

# Justifications for a Discontinuity Theory of Language Evolution

Callum Hackett

In Chapter 6 of *Biological Foundations of Language*, Lenneberg argues against continuity theories of language evolution, which claim that language evolved from simpler communication systems. Although Lenneberg was pessimistic about even discontinuity theories explaining how language evolved, discontinuity has become significant in the Minimalist program, which posits that our species' acquisition of Merge was the key discontinuity that made language possible. On the basis of a unified description of natural communication systems, I show that language is indeed based upon a cognitive discontinuity, which is moreover specific to linguistic ability. However, I argue that even Minimalist theories must recognise this discontinuity as the sensorimotor interface with syntax, rather than syntax itself. This ultimately supports the view that syntactic structures are structures of thought, but taking this claim seriously means reimagining how syntax relates to semantics and morphology, as the traditional 'lexical item' is no longer a tenable primitive of generative theory.

*Keywords:* continuity; communication systems; minimalism; syntax-semantics interface

## 1. Introduction

One thing that makes language so fascinating—and its origins so difficult to pin down—is that it is unprecedented in evolutionary history. This need not necessarily make it mysterious or problematic, as uniqueness is simply a corollary of diversity, which is plentiful in the natural world. It is therefore tempting to agree with Fitch, who says that

the fact that humans [alone] have this particular capacity is no more surprising [...] than other unusual features like the elephant's trunk, bat echolocation, or "radar" in electric fish. (Fitch 2010: 5)

Yet, the idiosyncrasies of other animals tend nonetheless to serve common biological needs, like ingestion, navigation, and object manipulation, while human language serves functions that no other species has been capable of, like creative and

---

I would like to thank Noel Burton-Roberts, Anders Holmberg, and Maggie Tallerman for many insightful conversations on the issues addressed here. I am also grateful to two anonymous reviewers for comments regarding clarity and precision—I trust that all the remaining bad ideas are cleanly articulated.



unbounded expression. These functions having moreover made civilization possible, it is the emergence of language—and not the elephant's trunk—that Maynard Smith & Szathmáry (1997) give as much weight to in their history of life as the origins of genes, eukaryotes and sexuality.

Of course, to talk of the origins of language, we must bear in mind that it is an integration of many traits with evolutionary histories that are largely independent of language and independent of each other. Some of these traits—such as our vocal physiology, motor control, imitative abilities, social intelligence, and so on—may have been selectively refined for the roles that they now play in language, while others may have been recruited by language mostly unchanged. Some may have homologues in closely (or even distantly) related species, while others may only have precursors in hominins, if any at all. Here I argue that, however this human tapestry was woven, a sensorimotor interface with syntactically structured thought is a core linguistic ability that is unique to humans, so the cognitive basis of language must have originated within our lineage during the past six million years (the time of our last common ancestor with chimpanzees; see Patterson et al. 2006).

Theories like this one, which claim that there is no evolutionary precedent for language cognition, are often taken by their critics as claiming that language was something that came from nothing (e.g., Deacon 2003, Lieberman 2015): Some uniquely human capacity, whatever that is, must have been bolted onto our biology by sheer good fortune, rather than having developed from it by any ordinary means like adaptive selection. If this is how we are to interpret 'discontinuity', then I agree that we should reject it—the only genuine possibility for something to have come from nothing is to be left to cosmologists. In life, all new traits, no matter how unique or how quickly acquired, arise through the variation and selection that acts upon existing organic material. In a very real sense, then, evolutionary discontinuity *must* be construed as just a special case of continuity, and this paper will in part explain how that can be sensibly done.

In section 2, I reformulate Lenneberg's discussion of continuity and discontinuity theories of language evolution, with a consideration for how the debate has progressed in the fifty years since *Biological Foundations of Language*. Owing to some deficiencies in the terms of that debate, in section 2.1 I introduce some clarifying definitions. With those in place, in section 2.2 I adapt the notion of intentionality to describe the necessary properties of all natural communication systems, and I discuss why the defining property of language—its mapping of words to meanings—must be accounted for separately. I then consider whether there is any evidence for this property or its potential in nonhuman species, and I conclude that there is not, so language must be an evolutionary discontinuity. In section 2.3, I consider the significance of this, and I demonstrate that there is no reason to think of discontinuities as 'miracles'.

Given the need for a discontinuity theory, in section 3 I investigate the role of discontinuity in the Minimalist program. I begin in section 3.1 with Chomsky's standard conception, which I defend as generally plausible, though I argue that it is on the whole unconvincing in its integration with all that we know about human and nonhuman cognition. In section 3.2 I propose an alternative Minimalist discontinuity theory, which does not stipulate that Merge is biologically recent and unique to humans. I claim instead that the only evolutionary innovation in language was

sensorimotor access to already available syntactic structures, which do not themselves provide linguistic ability. For this to be true, the input to Merge must be purely conceptual, not lexical, and I sketch some support for this by considering how syntactic and semantic structures might be maximally transparent. I end by highlighting some broader programmatic issues.

## 2. Justifications for a Discontinuity Theory

### 2.1. *The Terms of the Debate*

Lenneberg does not give especially principled definitions of the terms ‘continuity’ and ‘discontinuity’, essentially because he adopts them for rhetorical reasons.<sup>1</sup> At the beginning of his discussion in chapter 6, he lists examples of animal noise-making that have been considered primitive forms of communication and then summarises the rationale underlying continuity theories in this way:

Since Darwin has shown that man is not the product of special creation but that he descended from more primitive animal forms, neither his structure nor his behavior are special creations. His forms of communication must have descended from primitive animal forms of communication, and a study of the latter is likely to disclose that there is indeed a straight line of evolution of this feature. (Lenneberg 1967: 227–228)

After exposing several issues with these assumptions (developed in section 2.2), Lenneberg introduces the alternative of discontinuity in these highly subjective terms:

A discontinuity theory is not the same as a special creation theory. No biological phenomenon is without antecedents. The question is, “How obvious are the antecedents of the human propensity for language?” It is my opinion that they are not in the least obvious.

(Lenneberg 1967: 234)

Lenneberg did not need any more precise definitions than this, as his intention was simply to provoke a reassessment of the similarities between human and nonhuman forms of communication. However, this paper will range beyond that reassessment and we will benefit from some more thoroughly worked out terminology.

Minimally, ‘continuity’ and ‘discontinuity’ ought to specify something to do with how traits in different species are related to each other. As Lenneberg noted, ‘continuity’ has been used opportunistically for every kind of cross-species similarity, no matter whether there is any underlying connection, such as a shared genetics. This lax equivalence of continuity with similarity effectively means that any trait we identify as similar to any other is a continuity, while any trait we identify as unique to one species is a discontinuity.

<sup>1</sup> He was aware of this and consequently sought to clarify his usage of these terms in the preface to the Japanese translation (see the editorial of this special issue for brief discussion).

Yet, this precludes the use of ‘continuity’ and ‘discontinuity’ in theories of trait *evolution*, as a trait’s uniqueness is irrelevant to its origins and development. A trait might be shared by two species purely because it has evolved more than once through convergence, each time with a distinct genetic basis. In such cases, the sharing of the trait is a mere coincidence having nothing to do with its genesis and inheritance in any particular lineage. Just imagine that we wanted a theory of human bipedalism, for example. By examining closely related species, we can infer that the trait must have originated and developed in hominins, as it is a distinctively hominin characteristic among the larger family of hominids (Richmond et al. 2001). To understand that hominin heritage, it would be useless for us to catalogue instances of convergent bipedalism in, say, dinosaurs and birds, yet we would be compelled to do so to establish the trait’s continuity, if its continuity is its similarities.

Presuming that we do intend ‘continuity’ to specify trait relationships with respect to their evolution, this means that the term must be restricted at least to trait similarities that have a shared origin, so that they are related by genealogy. Still, even a shared origin is too coarse a measure for addressing another key relationship in evolution—the development of traits over time—as two traits having such a connection is coincidental to their development through descent. Thus, while there was likely one origin for bipedalism in all bipedal hominins, the evolution of the trait in the lineage that led to *Homo sapiens* was independent of its evolution in lineages that were not ancestral to our species.

As an evolutionary term, then, ‘continuity’ is best limited to shared traits that have shared origins and a shared line of descent. These qualifications have a very straightforward expression:

- (1) *Continuity*: The inheritance of a trait in an individual or population.

When a trait is inherited, it is passed on through descent, it extends from one origin, and it resembles its antecedent, permitting some variation. All of the apparent concerns of continuity are therefore contained in this one relationship, while none can be improperly singled out. So far, this definition simply ensures that any intuitions we have about continuity are made appropriate for an evolutionary context, but there are some difficulties in applying it to the evolution of language.

Crucially, to determine the origins of language, we must determine to what extent the various traits that make up linguistic ability have a continuous inheritance, yet this is difficult to do by any direct means, as we have little understanding of their genetics, they leave no environmental imprints, and their physical bases are challenging to identify in whatever dregs of our ancestors the earth occasionally preserves (Tattersall 2014). Instead, we can try to draw inferences about trait continuities by looking for homologues in other living species, with the cross-species reach of a homologous trait giving an indication of how deep into history its continuity likely runs (i.e. as far back as the common ancestor of the most distantly related species with the trait). But while this method is simple in its conception, what homology *is* exactly has been debated and revised in recent biological theory, so we must briefly examine the homology concept itself.

In line with the traditional notion, Mayr states that “ [a] feature in two or more taxa is homologous when it is derived from the same (or corresponding) fea-

ture of their common ancestor" (Mayr 1982: 45). With respect to the definition of continuity I gave in (1), the following is an equivalent restatement:

(2) *Homology*: The intersection of trait continuities.

If we know the continuities of two traits in two species, their intersection would be their point of origin in a common ancestor, while having no intersection would indicate an independent evolution. It is implicit for Mayr (and explicit for, e.g., Wake 1999) that this definition is unassailable, yet it has been challenged. As observed by Butler & Saidel (2000), the homology relation is meant to allow us to identify traits in different species as fundamentally the 'same', despite variation in form and function, but this 'sameness' cannot be established absolutely. Phenotypic traits are constituted at a minimum of genes, developmental processes, and morphology, with each level in this biological hierarchy having its own continuity, so there is an inherent ambiguity in the continuity and homology of traits.

For example, when two distantly related species have traits that are only functionally similar, with structural dissimilarities that do not share an origin, we generally categorise them as convergent. However, such traits can nonetheless share deeply conserved genetic regulation, as is the case with eye development in vertebrates and *Drosophila* flies—despite the profound differences between camera and compound eyes, as well as the structures' lack of common ancestry, they turn out to be regulated by homologous *Pax-6* genes (Callaerts et al. 1997; see Fitch 2011 for the potential relevance of this phenomenon to language evolution). Also, many anatomical structures across species that *are* clearly derived from a common ancestor may nonetheless be produced from different developmental mechanisms, owing to embryological epigenetics. This is the case with Meckel's cartilage in vertebrates, its formation being stimulated by different tissue interactions in amphibians, birds, and mammals (Hall 1989).

In more closely related species, genetic mechanisms can have an unbroken inheritance in several lineages and yet be expressed in morphology only sporadically, so that phylogenies can appear to exhibit multiple convergences that are in fact genetic homologues (e.g., the nucleus rostromedialis cell group found in some fish brains, as discussed by Saidel & Butler 1997). Moreover, even unequivocal instances of convergence might be best understood as arising from a common inheritance of design limitations, rather than from chance similarities in responses to selective pressures, as the regularity of some analogues is too improbable to be coincidence. Wake (1991) provides an example of this involving changes to digit numbers in *Plethodontidae* salamanders.

Some biologists have tried to impose order upon this chaos of continuities by introducing more fine-grained terms that apply at different levels of organisation (e.g., Butler & Saidel 2000), while others have sought to limit the domain of homology entirely to one level, that of developmentally individualised structures (e.g., Wagner 1989, Müller & Wagner 1996). This latter, so-called 'biological' homology concept treats many ordinarily analogous traits as homologous (such as camera and compound eyes), and its advocates have criticised the traditional homology concept, which they label 'historical'. But their criticisms, while valid, have been levelled particularly at the prioritisation of *genetic* history above all else—the basic temporal relation defined by (1) and (2), which does not specify any level of

organisation, remains fundamental to every definition of homology that has been proposed. Thus, we find in Wagner's 'biological' homology a criticism of Mayr's 'historical' conception and at the same time an inevitable resonance with it:

The homology concept is an attempt to identify and name the units of phenotypic organization (i.e. *those body parts that have historical continuity and that can often be found in many species derived from a common ancestor*).

(Wagner 2014: 44; emphasis added)

As in many terminological disputes, what we find advertised as a disagreement on the substance of a concept is really a disagreement on how we ought to use a label to carve up a theoretical space. At heart, attempts to constrain the homology concept to one level of organisation have been motivated by a desire to sift the ambiguous continuities of traits for those that seem somehow more *meaningful*, while an acceptance of a level-free homology concept is open to all kinds of observations that are simply more or less *informative* (often hardly at all). As Wagner points out (2014: 43f.), it would be fatuous for us to consider the continuity of a particular red blood cell, as opposed to the system that generates red blood cells, but it is in the end a matter for terminological hygiene (not biological theory) whether we rework the homology concept so that fatuous applications of it are inexpressible, when we could just as well choose to not apply it to fatuous cases.

The original, quick and dirty definition of homology therefore suffices, so long as we are careful to refer to a level of organisation when using it (Bolker & Raff 1996). If we are interested in the wings of bats and birds, for example, we must identify them as homologous forelimbs, but not homologous wings, as the common ancestor of bats and birds had forelimbs that were not wings (Dickinson 1995). Likewise, although we might expect to find conserved genetic mechanisms for their development, these would be homologues of forelimb (not wing) construction. As I detail further in section 2.3, continuity and homology are always partial in this way, so we must be sensitive to how traits integrate various interacting and non-interacting levels.

To apply these refined notions of continuity and homology to language, we must first observe a critical consequence of defining continuity as trait inheritance (and therefore discontinuity as the lack of it): *All* trait evolution is predicated on *both* continuity and discontinuity, and so all theories of trait evolution must likewise deal in both. A trait's discontinuity is simply its origin or end, at which time it must be created or eliminated rather than inherited, while its continuity is its stasis or change through inheritance. Of course, few traits originate in a single transition between generations or species, so we must not think of discontinuities as clean breaks in evolutionary history. Nonetheless, all traits originate somehow and we can call these origins 'discontinuities', whatever the details turn out to be. But how does this definition, which brings continuity and discontinuity together in all trait histories, relate to the debate on language evolution, which treats continuity and discontinuity as opposites? For this, we need to be certain what we mean by 'language'.

The target that we all share is the species' language faculty: A biological property comprising all the anatomical and cognitive features that make it possible for us to acquire and use languages, which are themselves individual, mind-internal

systems ('I-languages', after Chomsky 1986) that provide some means of mapping between sensorimotor expressions and conceptual content (i.e. sound and meaning). In this context, I will use 'ideas' to refer exclusively to the kinds of thoughts that I-languages access, and 'utterances' to refer to any, even non-linguistic sensorimotor expressions.

In principle, utterances and ideas need not be related at all. In humans, emotive vocalisations, like laughter and crying, are utterances that do not stand for any conceptual content, instead having only pragmatic interpretations, and it is clear from our own internal lives that ideas can form and go unuttered (often for the better). The same is true of other species—vocal productions in birds vary acoustically with no matching variation in meaning, such as there is any (Catchpole & Slater 2008), and at least socially intelligent species seem capable of complex thought without having a capacity to express it (Cheney & Seyfarth 2008).

Whether there are any nonhuman thoughts that are the same in kind as the ideas accessed by human I-languages is a more difficult issue that I return to in section 3.2. Yet, regardless of whether language introduced a new kind of thought, it is clear that the ubiquitous biological systems that are used to utter and to think exist independently of language and are insufficient for it, so an organism can only have language if it has mechanisms for mapping between these systems. This amounts to a species-level equivalent of Chomsky's (1965) distinction between competence and performance: To have a language faculty, a species must have competence mechanisms for doing the mapping between utterances and ideas, and performance systems for putting these mappings to use in actual utterances and ideas. As the fundamental property of language is the utterance-idea mapping, no matter what substance is mapped in particular individuals, I will refer to these competence mechanisms as the mapping mechanisms.

This distinction between the performance systems and their mapping gives theories of language evolution *two* histories to reconstruct. One is the history of how the language faculty as a whole came to be as it is, with its particular integration of all the traits that underlie it; the other is of how a language faculty became possible at all, through the first emergence of a mapping between utterances and ideas. There was no doubt co-evolution, with the performance systems adapting to make effective use of their mapping (e.g., with newly intricate control of vocal anatomy), and likely contributing to utterance structure where it is under-determined by the mapping mechanisms (e.g., Holmberg 2010). All the same, these are nuances rather than counter-examples to the simple observation that language consists of things that are mapped to each other and things that do the mapping.

With these qualified ideas of continuity, homology, mapping, and performance, we can try to determine the continuity of the language faculty by searching for homologues of the human performance systems and the mapping between them. Given that homologues must be sought at particular levels of biological organisation, we are of course constrained by having little understanding of what constitutes the language faculty at any levels besides cognition and behaviour, as its genetic and neurological underpinnings remain mostly obscure. Moreover, what we do know of language cognition is largely an inference from behaviour, which underlies almost all syntactic theory in the guise of acceptability judgements (even our own introspections are of course behavioural data, which we rightly try to ex-

plain with theories of cognition, but which do not themselves reveal our cognitive processes).

As far as nonhuman behaviours are concerned, there are no other living species that can learn and use languages, whatever talents some have to mimic fragments of them, but we can't infer any more from this than that there aren't any homologues of the entire language faculty, developed and integrated as it is in humans. Although our particular configuration of performance systems and their mapping must have originated in hominins, the date for the whole may not be the same as for its parts—it could be that the performance systems have a pre-hominin history and their mapping a hominin origin, or their mapping could be pre-hominin, with the inability of other animals to use language being due to severe performance limitations.

Setting aside polemic, both continuity and discontinuity theories recognise hominin discontinuities in performance. The most obvious example is speech, for which there are no homologues (there may be homologues of its genetic regulation, but the behaviour and much of its neuroanatomy is uniquely human; see Lieberman 2007). Recall that all trait origins are by definition discontinuities, so all that needs to be agreed upon for this is that speech arose in hominins. It should come as no surprise that we all embrace a discontinuity of *some* sort, otherwise it would be bewildering that humans have anything unique about them. However, even where there are disagreements about the inheritance of performance traits, these are unhelpful for categorising theories of language evolution more generally, as such disagreements regard the inheritance of traits that could have evolved differently while leaving our fundamental capacity for language unaffected, and it is the inheritance of that capacity that we are interested in. What matters instead, then, is what theories have to say about the homologues (and concomitant continuity) of the mapping between utterances and ideas, irrespective of the character of the performance systems involved.

Continuity theories are those that find these homologues in abundance. Corballis (2017), for example, suggests that nonhuman great apes use manual utterances to convey ideas, even though they are relatively imprecise and ambiguous. Though these homologues are rudimentary, their prevalence suggests a long, pre-hominin history of inheritance. Continuity theories do not deny that human language has remarkable properties, but all are argued to result from changes to the performance systems and their interactions with general intelligence (syntactic structure, for example, is often dealt with as an epiphenomenon due to constraints on learning and use; e.g., Hawkins 2004). It remains a truism that the utterance-idea mapping must have had *some* discontinuous origin, but continuity theories bury it deep in the histories of our species and others. Lieberman 2002 and Tomasello 2008 are other examples of continuity theories.

Discontinuity theories, meanwhile, go looking for nonhuman utterance-idea mappings and come back empty-handed. All behaviours that have been candidates are viewed as unconvincing abilities of some other sort, making it necessary to conclude that mapping had a hominin origin. There is still a place for continuity in these theories, as the many parts of the language faculty that are not unique to our species must have pre-hominin histories of inheritance, and even the parts that *are* unique have of course been continuous for the short time that they have been around. There



is almost as much variation among discontinuity theories as there is among theories of language cognition generally—Bickerton (2014), Bouchard (2013), and Jackendoff (2002) all argue for discontinuities of very different kinds.

In terms of continuity, then, what distinguishes theories of language evolution is the time that they propose for the origin of the human capacity to map between utterances and ideas, as determined by searching for nonhuman homologues. Either it originated in some pre-hominin ancestor, leaving many traces in many species, or it originated in hominins, leaving no traces in any species but our own (and possibly other extinct hominins). It has undoubtedly been misleading to use ‘continuity’ and ‘discontinuity’ as labels for this distinction, as it does not regard *whether* there was discontinuity, but rather *when* there was discontinuity of the relevant kind. It is only because these theories have such radically different answers, correlating with categorical claims about the existence or not of homologues, that these labels have stuck as shorthand for what is really a division over pre- versus post-hominin discontinuity.

Lenneberg (1969: 642) rightly warned that it would be futile to ever speculate on the evolutionary circumstances that gifted ancient hominins with language. The purpose of biolinguistics, informed though it must be by evolutionary theory, is not to tell stories about our ancestors, but rather to tell stories about ourselves, by giving answers to the sorts of straightforward engineering questions posed by Searle:

Subtract language from a species like us: What do you have? Now add  
language: What are you adding? (Searle 2009: 178)

Nonetheless, it is worth noting that the need for continuity and discontinuity in all traits makes it meaningless to say that some traits evolve ‘continuously’ and others ‘discontinuously’. These are not terms that apply to evolutionary processes. It is therefore wrong to think that continuity theories are gradualist theories, and discontinuity theories saltational. Although these pairings have dominated mainstream debate, the time of a trait’s origin is nonetheless independent of the mechanisms of its evolution. There is certainly a connection between the amount of time a theory allows for the evolution of a trait, the size of its phenotypic difference to be evolved, and the mechanisms that might reasonably bring about that difference in that time, but these are only implications to be explored. As such, many anti-generative theories, though gradualist, are also discontinuity theories (e.g., Deacon 1997), and there are now some saltational generative theories that are continuity theories (see section 2.2).

Thus, if we are led to a discontinuity theory when answering Lenneberg’s question, “how obvious are the antecedents of the human propensity for language?”, we are not automatically committed to the belief that a gradual evolution of language was impossible, or even unlikely. If we find that there are no obvious homologues of linguistic ability, this *could* suggest a sudden emergence of language’s entire complexity, but it might turn out that language has antecedents that are simply *not* obvious, and which required only small discontinuities in hominins for the language faculty as we have it to come into being. If so, its relatively recent origin would be compatible with gradual evolutionary change. Before considering that possibility, in the next section I consider the natures of linguistic and non-linguistic communication to demonstrate the basic implausibility of continuity theories.

## 2.2. *A Theory of Communication*

### 2.2.1. *Giving and Taking Meaning*

The major claim of continuity theories is that the human capacity to map between utterances and ideas has homologues in other living species, so it has a continuous inheritance extending further back than the first hominins (from here, I will frame the debate in terms of whether these homologues exist, but readers should recall from section 2.1 that this is a shorthand for whether the mapping is pre- or post-hominin). As the utterance-idea mapping is the fundamental property of language, continuity theories of course deny what I took to be a basic fact in my introduction, which is that language is biologically unprecedented. For these theories, human language really is just language as *humans* have it; it is only a hypertrophied form of widespread communicative behaviours that are simpler, but which have essentially the same mental basis.

One way to challenge continuity theories is to argue that the human utterance-idea mapping must be carried out by particular mechanisms, and that these mechanisms appear to be absent in other species. Generative theories in the Chomskyan tradition typically do this by arguing that our performance systems are mapped by a generative module that builds syntactic structures which each system must interpret, so an absence of syntactic structure in utterances means they mustn't be generated by the kind of module that is necessary for utterances to map to ideas. There is now a rich literature addressing the possibility for syntax in nonhumans, but most studies have been undermined by shortcomings in experimental design (Watumull et al. 2014), misinterpretations of data (ten Cate & Okanoya 2012), and a lack of clarity on what we should even be looking for (Rogers & Pullum 2011). They are further vexed by the systematic ambiguity with which 'syntax' is used to refer both to *any* regularity in utterance sequencing (whatever its computational complexity) and *only* regularities that match the particular complexity of human morpheme sequences. Nonetheless, while some nonhuman utterances (especially bird calls) do seem to be generated by cognitive rule-systems, their syntax is never more complex than that of human phonology, which has no mapping to ideas (Berwick et al. 2011).

Of course, it only follows *directly* from this that there is no nonhuman mapping between utterances and ideas if that mapping demands a syntax at least as complex as the one that structures human morphemes. Certainly, a mapping like ours, with hierarchical concept composition, requires a suitably powerful syntax, but it is at least conceivable that a much simpler generative module could map between non-compositional utterances and non-composed concepts. If so, hominins might have evolved only the syntactic complexity of language through some new computational apparatus, rather than the utterance-idea mapping itself, and we would require a continuity theory of language evolution. This is a latent possibility in Minimalist theories like Hornstein 2009 and Boeckx 2015, where the emergence of language is ascribed to computational discontinuities that are compatible with (though do not mandate) a continuity in mapping mechanisms (this is unlike Chomsky's theory, where a computational discontinuity *is* the mapping mechanism; see section 3.1).

Although these theories have much to recommend them, in what follows I seek to establish that such syntactic discontinuities are unnecessary and burden-

some to theories of language evolution, provided that a non-syntactic discontinuity is inescapable and is alone sufficient to account for human uniqueness. All the same, readers should take a moment to appreciate the significance of these theories having been developed at all. Generative linguists have ordinarily taken the evolutionary discontinuity of language for granted, owing to the richness of language-specific primitives that have seemed to be innate and which are transparently lacking in nonhuman communication systems (such as the principles relating to case, binding, argument structure, and so on). Now, however, some linguists have pursued the Minimalist reduction of language-specific primitives to such extremes that they flirt with the possibility of continuity, albeit a new kind, with continuity in the utterance-idea mapping and rapid discontinuity in its structural complexity. Proponents of this continuity Minimalism might object to it being considered a continuity theory at all, as a discontinuity in the structure of utterances and thoughts would make human language clearly distinct from all other animal communication systems. However, we have already seen that everyone acknowledges at least one discontinuity in the performance systems, so our strength of feeling about some specific discontinuity is irrelevant to its place in the overall debate about the origins of language. We must always maintain a distinction between the history of the particularities of *our* language faculty, and the history of the defining property of *any* language faculty, which is the utterance-idea mapping—any theory that couples a continuity in that mapping with a discontinuity in its structure (or in anything else) can only be identified as a continuity theory, though its proponents may set themselves apart in other ways from gradualists.

Altogether, this new theoretical landscape makes arguing for discontinuity much less trivial than it has often appeared, with many more consequences for linguistic theory. For now, I will delay further discussion of what constitutes the utterance-idea mapping in humans, as it will be more fatal to continuity theories to reject them without having any particular theory of our human faculty in mind. Instead, another means of challenging continuity theories—this one taken up by Lenneberg—is to show that *any* discontinuity theory is preferable because all continuity theories are inherently implausible. This would be the case if either: (i) none of the nonhuman utterance-idea mappings that are said to exist are of a type that the human mapping could have evolved from, or (ii) there are no nonhuman mappings at all. Justifying this second, stronger claim will be the focus for the rest of this section (Lenneberg's approach is quite different to the one I develop here, though I consider the importance of that difference in section 2.3).

It is vital to recognise what identifying utterance-idea mappings in nonhuman species demands. It is not enough to notice that an organism utters and thinks, or that it utters when it thinks, or even that it reliably utters one thing when it has a particular thought, as the mapping of utterances and ideas is more than their mere coincidence. If I see a leopard and yell in fear, my yell will be *related* to ideas, in that it will occur alongside my perception of danger and will allow others to infer that a danger exists, but my yell will have no mapping to ideas, in that it would not *represent* any of the ideas that it is related to. It might, for example, lead others to think that there is a predator on their trail, that there is a person in need of help, or that someone in a north-westerly direction is making an awful racket, but it would not stand for any of these thoughts, or any of the concepts that they contain. A yell

communicates simply *that* someone has yelled, with the association of yells and heightened emotional states allowing a certain variety of useful inferences about the yell's possible causes.

This is totally unlike genuine utterance-idea mappings, such as the word 'leopard' representing the concept LEOPARD in the expression of ideas that contain that concept. Of course, the relation between words and concepts is not straightforward. It has been clear since Grice 1957, 1969 that the interpretation of linguistic utterances is just as context-dependent as the interpretation of yells, as words do not denote concepts in a one-to-one correspondence. Moreover, some have argued that this is not just pervasive ambiguity, but rather that all words (as types) intrinsically under-specify the ideas that they can refer to (as tokens) in contexts of use (there is an interesting convergence on this point in both pragmatic and syntactic theories; see Recanati 2004 on the former and Borer 2013 on the latter). However, the fundamental difference between something like a yell and something like the word 'leopard' is that, *in context*, only word-like things stand for concepts.

It might therefore seem that we ought to search for nonhuman utterance-idea mappings by searching for nonhuman words, but this is ill-advised (if it is not simply tautology), as many of the properties that human words have in addition to their idea-mapping are not necessary for that mapping. For example, I have already dismissed syntactic and semantic compositionality, just in case nonhuman species have an unstructured, non-compositional mapping. We also cannot presume that meaningful units (i.e. morphemes) will decompose into strings of meaningless ones (i.e. phonemes), as this 'duality of patterning' may be necessary for large vocabularies of mappings (Nowak & Krakauer 1999), but it is not necessary for small ones (Sandler et al. 2011 argue that Al-Sayyid Bedouin Sign Language has no phonological patterning). Moreover, there are phonologically complex units, like 'hello' and 'abracadabra', that are not syntactically compositional and do not map to non-composed concepts, but which we nonetheless think of as words (notice that human language seems to lack syntactically non-compositional words that *do* map to concepts—I think this tells against the very possibility of it even in other species, but I leave it as an open question).

As such, we can *only* consider whether utterances have a mapping to ideas, ignoring all other incidental properties, but we face a difficulty here, as this is a mind-internal property that we cannot observe. While we can probe the invisible boundaries of our own linguistic capacity using circuitous introspective methods, we obviously cannot do the same with nonhuman species. Thus, without access to nonhuman minds, we are limited to a study of nonhuman behaviours, and so must examine their utterances for literally superficial signs of an underlying mapping to ideas. Yet, what visible features of utterances could possibly tell us that they have this invisible mental property? The classic inventory of utterance features is Hockett 1960, and though none of Hockett's features (as he defines them) is guaranteed of all and only utterances with a mapping to ideas, *en route* to such a feature, it is crucial to understand why it cannot be his most likely candidate: 'Semanticity'.

There are really two ways that people think about 'semantics', only one of which is sensible, though both are unhelpful to the task we are embarked upon. Hockett defines 'semanticity' in terms of "relatively fixed associations between elements in messages (e.g., words) and recurrent features or situations in the world

around us.” This is an unabashed behaviourist definition, which may be attractive for locating semantics in things we can observe, but, like classic behaviourism, these observables give no indication whatsoever of mind-internal properties. As Hockett notes, it is semantic by this standard for gibbons to produce alarm calls in alarming situations, though he stops short of recognising that such calls would *still* be semantic by this standard even if produced with no conscious control, and with no representation of conceptual content, just like my yell at the sight of a leopard. All that this semantics requires is a reliable association between an utterance (e.g., a yell) and its contexts of production (e.g., nearby predators), so that hearers can respond automatically (e.g., by running away), *just as though* they can recognise and react to a concept-representing statement like, “there’s a predator!”

Hockett’s kind of semantics was rebirthed as ‘functional reference’ by Marler et al. (1992), with their emphasis on functionalism suggesting that the term should be agnostic about mind-internal properties. So it should be, but it’s important to see that this is not because our knowledge of those properties is limited (though it is), but because functional reference is just irrelevant to them. Primate vocalisations have always been the troublesome case, as primates’ use of the vocal-auditory channel has a seductive familiarity that tempts us to treat their utterances as word-like, when we might treat them as yell-like if they were produced by humans (the classic example being vervet monkey alarm calls, which vary with particular predators, *just as though* they have words for each of them; Seyfarth et al. 1980). Luckily, we needn’t take sides on this issue yet, as all we need to notice is that functional reference *makes no distinction* between words and yells, whatever alarm calls happen to be more like—all it cares about is whether there is an observable association between an utterance and some salient element in its contexts of production, which is true of many kinds of utterances that nobody even suspects to have a mapping to ideas.

If we turn away from primate vocalisations, we find exactly this kind of meaningful association in avian courtship dances, where individuals interact with visual scenes, rather than acoustic ones, to make the same sorts of contextually relevant inferences. Of course, there is no suggestion that a dance might refer to a concept in the dancer’s mind, so functional reference has no grip on the issue (see Wheeler & Fischer 2012 for a detailed discussion of how permissive ‘functional reference’ is). In fact, all of this can be easily extended to non-biological phenomena, where there are no minds producing utterances at all, let alone ones that might possibly be producing utterances to represent concepts. Smoke clouds can be said to have Hockett’s kind of semantics, as their fixed association with fire can be learned by an organism, so that it can infer from any instance of smoke clouds that there must be a fire somewhere. While the notion of functional reference might seem to exclude these mindless cases, it does so only by stipulating an interest in evolved rather than non-evolved associations and responses—it would otherwise construe meaning identically in each scenario, and thus contributes nothing to our understanding of meaning’s mental character.

Having said all this, it is difficult to finally dispatch behaviourist semantics and its reincarnations, as it remains trivially true that a word can only represent a concept by somehow coming into an association with it. Quine (1960) showed that words having associations with observable features in the world is by itself

hopelessly insufficient for determining a word's meaning (see also Gleitman 1990 for a syntactic perspective), but it is at least a precondition for word meaning that a word's use should have some contiguity with the concept that it is meant to represent (i.e. for 'leopard' to come to mean LEOPARD, it is necessary, though insufficient, for 'leopard' to be used *alongside* thoughts of LEOPARD). As such, associations must have some role in a theory of semantics, and therefore in a theory of language origins, but only insofar as our interactions with smoke clouds have a role. What is at issue in all this is only the capacity that some animals have to perceive things as meaningful, rather than any capacity they may have for language in particular, which is an entirely different, derived phenomenon.

To put this more concretely, consider the extremes of a smoke cloud compared to some uttering of "there's a fire." In the right circumstances, both events could be *taken* to mean THERE IS A FIRE (i.e. an observer could have that thought as a consequence of observing the event), but only the linguistic utterance would *give* the meaning that is taken from it—the smoke cloud is not produced in order to specify the thought that the observer should have. What Hockett's semantics does is identify those associations that have a potential to be taken to mean something, and functional reference identifies those evolved behaviours that are in fact taken to mean something, but neither notion addresses whether these meanings are taken purely through observer inference, or through the meaning being given by what the observer observes. Moreover, note that if a meaning is always something like an idea in the mind of an observer (we will explore other possibilities, but the lesson here won't change), then 'given meaning' is simply a reformulation of the utterance-idea mapping, and as this is the defining property of language, language is thus beyond Hockett's semantics and functional reference. These ideas are not inadequate or wrong in theories of the nature and origins of language, they just have nothing to say about language at all, despite their use of linguists' jargon.

At this point, we can settle upon a conception of 'semantics' that is more sensible. Presuming that we do not want to say that smoke clouds have semantics, as this would empty the term of its purpose in a characterisation of language, we must concede that having semantics is not the same thing as having meaning, as smoke clouds can be meaningful. Semantics must therefore constitute some particular, limited contribution to meaning, in which case it could either be observers' inferences, which determine the meanings that are not given by an observed event, or it could be mappings to ideas, where the meanings are given. Of course, linguistic theory already has well-worn terms for these things: Observers' inferences are the subject of pragmatics, and so semantics simply *is* the utterance-idea mapping (that 'semantics' is therefore just another label for what we are trying to identify is why this more sensible conception is also of no help).

It follows from this that the utterances that animals take to be meaningful without there being any mapping to ideas ought to be classified as having pragmatic meaning (and no semantics). That 'pragmatics' is not commonly used this way is no doubt due to the false equivalence typically drawn between semantics and meaning, as well as the reputation that 'pragmatics' has for referring to somehow peripheral, ineffable, and even socially complex aspects of meaning determination, rather than to the simple process of inference from observation. However, as I will argue presently, post-Gricean, cognitive pragmatic frameworks (like Rele-

vance Theory after Sperber & Wilson 1986) provide a robust basis, even if not fully realised, for understanding pragmatics as inferential meaning *wherever* we find it in the natural world, even in socially unintelligent and non-conscious interactions. It's worth bearing in mind here that, because linguistic utterances depend (like non-linguistic ones) on pragmatics for their meaning, the 'utterance-idea mapping' is specifically the non-pragmatic element of linguistic meaning, as it is this alone that distinguishes it from non-linguistic meaning (Carston 2002 has an insightful discussion of these distinctions, though the finer details will not bother us here).

### 2.2.2. Communication Systems

There remains one other possibility for continued confusion that leads us now to undertake a more formal analysis of the distinctive properties of language. While we might be willing to accept that there can be meaning without semantics, and therefore that wholly pragmatic meanings in animal interactions are no more informative about the origins of language than wholly pragmatic meanings derived from smoke clouds and other such cues, it nonetheless seems that there is a boundary being crossed here: Even the most mindless animal interactions can be regarded as *communication*. Naturally, the notion of communication, and its intimations of information transfer, suggests something like the giving of meaning, even though there is no question that most communicative behaviours do not have semantics as I have described it. What we need in order to navigate the difficulties raised by these intuitions is a principled description of the necessary properties of animal communication, which can allow us to appreciate where the utterance-idea mapping fits in a grander evolutionary scheme.

Note that an essential premise of the description I will give is that there is no such thing as metaphorical communication—if we feel that we are licensed to describe something as communicative, that is because it is, even if we are describing a non-conscious interaction, like a poison dart frog dissuading a predator from eating it by having a vivid yellow colour that marks it as poisonous. For the moment, I will focus particularly on what unifies non-linguistic communication, and I will afterwards consider what must be added to form a minimal linguistic system.

Let's begin by focusing on the simplest case of communication, where there is one signaller and one receiver. Every instance of communication involves something, which we'll call the *object*, that the signaller wants the receiver to understand (I will use such mentalistic language as 'wanting' and 'understanding' for exposition, and will encompass non-conscious communication later, so the exposition should not be read as claiming a psychological reality for beliefs and desires in nonhuman communication). There are two important limitations on objects in communicative acts: An object must be something that the signaller and receiver *can* perceive, and something that the signaller *does* perceive. To put this impressionistically, a signaller can only communicate about something that both parties can think about, and the signaller must be in some sense thinking about that thing to communicate about it. Of course, by 'perceive', I mean something much broader than sensory contact with an external reality, as humans are exquisitely capable of talking about fictions, like fairies at the bottom of the garden, as well as about real things that are abstract or not present to the senses. Instead, I take something to be

perceptible so long as it can be thought about at all, and so something *is* perceived when it is thought about, in which case, it is a *percept* (I deliberately avoid ‘conceive’ for this, so that we can later have a different understanding of ‘concepts’).

We can be more precise about this if we consider the property of intentionality. For now, I don’t mean the term as it has generally been used in previous discussions of nonhuman communication, as there it often has only a common sense meaning (i.e. that of volition and desire), or it refers only to a capacity to represent others’ beliefs, following Premack & Woodruff (1978) and Dennett (1983). This latter sense preserves more of the philosophical meaning, but not quite enough of it. Instead, we must begin with intentionality as it was revived by Brentano (1874), and as explored by Searle (1983) and especially Dennett (1987). Quite simply, intentionality is the property of mental states and events that are *about* something, or directed towards something. Belief and desire are certainly the classic examples: If I desire an apple pie, my desire is an intentional state in being directed at the pie. If I believe that the pie is too hot, my belief is intentional in being about that particular proposition. However, these states do not exhaust the possibilities—if I am scared of clowns, my fear is an intentional state about clowns; if I am irritated by a fly, well, you see how the story would go. With this in mind, to say that the object in a communicative act must be a percept (i.e. that it must be something thought about) is to say that it must be something that an intentional state is directed towards. This is satisfactory for now, but we will later on break down this equivalence between thinking-about and intentionality, and it is vital that we understand percepts exclusively in terms of intentionality.

So far, so much stipulation, but why should any of this be true? It’s important to appreciate that the physical reality that is external to an organism never enters directly into communication any more than apple sauce gets smeared across my brain when I think about it. However accurately or inaccurately our thoughts represent the world, all that we know of the world is what we perceive of it through the prism of our biology, and everything that we can communicate about is derivative of this mental organisation of experience. Thus, when we communicate about apple sauce, we really communicate about our mental representations of it, rather than about the sticky stuff itself, and any such mental representation is by definition a percept when it is thought about. More generally, whenever we talk about having any intentional disposition towards something, the disposition must be mediated by a mental representation of the physical thing itself. As such, it is always that mental representation that is the percept, and this is why communication must always be about things as we think of them, rather than about things as they actually are (here lie the beginnings of a criticism of referential theories of meaning, which, like Chomsky 2000, I reject).

For exactly the same reasons that a signaller’s object must be a mind-internal percept, so a receiver’s understanding of the object must also be a mind-internal percept (a receiver is also a biological entity that communicates in mental representations). Of course, these percepts are not publicly accessible—even when we are in some sense thinking the same thing, we do not literally have the same thought, as thoughts are fundamentally physical properties of individual brains—so the percept that is the signaller’s object cannot be the same percept as the receiver’s understanding, which must instead have an independent existence in the receiver’s mind.



This second percept introduces a second intentionality into the communicative dynamic, and we shall call it the *interpretant* in imitation of Charles Sanders Peirce's philosophy, which has many important points of contact with this discussion that I won't develop explicitly (see Short 2007 for an introduction to Peirce's work). Note that communication is not perfect, and there are always two perspectives on interpretants: That of the signaller's intended meaning, and that of the receiver's actually taken meaning (e.g., I might point to something on the windowsill, and you might look through the window). As we are primarily interested in the mental basis and content of utterances, the utterer's desires are paramount here, so for us the 'interpretant' is the percept in the receiver as the signaller wishes it to be.

So far, we have a signaller and a receiver who are synchronous in perceptual abilities, and we have labels for the percepts that they have in a communicative interaction. For example, if I want to communicate with you about some book, we had both better be able to think about it, and, if we can, my thought of it will be the object, and the thought that I want you to have of it, the interpretant. However, we have not yet equipped signaller and receiver with a communicative competence, so that an object can in fact be related to an interpretant. For this, the signaller requires some kind of *sign*, such as an utterance, that can direct the receiver to have the interpretant that the signaller desires, as would be the case if I directed you to share my percept of some book by pointing at the book that my percept represents. Here, the use of a sign embodies a third intentionality, involving the desire that the signaller has towards provoking some particular interpretant. Now, if we have an intentionality, we must also have the percept that it is directed towards, but we must be careful here, as this percept is not quite the interpretant that the signaller desires. I could want you to notice a stain on the carpet, perhaps, and yet do absolutely nothing about that desire, in which case it would have no place in a story about communication. Instead, if I decide to use a sign, it is because I know the sign is a means of bringing about a response, so it is the whole complex of the sign *and* its interpretant that is the percept of this desire.

Signs do not have their meaning intrinsically, so they must have a *ground* that gives them some connection to their object. The grounds that are traditionally recognised are iconicity (resemblance of sign and object), indexicality (existential connection), and symbolism (arbitrary connection), with the three often mixed in one sign (Short 2007). The details of this taxonomy are unimportant here, except for the fact that all grounds acquire an additional property in communicative interactions that it is worth keeping in mind. A non-communicative sign (e.g., a rash) might be taken to represent its object (e.g., an illness) by being indexically linked to it, such that the connection will hold independently of any specific instance of the sign (rashes just signify illness generally by brute association, wherever we find them). A communicative sign, however, must always be grounded with some element of *self-reference*, in that its object is inevitably dependent upon a specific token of the sign. With pointing, for example, though we all have a general plan for how to interpret the gesture when we encounter it, it has no pre-determined object independent of its particular uses (there are as many objects as there are things you can point at). This is true even for symbols, which need not be linguistic—culturally variable hand gestures for approval, for example, do not signify approval abstractly, but rather approval as an attitude of the signaller (note that rashes do not signify

illness as a property of their signaller, as there *isn't* one—the rash is an effect of the illness, not produced by the person who is ill). The reason that this self-reference is necessary is simply that signs that are produced by signallers always have mind-internal objects, and so they are always used to represent properties of their users.

With the object, interpretant, sign, and ground, we have enough in place for many varieties of meaning, but not quite enough for communicated meaning. There are many interactions in which an object and an interpretant are brought into a relation by a sign, but which are *not* communicative. For example, I might have a piece of music in mind that I want you to think about, and I could achieve this by inconspicuously whistling its melodies while you're nearby, without you realising when you think of the piece that I intended you to think of it. Although you thus perceive the object and I get my desired interpretant by whistling, the sign has not been *mutually recognised* for what it is. Similarly, I might go about whistling with no intentions at all, and you might notice the whistling and interpret it as a sign of my good mood, though I had no desire to communicate that state of mind (although I could do just that if I whistled ostentatiously while kicking my heels). Thus, in order for the use of a sign to constitute a communicative act, there must be mutual recognition that the sign is intended to bring an object and an interpretant into a relation (I will comment on the obvious connection this has with Gricean pragmatics momentarily).

Here, we add yet further degrees of intentionality that we must be careful with. First, recognition is itself an intentional state, directed at whatever it is that is recognised. In this case, it is an intentional state in the receiver, directed at the signaller's *use* of the sign (i.e. it is recognition not only of what the sign means, but of the sign being meant to mean what it means). Second, to ensure mutuality, the signaller must have a desire for the receiver to recognise their sign as having communicative purpose, and the percept here is thus the complex of the sign and its recognition. As an example, if I point at something, in addition to noticing what I am pointing at, you must recognise that I am in fact pointing, rather than merely stretching out my arm, and in performing the action, I must want you to have that recognition, as it aids the interpretant I intend.

In the description up to this point, we have encountered all of the elements that we need. To recap, we have five intentional states, each directed at one or more of the object, the interpretant, and the sign (and it is the relations among these that collectively determine the sign's ground). Of the intentional states, two are simply the perceptions of the object and interpretant, respectively; the third is a signaller's desire directed at use of a sign to bring about its interpretant; the fourth is a receiver's recognition of this desire for the sign and its interpretant; and the fifth is the signaller's desire for this purpose in the sign to be so recognised. There are many interdependent elements to juggle in this account, and they are easily confused, but all can be arranged into a more palatable definition:

- (3) *Communication*: Use of a sign to produce an interpretant through mutual recognition of the sign as grounded to an object.

Human non-linguistic communication lends itself most readily to description in these terms. For example, if I communicate to you that someone else's cat is crouching in the garden again by looking wide-eyed from you to the patio door, a

fuller description we could give of this is that I use a facial gesture to produce (in you) a percept of the cat through mutual recognition of the gesture as standing (indexically) for (my) percept of the cat. My claim, however, is that this definition, with one important adjustment, can extend to all possible varieties of communication in the natural world, including those where all of the elements in the communicative act are non-conscious.

Returning to an earlier example, we essentially have to ask what it could mean for a poison dart frog to have communicative desires that it fulfils by signalling a property of itself. Although intentionality was originally believed by Brentano to be the hallmark of mental phenomena, it is not difficult to extend it to non-mental phenomena if we adopt the intentional stance advocated by Dennett (1971). We already do exactly this when we think, say, that a frog *wants* to make its poisonousness known in order to be avoided, and there are reasons why this doesn't seem so ridiculous, even though we don't believe that frogs have desires in the sense that humans do. To adopt the intentional stance towards a system (such as an organism, or even a computer) is to describe it as having beliefs and desires, with a view to predicting its behaviour. Thus, when we say that a frog *wants* to make its poisonousness known, we can predict the development of a sign to represent that property as a fulfilment of its desire.

Unsurprisingly, there has been considerable controversy about the adequacy of using these anthropomorphic descriptions for non-conscious systems, as well as about whether human beliefs and desires might in fact reduce entirely to non-conscious equivalents (see Dennett 1987 and Ross et al. 2000 for more discussion). However, the purpose of the intentional *stance* is not to make claims about the reality of beliefs and desires as psychological states, but rather to predict behaviours through the idiomatic ascription of such states, when this may be the best (and even only) means of understanding behaviour. In this context, we can adopt the stance just instrumentally, without concern for its metaphysics, and we need only think about organic systems.

The critical point is that it is in the nature of evolution by natural selection to produce organisms that can *all* be characterised as intentional systems. It is fundamental to the selective process that it favours those organisms that survive and reproduce, and it favours those that do so most effectively (note that the ubiquitous talk of natural selection as 'favouring' anything at all treats it as an intentional system), so all organisms that exist can be fruitfully described as having a *desire* to survive and reproduce, as well as *beliefs* about how best to go about it. Thus, it is not so startling to say that a poison dart frog *wants* to communicate its poisonousness because it *believes* that doing so will keep predators away, as this is just one way of employing certain vocabulary to talk about targets of natural selection. Of course, we still have all our work cut out for us in ensuring that the traits we talk about *are* so targeted, but, if they are, the intentional stance applies automatically as a consequence. One way of viewing this is to say that any attribute an organism has that it would want to have in aid of its fitness, if it could think about it that way, should be regarded as a percept of its survival intentions. Perhaps a less awkward way of thinking about this is simply to say that any adapted property of a species is a percept of selective intentions, and it is thus natural selection, rather than the frog, that perceives the benefits of the frog being yellow.

With this expanded notion of intentionality, we can understand percepts either as things thought about, or as targets of natural selection. This imposes some important limits on what might otherwise seem to be an unconstrained metaphor, in that our definition of communication requires its elements to be percepts, so if a property of an organism is *neither* a thought or an adaptation, then we should expect that it cannot enter into communicative acts. This moreover means that non-human communication cannot be distinguished from human language by thinking of the former as being externally referential, as the only things that nonhumans can communicate are things that are relevant to their fitness and are thus percepts of natural selection—there cannot be referential meaning, human or otherwise. Bearing all this in mind, we need not alter the definition of communication in (3) to have it apply to non-conscious varieties—all we have to do is be clear that the five intentionalities could be either psychological or selective. Compare the following descriptions of communicative interactions, for example:

- (4) a. Use of a facial gesture to produce a mental percept of a cat through mutual mental recognition of the gesture as grounded to a mental percept of the cat.
- b. Selection of yellowness to produce selection of predator avoidance through mutual selected recognition of the yellowness as grounded to a selected property of poisonousness.

The parallelism is exact even with regards to mutual recognition, as we would discount as non-communicative any interactions where a trait has a signification that evolution did not intend it to have (e.g., when the size of an animal is taken as a deterrent, though that is not its evolved purpose). In the intentional idiom, a frog's colour is rightly regarded as communicative because it is *meant* to be recognised as representing its poisonousness, and it *is* recognised as such, as evidenced by the avoidance response.

Before seeing where this leaves us with language, we should appreciate quite how general this account of communication is, given my claim that it can apply to any communicative system. There is in fact no element of this description that is novel, though it brings together several different perspectives. The triadic conception of the sign, object, and interpretant is of course Peirce's (though Peirce would have rejected such thorough mind-internalism), but, though the triad is excellent for capturing taken meaning generally, it is only by making all three elements intentional percepts that we limn communicated meaning especially. Of course, my particular use of intentionality is Dennett's, whose (1983) discussion of nonhuman intentionality was heavily influenced by Grice, though he construed Grice's approach more in terms of theory of mind than in terms of necessary elements in a communicative dynamic. More important than Grice's approach, however, is its development into the definition of ostensive-inferential communication in Sperber & Wilson 1986:

- (5) *Ostensive-Inferential Communication*: The communicator produces a stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of this stimulus, to make manifest or more manifest to the audience a set of assumptions **I**.

It is by deliberate design that every aspect of this definition has a correspondent in the definition of communication I gave, with the exception of the ground, which is an important but often overlooked interaction of the other elements. In translation, the stimulus is the sign, the set of assumptions is the object, their desired manifestness is the interpretant, and that manifestness is mutual. The definition also has a correlate of the signaller's desire towards the sign and interpretant complex (Sperber & Wilson call this the 'informative intent'), as well as to the signaller's desire towards the sign and its recognition (what they call the 'communicative intent'). The five intentionalities are all here, categorised in the same way, with no more and no less. The only limitation of this definition of ostensive-inferential communication is that it treats human interactions as possibly a special case, dependent upon the subtleties of our social intelligence. Not so—this simply *is* communication as it *must* be, and its ostensive-inferential character is simply the result of the informative and communicative intents being realised mentally, rather than selectively. Compare this with a completely different definition of communication, that of animal signalling given by Maynard Smith & Harper (2003):

- (6) *Signal*: Any act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver's response has also evolved.

Note first that the inclusion of 'structure' here allows physical traits, like skin colour, to be signals, which is what we allow by making it possible for signs to be percepts of natural selection. In this definition, we find that the sign corresponds to the act or structure, the interpretant to the altered behaviour, and the mutual recognition of the sign to the necessary co-evolution of signal and response. There are things missing in this definition, but they are all implicitly necessary—a signal can only have some effect on another organism if it has an object (or it would be meaningless), which it is grounded to (this relation being what underlies the well-studied problem of honest signalling; see Higham 2014 for a recent overview). Maynard Smith & Harper also circumscribe the intentionalities for an interpretant and for recognition by contrasting signals with other kinds of interactions: 'Cues' are those where the trait has a signification it has not evolved to have (i.e. there is no intention for the interpretant), and 'coercion' is where a response is forced without communicative engagement (i.e. there is no intention for recognition).

These approaches to communicated meaning appear very different on the surface as they were designed to deal with very different phenomena, but I find that they all present facets of one underlying dynamic that could be no other way. Once we properly identify the five intentionalities of communication and the various interactions of their percepts, we can abstract from their particular psychological or selective character in particular acts and species, and thereby account for every variety of communication in one like manner.

### 2.2.3. *Language is More than Communication*

Recall that, to establish the discontinuity of language, we must establish the non-existence of the utterance-idea mapping in nonhuman species. Our account of communication now allows us to ask more pointedly how utterance-idea mappings fit

in the overall scheme of utterances relating to ideas. Consider, then, the difference between our poison dart frog, and, say, one person uttering to another, “it’s poisonous”, to dissuade them from eating a death-cap mushroom. What are the differences, and which (if any) are especially linguistic?

The most immediately obvious difference is the mental, rather than selective, nature of the intentions in the communicative act. If language were selective like a frog’s colour, we would all be born speaking the same words, and speaking them automatically in the same situations. Fully selective vocal utterances are in fact what yells are: Selected calls that produce selected alertness in conspecifics through mutual selected recognition of the calls as grounded to a selected property of panic. Note how the fact that yells only *relate* to, rather than *represent*, ideas is partly captured in this by their objects and interpretants being affective states of the signallers and receivers; any ideas that the receivers have are inferences from this affective stimulation, which are not given by the sign.

We must therefore recognise the intentions in linguistic communication as mental, but which ones? So far, I have only discussed cases where all of the intentions are in harmony, but they need not be. For example, it is possible to make conscious *mental* use of a sign as standing for a *selected* object, to produce a *selected* interpretant. If I voluntarily fall about in tears (it being an act, rather than a physiological response), it will be to produce sympathy (a selected response to crying) through mutual selected (i.e. programmed) recognition of the crying as grounded towards an unhappy state of mind (states of mind without percepts also being products of organic design). We therefore have to ask whether all, or only some, of the intentions must be mental for utterance-idea mappings. It is certain that the object and interpretant must be, as mappings are always from and to meaning. Even when an intended interpretant is a behaviour, as in an imperative, it is always targeted indirectly through perception of the mental content specified by the utterance. It is also necessary that the signaller’s intention towards the sign and its interpretant is mental, as not having this mental intention would mean that the sign must be produced automatically, while linguistic utterances are voluntary.

The receiver’s recognition, and the signaller’s intention for it, are similarly clear-cut, though they are more nuanced. If there were a one-to-one correspondence between utterances and the ideas they refer to (i.e. if language were entirely unambiguous), then the receiver would not need to know of the signaller’s intentions towards the sign and its interpretant, as each interpretant could be derived without any understanding of the signaller’s psychology. However, this is nowhere close to reality—language is deeply ambiguous and, as previously discussed, pragmatic enrichment of utterances is an all-pervasive necessity, so there is no possibility for language unless a receiver has some mental recognition of the purpose to which a signaller is putting their utterance, and the signaller likewise intends this recognition. Unlike non-linguistic and non-conscious communication, any communication with a mapping of utterances and ideas therefore requires *all* the intentional components to be mental in character.

We reach an important juncture here, as a lesson we learn from this is that mutual recognition of communicative intentions (i.e. ostensive-inference) is a biological precondition for utterance-idea mappings. This mutual understanding of intentions is the property of ‘shared intentionality’ analysed by Tomasello et al.

(2005), and it is generally agreed to be a human development of theory of mind. In other words, we are all implicitly or explicitly in agreement that a precondition for utterance-idea mappings is not met by nonhuman species, so they cannot possibly have such mappings. The few scholars who suggest otherwise (e.g., Corballis 2017) largely focus on the gestural repertoires of other great apes, which are 'intentional' in the sense of being voluntary, but there is no evidence that they have *shared* intentionality (Byrne et al. 2017). Instead, apes gesture mostly unsuccessfully and repetitively until some behavioural goal is achieved, as their actions are structured to give cues for physical responses, rather than to communicate mental percepts (this is true whether they are interacting with conspecifics or humans). They realise that conspecifics are agents with abilities to satisfy desires, and they may even understand that others have beliefs and desires of their own, but they do not cognize that an individual's beliefs could be about another's, so they lack the leverage of shared intentionality to exchange mental content (note that their interactions can be described as communicative despite not having shared intentionality, as mutual recognition is still instantiated by natural selection, the gestures and their responses being innate). As See (2014) observes, even if we are generous about primate vocalisations, interpreting those as volitional and goal-directed rather than automatic and non-conscious, they would also have the same intentional limitations.

In a moment, I will describe a visible property of utterances that we could use to identify them as mapping to ideas, but so long as we accept the uncontroversial point that only humans have shared intentionality, we have already ruled out nonhuman mappings as an impossibility. Why is it, then, that we are not all discontinuity theorists? Well, we are interested in something more precise than mappings themselves—we are interested in the cognitive mechanisms that facilitate the development of mappings in the right circumstances. We know that shared intentionality is not itself this mechanism, as there is shared intentionality in much non-linguistic communication (i.e. communication that is all pragmatics and no semantics). Continuity theories are therefore those that claim that the mechanism that facilitates the association of utterances and ideas pre-exists shared intentionality, and the evolution of shared intentionality led inevitably to utterance-idea mappings through a drift to abstract linguistic signs (e.g., Tomasello 2008 and Scott-Phillips 2014).

Stating this in some other equivalent terms may set some alarm bells clanging: Continuity theories are those that claim that the mechanism underlying utterances' mappings to semantics pre-exists shared intentionality, and the evolution of shared intentionality led inevitably to semantics, *as just a highly derived form of pragmatics*. Discontinuity theories maintain a categorical difference between pragmatics and semantics, and claim instead that the mechanism for semantic mappings came after shared intentionality and is distinct from it. We therefore need to consider whether this categorical distinction is justified.

We have a framework in place for the meanings involved in pragmatic communication: They are mental or selected percepts, and, as semantic meanings are never selected meanings, we need only be concerned with mental ones. These percepts are *individual* in two very important respects. First, they are mental representations of individuated perceptible entities, whether real or fictional. For example, I can think about and point at some particular chair, but I cannot gesture non-linguistically at CHAIR, the abstraction of all things that can count as some

kind of chair. These meanings are also individual in the sense of their grounds being self-referential—the chair that I point at can only be identified as the object of my pointing by recognising my signaller's intention and following the line of my arm.

Utterances with semantics are also partially individual, as they all require pragmatic interpretation. It is the purpose of any linguistic utterance, like many non-linguistic ones, to specify an individuated mental percept, and this requires an amount of self-reference for ambiguity resolution. However, the essence of utterance semantics is what is *general*, rather than individual. Although I have no non-linguistic gesture for CHAIR, I have a word for it, and I have words for all kinds of properties that cross-cut individuated mental representations—how could one non-linguistically gesture at the colour of a wall, or at the exact number of a heap of stones without counting them out individually? All of these things are *concepts* rather than percepts, and it seems that concepts cannot be percepts, in that we cannot bring concepts before our mental attention. Although we may try to think of our concept BIRD, say, what we will think *about* is some particular prototypical bird standing for the properties that our concept organises, though that prototype cannot be the concept itself (see Armstrong et al. 1983, Connolly et al. 2007). We cannot have a mental perception of the concept as it subsumes everything from the peewit to the penguin, and if we otherwise tried to describe it, we would resort to using other words for other concepts as periphrastic definitions. Concepts are thus like 'filenames' that organise our thoughts (Fodor 2008), but the only things we can actually think *about*, in intentional terms, are perceptible file contents.

While the individuality of pragmatic communication requires the ground of a sign to refer to a signaller, the generality of semantic communication allows non-self-referential grounds, as concepts are not specific to individuals—concepts are relations among percepts, and though such relations must have individual neural instantiations, they are also abstractions over individuals, in just the same way that we must all have different neural instantiations of structurally identical sentences. We can therefore identify utterances with semantics by observing whether their use has a non-self-referential ground (i.e. an independence of signallers' communicative intentions). Of course, this is impossible to observe with any *single* utterance, as all utterances have a pragmatic element, but if we look at *patterns* of use, semantic utterances ought not to be correlated with characteristic intentions for recognition and interpretation:

- (7) *Non-Self-Referential Grounding*: Utterances that map to ideas can be identified by a pattern of use that is independent of particular communicative intentions on occasions of use.

This applies quite neatly to humans, as we can use it to distinguish words that do and do not have a mapping to ideas. For example, 'hello', which is not syntactically or semantically compositional, is strongly correlated with a specific intention to be recognised as signalling a greeting. Take semantically contentful words, however, and they have no strong correlation with specific intentions (we put other words to use for all kinds of different purposes). In case we are not convinced of other species' deficiency in shared intentionality, we can instead observe that none of their utterances exhibit this patterning—the most advanced vocal and



manual gestures in closely related primate species are all directed towards specific communicative goals, and are thus fully pragmatic. This evokes some of the old arguments against behaviourism, the first page of Fodor 1975 stating:

the contribution of [an] organism's internal states to the causation of its own behavior seems sufficiently undisputable, given the spontaneity and freedom from local environmental control that behavior often exhibits. (Fodor 1975: 1)

Here, we could just as well say that an organism's internal mapping of utterances to ideas is sufficiently undisputable when the utterances have a spontaneity and freedom from local environmental control.

This much establishes that semantic and pragmatic meanings are mentally different in kind, but mightn't shared intentionality nonetheless be sufficient for both, as continuity theories claim? There are two points to consider here. In a semantically compositional system, where composition is driven by syntactic structure, accepting standard arguments for a poverty of the stimulus entails that there must be some innate syntactic capacity over and above our pragmatic abilities that is responsible for the semantic composition of morpheme sequences. One can try to get around this by denying the poverty of the stimulus and arguing that syntactic structure is a byproduct of cognitive efficiency or learning constraints, but in setting out the theoretical landscape, I will simply assume the generative position that these arguments are fundamentally misguided.

Earlier, however, I stated that compositionality (and thus syntax) is not an obviously necessary component of the utterance-idea mapping. We can at least imagine a scenario in which shared intentionality at first permits holistic utterances to be associated with situation-specific pragmatic meanings, with forces in cultural evolution later leading to holistic utterances being associated with abstract concepts, thus being more recognisable as morphemes with semantics. To get compositionality, an innate syntactic faculty would still need to be added afterwards, as there is no route from holism to compositionality without it (Tallerman 2007), but still, some form of utterance semantics might have preceded syntactic structure. At heart, what this claim amounts to is that there is no cognitive, mechanistic difference between acquiring the purely pragmatic significance of a word like 'hello' and acquiring the semantic significance of a word like 'cat'—in both cases, shared intentionality, mixed with domain-general associative learning mechanisms, are sufficient for the morphemes to get their meanings, even though these abilities would be insufficient for morpheme *sequences* to get *compositional* meanings.

Though this seems admittedly more plausible, I think it has to be false. As with syntactic structure, the problem is again a poverty of the stimulus, though this time with the non-self-referential grounding that characterises semantic utterances. Utterances with pragmatic meanings, like 'hello', are by their nature identifiable with the intentions of their users—what their users intend by them on individual occasions simply *is* constitutive of their meaning, and associative learning can handle this without issue. However, semantic utterances are *more* than their intended use on individual occasions. It is not just that semantic utterances refer to abstract concepts that are context-independent, but also that their use extends beyond the conceptual referents to which we are exposed, such that there are cases in which

it seems implausible to account for the novel use of a word as being a pragmatic modulation of some meaning that we have observed it to have in a prior instance.

The most obvious examples are metaphors. Consider the word, 'open', for example. It has various ordinary uses, as in 'open the door', 'open the letter', 'open the shop', and 'open the wound'. Each of these uses requires a contextual interpretation—taking a scalpel to a door would not satisfy the meaning as it would for a wound (Recanati 2004)—but it superficially seems that there is some highly abstract sense of 'opening' that applies to them all. Yet, consider the use of the word in Bellow's *The Dean's December*, where a dog, barking as if in protest at the limits of its experience, is imagined to be saying, 'open the universe a little more!' There is no difficulty in assigning an interpretation to this unusual use of the word, but what schematic sense of 'open' applies equally to this use as it does to all the others? It seems rather that 'open' has no core conceptual meaning that is modulated situationally, but rather that it can refer to entirely novel concepts, so long as these concepts have a sufficient resemblance to those of previous uses (Carston 2013). If this is true, then the meaning of a word is strictly speaking not a concept or even a range of possible concepts, but rather some network of conceptual features that implicate concepts as more or less likely to be referred to by the word. As these features and their connection to morphemes is a form of tacit knowledge not supported by experience, there must be some domain-specific mapping capacity that is responsible for organising an individual's lexicon in this way, and this would have to be an evolutionary discontinuity that post-dates shared intentionality. Exactly what this lexical capacity consists in is as yet unclear, but in §3 I will consider some of its broader implications for a Chomskyan theory of syntax. First, however, a few brief comments are warranted on the overall significance of such a discontinuity.

### 2.3. *The Significance of Discontinuity*

In section 2.1, I stated that defining discontinuity as a lack of inheritance is compatible with evolutionary gradualism, so the fact that a discontinuity theory is necessary should not dishearten those who feel that an adaptive explanation of language origins is essential. To understand the relation between discontinuity and gradualism more precisely, we first need to acknowledge that choosing traits to study can be somewhat artificial if it is not done carefully, as we are led to deconstruct whole, integrated organisms into collections of interacting parts. Often, we are biased in this by how we, as humans, have evolved to categorise the things we perceive in the world, and this need not relate in any informative way to the actual genetic, developmental and structural organisation of organisms. We therefore tend to end up with false impressions of traits as being somehow isolated from the rest of the organism in which they are incorporated, suggesting that they have been targeted by selection pressures, even though they may be nothing more than non-selected byproducts of other traits, or 'spandrels' (Gould & Lewontin 1979).

The 'biological' conception of homology that I discussed in section 2.1 has tried to counteract these problems by focusing on just those traits that are developmentally individualised, with the expectation that their evolutionary development will likewise have been independent of much of the rest of the organism (Wagner 2016). However, even if we succeed in identifying such traits, this ontogenetic and

phylogenetic independence does not eliminate a fundamental reliance that all traits have upon pre-existing organic material to come into existence at all, and this basic fact grounds all discontinuities in continuity.

To demonstrate this, it is perhaps unhelpful to search for developmentally individualised traits, as we know far too little about how language cognition individuates in this way. Instead, we can learn from the similarly messy, integrative trait of bipedalism. There are many other traits besides bipedalism with longer genetic histories that had to exist prior to its development—some are obviously necessary precursors, like having at least two legs, while others are more subtle, like having a particular configuration of the legs, hips, and spine. All of these traits would have evolved for their own independent benefits, and none with any foresight for making bipedalism inevitable or even possible, but only when they came together in conspiracy could a rudimentary form of bipedal behaviour be experimented with, and subsequently refined. In a sense, though we can identify the discontinuous origin of bipedalism with the narrow window of time in which something resembling bipedal behaviour began, the *possibility* of its origin was much more deeply rooted in the history of life. Every trait has these *substrates*: Traits that are the basis for other traits to evolve.

When we identify a trait of interest, then, establishing some vague time for its origin is only the first step in an explanation of its evolution, and the easiest. True understanding comes with identifying a trait's substrates in antecedent species, before the interactions of variation and selection created it in some primitive form and then developed it through descent. Of course, all substrates are themselves traits with their own substrates, each one implicating more and more, until the entire organism and its evolutionary history are involved. Part of the challenge we face in this is to delineate traits and trait relationships in ways that can elucidate an organism's functions and their history. With regards to language evolution, a useful example is speech physiology. Our species is capable of producing a variety of speech sounds, and we know from the highly limited abilities of other apes that this capacity arose as a discontinuity in hominins, but how did its substrates lead to the possibility of that origin?

The production of speech sounds has many phenotypic and genotypic substrates, minimally and most importantly including the anatomy of the vocal tract and its integration with fine motor control. Consider, then, a hypothetical scenario in which some hominin ancestor had a vocal tract configured roughly like that of modern humans, but without the neural integration necessary for any actual speech-making behaviour. Tattersall (1998) suggests that this is in fact close to the truth, and macaques provide an intriguing parallel—a study by Fitch et al. (2016), revising the methods and conclusions of an influential paper by Lieberman et al. (1969), showed that macaques have vocal tracts that are physically adequate for a variety of speech sounds, though the monkeys don't produce any, presumably because of a lack of neural control, rather than motivation. Given a scenario like this in hominins, we would obviously conclude that the ability for speech was brought about by modifications to our ancestors' motor control of the vocal tract. Thus, the discontinuous origin of one trait—speech—would have been predicated on underlying continuities in its substrates—vocal tracts and motor control—which

underwent a gradually transitioned integration (see Ackermann et al. 2014 for one such story).

Ultimately, what this means is that it is never quite right to talk of a trait discontinuity as a discontinuity *tout court*. All traits are the integration of their substrates, and all trait continuities are the integration of their substrates' continuities. Moreover, as the evolutionary process is enormously conservative—recall that the average genomic difference between humans and chimpanzees is around 1–2% (Chen & Li 2001)—every discontinuity arises in a context of massive substrate continuity. It is in this sense that all discontinuities can and should be construed as special cases of continuity: All are tips of innovation on icebergs of conservation, with the tip unable to exist without its supporting structure.

Bearing these aspects of discontinuity in mind, Lenneberg's approach differed considerably from the one I developed in the previous section, as he regarded the discontinuity of language in a deliberately trivial way, noting only that every species is by definition discontinuous with every other, and he felt that the unique properties of language ought to be explained as much as possible through developmental reorganisations of commonalities (see Boeckx & Longa 2011 for a detailed discussion of this point from a more modern perspective). Thus, while Lenneberg did not deny that language is unique among animal communication systems, he avoided talk of language as an integration of other, independent traits (and especially the idea that some might be *sui generis*), while trying to explain characteristics like syntactic structure as developmentally complex expressions of widely shared conceptual capacities for categorisation and differentiation.

Initially, this seems antagonistic to any attempt to isolate a constituent property of language and claim it as an evolutionary capstone that brought other traits together in a new system of linguistic behaviour. However, note how the hypothesised discontinuity in speech described above was behavioural, and predicated on continuities in traits at the lower anatomical level. Just as was the case in the discussion of homology (and drawing parallels with Marr's 1982 plan for understanding the visual system), if we are careful to not mix up different levels of analysis, we will find that these different approaches to trait evolution are complementary.

By simply identifying a unique behavioural or cognitive feature, we are not forced to make any claims about it corresponding exactly and exclusively to some individualised developmental process. It could well be the case that the utterance-idea mapping, or Merge, is nothing more than a theoretical abstraction that labels a phenomenon which emerges from our complex biological design, but which has no encapsulated existence at a physical level. As such, Lenneberg's program is crucial for identifying how these sorts of discontinuities can result from gradual descent with modification. However, just as we mustn't take 'biological' homology too far in its obsession with developmental individuation, so we mustn't do the same here—if we don't develop the theoretical abstractions in the first place, we will not even identify what it is that is to be explained in terms of its physical implementation. Indeed, approaches that repudiate these abstractions inevitably have their own; they are just either so mindless or simplistic that they are not seen for what they are. Of course, our abstractions must be open to significant revision upon discoveries in neuroscience and developmental biology, but we will end up

understanding nothing if we pretend that we can understand everything from the bottom up (Krakauer et al. 2017).

In sum, there is nothing upsetting about discontinuity, and nothing inherently saltational about it, as all trait origins are discontinuities, and all discontinuities have continuous substrates. Moreover, the task I have been engaged in is specifically the identification of a cognitive discontinuity, independent of any claims about its developmental organisation. Of course, the real test for such a discontinuity, as I will explore in the next sections, comes with actually explaining its origin. What discontinuity theories of language evolution must do is make clear what substrates preceded linguistic ability so that it could have had a hominin origin. As stated previously, there are many potential varieties of discontinuity theory, but I will only be concerned with Minimalist ones here.

### 3. Discontinuity in Minimalism

#### 3.1. *The Standard Account*

In the context of language evolution, the main attraction of Minimalism—and admittedly its implausibility to some (e.g., Pinker & Jackendoff 2005)—is its offer of a clear response to what has been dubbed ‘Darwin’s problem’. This is simply the question, “how could language have evolved?”, only asked quizzically because we appear to be dealing with an inconceivably speedy evolution of a biologically complex system. As Lieberman (2015) states, Darwin’s problem does not exist in continuity theories, which hypothesise an evolution over many millions of years, but we’ve seen that continuity theories have more serious things to worry about.

Chomsky has also in various places (e.g. Chomsky 2010) referred to this as Wallace’s problem to emphasise the doubt he shares with Wallace about the mechanism of natural selection being able to produce linguistic competence (Wallace 1869). However, Wallace also doubted that natural selection could be responsible for human hands and hairlessness, so his was not a cogent reflection on the relationship between human cognitive abilities and evolutionary processes. Ultimately, whether it’s Wallace arguing incredulously for a guiding divinity or, say, Pinker & Bloom (1990) flogging adaptationism as an elixir for complex design, too many hypotheses about language’s evolution begin with astonishment and then search for mechanisms that can create astonishing things. I think we are safer starting with biological properties.

Biologically, then, the core motivation of Minimalism is to narrow the difference between humans and other animals while maintaining a qualitative separation in linguistic competence, so that the origin of language can happily fit into the available time. Ideally, this should be accomplished by some very small genetic changes with large phenotypic effects, so that there would be no need for inconceivable amounts of mutation and selection. If this were possible, accusations of fanciful saltationalism would then fall flat.

Linguistically, Minimalism aims to do this by unifying the diverse structural phenomena that were at the heart of Government and Binding theory (GB), such as binding relations, case assignment, and argument structure, with more basic principles of syntactic derivation. As Hornstein (2009: 7f.) notes, the GB vision of the

language faculty was internally modular, its phenomena having their own independent primitives, and this is a highly suspect property for the language faculty, as it is unlikely that such modules could have evolved and integrated in so short a time. Indeed, even over many millions of years, it is difficult to see how these modules could have developed as basic units of language ability, rather than their properties being emergent (see Poeppel & Embick 2005 on the ‘mismatch problem’ between the primitives of linguistic theory and cognitive neuroscience).

There are now many approaches within the Minimalist framework, but all try to stray as little as possible from the Strong Minimalist Thesis, which claims that all the structural properties of languages are derived from the binary set-forming operation, Merge, taking lexical items as input, building them into recursive hierarchical structures while constrained by third factor principles of efficient computation (Chomsky 2005), before converging at interfaces with the sensorimotor and conceptual-intentional systems.

Of many expositions, Berwick & Chomsky (2016), hereafter B&C, give a concise summary of how the Strong Minimalist Thesis ought to lend shape to our biological history:

In some completely unknown way, our ancestors developed human concepts. At some time in the very recent past, apparently some time before 80,000 years ago if we can judge from associated symbolic proxies, individuals in a small group of hominids in East Africa underwent a minor biological change that provided the operation Merge—an operation that takes human concepts as computational atoms and yields structured expressions that, systematically interpreted by the conceptual system, provide a rich language of thought. These processes might be computationally perfect, or close to it, hence the result of physical laws independent of humans. The innovation had obvious advantages and took over the small group. At some later stage, the internal language of thought was connected to the sensorimotor system, a complex task that can be solved in many different ways and at different times.

(Berwick & Chomsky 2016: 87)

Although I ultimately favour Minimalist descriptions of syntactic derivation, I think this account has some irreparable conceptual flaws, deeply rooted in a certain view of language that we must take pains to pick apart. First, though, let me discuss what is *not* wrong with this description.

An architecture consisting of Merge (or something like it), with some interface and third factor constraints, I fully take for granted. There is no issue with positing that a minor biological change could have produced Merge (though the talk of its adaptive benefits is perfunctory), as it is computationally basic. It may be a composite operation of, say, concatenation and labelling (Hornstein & Pietroski 2009), but, however it is implemented, a capacity to compute recursive set formation is an all-or-nothing deal.

Of course, a consequence of accepting both that syntax is fully captured by Merge and that Merge evolved almost instantaneously is that all syntactic phenomena ought to have appeared instantaneously too. At risk of contradicting myself, I share the doubts about this (e.g., Tallerman 2014), but in a specific weak sense. It

has been the traditional approach in generative linguistics to have syntactic derivations fully determine utterance output. This is exemplified by Kayne 1994, where the Linear Correspondence Axiom was intended to ensure a one-to-one correspondence between the left-to-right sequence of morphemes in an utterance and their top-down sequence in a syntactic tree (this approach had a notable cross-linguistic application in Julien 2002). However, it is now common practice, if not universal, for syntactic derivations to *under*-determine utterance output, so that many features of left-to-right morpheme sequences must be imposed by the sensorimotor systems, creating many language-particular discrepancies with the underlying syntax. Although this inevitably makes the connection between utterances and syntax more oblique, so that it is harder to work out what syntax is doing (Anderson 2013), it means that many traditionally 'syntactic' word order phenomena may not be 'syntactic', in the sense that they arise after mental computations have finished and been spelled out. Thus, the full panoply of traditionally 'syntactic' (i.e. word order) phenomena would not have arisen instantaneously with Merge but this is because they are not due to Merge, in which case they are after all not 'syntactic' in the relevant sense (i.e. computationally derived). The small-print of this Minimalist offering is that Merge will give you everything, but everything isn't as much as it used to be.

Also unproblematic is the suggestion that concepts precede language, and the structures built by Merge are primarily connected to the conceptual system, rather than the sensorimotor systems. The most obvious reason for this is that there simply isn't any inherent hierarchical structure in utterances, as they are one-dimensional acoustic or visual strings. We are only able to infer that this structure is supposed to be represented in utterances, and perhaps form our expressions with prosody that respects and suggests it, because we have an innate mental prejudice to discover it, given that utterances map to semantic compositions that *are* syntactically structured by necessity. Moreover, while hierarchical structures have clear benefits in a system of thought, allowing the generation of internally complex ideas, there has been no convincing demonstration that a non-hierarchically structured signalling system would be improved by this additional complexity (Számádó & Szathmáry 2006)—unless, of course, it was to represent ideas that already possessed that structure.

The fact that utterances have syntactic structure only secondarily often creates a confusion about what exactly linguistic competence is, given that the syntactic structure of concepts is a precondition for anything we would identify as linguistic expression. Does this mean that linguistic competence after all cannot be the ability to map between utterances and ideas, as I have described it, and must instead be the ability to generate syntactic structures, even without actual or possible externalisation? While it is undeniable that linguistic *structure* is rooted in the conceptual system, it is stubborn to insist that linguistic *competence* is merely that structure-building capacity. On the standard Minimalist account, there was a time, however brief, when hominins had syntactically structured ideas that were not externalised, as the sensorimotor systems had not yet been integrated, and it would be foolish to call these intelligent but mute hominins linguistically competent. Allowing for the possibility that syntax originated in ideas therefore does not affect my statements characterising linguistic ability as the utterance-idea mapping. Competence is therefore not the same thing as the 'narrow' language faculty limned by Hauser

et al. (2002), which may in any case not be helpfully defined (see especially Boeckx 2013 on this).

Finally, there is mounting evidence that refinement of the sensorimotor systems, with neural integrations allowing complex acoustic signals and a vocal learning capacity that far surpasses that of other primates, would have had sufficient time to evolve, even if the developments were entirely hominin. As stated already, there is evidence that our vocal tracts were speech-ready long before speech use, even if not quite so exquisitely (Tattersall 1998), and there appear to be deeply conserved genetic bases for vocal learning that may find a rapid implementation in the right circumstances (Pfenning et al. 2014).

Now to the problems. As stated in section 2.3, identifying the trait (Merge) and the time of its origin (in hominins) is only the easy work of setting the parameters for the question (perhaps the arrival of the Minimalist program over a century after *On the Origin of Species* should be a warning to us). What an actual theory of language evolution needs to do is identify the substrates of the identified trait in antecedent species to explain how its origin and spread was possible. For the language faculty, the substrates that we might identify could be genetic, anatomical, neurological, cognitive, behavioural, and even cultural. Although it is encouraging that B&C and Boeckx (2017) have attempted to identify a neurological basis for Merge (for a short review of neuroimaging data also see Friederici, this issue), in order to make it seem less like a bolt from the blue, I think there is still so much to be said about the computational substrates of the language faculty (again setting aside their developmental basis as a separate issue) that we don't yet need to venture into other territories.

Although Merge would have appeared rapidly, its first interactions would have been with the conceptual system, and the sensorimotor systems could have been refined quickly, what seems entirely unreasonable is that Merge's integration with the rest of our cognitive faculties had to come for free. Bear in mind, Merge is not just an operation with some output, it is an operation with very specialised input ('conceptual atoms', which are supposed to be unique to humans), as well as outputs that have to be interpreted by interfaces with independent cognitive modules of phonology and semantics.

Even if we can explain how syntactic computation originated in hominins (ignoring the mystery of its conceptual input), we would still have to account for the cognitive mechanisms that translate its structures into forms that are intelligible to these other mental systems. For this, it does not matter that the sensorimotor and conceptual systems are comparatively ancient—syntactic structures are purported to be generated in a novel autonomous module that these other systems now interface with, so their mechanisms for interpreting syntactic structures must be every bit as novel as the syntactic process itself. For example, if the meaning of some syntactically structured utterance depends upon interface rules that translate syntactic configurations into logical forms, those translation rules must surely post-date the origin of syntax. What would it mean to say that the conceptual-intentional system was able to interpret syntactic structures before there were any such structures? Did Merge bring its own interface rules, or did the interfaces evolve their rules independently? Merge is generally thought to be constrained in its derivations by conceptual impositions like the functional hierarchy, but if Merge is just



a set-forming operation whose autonomous structures are built for interpretation, how did the conceptual-intentional system ever come to impose these constraints? To date, the most prominent answer is only that Merge is not so constrained (e.g., Chomsky 2004) and that the interfaces filter out sequences they don't like. As I will come to explain more fully, these *ad hoc* stipulations are the result of an unnatural severance of a syntax that should be closely intertwined with its semantics.

To make a start at this resolution, let's return to the three as yet unexplained hominin discontinuities that the standard Minimalist account proposes. Paraphrasing the earlier quotation from B&C, these are (in order of origin):

- A. The origin of (some relevant notion of) atomic concepts.
- B. The origin of Merge within the conceptual system.
- C. The origin of sensorimotor integration with the output of Merge.

Immediately, we can note that this improbable sequence must be incorrect if all of (A)-(B) are hominin origins *and* independent of each other. Although language, as an autapomorphy, is simply biologically improbable, if (A) does not predispose a species to (B), and (B) does not (C), it is inconceivable that hominins should have accrued all of these fortunes while nonhumans have been blessed with none. Nonetheless, it is basically correct to characterise the language faculty as constituted somehow of the elements in (A), (B) and (C), and in that order, so we have two routes to plausibility: Either these are *not* all hominin in origin (i.e. other species possess at least concepts, and perhaps Merge), or these are not independent.

The approach taken in B&C entails that language evolution, though described in three stages, after all required only *two* innovations—concepts and Merge—with sensorimotor integration in some sense inevitable or free (and thus not an independent development). What this means is either that sensorimotor integration is so cheap and valuable after Merge evolves that it is unavoidable, so a species simply cannot find itself at (B) and not proceed to (C), or the sensorimotor systems have always had some kind of access to the content of the conceptual system, so that (C) was not the evolution of lexical and phonological *capacities*, but was rather a mere refinement of sensorimotor physiology to make the externalisation of an unbounded system of ideas more effective. In either case, for B&C other species cannot have Merge because its presence is sufficient for its externalisation.

B&C in fact seem to believe that, whatever integration of the sensorimotor systems was necessary, sensorimotor access to concepts is ancient (and therefore free, in the sense above), as they take utterances in every species they consider to be straightforward reflections of conceptual ability (see Tallerman 2014: 208f. for more discussion of this). This effectively means that they draw an equivalence between words and concepts, or, more weakly, that they presume conceptual repertoires to guarantee lexical ones, so that the lack of words in other animals (which is undoubted for the reasons given in section 2.2) means that animals have no concepts. They do state (p. 84) that other primates have “conceptual structures”, but by this they seem to mean that primates can make relational distinctions between things like actors and actions, singular and plural, without anything like property categorisations.

There are bizarre consequences to claiming that animals do not have concepts, as demonstrated by a discussion of the chimpanzee, Nim, where B&C cite Pettito (2005: 86f.) before concluding that he had neither words nor concepts:

Chimps, unlike humans, use [...] labels in a way that seems to rely heavily on some global notion of *association*. A chimp will use the same label apple to refer to the action of eating apples, the location where apples are kept, events and locations of objects other than apples that happened to be stored with an apple (the knife used to cut it) [...] all simultaneously and without apparent recognition of the relevant differences or the advantages of being able to distinguish among them [...] Surprisingly then, chimps do not really have “names for things” [...] They have only a hodge-podge of loose associations with no Chomsky-type internal constraints or categories and rules that govern them.

(Berwick & Chomsky 2016: 146)

Nim clearly did not use anything in a word-like way, but if the sign he used was used of eating apples, apple locations, apple actions, apple tools, and so on, then he likely had the *concept* of an apple, as that is the only psychological constant that relates all of these disparate references. However, by allowing him the concept, we are not compelled to somehow interpret his use of the sign as being intended to mean APPLE, which I agree is impossible. Rather, his haphazard pragmatics seem very much like a child at the one-word stage, where single words can represent complex thoughts that include the referent—for example, *doggie* for, “there’s a dog!”, “can I pet the dog?”, and “the dog is biting me!” (Jackendoff & Wittenberg 2014). What Nim lacked may not have been the concept that his trainers wanted the word to signify, but rather an ability to form systematic associations between repertoires of concepts and signs, so that he could pass the one-word stage and begin expressing structured ideas in a way that respects things like lexical categories and functional relations. In other words, this is just what sign-use looks like when you have concepts and a reasonable pragmatic ability, but no mechanism for forming semantic mappings.

Moreover, just as Nim would have been incapable of mentally representing his handlers’ intentions, his handlers appear to show a systematic oversight of Nim’s own intentional competence. As we know chimpanzees to be limited by their lack of shared intentionality, we can only expect that, even if they have concepts, they must fail to communicate them, as being able to do so requires a degree of theory of mind that they don’t possess. Indeed, Nim’s behaviour is remarkably similar to the goal-orientated activity of chimpanzees in the wild, whose behaviour is highly likely to be structured by competences to categorise objects and properties in the world as concepts (see Fitch 2017 for more discussion of the false conclusions that B&C draw from limitations on communication of ideas).

The identity between word and concept that B&C tacitly assume is particularly evident in their conclusion from that same passage:

It appears that chimpanzees are perfect examples of pure “associationist learners”—what they seem to have are direct connections between particular external stimuli and their signs. *They do not seem to regard the apple*

*they see in some mind-dependent way [...]* Rather, they have stored a list of explicit, *mind-independent* associations between objects in the external world and the ASL signs for them.

(Berwick & Chomsky 2016: 146–147; emphasis added)

Again, rather than Nim using associative intelligence, B&C interpret Nim *as* an associative intelligence, his psychology a repository of statistical correlations between stimuli, with no conceptual understanding of types of objects and events that he encountered—Nim was just one of Skinner’s superstitious pigeons in an ape costume. I in fact agree that Nim stored an explicit list of associations between external objects and signs (i.e. he had only Hockett’s semantics, not the real deal), but the portion I italicised is a non-sequitur: Just because Nim was unable to form an association between utterances and concepts does not mean that Nim had no such concepts, as there are cognitive capacities *besides* having concepts that are necessary to form systematic sensorimotor representations of them. B&C are thus led to a catastrophic vision of human evolution, with language not just a capacity to communicate structured ideas, but the primary source of ideas, fundamentally reconstituting the way that our brains organise and process information:

Human cognoscitive powers provide us with a world of experience, different from the world of experience of other animals.

(Berwick & Chomsky 2016: 86)

Their continued underlying assumption is that utterances across species are a direct insight into their mental life. All one has to do is listen, and one knows what it is like to be a bat (see Nagel 1974).

Against these ideas, not only is there substantial empirical evidence that non-human species have concepts as humans have them (see Hurford 2007 for a comprehensive review), it also seems cognitively necessary, in that the same associative mechanisms that were rejected as a theory of utterance-idea mappings in the rejection of behaviourism are just as incapable of explaining *non-linguistic* behaviour—nonhuman species would simply not be capable of their richly structured behaviours were it not for a capacity for forming mental abstractions of sensory experience. Dupré (1996) has correctly pointed out that to assume otherwise is to be beholden to the Cartesian view that there is a gulf between humans and nonhumans, the latter being unthinking automata.

Importantly, if we are compelled to believe that other species have concepts, then, as they don’t have words, words and concepts must be fundamentally dissociated, and this dissociation means that B&C must be wrong to think that sensorimotor access to the conceptual system was cheap or free—if so, animals with even the most basic conceptual systems should be capable of the most basic meaningful expressions, and this is not what we find. As explored in section 2.2, despite the abundance of nonhuman conceptual abilities, the *representation* of concepts in non-human utterances isn’t tiny, it’s nil, as it’s impossible. As sensorimotor access to concepts must be secured by a specific capacity that uniquely evolved in humans, it’s absolutely essential that we understand the role of Merge in this, as if Merge in addition to concepts is sufficient for their externalisation, then other animals ought not to have Merge, but if Merge is genuinely a conceptual tool, and concepts have

no automatic connection with sensorimotor representation, then Merge might be found in non-speaking animals.

Though evidence of syntactically structured thought in nonhuman species is slim, and we perhaps don't even know how to confirm or deny it convincingly, it would arguably be a parsimonious outcome, as other animals could be said to possess the sorts of ideas that Merge can generate (taking concepts rather than words as input), and many species seem capable of mentally representing relationships like *who* did *what* to *whom*, with the kinds of argument structures that are secured by a hierarchical syntax (see Carruthers 2002 and Gallistel 2011 for discussion). What prevents them from having the same cultural and symbolic intelligence would therefore not be a lack of Merge, but a lack of an ability to externally represent what Merge generates, so that they could acquire more complex conceptual structures by building a shared lexicon (Tallerman 2009). Likewise, what caused the flourishing of symbolic behaviour in human societies around 100,000 years ago may not have been our species' acquisition of Merge, but rather our acquisition of a capacity to express its structures. Of course, if feasible, this would also be attractive within the Minimalist aesthetic, as it would reduce the chasm between humans and other species, while still maintaining some domain-specific linguistic capacity.

Arguably, a key motivation for B&C to hold to their position is that if we allow Merge to exist without externalisation, it is more literally a device for structuring thoughts than they ever intended, as syntax could no longer be understood as a fully autonomous module facilitating interaction between the performance systems—Merge would just build ideas, and whatever maps those ideas to utterances is a mystery again. Although the hierarchical structure of utterances and ideas would be as central to linguistic competence as ever, *narrow* syntax would be gone, and, as hinted at twice already, much of what we previously believed to be determined by mental computations would instead be effects of externalisation processes. In the remainder of this paper, I will briefly outline some ideas that lend support from a more linguistic perspective.

### 3.2. *There's Minimalism and then There's Minimalism*

If it is true that Merge has no necessary connection to utterances and can exist in many species purely to structure ideas, then there is no prior reason for the objects it derives to feature elements or relationships that have no semantic interpretation. Moreover, though syntax might in the end be imperfect and generate some things that are not semantically interpreted, none of its redundancies should be coincidentally targeted for sensorimotor legibility. For example, in a standard T-model of the language faculty, syntax may undertake a movement operation to ensure the left-to-right order of words in an utterance. However, if this movement operation has no purpose at logical form, it ought to be prohibited from syntax and the displacement explained by externalisation mechanisms (see Tallerman 2014 for some possibilities).

This suggests a broader research program to demonstrate that Merge can indeed exist as a conceptual device with no necessary sensorimotor connection, or, in other words, to demonstrate that Merge only needs and uses atomic concepts instead of lexical items (see Burton-Roberts & Poole 2006 for some reasons why this

undertaking is attractive for reasons independent of the arguments presented here). Whereas the Minimalist program proffers the Strong Minimalist Thesis, claiming that Merge is an optimal solution to the needs of two interfaces with the performance systems, we can test what I will call, only half-facetiously, the Minimalist Minimalist Thesis, which claims that Merge meets only the needs of semantics, with all utterance phenomena that are not reducible to the derivation of a logical form being explained by externalisation pressures imposed upon completed derivations. Testing this properly is beyond this paper, but it has some initial promise.

We can approach the core issues by thinking about the ways in which syntactic structure can relate to semantics. If what we are after is some kind of equivalence between syntax and semantics, then it is immediately tempting to reconsider generative semantics, but I take it that that paradigm failed for good reasons (Newmeyer 1980). However, we should bear in mind that the fault with generative semantics may not have been its proposal that semantics should be homomorphic with syntax, but rather that it attempted its homomorphism *pre-derivationally*. Let us think about various possible derivational relations, then:

1. Configurational interpretation
2. Rule-to-rule interpretation
3. Piggybacking
4. Derivational equivalence

The first two terms are due to Bach (1976). In a configurational system, Bach says, interpretative rules take “a [syntactic] structure of such and such a form [and] translate the structure into an expression in the interpreted language of such and such a form” (Bach 1976: 184). In such a system, rules of arbitrary complexity map completed syntactic structures onto logical forms, with any input suitable for any output. Although this sounds unconstrained, limits are of course imposed by explanatory elegance and it has perhaps been the most common approach.

There is some variation in what rule-to-rule interpretation is supposed to mean, but the interpretative rules ought to be in some sense less arbitrary, with every application of some well-defined syntactic rule having a direct semantic consequence. Lohndal (2014) develops a version of this where every cyclical Spell-Out of syntactic structure creates a conjunct in a neo-Davidsonian semantics (Davidson 1967, Pietroski 2005). Conversely, this means that every conjunct in a logical form ought to correspond to a Spell-Out domain.

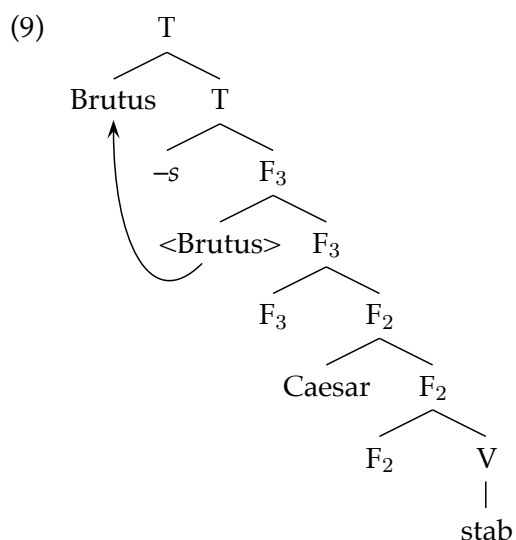
What I call ‘piggybacking’ is just a special type of rule-to-rule interpretation that is maximally transparent. For this, *every* syntactic operation ought to have an immediate logical interpretation. One such system is developed by Hornstein & Pietroski (2009), where Merge (renamed Combine) is first decomposed into the structure-building operations Concatenate and Label, and then all concatenations signify conjunction of monadic concepts, and all labelling signifies the building of thematic relationships.

These three approaches are all interpretative: Syntax first creates a structure on its own and, however soon afterwards, translation rules apply to that structure to create a logical form. However, abandoning any notion that the output of Merge

must interface with a sensorimotor system means that having an independent syntactic module with just a semantic interface is unmotivated. Such an interface is only substantive when there is a many-to-one relationship between syntactic structures and logical forms, which is what we find in architectures where syntax takes lexical items (not concepts) as input, with that relationship holding of cross-linguistic differences in utterances that express synonymies. However, if atomic concepts are the input to syntax and, by extension, the syntactic process is *not* affected by its input, then there must be a one-to-one correspondence between structure and meaning, making an ‘interface’ vacuous. In essence, separating syntax from semantics is only credible if syntax takes input with non-semantic features, or in some other way interfaces with non-semantic systems. Especially given the possibility for there to be a maximal rule-to-rule correspondence between operations in syntax and semantics, if syntax has no necessary connection to any other cognitive faculties, then syntactic rules simply *are* semantic rules, and the two are derivationally equivalent.

Having talked about ‘syntax’ now in many different ways, a concrete example may help illustrate these points. Consider the utterance in (8), the simplified syntactic structure in (9), and the logical form in (10) (for the purposes of this discussion, we do not need a precise theory of the functional sequence, so I do not specify all functional heads):

(8) Brutus stabs Caesar.

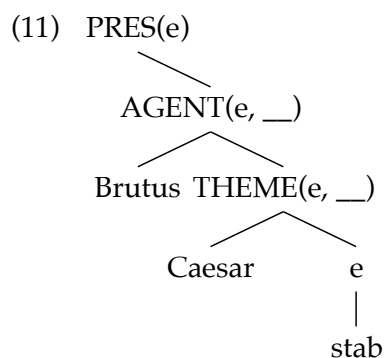


(10)  $\exists e[\text{stab}(e) \ \& \ \text{THEME}(e, \text{Caesar}) \ \& \ \text{AGENT}(e, \text{Brutus}) \ \& \ \text{PRES}(e)]$

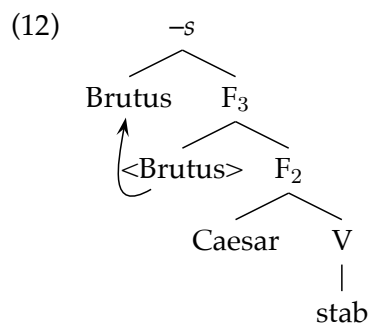
The claim of the Minimalist Minimalist Thesis is that Merge, taking concepts as input, is directly responsible for the derivation of (10). That is, the logical form is not created by an interpretation of a previously built syntactic structure; the logical form *is* a syntactic structure, from which the utterance ‘Brutus stabs Caesar’ ought to be derived. As such, the representation in (9) is equivalent to (10) to the extent that what it represents is also represented in (10), and whatever other features (9)

has that correspond only to the utterance (e.g., agreement and possibly movement) need to be explained as artefacts of externalisation—they are not genuine features of a syntactic derivation.

While the transparency between syntactic and semantic forms means that syntactic trees as normally drawn contain morpho-phonological information that is not truly part of syntax, we ought nevertheless to recognise richer structure in logical forms. To reiterate, I argue that logical forms themselves are produced by derivation, not by interpretation of prior derivations, so logical forms have derivational histories due to Merge, making them recursive hierarchical trees. Though that may not be clear from the appearance of (10), we can give it a better representation:



Although this looks simpler than the earlier syntactic tree, it contains the same structural, functional information and, in a more developed theory, should be sufficient to determine externalisation. Readers may notice that (11) strongly resembles Brody's (2000) Mirror theoretic representations of syntactic structures, which he argues for from entirely different motivations. In that system, (9) would look like this:



Clearly some questions immediately arise, presuming that a system like Hornstein & Pietroski (2009) can indeed establish a derivational equivalence between these semantic and syntactic forms: What exactly is the relation between the functional heads and the thematic roles (Borer 2005b, for example, replaces these thematic roles with ones that specify event structure)? How might movement be motivated at the sensorimotor interface, so that *Brutus* is linearised as the specifier of Tense? And is it possible to dispense with other canonical syntactic relations, like specifier-head-complement order, feature-valuing and variable argument structure?

It should be noted, of course, that syntactically structured logical forms undermine any potential objection that they are not suitable for externalisation, as such objections fail to treat logical forms as inherently structured. For example, one might say that because conjunction is commutative—AGENT & THEME = THEME & AGENT—pronunciation order is impossible to determine. However, as a semantic derivation by Merge has a temporal history and is constrained in its construction by whatever usually constrains Merge, like the functional hierarchy (Cinque 1999, Ramchand & Svenonius 2014), the proper shape for a derived logical form is the one in (11), which is suitable for linear ordering.

The most obvious difficulty for these ideas is that Merge taking atomic concepts, rather than lexical items, as input means that every identical meaning across languages ought to have an identical syntactic representation despite massive surface variation. One could argue that, once concepts have been connected with words, Merge could take feature-rich lexical items as input instead of concepts. This is essentially what Chomsky claims (with the presumption that the connection to words came for free), as he argues that Merge began with concepts, though the entire basis of parameterised language acquisition according to the Borer-Chomsky conjecture (Baker 2008) is that the input to syntax can be marked with syntax-affecting features (thus, they must be lexical items, rather than concepts, which do not bear such features). However, I believe this defeat can be avoided if we consider what it means to get rid of the lexical and morpho-phonological features that are normally associated with the lexical input to syntax.

In mainstream Minimalism, the secondary nature of externalisation is usually taken to mean that syntax is blind to morpho-phonological features, if they are present at all, though parameterised syntactic features are of course only possible if there are sensorimotor forms from which they can be inferred during acquisition. In other words, syntactic features are properties of lexical items that index concepts; they are not properties of concepts themselves. However, if concepts do not automatically entail lexical items, this system faces a fundamental problem, as the syntactic process at the time before externalisation either has unvalued features that cannot be valued, or if it can get by without feature-valuing, the origin of features and a sensitivity to them is unexplained. This much suggests that we should be aiming to get rid of non-semantic features if we believe that morpho-phonological features are irrelevant to syntax, but we need to approach the issue from another direction to really substantiate this.

Exo-skeletal (XS) syntax, following (Borer 2005a,b), has in the past ten years had great success with undermining the traditional view that syntactic structure projects from lexical specifications (with, for example, argument structure being licensed by features on verbs specifying how many and which theta-roles they must assign). Instead, XS syntax develops a framework where structure is templatic, universal, and (in a T-model) imposed upon lexical items, with meaning turning out to be variously acceptable, coerced or unacceptable (often with semantic unacceptability better explaining 'ungrammaticality'). Moreover, there are programmatic reasons for wanting to do away with lexical features even without pursuing a derivational semantics (e.g., Boeckx 2015), and there is interesting work being done to reformulate phenomena like case assignment in terms of structural configurations rather than feature-valuing (e.g., Preminger 2014).



For now, if we take it for granted that lexical features can be dispensed with, it becomes much more difficult to do away with morpho-phonological ones, and it is seriously tempting to conclude from studies like Borer (2013) that syntax is simply not possible without morpho-phonological input leading derivations, which seems counter to the basic intuition that syntactic structure is the structure of ideas. Perhaps we will one day have to concede that syntactically structured thoughts were only possible once spoken communication provided a means of organising conceptual structure, but this is an extraordinary claim requiring extraordinary evidence.

The main obstacle to reconciling XS syntax with a derivational semantics is the very clear mismatch between the units of morphology and units of meaning. In particular, basic (i.e. non-composed) meanings are sometimes matched with already syntactically structured words and word sequences, so the apparently necessary conclusion that has been reached by Borer is that morphemes are the true atomic input to syntax, with super-morphemic Spell-Out domains triggering conceptual access. This moreover means quite strikingly that morphemes (or 'roots') as listed in the lexicon are not associated with conceptual content, as they acquire it only when integrated within functional structure.

Of course, this must all be turned on its head to maintain the claims that other species have Merge, and that Merge structures concepts alone, and the only way to do this is to challenge the belief central to most syntactic theory that syntactic terminals correspond to actual or possible morphological units. With the morphology-concept mismatch, if concepts are to be the input to syntax, then syntactic terminals must be in many cases larger or smaller than morphemes, and morphemes must in many cases be externalisations of multiple terminals. This novel approach has had its first explorations in 'spanning' (Svenonius 2016), an outgrowth of Nanosyntax (Starke 2009) and Distributed Morphology (Halle & Marantz 1993), although these approaches still retain language-specific elements as the input to syntactic computation, rather than human-universal conceptual atoms.

To end this sketch, I would like to offer a metaphor for envisioning exactly the kind of system that is required, as technical implementations could have many guises, and the real sticking point here is what the *input* to syntactic computation must be. Especially as Chomsky already regularly claims that syntactic derivation is optimised for semantic interpretation, how is any of the preceding discussion genuinely different from his proposals? It is the input that matters, as syntactic analysis since the earliest days of generative grammar has always assumed, to a greater or lesser degree, an isomorphism between syntactic terminals and the morpho-phonological units of particular languages, whereas it is essential to this system that syntactic terminals are the atomic concepts that can enter into syntactic composition, which are *universal* for all speakers of all languages (and such concepts must exist if concepts and their composition are prior to linguistic ability). In other words, the 'lexical item', defined as a conjunction of phonological, syntactic, and semantic content, has no basis in the theory. Yet, if syntactic structures are so regularised, and if they do not structure morphemes, how is language variation possible? Note that frameworks like Distributed Morphology (Halle & Marantz 1993), which posit the phonological emptiness of syntactic roots, are not answers to this, as they still claim that roots are language-*particular*, and thus the invisibility of the phonology is sleight of hand.

Let's address this by asking quite simply: How do morphemes get their syntax if not from Merge, which is said to structure only concepts, prior to any morphological representation? Imagine first of all that you are a child again, and, presuming that your childhood was at all like mine, that you're on a school trip to the woods. You have in your hands a sheet of paper and a wax crayon, and you walk up to a tree, hold the paper against the trunk, and rub the crayon across the sheet. What you now hold in your hand is a two-dimensional picture of the three-dimensional trunk that was underneath. Of course, even though it is two-dimensional, there is sufficient contrast in light and depth of colour that you can look at it and accurately reconstruct an idea of how the tree must have been. It is a property of our visual systems, as with any painting or photograph, that we can look at a two-dimensional object and *see* it as three-dimensional. However, just because your tree rubbing has the appearance of three-dimensionality to you, you do not conclude that it must have been produced by a three-dimensional process. Whatever constructed the tree's trunk from carbon and nutrients in the soil did not construct your picture from the wax crayon. You placed against the tree an independent interface, and applied a different substance to represent the tree in lesser dimensions.

In the mental woodlands of our language faculty, we have syntactic trees, which are logical forms constituted of concepts, that are two-dimensional, and somehow they correspond to sequences of morphemes. Of course, when we produce an acoustic signal, the signal is entirely one-dimensional, though we are capable of identifying the two-dimensional mental object that it represents. Perceiving this correspondence, however, does not mean that the substance of the one-dimensional representation must be brought into sequence by the same process that is responsible for the two-dimensional structure it represents. Instead, it is perfectly possible that the two-dimensional object has a prior, independent existence, and the function of language—the utterance-idea mapping—is to hold up a one-dimensional interface against that structure and cover it with morpho-phonological substance in ways that are highly variable across languages. This would in part explain why it is that words do not straightforwardly associate with concepts, and it would ultimately support the view of Fodor (1998) that language is a non-compositional representation of a compositional Mentalese. It would moreover defuse the criticisms of Fodor's position in Pietroski 2006, as there would be no second morpho-syntax to account for, independent of the syntax of Mentalese—the structure of morpheme sequences comes not from a syntactic operation, but from an externalisation module. There are of course many intricacies in cross-linguistic variation that it would take detailed study to reconcile with these suggestions, but it seems more and more that independent developments in pragmatics, semantics, and syntax are converging to make this a possibility worth serious attention.

## References

- Ackermann, Hermann, Steffen R. Hage & Wolfram Ziegler. 2014. Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective. *Behavioral and Brain Sciences* 37(6). 529–604.
- Anderson, Stephen R. 2013. What is special about the human language faculty and how did it get that way? In Rudolf Botha & Martin Everaert (eds.), *The Evolutionary Emergence of Language: Evidence and Inference*, 18–41. Oxford: Oxford University Press.
- Armstrong, Sharon Lee, Lila R. Gleitman & Henry Gleitman. 1983. What some concepts might not be. *Cognition* 13(3). 263–308.
- Bach, Emmon. 1976. An extension of classical transformational grammar. In Jerry M. Sadock (ed.), *Problems in Linguistic Metatheory*, 183–224. East Lansing: Department of Linguistics, Michigan State University.
- Baker, Mark C. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The Limits of Syntactic Variation*, 351–373. Amsterdam: John Benjamins.
- Berwick, Robert C. & Noam Chomsky. 2016. *Why Only Us: Language and Evolution*. Cambridge, MA: MIT Press.
- Berwick, Robert C., Kazuo Okanoya, Gabriel J. L. Beckers & Johan J. Bolhuis. 2011. Songs to syntax: The linguistics of birdsong. *Trends in Cognitive Sciences* 15(3). 113–121.
- Bickerton, Derek. 2014. *More Than Nature Needs: Language, Mind, and Evolution*. Cambridge, MA: Harvard University Press.
- Boeckx, Cedric. 2013. Biolinguistics: Forays into human cognitive biology. *Journal of Anthropological Sciences* 91. 63–89.
- Boeckx, Cedric. 2015. *Elementary Syntactic Structures: Prospects of a Feature-Free Syntax*. Cambridge: Cambridge University Press.
- Boeckx, Cedric. 2017. A conjecture about the neural basis of recursion in light of descent with modification. *Journal of Neurolinguistics* 43. 193–198.
- Boeckx, Cedric & Víctor M. Longa. 2011. Lenneberg's views on language development and evolution and their relevance for modern biolinguistics. *Biolinguistics* 5(3). 254–273.
- Bolker, Jessica A. & Rudolf A. Raff. 1996. Developmental genetics and traditional homology. *BioEssays* 18(6). 489–494.
- Borer, Hagit. 2005a. *Structuring Sense, Vol. 1: In Name Only*. Oxford: Oxford University Press.
- Borer, Hagit. 2005b. *Structuring Sense, Vol. 2: The Normal Course of Events*. Oxford: Oxford University Press.
- Borer, Hagit. 2013. *Structuring Sense, Vol. 3: Taking Form*. Oxford: Oxford University Press.
- Bouchard, Denis. 2013. *The Nature and Origin of Language*. Oxford: Oxford University Press.
- Brentano, Franz Clemens. 1874. *Psychologie vom empirischen Standpunkt*, vol. 1. Leipzig: Duncker & Humblot.
- Brody, Michael. 2000. Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry* 31(1). 29–56.

- Burton-Roberts, Noel & Geoffrey Poole. 2006. 'Virtual conceptual necessity', feature-dissociation and the saussurian legacy in generative grammar. *Journal of Linguistics* 42(3). 575–628.
- Butler, Ann B. & William M. Saidel. 2000. Defining sameness: Historical, biological, and generative homology. *BioEssays* 22. 846–853.
- Byrne, Richard W., Erica Cartmill, Emilie Genty, Kirsty E. Graham, Catherine Hobaiter & Joanne Tanner. 2017. Great ape gestures: Intentional communication with a rich set of innate signals. *Animal Cognition* 20(4). 755–769.
- Callaerts, Patrick, Georg Halder & Walter J. Gehring. 1997. PAX-6 in development and evolution. *Annual Review of Neuroscience* 20(1). 483–532.
- Carruthers, Peter. 2002. The cognitive functions of language. *Behavioral and Brain Sciences* 25(6). 657–674.
- Carston, Robyn. 2002. *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.
- Carston, Robyn. 2013. Word meaning, what is said, and explicature. In Carlo Penco & Filippo Domaneschi (eds.), *What is Said and What is Not: The Semantics/Pragmatics Interface*, 175–203. Stanford, CA: CSLI Publications.
- Catchpole, C. K. & P. J. B. Slater. 2008. *Bird Song: Biological Themes and Variations*. Cambridge: Cambridge University Press 2nd edn.
- ten Cate, Carel & Kazuo Okanoya. 2012. Revisiting the syntactic abilities of non-human animals: Natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 367(1598). 1984–1994.
- Chen, Feng-Chi & Wen-Hsiung Li. 2001. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *American Journal of Human Genetics* 68(2). 444–456.
- Cheney, Dorothy L. & Robert M. Seyfarth. 2008. *Baboon Metaphysics: The Evolution of a Social Mind*. Chicago: University of Chicago Press.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.
- Chomsky, Noam. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, 104–131. Oxford: Oxford University Press.
- Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.
- Chomsky, Noam. 2010. Some simple evo devo theses: How true might they be for language? In Richard K. Larson, Viviane Déprez & Hiroko Yamakido (eds.), *The Evolution of Human Language: Bilingual Perspectives*, 45–62. Cambridge: Cambridge University Press.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Connolly, Andrew C., Jerry A. Fodor, Lila R. Gleitman & Henry Gleitman. 2007. Why stereotypes don't even make good defaults. *Cognition* 103(1). 1–22.

- Corballis, Michael C. 2017. The evolution of language: Sharing our mental lives. *Journal of Neurolinguistics* 43. 120–132.
- Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (ed.), *The Logic of Decision and Action*, 81–95. Pittsburgh: University of Pittsburgh Press.
- Deacon, Terrence W. 1997. *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: W. W. Norton & Company.
- Deacon, Terrence W. 2003. Multilevel selection in a complex adaptive system: The problem of language origins. In Bruce H. Weber & David J. Depew (eds.), *Evolution and Learning: The Baldwin Effect Reconsidered*, 81–106. Cambridge, MA: MIT Press.
- Dennett, Daniel C. 1971. Intentional systems. *The Journal of Philosophy* 68(4). 87–106.
- Dennett, Daniel C. 1983. Intentional systems in cognitive ethology: The “Panglossian paradigm” defended. *Behavioral and Brain Sciences* 6(3). 343–390.
- Dennett, Daniel C. 1987. *The Intentional Stance*. Cambridge, MA: MIT Press.
- Dickinson, W. J. 1995. Molecules and morphology: Where’s the homology? *Trends in Genetics* 11(4). 119–121.
- Dupré, John. 1996. The mental lives of nonhuman animals. In Marc Bekoff & Dale Jamieson (eds.), *Readings in Animal Cognition*, 323–336. Cambridge, MA: MIT Press.
- Fitch, W. Tecumseh. 2010. *The Evolution of Language*. Cambridge: Cambridge University Press.
- Fitch, W. Tecumseh. 2011. “Deep homology” in the biology and evolution of language. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The Bilingualistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, 135–166. Oxford: Oxford University Press.
- Fitch, W. Tecumseh. 2017. On externalization and cognitive continuity in language evolution. *Mind & Language* 32(5). 597–606.
- Fitch, W. Tecumseh, Bart de Boer, Neil Mathur & Asif A. Ghazanfar. 2016. Monkey vocal tracts are speech-ready. *Science Advances* 2. doi:10.1126/sciadv.1600723.
- Fodor, Jerry A. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.
- Fodor, Jerry A. 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.
- Fodor, Jerry A. 2008. *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.
- Gallistel, C. R. 2011. Prelinguistic thought. *Language Learning and Development* 7(4). 253–262.
- Gleitman, Lila. 1990. The structural sources of verb meanings. *Language Acquisition* 1(1). 3–55.
- Gould, Stephen J. & Richard C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society B: Biological Sciences* 205(1161). 581–598.
- Grice, H. P. 1957. Meaning. *The Philosophical Review* 66(3). 377–388.
- Grice, H. P. 1969. Utterer’s meaning and intention. *The Philosophical Review* 78(2). 147–177.

- Hall, Brian K. 1989. Genetic and epigenetic control of vertebrate embryonic development. *Netherlands Journal of Zoology* 40(1). 352–361.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & S. J. Keyser (eds.), *The View from Building 20*, 111–176. Cambridge, MA: MIT Press.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 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: Oxford University Press.
- Higham, James P. 2014. How does honest costly signaling work? *Behavioral Ecology* 25(1). 8–11.
- Hockett, Charles F. 1960. The origin of speech. *Scientific American* 203(3). 88–96.
- Holmberg, Anders. 2010. Parameters in minimalist theory: The case of Scandinavian. *Theoretical Linguistics* 36(1). 1–48.
- Hornstein, Norbert. 2009. *A Theory of Syntax*. Cambridge: Cambridge University Press.
- Hornstein, Norbert & Paul Pietroski. 2009. Basic operations: Minimal syntax-semantics. *Catalan Journal of Linguistics* 8. 113–139.
- Hurford, James R. 2007. *The Origins of Meaning*. Oxford: Oxford University Press.
- Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jackendoff, Ray & Eva Wittenberg. 2014. What you can say without syntax: A hierarchy of grammatical complexity. In Frederick J. Newmeyer & Laurel B. Preston (eds.), *Measuring Grammatical Complexity*, 65–82. Oxford: Oxford University Press.
- Julien, Marit. 2002. *Syntactic Heads and Word Formation*. Oxford: Oxford University Press.
- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Krakauer, John W., Asif A. Ghazanfar, Alex Gomez-Marin, Malcolm A. MacIver & David Poeppel. 2017. Neuroscience needs behavior: Correcting a reductionist bias. *Neuron* 93(3). 480–490.
- Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: John Wiley & Sons.
- Lenneberg, Eric H. 1969. On explaining language. *Science* 164(3880). 635–643.
- Lieberman, Philip. 2002. On the nature and evolution of the neural bases of human language. *Yearbook of Physical Anthropology* 45. 36–62.
- Lieberman, Philip. 2007. The evolution of human speech: Its anatomical and neural bases. *Current Anthropology* 48(1). 39–66.
- Lieberman, Philip. 2015. Language did not spring forth 100,000 years ago. *PLOS Biology* 13(2). doi:10.1371/journal.pbio.1002064.
- Lieberman, Philip, Dennis H. Klatt & William H. Wilson. 1969. Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science* 164(3884). 1185–1187.
- Lohndal, Terje. 2014. *Phrase Structure and Argument Structure: A Case Study of the Syntax-Semantics Interface*. Oxford: Oxford University Press.

- Marler, Peter, Christopher S. Evans & Marc D. Hauser. 1992. Animal signals: Motivational, referential, or both? In Hanuš Papoušek, Uwe Jürgens & Mechthild Papoušek (eds.), *Nonverbal Vocal Communication: Comparative and developmental approaches*, 66–86. Cambridge: Cambridge University Press.
- Marr, David. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W. H. Freeman and Company.
- Maynard Smith, John & David Harper. 2003. *Animal Signals*. Oxford: Oxford University Press.
- Maynard Smith, John & Eörs Szathmáry. 1997. *The Major Transitions in Evolution*. Oxford: Oxford University Press.
- Mayr, Ernst. 1982. *The Growth of Biological Thought*. Cambridge, MA: Harvard University Press.
- Müller, Gerd B. & Günter P. Wagner. 1996. Homology, *Hox* genes, and developmental integration. *American Zoologist* 36(1). 4–13.
- Nagel, Thomas. 1974. What is it like to be a bat? *The Philosophical Review* 83(4). 435–450.
- Newmeyer, Frederick J. 1980. *Linguistic Theory in America: The First Quarter-Century of Transformational Generative Grammar*. New York: Academic Press.
- Nowak, Martin A. & David C. Krakauer. 1999. The evolution of language. *Proceedings of the National Academy of Sciences of the United States of America* 96(14). 8028–8033.
- Patterson, Nick, Daniel J. Richter, Sante Gnerre, Eric S. Lander & David Reich. 2006. Genetic evidence for complex speciation of humans and chimpanzees. *Nature* 441(7097). 1103–1108.
- Pettito, Laura-Ann. 2005. How the brain begets language. In James McGilvray (ed.), *The Cambridge Companion to Chomsky*, 84–101. Cambridge: Cambridge University Press.
- Pfenning, Andreas R., Erina Hara, Osceola Whitney, Miriam V. Rivas, Rui Wang, Petra L. Roulhac, Jason T. Howard, M. Arthur Moseley, J. Will Thompson, Erik J. Soderblom, Atsushi Iriki, Masaki Kato, M. Thomas P. Gilbert, Guojie Zhang, Trygve Bakken, Angie Bongaarts, Amy Bernard, Ed Lein, Claudio V. Mello, Alexander J. Hartemink & Erich D. Jarvis. 2014. Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* 346(6215). doi:10.1126/science.1256846.
- Pietroski, Paul. 2005. *Events and Semantic Architecture*. Oxford: Oxford University Press.
- Pietroski, Paul. 2006. Character before content. In Judith Thomson & Alex Byrne (eds.), *Content and Modality: Themes from the Philosophy of Robert Stalnaker*, 34–60. Oxford: Oxford University Press.
- Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13(4). 707–784.
- Pinker, Steven & Ray Jackendoff. 2005. The faculty of language: What's special about it? *Cognition* 95(2). 201–236.
- Poepfel, David & David Embick. 2005. Defining the relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-first Century Psycholinguistics: Four Cornerstones*, 103–118. Mahwah, NJ: Erlbaum.

- Premack, David & Guy Woodruff. 1978. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 1(4). 515–526.
- Preminger, Omer. 2014. *Agreement and its Failures*. Cambridge, MA: MIT Press.
- Quine, W. V. O. 1960. *Word and Object*. Cambridge, MA: MIT Press.
- Ramchand, Gillian & Peter Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46. 152–174.
- Recanati, François. 2004. *Literal Meaning*. Cambridge: Cambridge University Press.
- Richmond, Brian G., David R. Begun & David S. Strait. 2001. Origin of human bipedalism: The knuckle-walking hypothesis revisited. *Yearbook of Physical Anthropology* 44. 70–105.
- Rogers, James & Geoffrey K. Pullum. 2011. Aural pattern recognition and the sub-regular hierarchy. *Journal of Logic, Language and Information* 20(3). 329–342.
- Ross, Don, Andrew Brook & David Thompson (eds.). 2000. *Dennett's Philosophy: A Comprehensive Assessment*. Cambridge, MA: MIT Press.
- Saidel, William M. & Ann B. Butler. 1997. An atypical diencephalic nucleus in actinopterygian fishes: Visual connections and sporadic phylogenetic distribution. *Neuroscience Letters* 229(1). 13–16.
- Sandler, Wendy, Mark Aronoff, Irit Meir & Carol Padden. 2011. The gradual emergence of phonological form in a new language. *Natural Language and Linguistic Theory* 29(2). 503–543.
- Scott-Phillips, Thom. 2014. *Speaking Our Minds: Why Human Communication is Different, and how Language Evolved to Make it Special*. New York: Palgrave Macmillan.
- Searle, John R. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Searle, John R. 2009. What is language? Some preliminary remarks. *Etica & Politica / Ethics & Politics* 11(1). 173–202.
- See, Adam. 2014. Reevaluating chimpanzee vocal signals: Toward a multimodal account of the origins of human communication. In Marco Pina & Nathalie Gontier (eds.), *The Evolution of Social Communication in Primates: A Multidisciplinary Approach*, 195–215. New York: Springer.
- Seyfarth, Robert M., Dorothy L. Cheney & Peter Marler. 1980. Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science* 210(4471). 801–803.
- Short, T. L. 2007. *Peirce's Theory of Signs*. Cambridge: Cambridge University Press.
- Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1). 1–6.
- Svenonius, Peter. 2016. Spans and words. In Heidi Harley & Daniel Siddiqi (eds.), *Morphological Metatheory*, 199–220. Amsterdam: John Benjamins.
- Számadó, Szabolcs & Eörs Szathmáry. 2006. Selective scenarios for the emergence of natural language. *Trends in Ecology & Evolution* 21(10). 555–561.
- Tallerman, Maggie. 2007. Did our ancestors speak a holistic protolanguage? *Lingua* 117(3). 579–604.



- Tallerman, Maggie. 2009. The origins of the lexicon: How a word-store evolved. In Rudolf Botha & Chris Knight (eds.), *The Prehistory of Language*, 181–200. Oxford: Oxford University Press.
- Tallerman, Maggie. 2014. No syntax saltation in language evolution. *Language Sciences* 46. 207–219.
- Tattersall, Ian. 1998. *The Origin of the Human Capacity*. New York: American Museum of Natural History.
- Tattersall, Ian. 2014. Communication and human uniqueness. In Marco Pina & Nathalie Gontier (eds.), *The Evolution of Social Communication in Primates: A Multidisciplinary Approach*, 219–227. New York: Springer.
- Tomasello, Michael. 2008. *Origins of Human Communication*. Cambridge, MA: MIT Press.
- Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne & Henrike Moll. 2005. Understanding and Sharing Intentions: The Origins of Cultural Cognition. *Behavioral and Brain Sciences* 28(5). 675–735.
- Wagner, Günter P. 1989. The biological homology concept. *Annual Review of Ecology and Systematics* 20(1). 51–69.
- Wagner, Günter P. 2014. *Homology, Genes, and Evolutionary Innovation*. Princeton: Princeton University Press.
- Wagner, Günter P. 2016. What is “homology thinking” and what is it for? *Journal of Experimental Zoology B: Molecular and Developmental Evolution* 326(1). 3–8.
- Wake, David B. 1991. Homoplasy: The result of natural selection, or evidence of design limitations? *American Naturalist* 138(3). 543–567.
- Wake, David B. 1999. Homoplasy, homology and the problem of ‘sameness’ in biology. In Gregory R. Bock & Gail Cardew (eds.), *Novartis Foundation Symposium 222: Homology*, 25–46. New York: John Wiley & Sons.
- Wallace, Alfred Russel. 1869. Sir Charles Lyell on geological climates and the origin of species. *Quarterly Review* (April). 359–392.
- Watumull, Jeffrey, Marc D. Hauser & Robert C. Berwick. 2014. Conceptual and methodological problems with comparative work on artificial language learning. *Biolinguistics* 8. 120–129.
- Wheeler, Brandon C. & Julia Fischer. 2012. Functionally referential signals: A promising paradigm whose time has passed. *Evolutionary Anthropology* 21(5). 195–205.

Callum Hackett  
Newcastle University  
School of English Literature,  
Language, and Linguistics  
Newcastle upon Tyne  
NE1 7RU  
United Kingdom  
[callum.hackett@newcastle.ac.uk](mailto:callum.hackett@newcastle.ac.uk)