都留文科大学大学院紀要　第26集（2022年３月）
*THE TSURU UNIVERSITY GRADUATE SCHOOL REVIEW,*
No.26（March, 2022）

# Language evolution, semantic domains, and the computation delusion

Hywel EVANS

## Abstract

While the Chomskyan revolution has delivered many benefits, the preoccupation with scientific-sounding explanations centered on linear computation may also have taken us down one or two wrong rabbit-holes. This is true not just for mainstream theoretical linguistic research but for all language-related endeavors. There are severe problems with the Chomskyan account, which is now the minority position. It is therefore time to consider the possibility that computational accounts involving abstract, discrete syntactic features will prove an inadequate basis for language-related research and practice. Our focus moves, once more, away from syntax to cultural and historical meaning. Just as language evolution relates to the evolution of bipedal apes over millions of years, language structure relates to domain-general cognitive processes that continue to develop in the cultural experience of the individual. Evidence is offered that the influence of certain kinds of local semantic domains has not been sufficiently acknowledged, perhaps because doing so undermines the institutional fantasy that language may be adequately characterized in terms of linear code. We stand at a crossroads, at a time of true paradigm shift. For language-related studies in general, we will see a movement away from the view that language is primarily a syntactically computational system, the accidental product of random mutation, primarily of importance in relation to abstract thought processes, with communication a mere by-product. We will increasingly come to credit the invention of language to our bipedal hominin ancestors, with the instinct to cooperate and communicate meaning as the central driving dynamic. Language is our most important tool, for thinking as well as communication. It has made it possible to evolve to the point where we are masters of the world, able to forge a bright future or lay waste to all around us. Understanding the evolution of language will help us understand who we are and where we are going.

Keywords: Evolution, Linear code; Semantic Frames; Semantic Domains

## Introduction

We have been living through a ‘Chomskyan revolution’ (Newmeyer, 1986; Searle, 2002) in

modern linguistics, one that has rumbled on over sixty years, often without question. This is firmly established, in the popular imagination at least to the extent that most people know the name of only one linguist. It is difficult to deny that this has been inspirational for generations of people working in language-related fields. As a result of this revolution, language has come to be seen as a mathematically based scientific endeavor, much closer to physics than it is to earlier studies of language. The acquisition of language has come be seen as computational, and expressible in terms of abstract syntactic features, including functional categories not assumed in any other theories of language.

This has had consequences. The preoccupation with scientific explanations in linguistics has carried over to second language acquisition (SLA) theory, although this has given way somewhat in recent years to a more culture-based view of language acquisition involving an array of cognitive domains (Ibbotson, 2020; Ibbotson & Tomasello, 2016). Proponents of sociocultural SLA (Zuengler & Miller, 2006) have accused mainstream SLA researchers of 'physics envy,' driving a desperate need to appear 'scientific' at all costs. Appearing to unintentionally confirm this, Long's (2007) polemic against theory proliferation famously rails against 'nebulous' sociocultural theory in favor of 'scientific' investigations that assume the existence of a language faculty, a discrete, syntactic endowment housing rules of universal grammar (UG), that can be characterized in precise mathematical terms. Long refuses to consider the possibility that such a faculty does not exist. For true believers, faith in the existence of UG persists even when considering the case of "learners in the very early stages who fail to produce overt verbal inflectional morphology associated with such functional categories as AgrP and TP" (Long, 2007, p. 23). In other words, the existence of the putative computational system which justifies formal computations is assumed without question, even when considering why there is no evidence for it!

In slight contrast, Slabakova (2016) does acknowledge that the necessity for UG is open to debate (p.7). However, Slabakova justifies completely ignoring non-UG *emergentist or usage-based* accounts of language acquisition (Tomasello, 2003) simply because UG-based linguistics currently constitutes the mainstream. In fact, a strong case could be made that it has quickly become the minority position (Evans, 2016; de Bot, 2015; Ibbotson & Tomasello, 2016), both in theoretical and applied linguistics. Tomasello's seminal work is not even included in Slabakova's references.

For UG-based SLA researchers, then, it seems that a formal, computational linguistic account must be preferred to any other as being more scientific, no matter what evidence emerges from other fields. In the extreme case there is no need to provide any evidence that such a system exists or to even consider the possibility that it *does not* exist. Related to this issue, the use of first language (L1) in SLA has long been condemned (Howatt, 1984, 2004) without any compelling rationale. The closest one gets to a rationale is the argument from 'naturalism' (Cook, 2010, p. 8) implying that the language faculty, hopefully specifiable in formal terms, is not susceptible to normal learning, and only responsive to target language input. However, undermining such faith, the use of L1 is gaining increasing support (Butzkamm

& Caldwell, 2009; Cook, 2001; Laviosa, 2014; Malmkjaer, 2004; Widdowson, 2003), with compelling evidence that *translanguaging* (Garcia & Wei, 2014; Lewis, Jones, & Baker, 2012) approaches have been beneficial to students. Indeed, instruction in the Welsh language, in which *translanguaging* has played a pivotal role, is now regarded as offering a blueprint for the recovery of endangered languages (Plackett, 2021). Therefore, while acknowledging the achievements of the Chomkyan revolution, we should also consider the possibility that there have been adverse effects. It is possible that we are sometimes blinded by science and seduced by the notion of syntactic computation.

It is now conventional to suggest that devotion to UG theory and 'naturalistic' methodology may be best understood as a cover for powerful corporate interests that favor global solutions and the mass deployment of native speaker teachers (Canagarajah, 1999; Pennycook, 1998; Phillipson, 1992; Phillipson, 2010). Disturbingly, in this context, Knight (2016) argues forcefully that research in theoretical linguistics has been driven to some extent by military goals. According to this rather incendiary view, modern research projects in linguistics have been at least partly motivated by the desire to develop 'command and control' computer systems related to machine translation, that were expected to improve the efficiency of military operations. Encouraged by successes in reaching digital solutions via distinctive features in phonology (Jakobson *et al,* 1951), such funding biases favor a computational theory of mind in linguistic research, adding to the appeal of notions such as 'deep structure' (Hockett, 1958) and UG (Chomsky, 1965). This accusation is, in part, supported by the work of Golumbia (2009), who situates such goals more broadly within the context of the west's rather unhealthy dreams of technological supremacy, wielded computationally via centralized, hierarchical control.

For a researcher in linguistics or SLA, this is disquieting, to say the least. If such a view is to be taken seriously, it appears that the goal of reducing language to a device specifiable in precise mathematical terms was not merely unrealistic but also perhaps corrupt from the outset. It is difficult to dismiss these claims entirely as empty conspiracy theories. On January 17th, 1961, President Dwight D. Eisenhower warned us against the malign influence of the military-industrial complex. It is difficult to argue that we have successfully dealt with that threat.

Writing in September 2021, we have recently witnessed the debacle of the western powers' ignominious retreat from Afghanistan, accompanied by desperate retaliatory drone attacks remotely controlled by computers. The war itself was an unmitigated disaster and utter failure at the cost of thousands of lives on all sides and huge amounts of taxpayers' money, with billions of dollars-worth of military hardware abandoned to the triumphant Taliban government. It is difficult to imagine a more complete debacle. However, there is still every reason to believe that the military-industrial complex, in the form of corporate defense contractors, has done very well from many years of prolonged conflict (Schwarz, 2021). Therefore, while one should be careful not to give way to conspiracy theories, one does feel the need to reconsider certain theoretical assumptions in linguistics and implications for SLA.

The great English poet John Keats coined the expression 'negative capability' to refer

to the perceived quality of being able to live creatively with mysteries and acknowledge that there are some things we do not yet understand and may never understand, rather than grasping after an unattainable philosophical certainty. The general line of analysis presented here is that certain aspects of meaning are handled more sensibly by, as a starting point, acknowledging their inherent mystery rather than by attempting to simply tidy them away by giving them formal expression. Feldman (2008, p. 7) suggests that human language and thought are "not best studied as formal mathematics and logic, but as adaptations that enable creatures such as us to thrive in a wide range of situations." Crucial to this view is the notion that language and thought are products of our bodies as well as our brains (Anderson, 2003; Bergen, 2012; Brooks, 1999; Dourish, 2001; Gallagher, 2005; Lakoff, 1987; Lakoff & Johnson, 1980, 1999; Lakoff & Núñez, 2000). As a starting point, then, one might question the wisdom of treating language as a 'ring-fenced mental processor' (Ibbotson, 2020) characterizable in its entirety in terms of formal operations on abstract symbols. A reconsideration of some of the more obvious products of the computational model of mind and its theoretical progeny is, therefore, in order.

As suggested, one of the concerns here will be with the influence of linguistics on SLA. Indeed, the seminal influence of the great linguist Noam Chomsky (1957, 1965) has been huge. However, there is no doubt that many ideas in mainstream linguistics have recently been questioned by those working in related fields. There is evidence, for example, that the grip of UG on applied linguists and second language learning researchers has been significantly weakened in recent years, with some researchers suggesting that the whole UG enterprise was a complete waste of time (de Bot, 2015, p. 60). There is a danger of throwing the baby out with the bathwater, rejecting valuable findings as well as those which are unhelpful. One of the goals here, then, is to offer a reappraisal of what truly did have value in the Chomskyan revolution, that is still likely to be of help going forward.

## The end of the Chomskyan era

Chomsky (1988, p.4) famously invokes Plato via Bertrand Russell (1945) to relate language acquisition to Socrates' conclusion that 'there is no teaching, but only recollection.' In *The Meno*, Socrates demonstrates that a slave boy somehow knows the principles of geometry, despite never having been tutored. The conclusion that is derived from this is that certain aspects of our knowledge and understanding of things in general are innate and genetically determined. Our biological endowment means that we are destined to grow arms and legs rather than the wings of a sparrow, and human brains rather than those of rabbits, for example. Being in possession of human brains, we are similarly expected to possess certain cognitive abilities, for example being able to figure out the principles of geometry and how language works.

So far, there is little that could be called controversial. Basically, we have evolved as human beings, with human bodies and minds. No one disagrees with any of this. Human beings do have a unique, species-specific ability to learn and process language. However, we must be

careful about what conclusions we can draw from this fundamentally mundane (Hurford, 1995) observation. Regarding the history of modern mainstream linguistics, the important point is that Chomsky took this as a starting position from which to argue that human beings are in fact born with specifically syntactic, modularized knowledge, that functions as a discrete computational system autonomous of meaning and social interaction.

Feldman (2008) suggests that it is widely believed that syntacticians welcomed the (arguably preposterous) notion of the autonomy of syntax because it acted as a cover for the autonomy of linguistics from the findings of other disciplines. Golumbia (2009) and Knight (2016) argue that the computational, modularized conception of UG would have been strongly influenced by the availability of research funds and other forms of support from powerful sections of the military and industrial establishment. From this perspective, corporate America's impatience with Behaviorist theory (Skinner, 1957) paved the way for the cognitive revolution that was to reinvent the human mind as a digital computer or information-processing device. The position taken in this paper is that it is, perhaps, time to reunite language and general cognition.

To be clear, no one would dispute the fact that possession of human brains and minds means that we are able to learn human languages. Gorillas are similarly uniquely well-placed to learn how to communicate with other gorillas. However, the issue is really about the kind of information that must come built into our brains at birth and the form in which it is packaged. Chomsky (1980) claimed that general cognitive abilities were insufficient to explain the speed with which children acquire languages and that certain language-specific information, indeed syntax-specific information, must be innately modularized. This became known as the 'poverty of the stimulus' (POS) argument, that came to be used as evidence for the existence of UG. As mentioned, this is characterized as an innate biological endowment, replete with strictly syntactic information and entirely autonomous of meaning or anything else. In other words, it is like a little digital syntax computer hard-wired in the brain at birth. When we use language, this syntax computer takes just the syntactic information and processes it for us, eventually sending this off for semantic interpretation. As we are hard-wired with this faculty at birth, the process of language learning is made a lot easier for us.

One can see, therefore, that this takes us quite a bit beyond Plato and his untutored slave. If we were to claim, for example, that evolution has yielded general human cognitive abilities that underlie both geometry and language, then no one would really be arguing about anything. However, to claim that there is an independent computational module dedicated to processing syntactic information is quite another thing. If one were to claim that there is a discrete computational module dedicated to geometry (or a modularized *part* of geometry), for example, one would certainly be dismissed as a lunatic.

One should note that the POS argument does not go unchallenged (Clark & Lappin, 2011; Pullum & Scholz, 2002). It should also be noted that there are other ways of explaining the linguistic skills of humans in general, as well as children's incredible ability to learn language so quickly with so little effort. For one thing, children apparently get as much as ten

thousand hours of speaking practice by the time they are six years old (Anderson, 1995). Also, we should not forget that human beings have a long history of evolution. Pinker and Bloom (1990) offer a Neo-Darwinist model for the evolution of language ability by natural selection. They suggest that Australopithecines may have been the first speakers, which would give language as much as 5 million years to evolve. They point out that Broca's area is visible in cranial endocasts of two-million-year-old hominid fossils, citing Falk (1983) and Tobias (1991). Everett (2012, 2017) offers convincing arguments to support the view that Homo Erectus had some form of language.

If language has been co-evolving with the human brain over millions of years, as suggested by Deacon (1997), then one would certainly expect there to be similarities between the way humans conceptualize the world and the way that we structure language. It would be difficult to imagine otherwise. If the elements of language, such as verbs and nouns, or their analogs, map on to our mental structures, then one would certainly expect that this would be of help when it comes to learning language. You could, if you wished, argue that this mapping is what we *actually mean* when we talk about UG or the language of thought (Fodor, 1975). From this, we would expect fundamental core similarities between mental representation and language structure. Again, for emphasis, it would be surprising if that were not the case given perhaps five million years of co-evolution of language and the brain. If we assume that we are born with this general mapping between language and mental representation, and that children have the considerable advantage of learning language as they are bodily engaged in learning culture, we have quite enough to explain how children learn languages so well. In sociocultural terms (Vygotsky, 1978, 1986), language and brain have co-evolved in the phylogenetic domain while language and culture repeatedly co-evolve in the ontogenetic domain. Both these co-evolutionary processes would be subject to the everyday constraints of physical reality, of course.

Regarding the question of why children are so good at language learning, Deacon (1997) suggests that human languages may themselves have evolved to be particularly easy for children to learn. Certainly, humans have evolved such that our children are brilliant natural students of language, uninhibited, and resourceful. Also, if we are to believe Lakoff (2009), unused neural connections that we are born with die off by the time we are five years old. This means that young children have twice as many neural connections available to dedicate to the problem of cultural learning, with neural activity crucially connected to the body. Given these facts, it is hardly surprising that language learning is easier for young children.

It should be noted that, while a gradualist perspective on the evolution of human language is broadly consistent with evidence from the historical record, the UG account is not. The UG 'saltation' account (Berwick & Chomsky, 2015) pictures the human language faculty, not to mention humanity itself, as a historical anomaly, strangely isolated from millions of years of evolution. In this view, the language faculty appears abruptly, purely by random mutation. It is enormously difficult to resolve this with the scientific consensus on language and evolution (Böhm, 2020; DeSilva, 2021; Pääbo, 2014; Sykes, 2020; Stringer, 2012). To take one example,

Wrangham (2009) offers compelling evidence that early humans (Homo Erectus and before) learned how to control fire and, soon after, to cook their food, roughly two million years ago. This is hypothesized as having had a huge evolutionary effect as it enabled hominins, over time, to get more nutrition from their food with less time spent chewing and, hence, less evolutionary pressure to maintain powerful ape-like jaws that would have been unhelpful when trying to communicate via speech.

As it was easier to get nutrition, this would have made evolutionary increases in brain size and decreases in gut, jaw, and tooth size possible. Our early tool-using ancestors successfully abandoned life in the trees, eventually using fire as protection from animals while they lived on the ground. As a result, these ancestral bipedal hominins were able to develop altruistic tendencies (De Silva, 2021) that would have been supported by cooperative communicative behaviors including the evolving ability to speak. They were able to engage in hunting activities that contributed to division of labor and greater social complexity, including the development of family life. Indeed, Wrangham (2019) sees language as crucial in the process of self-domestication in human beings, which assumes relatively sophisticated language use from about 765,000 years ago (p. 164). Such complex language skills would also be expected to have emerged gradually, together with generally more sophisticated cognitive abilities, and those individuals who possessed such skills, particularly in infancy, would have held an evolutionary advantage. In other words, language ability is part of a shift in general cognitive strategy among ancestral humans, driven by bipedalism, tool use, and cooking. Random mutation always plays a part in evolution, in this case the gradual emergence of improved cognition and greater control over speech organs adapted for speech. However, our first assumption should not be the sudden appearance of a computational syntax-engine that in no way fits with the historical record.

Shipton (2010) argues that evidence from Acheulean (1.76-0.13 Mya) tools indicates that unique human cognitive abilities evolved much more gradually, emerging in a propensity for imitation and shared intentionality (Hrdy, 2009; Tomasello, 1999; Tomasello & Rakoczy, 2003) with roots developing from at least 2 million years ago. Stout et al. (2008) suggest that increased sophistication in toolmaking ability from the Early Stone Age (2.6-0.25 Mya) would likely have proceeded in tandem with the development of language ability and would probably have involved overlapping mental development. Shipton et al. (2013) argue that generativity, the ability to create new forms out of previously existing elements, is evident in Acheulean technology, and may have emerged with the evolution of improvements in working memory.

In other words, there is an abundance of evidence to support the view that complex language skills evolved gradually (Shipton, 2019) as part of a natural evolution of general cognition in sociocultural activity. By contrast, there is no reason to imagine that, over millions of years, an autonomous computational syntax-engine would have gradually evolved in isolation from other cognitive faculties.

Indeed, UG theorists (Berwick & Chomsky, 2015) must assume that the core language-specific abilities emerged rather recently, in a single genetic mutation and in a single individual

(often referred to as Prometheus) somewhere in Africa sometime over 70,000 years ago. It is not explained why, given that human beings would have been indisputably anatomically modern at that time, Prometheus's speech organs were already fully evolved! In any case, this core endowment that led to the unique ability of humans to use language is referred to as Merge (Berwick & Chomsky, 2015), a fundamental explanatory mechanism in the Minimalist Program (Chomsky, 1995), with all other candidate syntactic rules of UG now abandoned. Merge simply means that two mental (purely syntactic, in this case) objects may be combined to form a new structure that may in turn be characterized as a set of mental objects. An important feature of Merge must be that it is recursive in the sense that it must at least be possible to apply the Merge operation to the output of Merge.

An obvious problem with this is that it leaves poor Prometheus isolated in a non-linguistic world with a highly abstract, modularized, specifically syntactic computational ability that suddenly appeared from nowhere for no reason. Clearly, this would confer no evolutionary advantage. Therefore, it must be conceded that Merge did confer a more general cognitive advantage from which Prometheus was able to benefit and which she was able to pass on to her children. This forces us to accept that language evolved for assisting thought rather than for communication. There are at least two huge problems doing so. First, it is not clear that an abstract syntactic endowment really would have assisted thought in the absence of language to structure it. Second, it is now not clear why Merge should be considered a specifically linguistic, let alone syntactic, endowment. Indeed, Merge seems to be so general as to be observable literally everywhere in the universe, even if we avoid mention of the cognitive domain (Fauconnier & Turner, 2002; Hofstadter & Sander, 2013; Holyoak & Thagard, 1995; Koestler, 1964; Lakoff, 1987; Lakoff & Johnson, 1980, 1999; Murphy, 2015; Turner, 1996, 2001, 2015).

Therefore, we can say that there are good reasons to doubt the existence of a discrete syntactic endowment that may be characterized in terms of a combinatorial operation. At the same time, we may be justified in harboring at least some suspicions regarding how belief in such an object became the utterly dominant mainstream view. Even so, there is a danger of 'Chomsky bashing' and concluding that, if you no longer believe in UG, everything will be fine. De Bot (2015, p.70) claims that there is a clear trend in Applied Linguistics (AL) "away from formal theoretical linguistics, in particular UG, to more socially oriented and usage-based approaches." Given the arguments presented above, this would seem to be a welcome development, in line with the more scientific consensus. De Bot (2015) suggests that increasing numbers of AL specialists see UG-based Generative Grammar approaches as having failed in SLA, having not provided any explanations for language learning-related phenomena. However, we should still be seeking out universals that may be helpful for consideration by language teachers. What follows will amount to an attempted synthesis of UG and competing theories of formal linguistics with a view to salvaging what is valuable in a field undergoing a paradigm shift.

## Explaining structure

Chomsky's early work (Chomsky, 1955; 1957) in the generative grammar (GG) tradition may be regarded as revolutionary due to two highly influential proposals. One, truly groundbreaking, was to take phrase-structure trees very seriously. This successfully captured the intuition that sentences are hierarchically organized into phrases and that phrases are linearly ordered. This proposal is almost universally followed in one way or another in modern linguistics. The other, much more controversial, was to soon kick off the so-called "Linguistics Wars" (Harris, 1993). Based on the (arguably very flimsy) evidence of Chomsky's most famous and putatively meaningless sentence, (*Colorless green ideas sleep furiously.*), this was the view that a formal rule-based account of syntax can proceed independently of semantics. There have been many changes in the historical development of GG, and many challenges to this notion, but the core assumption that we are dealing with purely syntactic features, autonomous of semantics, has remained.

To illustrate, phrase structure rules originally followed the pattern indicated in (1) below:

1. A → B C

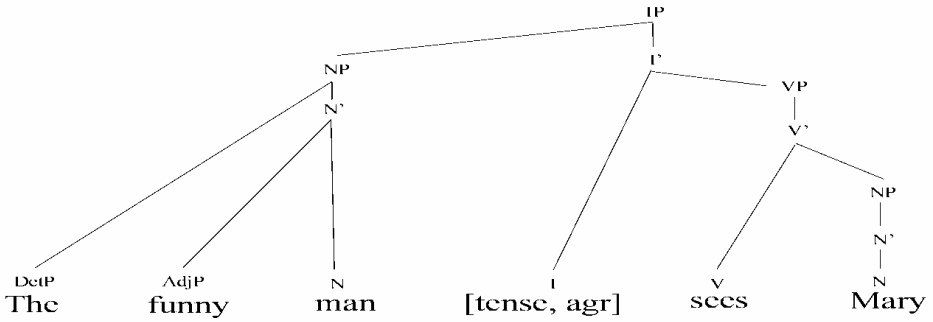
Constituent A may be understood as being separated into subconstituents B and C. Of course, this makes sense if you restrict your attention to English and consider A as a sentence, B as a noun phrase, and C as a verb phrase, for example. However, it is important to remember that, in the Chomskyan tradition, we are dealing with unstructured, abstract syntactic features, not the signs themselves. In other words, the language faculty and UG form a stand-alone system that has no place for sound or meaning experienced by human beings. To emphasize, the language faculty is understood as a *discrete syntactic* device, processing syntactic features. As such, the language that people use in everyday life could be construed as behaving like mathematically based languages used in the newly emerging field of computer science. Also, such constructs are characterizable in terms of linear formulae, as with computer code.

In the GG tradition, a system of rules generates grammatical combinations of constituents. Transformational Grammar (TG) forms part of this theory (Hjelmslev, 1961), and involves certain movement operations that can produce new grammatical sentences from existing ones. In X-Bar Theory (Chomsky, 1970; Jackendoff, 1977), it was hypothesized that linguistic categories follow certain category formation rules that allow intermediate constituents projected from a head. In Government and Binding Theory (GB), which Chomsky developed from the early 1980s (Chomsky, 1981; 1982; 1986a), there was a focus on the structural relations between certain syntactic categories.

By way of illustration, in (2) below, it was hypothesized in X-Bar Theory that all categories expand from heads through intermediate levels to maximal phrases. This delivered neat uniformity but generated a mass of over-complicated and redundant structure. It was abandoned in the MP in favor of a bare phrase structure formulation. For example, as can

be seen in (2), the intermediate structure for NP *Mary* does nothing at all. Also, it would be expected that AdjP and DetP have intermediate projections (omitted here), for no obvious reason. IP is the functional category Inflectional Phrase, a putative feature bundle including tense and agreement features associated with verbs. It was assumed that these were required to move to the appropriate projections. In the MP, IP is often explicitly bifurcated to Tense Phrase and Agreement Phrase. As movement operations (reformulated as variants of Merge in the MP) remain the fundamental explanatory mechanism in the GG tradition, functional heads are required to generate landing sites for moved items. This is illustrated below with a familiar tree structure, followed by a linear code version.

2.

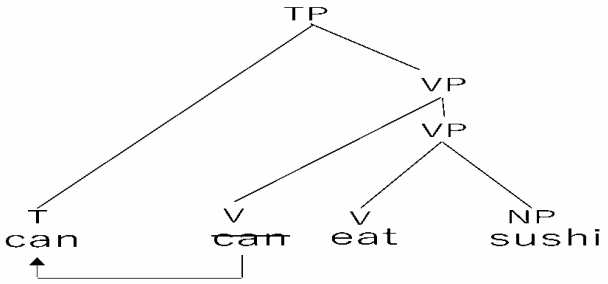



[IP [NP[DetP *the*][N'[AdjP *funny*][N *man*]]] [I'[I +*tense*+*agr*] [VP[V'[V *sees*] [NP *mary*]]]]

The MP (Chomsky, 1995), therefore, may be considered the latest in a long line of theories emerging within the GG tradition. The main innovation in the MP was that the language faculty was to be investigated under the assumption that UG constitutes a perfect design, hence is subject to certain optimality conditions. The MP has been widely criticized as both unscientific and undemocratic (Johnson & Lappin, 1997; Lappin, Levine, & Johnson, 2000; Lappin, Levine, & Johnson 2001; Seuren, 2004), essentially unleashed by decree without normal academic debate.

Be that as it may, the theory of the autonomy of syntax continues to drive a two-dimensional characterization of language structure. For example, making allowances for regular changes in terminology, in a sentence such as *John can eat sushi,* the auxiliary verb is assumed to start off (is base-generated or externally merged) inside some kind of verbal structure, then moved (internally merged) to a landing site in a higher functional structure (leaving behind a trace) that more naturally houses such moved constituents, the tentative suggestion here being a tense phrase, as indicated in (3) below.

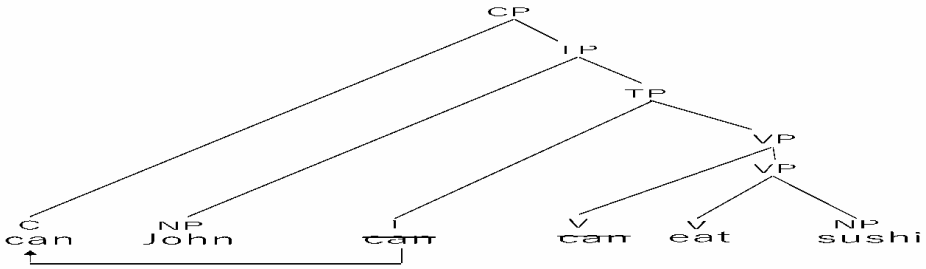
3.




[TP [T *can*$_1$] [VP[V $t_1$][VP[V *eat*] [NP *sushi*]]]]

Now, while Chomskyan linguistics has been overwhelmingly dominant, it should be noted that the assumption of the existence of functional, inflectional categories such as tense or agreement is decidedly anomalous. From a commonsense perspective, for example, one might question whether tense features are truly syntactic. Even so, the original conception of deep and surface levels (also known as D-structure and S-structure) of representation, here indicated via the relationship between the original position for *can* and the position to which it is moved, was hugely influential from the early 1960s. Later, two more levels of representation, logical form (LF) and phonetic form (PF), were introduced, then the theory was pared back to just PF and LF with the advent of the MP. If subjects adjoin to TP, and if we allow the projection of a further functional category (suggested here in (4) in classical terms as complementizer C), a cyclical movement operation allows a question sentence to be formed.

4.




[CP [C *can*$_1$] [TP [NP John] [TP [T $t_1$] [VP [V $t_1$][VP[V *eat*] [NP *sushi*]]]]]]

## Unification-based accounts and binding

While movement operations (internal Merge in the MP) have analogs in rival theories, these do not generally require functional categories such as tense or agreement. For example, in Head-Driven Phrase Structure Grammar (HPSG), where lexical items carry rich feature specifications, auxiliary verbs such as *can* are assumed to SUCATEGORIZE (SUBCAT) for a subject NP and an untensed base verb form which itself subcategorizes for the same subject NP. SUBCAT

(alternatively referred to as VALENCY) features for *can* are shown in (5) below.

5.

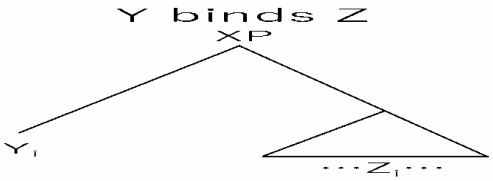
$$\begin{bmatrix} \text{HEAD} & \text{verb[fin, +AUX]} \\ \text{SUBCAT} & \langle \boxed{1}\text{NP}[nom], \text{VP[bse]: } [\text{SUBCAT}\langle \boxed{1} \rangle] \rangle \end{bmatrix}$$

In HPSG (Pollard & Sag, 1994) *can* is a finite auxiliary verb that takes a nominal subject, like all finite verbs, but also takes a non-finite verb phrase as its complement, the subject of which is structure-shared with the main subject. As auxiliaries such as *can* may bear the specification [+INV], it is licensed by a schema constraint to have both subject and object appear after the auxiliary verb in question forms. Thus, the illusion of movement (internal Merge) is handled via structure sharing and schema constraints, and there is no need to invoke movement operations or functional categories. In essence, the facts are handled by allowing rich, highly structured lexical specifications. Ironically, given that HPSG is widely used in computational linguistics, one could argue that it is a much more traditional explanation (tensed auxiliary/modal verbs like *can* take a subject and an untensed verb phrase with an unrealized subject) than the UG-variant.

However, the MP retained movement operations and functional categories and, at the same time, shifted the focus to a search for evidence that language is optimal in design. Thus, GB (Chomsky, 1986b) was essentially deprioritized. This could be considered rather odd, in that GB offered arguably the most valuable discoveries made in modern linguistics.

The key idea of Chomsky's binding theory was that referentially dependent anaphoric elements, such *as himself* or *each other,* must be bound within certain local domains of syntactic structure, illustrated in (6).

6.




7. Y binds Z just in case:
a. Y and Z are coindexed; and
b. Y c-commands Z

C-command is a configurational notion defined in (8).

8. Y c-commands Z just in case:
a. Z is contained in the least maximal projection containing Y; and
b. Z is not contained in Y.

In general terms, one could say that a node c-commands its sisters and their descendants.

9. a. John$_i$ likes himself$_i$.
b. The children$_i$ like [NP[each other's]$_i$ friends]].
c. *John$_i$ knows Bill$_j$ likes himself$_i$.

In 9(a-b), the subject binds (c-commands and is coindexed with) the anaphoric element. 9(c) is ruled out as *Bill* is a nearer potential binder. While there are several problems with binding theory, to be discussed later, it was undeniably a breakthrough in that it clearly indicated that human beings are subconsciously sensitive to linguistic structure. Phrase structure rules are real, not just a silly game played by linguists. However, the obvious problem from a GG perspective is that binding theory makes very clear reference to elements that are semantic (referential indexes), not syntactic. In other words, the jewel in the crown of GG is far from providing compelling evidence that syntax is autonomous of meaning. Also, binding theory provides no supporting evidence for either movement operations or the existence of functional categories.

Indeed, HPSG takes up the challenge of dealing with some of the problems inherited from binding theory. Pollard and Sag (1994) offer a non-configurational binding theory that relies on the notion of *obliqueness-command* or *o-command*. Obliqueness simply refers to linear order in a SUBCAT (VALENCY) list of some lexical head or other. In (10) below, $NP_x$ comes before $NP_y$ in the VALENCY list of *eat*, so is considered less oblique and, therefore, $NP_x$ locally o-commands $NP_y$. To illustrate this, I will make use of the equivalent, but intuitively somewhat simpler, Sign-Based Construction Grammar (SBCG) reformulation offered in Boas and Sag (2012). In (10), a verb such as *eat* (a lexical head) is expected to carry VALENCY and SEMANTIC specifications as follows:

10.

$$\begin{bmatrix} \text{FORM} \langle \text{eat} \rangle \\ \text{ARGUMENT-STRUCTURE} \langle \text{NP}_x, \text{NP}_y \rangle \\ \text{SYNTAX} \left[ \text{VALENCY} \langle \text{NP}_x, \text{NP}_y \rangle \right] \\ \text{SEMANTICS} \left[ \text{FRAMES} \langle \begin{bmatrix} \text{eat-frame} \\ \text{EATER } x \\ \text{FOOD } y \end{bmatrix} \rangle \right] \end{bmatrix}$$

FORM can be understood as fundamentally equivalent to the phonological features (not given a detailed treatment in HPSG or SBCG and essentially a placeholder in this case) of the verb *eat*, while ARGUMENT-STRUCTURE is essentially a convenient mirror of VALENCY features. While VALENCY (SUBCAT in HPSG) is a syntactic feature structure, the indexical information related to the subject and object is structure-shared in SEMANTICS of verbs and nominals, therefore at least partially semantic. Indeed, FRAMES (Fillmore, 1982; 1985) can be understood as the encyclopedic information related to the meaning of *eat*, essentially all the knowledge that relates to the word, structured via cultural experience. In other words, the concept of FRAMES is very far removed from both the notion of a discrete, autonomous syntactic endowment and

a linear coded characterization of its computational operations. To emphasize, SBCG makes no assumption of an innate grammar module and, as a construction grammar, explicitly should assume domain general cognitive processes in semantics. Participants referenced in the *eat* frame can be understood, rather straightforwardly, as the eater and the entity that is eaten. In this way, the subject and object of the verb referentially identify, via their INDEX information (indicated by shorthand subscript x and y), the participants in the eating event.

The noun *food* can be illustrated as follows:

11.

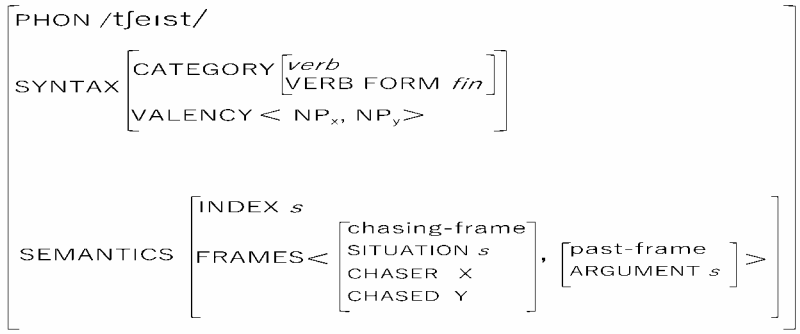

Here we can say that a noun such as *food* optionally takes a determiner phrase (as in *the food*), while the INDEX referentially identifies what is being eaten. Clearly, it is by no means too much of a stretch to think of INDEX features, while specifiable in SYNTAX, as at least partly semantic. Hence, the quandary facing the GG regarding binding theory, as semantics seems to play a big part in our subconscious awareness of syntactic structure.

In fact, Pollard and Sag (1994), with no need for functional categories, reformulate binding theory so that an anaphor must be bound by the index of any element that appears before it in a lexical head's VALENCY features (local o-command). Hence, a sentence such as *Jellyfish eat themselves* is grammatical because the index relating to *themselves* may be understood as appearing after the index relating to *jellyfish* in the VALENCY list for *eat*. A sentence such as *John's mother likes himself* is ruled out because the required antecedent *John* is not the indexed element in the VALENCY list, *John's mother* is.

Thus, binding theory is reformulated in terms of the index of arguments in a VALENCY list. English speakers look back to a less-oblique argument in a VALENCY list to find a semantic reference for anaphors. If there is none to be found, the sentence is ungrammatical. This reformulation allows for the grammaticality of *John$_i$ believes Mary bought pictures of himself$_i$ to paste up on the wall*. In this sentence, *pictures of himself* is an object, while *himself* is the least oblique (in fact, the only) member of the VALENCY list for the prepositional lexical head *of*, so it has no local o-commanding antecedent, and is free from strictly local binding constraints. The coindexing is therefore contextually, not syntactically, determined.

To illustrate further, the features for a verb such as *chased* are expected to include the following (12):

12.



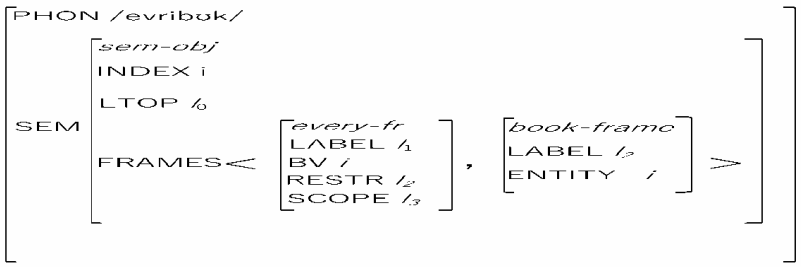

As we can see from (12), the verb subcategorizes for two noun phrases, indexed as *x* and *y*, while the semantic information indicates that it is understood as a kind of situation, that happened in the past. The *past-frame*, also contributed by the verb *chased*, takes this situation as its argument, which simply means that the situation will be understood as having taken place in the past. In the sentence *Fido chased himself*, the INDEX features *x* and *y* will be enriched in accordance with the INDEX features carried by the NPs to ensure that the features for both Fido and himself are identical. In other words, both x and y will end up carrying features for 3rd PERSON, SINGULAR, and MASCULINE. If the verb were *chases*, the INDEX information contributed by the verb would specify that the subject is 3rd PERSON, SINGULAR (Hence, the ungrammaticality of a sentence such as *You chases yourself*).

What is interesting about this, so far, is that the non-configurational account of a phenomenon which is of canonical importance in modern linguistic theory makes a crucial appeal to indexical linking between arguments in a VALENCY list. In other words, if an index has a less oblique antecedent in a VALENCY list, then there must be an indexical relationship between them. In other words, English speakers are sensitive to syntactic argument structure and semantic reference, while looking back to find an appropriate antecedent.

## Quantifiers and anaphors in semantic domains

Now, having established that there is cross-theoretical support for the view that antecedents play a role in semantic reference, we will turn to the issue of quantification to propose a radically simplified explanation for scope ambiguity. Boas and Sag (2012) claim that a quantificational NP such as every book must contain all the information in (13) below:

13.




Consider the ambiguous sentence in (14):

14. A student took every test.

As a cognitive linguistic theory of construction grammars, we do not expect movement operations (Aoun & Yen-hui, 1993; May, 1993) to play any part as a syntactic precursor to semantic interpretation in SBCG. Indeed, Boas and Sag abandon HPSG'S (Pollard & Sag, 1994) QSTORE analysis, somewhat analogous to movement, meaning that there is no need now for determiners to select the nominals they combine with. Instead, in the SBCG approach, new features are introduced in an ad hoc manner to handle scope ambiguity. LABEL and SCOPE features are employed to determine relative wide or narrow scope. However, this is still assumed to be feeding into logical formula in a manner analogous to syntactic movement that then undergoes semantic interpretation. Clearly, this is unnatural given that syntactic and semantic features operate together within the sign, and that there is no reason to think that a logical formula would play any part in semantic interpretation.

Instead, let us assume that scope ambiguities are related to the matter of antecedent reference already well established and unanimously accepted cross-theoretically. We assume that a noun phrase such as *some student* allows a mental simulation involving a single arbitrary referent taken contextually salient students. In other words, the *some-fr* allows a mental simulation for a single arbitrary referent of the appropriate kind. This single referent will be linked to the mental simulation for each referent within the domain of contextually salient tests. This will give us the wide scope interpretation for the existential NP in 14.

Similarly, consider the ambiguous sentence in (15):

15. Every student took a test.

The universal NP *every student* licenses a mental simulation for each of the referents taken from the set of salient students. In other words, the *every-fr* allows a mental simulation for each of the referents in the appropriate domain of contextually salient individuals. Each of these referents will be linked with the appropriate referent for a test, again an arbitrarily chosen individual test. Once more, this will yield the ordinary wide scope reading for the subject.

So how do we get the wide scope reading for the object in these examples? Well, the first assumption may be that something like local o-command plays a part in semantic interpretation, just as in the case of anaphoric reference. As there is a relationship of this sort between the subject and object in both (14) and (15), we might propose that another interpretation becomes available. In that case, an "antecedent interpretation" becomes available that gives rise to the wide scope reading for the object. This is essentially identical to the quantificational account provided by movement accounts that appeal to logical formulae. However, as we have seen, there is independent support for the view that antecedent relations play an important part in semantic reference within certain syntactico-semantic domains.

16. One student believes that John took every test.

In (16), there is no local o-command relation between the matrix subject and the embedded object, so we do not expect the antecedent interpretation to be available. Indeed, the wide scope reading for *every test* is much more difficult to access. Boas and Sag's account, relying on the arbitrary interplay between LABEL and SCOPE features, provides no straightforward explanation for such facts.

Similarly, one might consider examples such as (17):

17. Two auditors interviewed a representative from every company.

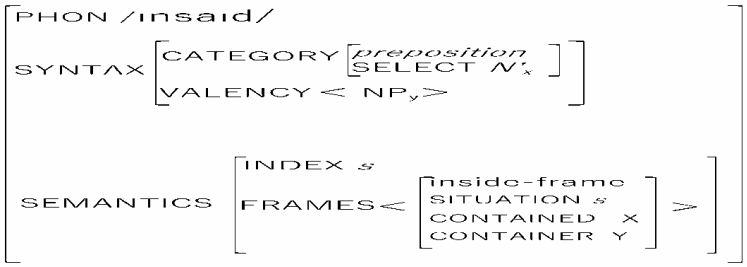
In (17) above, the NP-internal *every company* may easily be interpreted with wide or narrow scope in relation to the subject *two auditors*. However, it is very difficult to interpret *a representative* with wide scope in relation to *every company*. While it is certainly true that our knowledge of the world includes an awareness that one individual may not usually be in every company, it is not clear how Boas and Sag's account may handle such facts.

However, if we look at the data, we see that NPs modified by prepositional phrases form very strong antecedent-reference domains.

18. a. A $\text{bag}_i$ inside $\text{itself}_i$ would seem to be a physical impossibility.
b. A $\text{book}_i$ about $\text{itself}_i$ doesn't sound very interesting.

In (18), the syntactic interpretation for *itself* must be the bag, even though the interpretation is extremely unusual. Hence, it is not difficult to imagine that the NP forms a strong antecedent domain. This becomes even more compelling when one considers the possibility that it is the prepositional frame that provides the antecedent-reference domain. In other words, we expect our *inside-frame* to include, at the very least, a *container* entity and a *contained* entity.

19.




Here, we hypothesize that the meaning of *inside* may be understood in situational terms as a mental schema built up from our cultural experience of such container relations. One entity would be the reference of the NP or N' selected by the preposition, and the other would be the reference of the NP in VALENCY. Again, there seems no reason to doubt that such a mental schema would have to feature two such entities. This would certainly explain the availability of the relevant readings for the examples in 18.

Interestingly, it seems that the *pictures of* examples may be understood in the same terms.

20. a. [Paintings by him] are familiar to us today, but Hokusai's influence on European artists was revolutionary.
    b. *[Paintings by himself] are familiar to us today, but Hokusai's influence on European artists was revolutionary.
    c. [Photos of him] lay strewn across the floor, but Arnold remained impassive.
    d. [Photos of himself] lay strewn across the floor, but Arnold remained impassive.
    e. *[Clones of him] ran rampant through the citadel, but Arnold remained impassive.
    f. [Clones of himself] ran rampant through the citadel, but Arnold remained impassive.

It is very clear from the examples in 20 (c-f) that the subject NP must form a special kind of semantic domain, one involving *likeness* between entities, projected by a *likeness-frame*, for example. In (20 a-b), due to the nature of the *by-frame*, there is no assumption that the painter and painting will be alike in any way. Therefore, the anaphor is very strongly ruled out. This cannot be explained in configurational terms or with reference to VALENCY lists. The examples in (20 c-d) indicate a relatively weak likeness between entities, allowing both *him* and *himself*. (20 e-f) clearly show that the very strong likeness assumed between clones and originals, part of our real-world understanding, means that only the anaphoric form is felicitous.

## Conclusion

A semantic *likeness-frame* seems to play a part in the resolution of utterances involving anaphors. This forces us to consider the influence of specific kinds of semantic domain, as

mental schema constructed from complex real-world knowledge. The examples in 20 above clearly show that attempts to find a syntactic, linear computational solution will ultimately fail. The quantificational accounts have traditionally appealed to mathematical formulae to disambiguate sentences. We might well suspect that there has been a bias in linguistic research in favor of computational accounts, but the introduction of semantic frames means that such accounts will remain untenable.

The astonishing ability of human beings to manipulate and understand language is rooted in millions of years of use in social activity. Language was crucial in allowing bipedal apes to create more cooperative, altruistic communities, to survive and ultimately thrive. This has brought us to the point that we currently dominate the planet. We developed an instinct to cooperate and communicate, and language served a vitally important purpose within that dynamic. Language turned out to be useful for thinking as well as communication, and brain and language coevolved over a very long span of time.

The experience of children learning their mother tongue parallels the historical experience of our ancestors, from bipedal apes to modern *homo sapiens*. The difference is that, while our ancestors delivered the evolution of language and brain over millions of years of history, children learn their language and develop domain-general cognitive abilities within cultural activity in their early lifetimes. Part of the reason language is acquired "effortlessly" is that children learn their mother tongue and their mother culture at the same time. Language learning is locked in with cultural learning while children have more mental resources to dedicate to the task. As such, the lexicon connects to cultural meaning, with semantic frames contributing to an astonishing complex of real-world knowledge.

To understand language, we need to relinquish the fantasy that linguistic computations may be handled in terms of linear coding. It seems, in any case, to be an unworthy dream of technological supremacy and centralized, hierarchical control for the few. We need to reconsider language as a multi-dimensional system with a truly ancient historical context, inseparable from global thought processes and cultural experience. The dream that is truly worthy of human beings is one that acknowledges our debt to our ancestors and hopes for the fruition of our shared instinct to cooperate and communicate.

## References

Anderson, J. R. (1995). *Learning and memory: An integrated approach*. Wiley.

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, 149, 91-130.

Aoun, J., & Li, Y. A. (1993). *Syntax of Scope*. MIT Press.

Bergen, B.K. (2012). *Louder than words: The new science of how the mind makes meaning*. Basic Books.

Berwick, R. C., & Chomsky, N. (2015). *Why only us: Language and evolution*. Cambridge, Mass: MIT Press.

Boas, H. C., & Sag, I. A. (2012). *Sign-based Construction Grammar*. Stanford, California: CSLI Publications.

Böhm, M. (2020). *Ancient bones: Unearthing the astonishing new story of how we became human*. Greystone.

Brooks, R.A. (1999). *Cambrian intelligence: The early history of the new AI*. Bradford Books.

Butzkamm, W., & Caldwell, J. A. W. (2009). *The bilingual reform: A paradigm shift in foreign language teaching*. Germany: Narr Studienbucher.

Canagarajah, A. S. (1999). *Resisting linguistic imperialism in English teaching*. Oxford: Oxford University Press.

Chomsky, N. (1955). The logical structure of linguistic theory. Mimeographed, MIT Library, Cambridge. Retrieved from http://alpha-leonis.lids.mit.edu/wordpress/wp-content/uploads/2014/07/chomsky_LSLT55.pdf

Chomsky, N. (1957). *Syntactic structures*. The Hague/Paris: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Chomsky, N. (1970). Remarks on nominalization. In: R. Jacobs and P. Rosenbaum (eds.) *Reading in English Transformational Grammar*, 184-221. Waltham: Ginn.

Chomsky, N. (1980). On cognitive structures and their development: A reply to Piaget. In M. Piattelli-Palmarini (ed). *Language and learning: The debate between Jean Piaget and Noam Chomsky*. Harvard University Press.

Chomsky, N. (1981). *Lectures on government and binding: The Pisa Lectures*. Mouton de Gruyter.

Chomsky, N. (1982). Some concepts and consequences of the theory of government and binding. *Linguistic Inquiry Monograph 6*. MIT Press.

Chomsky, Noam (1986a). Barriers. *Linguistic Inquiry Monograph 13*. MIT Press.

Chomsky, N. (1986b). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Chomsky, N. (1988). *Language and problems of knowledge: The Managua lectures*. The MIT Press, Massachusetts.

Chomsky, N. (1995). *The minimalist program*. Cambridge, Mass: The MIT Press.

Clark, A., & Lappin, S. (2011). *Linguistic nativism and the poverty of the stimulus*. Wiley-Blackwell.

Cook, G. (2010). *Translation in language teaching: An argument for reassessment*. Oxford: Oxford University Press.

Cook, V. (2001). Using the first language in the classroom. *Canadian Modern Language Review*, 57(3): 399-423.

de Bot, K. (2015). *A history of applied linguistics: From 1980 to the present*. London and New York: Routledge.

Deacon, T.W. (1997). *The symbolic species: The co-evolution of language and the brain*. W.W. Norton & Company.

DeSilva, J. (2021). *First steps: How upright walking made us human*. New York: HarperCollins.

Dourish, P. (2001), *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge,

MA: MIT Press.

Evans, V. (2016). Review of Why only us: The language paradox. New Scientist. Retrieved from https://www.newscientist.com/article/2078294-why-only-us-the-language-paradox/

Everett, D. L. (2012). *Language: The cultural tool*. New York:Pantheon Books.

Everett, D. L. (2017). *How language began: The story of humanity's greatest invention*. Liveright (W.W. Norton).

Falk, D. (1983) Cerebral cortices of East African early hominids. *Science* 221: 1072-1074.

Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. Basic Books.

Feldman, J. A. (2008). *From molecule to metaphor: A neural theory of language*. Cambridge, MA: MIT Press.

Fillmore, C. J. (1982). Frame semantics. *Linguistics in the Morning Calm*, 111-137. Seoul: Hanshin Publishing Co.

Fillmore, C.J. (1985). Frames and the semantics of understanding. *Quaderni di Semantica* 6:222-354.

Fodor, J. A. (1975). *The Language Of Thought*. Crowell Press.

Golumbia, D. (2009). *The cultural logic of computation*. Harvard, MA: Harvard University Press.

Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Oxford University Press.

Garcia, O., & Wei, L. (2014). T*ranslanguaging: Language, bilingualism and education*. London, UK: Palgrave Macmillan.

Harris, R.A. (1993). *The Linguistics Wars*. Oxford: Oxford University Press.

Hjelmslev, L. (1961). *Prolegomena to a theory of language*. Madison: University of Wisconsin Press.

Hockett, C. (1958). *A course in modern linguistics*. Pearson College Division.

Hofstadter, D., & Sander, E. (2013). *Surfaces and essences: Analogy as the fuel and fire of thinking*. Basic Books.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Howatt, A. P. R. (1984). *A history of English language teaching* (1st edition) Oxford: Oxford University Press.

Howatt, A. P. R. with Widdowson, H.G. (2004). *A history of English language teaching* (2nd edition) Oxford: Oxford University Press.

Hurford, J.R. (1995). Nativist and functional explanations in language acquisition. In I.M. Roca (ed.). *Logical Issues in Language Acquisition*. Dordrecht, Rhode Island: Foris Publications.

Hrdy, S.B. (2009). *Mothers and others: The evolutionary origins of mutual understanding*. The Belknap Press of Harvard University Press.

Ibbotson, P. (2020). *What it takes to talk: Exploring developmental cognitive linguistics*. Mouton De Gruyter.

Ibbotson, P., & Tomasello, M. (2016). Language in a new key. *Scientific American*, Oct 18, 315(5): 70-75. Retrieved from https://www.researchgate.net/publication/311450032_Language_in_a_New_Key/link/5b18f59aaca272021ceeab3d/download

Jackendoff, R. (1977). X-bar-syntax: A study of phrase structure. *Linguistic Inquiry Monograph 2*. Cambridge, MA: MIT Press.

Jakobson, R., Gunnar, C., Fant, M., and Halle, M. (1951). *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.

Johnson, D. E. and Lappin, L. (1997), A Critique of the Minimalist Program. *Linguistics and Philosophy* 20, 273–333.

Koestler, A. (1964). *The act of creation*. Last Century Media.

Knight, C. (2016). *Decoding Chomsky: Science and revolutionary politics*. New Haven and London: Yale University Press.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. The University of Chicago Press.

Lakoff (2009). *The neural theory of metaphor*. Retrieved from http://www.neurohumanitiestudies.eu/archivio/SSRNid1437794The_Neural_Theory_of_Metaphor.pdf

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.

Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic Books.

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. Basic Books.

Lappin, Levine, S. & Johnson, D.E. (2000). The structure of unscientific revolutions. *Natural Language and Linguistic Theory*, 18, 665–771.

Lappin, Levine, S. & Johnson, D.E. (2001). The revolution maximally confused. *Natural Language and Linguistic Theory*, 19, 901-919.

Laviosa, S. (2014). T*ranslation and language education: Pedagogic approaches explored*. London: Routledge.

Lewis, G., Jones, B. & Baker, C. (2012). Translanguaging: Origins and development from school to street and beyond. *Educational Research and Evaluation*, 18 (7), 641–654.

Long, M. H. (2007). *Problems in SLA*. Lawrence Erlbaum Associates Publishers.

Malmkjaer, K. (2004). *Translation in undergraduate degree programs*. Philadelphia, PA: John Benjamins.

May, R. (1993). *Logical form: Its structure and derivation*. MIT Press.

Murphy, E. (2015). Label, cognomes, and cyclic computation: an ethological perspective. *Frontiers in Psychology, 6*, 1-18.

Newmeyer, F.J. (1986). Has there been a ‘Chomskyan revolution’ in linguistics? *Language* 62 (1): 1-18.

Pääbo, S. (2014). *Neanderthal man: In search of lost genomes*. Basic Books.

Papagianni, D., & Morse, M. (2013). T*he Neanderthals rediscovered: How modern science is rewriting their story*. Thames & Hudson.

Pennycook, A. (1998). *English and the discourses of colonialism*. London: Routledge.

Phillipson, R. (1992). *Linguistic imperialism*, Oxford University Press.

Phillipson, R. (2010). *Linguistic imperialism continued*. Routledge.

Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences* 13 (4): 707-784.

Plackett, B. (2021). Why are so many languages dying out? *Discover Magazine*. September 8th, 2021. Retrieved from https://www.discovermagazine.com/planet-earth/why-are-so-many-languages-dying-out

Pollard, C., & Sag, I. A. (1994). *Head-driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.

Pullum, G.K., & Sholz, B.C. (2002). *Empirical assessment of stimulus poverty arguments*. The Linguistic Review. 18(1-1): 9-50.

Russell, B. (1945). *The history of Western philosophy*. New York, NY: Simon and Schuster.

Schwarz, J. (2021). $10,000 invested in defense stocks when the Afghanistan war began now worth almost $100,000: Was the Afghanistan War a failure? Not for the top five defense contractors and their shareholders. *The Intercept*. August 17th, 2021. Retrieved from https://theintercept.com/2021/08/16/afghanistan-war-defense-stocks/

Searle, J.R. (2002). End of the revolution. New York Review of Books. February 28th, 2002. Retrieved from https://www.nybooks.com/articles/2002/02/28/end-of-the-revolution/

Seuren, P. A. M. (2004). *Chomsky's Minimalism*. Oxford University Press USA.

Shipton, C. (2010). Imitation and shared intentionality in the Acheulean. *Cambridge Archaelogical Journal*, 20(2), 197-210.

Shipton, C. (2019). Three stages in the evolution of human cognition: normativity, recursion, and abstraction. *Handbook of Cognitive Archeology*, 153-173.

Shipton, C., Clarkson, C., Pal, J. N., Jones, S. C., Roberts, R. G., Harris, C., Gupta, M. C., Ditchfield, P. W., & Petraglia, M. D. (2013). Generativity, hierarchical action and recursion in the technology of the Acheulean to Middle Palaeolithic transition: A perspective from Patpara, the Son Valley, India. *Journal of Human Evolution, 65*(2), 93-108.

Skinner, B.F. (1957). *Verbal behavior*. New York: Appleton Century Crofts.

Slabakova, R. (2016). Second language acquisition. Oxford University Press.

Stout, D., Toth, N., Schick, K., & Chaminade, T. (2008). Neural correlates of Early Stone Age toolmaking: technology, language and cognition in human evolution. P*hilosophical Transactions of the Royal Society, 363*, 1939 -1949.

Stringer, C. (2012). *Lone survivors*. Times Books.

Sykes, R.W. (2020). *Kindred: Neanderthal life, love, death and art*. Bloomsbury Sigma.

Tobias, P. V. (1981) The emergence of man in Africa and beyond. *Philosophical Transactions of*

*the Royal Society of London* B 292: 43-56.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared collective intentionality. *Mind and Language, 18*, 121- 47.

Turner, M. (1996). *The literary mind: The origins of thought and language*. Oxford: Oxford University Press.

Turner, M. (2001). *Cognitive dimensions of social science*. Oxford: Oxford University Press.

Turner, M. (2015). *The origin of ideas: Blending, creativity, and the human spark*. Oxford: Oxford University Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L.S. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Widdowson, H.G. (2003). *Defining issues in English language teaching*. Oxford: Oxford University Press.

Wrangham, R. (2009). *Catching fire: how cooking made us human*. Basic Books.

Wrangham, R. (2019). *The goodness paradox: The strange relationship between virtue and violence in human evolution*. Pantheon.

Zuengler, J., & Miller, E. R. (2006). Cognitive and Sociocultural Perspectives: Two Parallel SLA Worlds? *TESOL Quarterly*, 40(1), 35-58.

Received: November 24, 2021
Accepted: November 25, 2021