the brain (before age five), working-memory capacities are actually at their lowest, as has also been suggested elsewhere (e.g., Sekerina et al., 2004). Similarly, older learners’ increased use of declarative memory is generally thought to be a function of high levels of exposure to certain common forms (Ullman, 2006) rather than having anything to do with maturation. In other words, people retrieve full forms of information they use a lot as a kind of cognitive “short-cut” that means they do not have to process familiar details procedurally (i.e., with internal structure). This is similar to the idea that older children “entrench” their earlier learning (e.g., Elman et al., 1996; MacWhinney, 2008). There is no reason, however, to think procedural memory is actually reduced in adults, especially since they are able to compute novel and rare forms that do require detailed internal structures, just because they may rely more than children on declarative memory. It is not clear, therefore, that increased use of declarative memory is the cause of the maturational timetable either.

D acknowledges the possibility that maturation effects in language development may relate to ‘UG becoming inaccessible’. To refine this statement a little, it could be that the neural plasticity that allows one to configure UG into different l-languages reduces with age.

### 3.7 Uniformity

D counters the long-held perception in Generative Linguistics that language development unfolds uniformly amongst humans by citing three non-uniform dimensions of it: cross-linguistic variation in development patterns, individual differences in development patterns, and different learning mechanisms. As D acknowledges, cross-linguistic variations in development trends are related to the differences in complexity and usage of certain phenomena across languages. If there are substantial differences in exposure to given phenomena cross-linguistically, it is unsurprising that this would be reflected in the timing of their acquisition. If a phenomenon is rare or inaccessible due to its complexity, it may take longer for the emerging l-language to reach a state where the complete phenomenon is interpretable to the brain, or for it simply to become aware that the relevant properties of UG apply to the target language. Cross-linguistic differences are not incompatible with UG. Meanwhile, the fact that there are plenty of broad trends in the sequence and timing of the development of many different languages does surely speak in favor of UG (see Guasti, 2002 for cross-linguistic examples of developmental trends in language).

It should not surprise us, either, that there are individual differences during language acquisition. After all, there are in any developmental curve one cares to look at. Take motor development. Children tend to role over (2-7 months), sit up (5-9 months), crawl (5-11 months), pull-to-stand (5-12 months), stand alone (9-16 months), and walk unaided (9-17 months) in this order (e.g., Alexander et al., 1993).
considered to be a genetically predetermined maturation of the brain. Notice, however, that the estimates of the timing at which each stage can occur are broad and overlapping. According to these estimates, one could expect to find a nine month old at any step except the first one, for instance. Developmental variance is inevitable. The relation between neural substrate and processing function, and the exact organization of neural connections, are not established as exactly the same in any two brains (e.g., Pinker, 1994). Sure, there are broad trends in these regards, but no two brains are precisely alike (e.g., Grodzinsky, 2000). That being the case, the development of different people’s brains, and of any of their brains’ functions, will not be precisely alike either, even though there is clearly genetic underpinning for much of it (e.g., Marcus, 2004). In this way, it is simply unreasonable for D to expect a “uniform” developmental curve to be precisely uniform for each human; that is just not how brain development works -- there is always variance within any of its developmental learning curves. The big picture here is that there are broad developmental patterns in the way linguistic phenomena emerge (e.g., Guasti, 2002).

Turning to D’s suggestion that children also differ from each other with respect to whether they use “top-down” holistic learning mechanisms, or “bottom-up” analytic ones, D may be making an erroneous assumption that these differences relate to the nature of the mental grammar. In fact, it is known that real-time sentence processing can unfold in adults using either top-down or bottom-up mechanisms too (e.g. Townsend & Bever, 2001; Bever & Sanz, 1997; McElree & Bever, 1989). It is not inconceivable that the use of these different mechanisms affects access to, and usage of, the grammatical mechanisms during language development as well. This requires testing in future research, but the fact that there is a clear (and likely) non-grammatical account of the differences in “learning style” means we should exercise a great deal of caution before accepting this as evidence against the underlying uniformities of grammar that have been observed in language development. This also highlights the importance of considering the role of real-time processing strategies alongside I-language and UG (see the discussion section below, §4).

### 3.8 Convergence

Learner communities are thought to converge on the same target language, and consequently use language in the same way as each other. This is often cited in support of UG underpinning the homogeneity of the language we speak. However, D discusses individual differences in the grammars that emerge out of supposedly homogeneous speech communities as evidence against this argument. D misses the point here. In fact, everyone does converge on the language of their speech community. It is self evident that a child growing up in Japan will acquire Japanese whilst a child growing up in the US will acquire English. However, the exact form of the grammar will not be identical for every speaker of Japanese or English since each person has a different stimulus from the next. Sociolinguistic studies reveal time and again that there are a multitude of factors that are known to affect the exact configuration of an individual person’s grammar. For instance, one can find systematic linguistic patterns determined by geography, ethnicity, sex, gender and class, amongst others. Furthermore, an individual will be influenced by the people they aspire to emulate or dissociate themselves from, both in real-life relationships and in the media (see e.g., Meyerhoff, 2006 for review). The exact intersection of all these language-external factors, and others, will be unique for each person, meaning that the influences in the stimulus upon which each person bases their I-language will be unique for them. As a result, each of us builds our own individual grammar -- an idiolect that will fit broadly into certain broader categories of dialect(s) and language(s) but that is, in fact, not a perfect idealized version of either. Furthermore, as we have seen, one must use a biological program to upload the stimulus into the brain and extrapolate generalized rules from it so as to make the grammar internalized and available for making new and novel outputs. That is, not only is the grammar individual, it is also internal. This is exactly what was meant by I(internal-individual)-language (Chomsky, 1986) in the first place. Individual differences, then, are entirely in keeping with Generative Linguistics.
4 General discussion

In the current paper, the counterarguments that D has levied against UG have been described (§2) and rebutted (§3). However, the case for UG is stronger than merely rebutting D’s counterarguments. In the discussion that follows, the above arguments against D will be situated in a broader context of further support for the concept of UG. Meanwhile, it will also be argued that the most productive outcome of D’s paper is that, coming from the school of Cognitive Linguistics that acknowledges an innate “language making” capacity, the space between Generative and Dabrowskian Cognitive Linguistics is made much smaller than for the Behaviorist or usage-based quarters of Cognitive Linguistics. This is especially important given developments in Minimalist Generative Linguistics that make significant space for the roles of domain-general cognition in language.

4.1 Neuroscience of language

In addition to the above responses to D’s comments about interactivity and the broader functional networks of the brain, there are further neuroscientific arguments that point towards a UG. First, neuroimaging studies on recursion show that linguistic recursion uses unique neural substrates (e.g. Dehaene, 2012). That is, the repeated use of a rule like Merge in language will trigger an increasing level of activation for each newly embedded constituent. This was also true for similar embeddings in mathematical and musical structures too, except the language modality alone involved neural substrates that the others did not. This suggests that language seems to have an independence from the other recursive modalities, maybe as a result of possessing additional complexity and intricacy within its structure. Likewise, Ding et al. (2015) recently found neuroimaging evidence that supports the concurrent processing of hierarchical elements of language: individual words, sentence-internal phrase structures, and complete sentence structures. The study tracked cortical activation for the statistical predictability of linear word sequences by comparing predictable, less predictable and non-predictable word orders, and they also looked at prosodic cues. Even taking account of statistical and prosodic cues, independent neural activations were observed for grammatical hierarchy whereby words, phrases and sentences are concurrently processed and integrated together. This was taken to indicate the presence of a hierarchical mental grammar within the structure of the brain that builds words into phrases, and phrases into sentences.

Similar evidence of linguistic neural autonomy comes from selective disorders of language, or aphasia. Such disorders can emerge as a result of damage to particular neural substrates, in particular the Perisylvian cortex or its subcomponents. Depending on the exact brain lesion, a range of linguistic symptoms may emerge including difficulty resolving dependencies like which girl did the boy chase_? or interpreting functional word classes like determiners (e.g. the, some) or prepositions (e.g. in, on). In these cases, language production is usually very slow. Other patients may exhibit syntactic structure that is separated from meaning such that they struggle to understand the content of what they hear or read and they produce phrases with seemingly meaningless, but grammatical, combinations of words (see, e.g., Grodzinsky, 2000 for a review). We should note that other lesions, combined with the high degree of interactivity across the brain that includes the language centers, may lead to other cognitive problems appearing in addition to the language deficiencies. We would not expect language always to appear as selectively impaired in brain-damaged patients, but the existence of neurologic language centers that typically elicit common symptoms when damaged is certainly consistent with UG.

4.2 Genetics of language

D’s discussion centered entirely on questioning the double dissociation of language and other cognitive capacities that have been argued for between SLI and WS patients. In spite of the justification for questioning that particular double dissociation, D largely ignores the bigger picture that Specific Language Impairment
(SLI) results from a mutation of a certain gene, thereby strongly implicating the genetic underpinning of language. A single substitution of a nucleotide (G-to-A) on exon 14 of chromosome seven (Lai et al., 2001), otherwise known as the SPCH1 or FOXP2 gene, appears to affect language function from birth (e.g. Fisher et al., 1998). It should be noted that whilst some researchers have claimed such mutations uniquely affect language, others have reported evidence that other cognitive domains including non-verbal reasoning are affected by the mutation alongside language (Vargha-Khadem et al., 1995). In fact, this really should be unsurprising. FOXP2 contains genetic code for a protein whose function is to express other genes -- a “forkhead binding domain”. This means that the correspondence between the gene and the proteins that are affected by a mutation of it is very complex indeed, and the range of symptoms may vary as a result. The important point to note is that SLI patients have shown the same mutation as each other whilst unimpaired controls do not exhibit it -- there is a clear genetic etiology for a pathology that is, even in D’s estimation, ‘a primarily linguistic impairment’. And the fact that in some studies it has indeed been uniquely language that seems to have been affected by the mutation is surely even more encouraging (e.g., Lai et al., 2001). Moreover, some work has identified very specific linguistic symptoms that relate to the genetic mutation. For instance, it has been claimed that morphosyntactic dependencies are deficient in these patients. They may make very selective errors in phenomena like subject verb agreement (e.g., he walk to town instead of he walks to town) or concord agreement (e.g., a buses instead of a bus or the buses) (e.g., Clahsen & Dalalakis, 1999).

4.3 Evolution of language

Evolutionary evidence is largely glossed over in D’s article. On the issue of species specificity (i.e. the fact that only humans have evolved language), she says that ‘arguments for the innateness of language in a general sense (what Scholtz & Pullum, 2002 call “general nativism”) do not constitute arguments for the innateness of UG (“linguistic nativism”) if UG is taken to be a specific body of linguistic knowledge,’ (p.3). In other words, she agrees that humans have evolved a unique mental capacity that includes language, but this may be domain-general and not unique to language. However, a closer look at evolutionary evidence points to language evolving separately from developments in conceptualization capacity and other aspects of “general nativism”.

Imagine a baby chimpanzee and a baby child, two close evolutionary cousins, growing up together. They would develop quite similarly with regards to vision, motor activity and, up to a point, conceptual structure. Only one of the two, however, would ever learn to talk: the human (e.g. Gardner et al., 1989; Terrace, 1979). Interestingly, chimps have been shown to have the capacity for sophisticated conceptual abilities that can be drawn out by teaching them basic gestural signs. A chimp is able to take a sign for a noun (like cat) and predicate an attribute onto it (like dead) to access a combinatorial meaning amounting to the cat being part of the set of entities that has the property of not being alive (e.g. Gardner et al., 1989; Terrace, 1979; Lyn et al., 2011). This form of conceptual “meshing” is quite unique to humans, chimps and maybe some other higher primates, and therefore seems to have evolved within the particular neural hardware of the higher primate family. To be more precise, the chimp is able to evaluate the “affordances” of a cat -- that is, a list of which attributes it can plausibly take on in the real world. Similarly, the chimp can evaluate the affordances of which entities the property of deadness can be applied to in the real world, and then bring those two sets of evaluations together in the form of a combined concept that might be paraphrased as the cat is dead. It has also been demonstrated that chimp can conceptualize the difference between seeing something in real time and possessing some static knowledge about it stored in long term memory (Heyes, 1998; Hare et al., 2000; Premack & Premack 2002). Taken together, this is remarkable conceptual sophistication -- to my knowledge it is not found elsewhere in the animal kingdom. However, what is especially interesting is that chimps do not make use of this conceptual sophistication for their own communication (e.g., Crockford & Boesch, 2005). Their communication system is actually quite typical for the more intelligent animal species -- namely, a one-to-one mapping between a simple atomic concept and a given call or gesture. This critically allows us to draw two conclusions: (i) that such increased conceptual sophistication emerged at around
the time the common ancestor of modern humans and chimps lived, which was approximately six million years ago; (ii) a language system good enough to communicate such meaning was clearly not possessed by this ancestor species, and must have emerged later, presumably within the hominid genus. The conceptual capacity of chimps is unlikely to be equivalent to that of modern humans, although for all we know it may simply be limited by the absence of the language faculty that enables humans to make thoughts tangible, thereby communicating and developing them further with each other. The main point to take from these findings is that conceptual structure was evolving in the primate mind ahead of a linguistic structure that could communicate those concepts. Indeed, logically, it is surely a prerequisite for the evolution of language that there is already the capacity to form messages that are of a sufficient complexity to require encoding into a tangible or expressible format in the first place. It could even be that the linguistic Merge that emerged in the hominids is a language-specific adaptation of the conceptual Meshing function. Taken together, then, chimpanzees indicate that emerging conceptual capacities are necessary but not sufficient to produce language. The language system must therefore be a more specific and unique capacity, contrary to the fundamental tenet of Cognitive Linguistics that fails to distinguish between language and conceptual abilities (see also Bierwisch 1982, 1989; Bierwisch & Schreuder, 1992).

There is a range of genetic, neurologic and archaeological evidence that broadly fits with, and points towards, this picture of the evolution of language. First, there was a substantial reorganization of the brain and tripling of brain size around 1.5 million years ago, within the hominid genus, particularly affecting the anterior regions where a great deal of the higher cognitive processing is centered, including some language centers (Bradbury, 2005; Holloway et al., 2004). Preceding this by about a million years, however, was the emergence of the first clothes, personal living space, and ritual behavior that indicates a prior high level of conceptual processing before language (Ambrose, 2001).

The oldest paintings seem to date only from very recent times in evolutionary terms, around 40,000 years ago (e.g. Clottes, 2001; Than, 2012); and likewise, complex tools like flint stone axes emerged only about 700,000 years ago (Plummer, 2004). These latter more complex structured artifacts may well reflect the presence of some kind of “inner speech”. Indeed, Stout et al. (2008) showed that the act of making such stone tools engages the main language centers of the brain, namely the left inferior frontal gyrus and the superior temporal gyrus.

Finally, it has been demonstrated that the FOXP2 gene of modern mice differs from primates (namely chimps, gorillas and rhesus monkeys) by one amino acid, whilst modern humans’ FOXP2 differs from modern chimps by two further amino acids (Enard et al., 2002). In evolutionary terms, this means that since the common ancestor of modern humans and mice, who lived around 150 million years ago, there have been three mutations of FOXP2. Two of those mutations have occurred since the more recent ancestor between modern humans and chimps, who lived a mere six million years ago. In short, a gene that has been strongly linked with language capacity (albeit maybe only indirectly) has been shown to have undergone much of its evolution very recently, and specifically within the hominid genus no less. Together, this range of evidence is consistent with the idea that language emerged very recently and quickly in humans, and certainly atop whatever conceptual development can be seen in the other higher primates. It is also consistent with the idea that language is specifically grounded in certain unique genetic and neurological materials, a la “linguistic nativism” rather than merely “general nativism”.

### 4.4 Going beyond UG: Including the rest of cognition in our model of language

The above review of D’s paper highlights that the UG claim is far from dead, contrary to her concluding remarks. There is a wide range of evidence that convincingly points towards a genetically endowed language-specific neural substrate. I do agree with her, however, that Cognitive and Generative Linguistics are moving closer together. Cognitive Linguists assume language arises out of domain-general cognitive capacities, yet D’s argument that there is some innate biological “language making” capacity is a far cry from the Behaviorist or Connectionist approaches many other Cognitive Linguists have appealed to. Under that approach, a good (or not so good) simulation of behavior (or unanalyzed statistical usage data) is
cyclically treated as if it were an explanation for the behavior itself, and as such is completely disconnected from the principles and mechanisms of the actual human brain that are responsible. Meanwhile, as Generative Linguistics has minimized its schematics for UG (maybe even all the way down to Merge alone), it leaves room for the other cognitive domains to account for the resulting vacuum in explanatory theory for the properties of language outside core grammar.

This is an exciting development, and is the biggest reassurance that I take from the “general nativist” overtures of D’s paper. For too long Generative Linguists have made the “grand assumption” that everything to do with language is grammatical, overlooking its broader role within Cognitive Psychology. They have neglected the roles of conceptual structure and domain-general capacities like (working) memory, information activation mechanisms, information suppression mechanisms and predictive mechanisms, claiming that linguistics is about “linguistic competence” and not real-time performance. But, of course, conceptualization and the processing systems are just as important for actually comprehending and producing language as the underlying rule system for language encoding. Cognitive Linguists, meanwhile, are often guilty of the reverse “grand assumption”. Grammar, if acknowledged at all, is reduced to a meaningless theoretical postulate and all the focus is given, more often than not, to conceptual structure, which is usually conflated with linguistic semantics. In fact, as seen in §4.3, conceptual structure is best viewed as its own domain-specific intelligence that creates the messages language attempts (sometimes even unsuccessfully) to encode as accurately as possible (e.g. Bierwisch 1982, 1989; Bierwisch & Schreuder, 1992). The main point is that these developments in Cognitive and Generative linguistic theory make it more likely that the nature versus nurture debate in language can now take its place alongside a second (equally important) debate that situates Linguistics within Cognitive Psychology, concerning how far each linguistic phenomenon exists either because of static encoding rules and constraints or because of the dynamic real-time processing mechanisms and resources that act upon them.

As language scientists begin teasing apart the contributions of the static grammatical rules from the dynamical processing mechanisms, we will be able to develop a much more refined theory of both grammar and processing, and as a result, of language as a whole. For instance, we would be able to discuss answers to questions such as whether or not postulated transformational movements and phases in syntax are really working-memory limitations (e.g. Berwick et al., 2011); if children really have a grammatical “Delay of Binding Principle B” effect or if they simply lack the working-memory and information activation or attention resources to use a sentence external referent for a personal pronoun when they could just recycle a previously encountered sentence-internal antecedent (e.g. Sekerina et al., 2004); and whether or not children have different grammatical learning strategies or if differences in their behavior are symptomatic of top-down or bottom-up processing (as discussed earlier, §3.7). In short, it will only be by considering the brain’s ability to navigate through language structure, together with its ability to encode it, that we can hope to gain an accurate understanding of the complete language system.

Finally, we should note that the issue of static encoding and dynamic processing mechanisms and resources is not disconnected from the nature versus nurture debate that UG sits at the heart of. Indeed, the two debates intersect neatly across one another, as in Figure 1.

Each quadrant of Figure 1 describes a different class of phenomena, which can apply over any domain-specific information type such as linguistic or conceptual structure. Taking them in turn, they are described below (notice that the nurtured phenomena in (6-7) presuppose the natural capacities in (4-5)):

(4) Natural structural encoding rules & constraints: These are structural principles determined by the structure of the brain that are in turn determined by the genes that synthesize the neural substrate, and are in turn selected by evolutionary natural selection. This is where one would find domain specific rules like UG or Merge (and some equivalent for conceptual structure).

(5) Natural structural navigation: These are the brain’s genetically endowed capacities for processing structure in real-time. They include domain-general abilities like working-memory, the ability to activate relevant information and suppress irrelevant information, and the ability to predict upcoming structure.
368

5 Concluding remarks

In spite of disagreeing with most of D’s arguments against UG, I think her paper makes the more general contribution of highlighting the increased common ground between Cognitive and Generative Linguistics. This is in part a result of being one of the more thoughtful and well-informed critiques of the Generative Program that takes account of a large array of experimental evidence that seeks to test theoretical principles rather than merely relying on the simulation of unanalyzed statistical usage data. It is this approach that led D to argue for an innate “language making” capacity -- a “general nativist”, or domain-general, endowment from which language emerges as one of many cognitive phenomena that works off of the same infrastructure. Nonetheless, in this response paper, I have attempted to show why her anti-UG position is not tenable, either logically or evidentially. I argued that some of her critiques, particularly those surrounding the difficulty in isolating absolute universals cross-linguistically, and in defining UG itself, result from misinterpreting UG as I-language. Relatedly, her statement against the Poverty of the Stimulus seems to be a straw man since it is shown the Poverty of the Stimulus only supports the presence of a mental grammar, and never did constitute direct support for any underlying UG, it turns out. Meanwhile, the paper attempted to show how much of D’s discussion draws conclusions that are an over-reach or misinterpretation of the available evidence in many fields of (psycho-)linguistic inquiry, including language development, neurolinguistics and evolutionary linguistics.
Similar “general nativist” arguments to those of D were also put forward by Cognitive Linguist Vyvyan Evans in his 2014 book *The Language Myth: Why language is not an instinct*. The book has been the subject of very heated criticism and debate (e.g., Adger, 2015), demonstrating, if nothing else, how contentious these issues remain to the field. Whilst the current paper is not a response to Evans (2014), it does address some of the concerns raised by it. Most significantly for the current paper, in a reply to Adger (2015), Behme & Evans (2015) throw down the following gauntlet for ‘anyone who wishes to defend the Chomskyan framework’:

‘...address the following questions: (i) what are the specific theories Chomskyans are currently committed to? (ii) which concrete findings from developmental psychology and neurobiology support the Chomskyan framework? (iii) how can the Chomskyan paradigm overcome the familiar, long standing challenges stated in the technical literature...?’

The current paper has attempted to define UG and I-language clearly, and to show their role in the contexts of broader cognitive capacities. It does so on the basis of a large array of experimental evidence from Developmental and Cognitive Psychology, Neurobiology and Evolutionary Psychology. By these criteria, then, the current paper should go some way to making the contribution that Behme & Evans (2015) invite to the general debate on the place of UG in the human mind.

More importantly, it is the implications of the decreasing distance between the Cognitive and Generative traditions that should be a focal point across the language sciences. The reduction in the postulated richness of innate Generative grammar leaves an explanatory vacuum for domain-general cognition to fill. Whilst the two schools of thought do not completely agree, the current paper argued that the domain-specific Generative enterprise needs to be integrated fully with the rest of Cognitive Science and Psychology not just with regards to the nature versus nurture debate, but also with respect to the encoding and processing of language. This would surely be the most productive direction-of-travel for a fuller understanding of human language.

**Acknowledgements:** I am grateful to three anonymous reviewers for their comments on earlier versions of this manuscript.

**References**


