Words as Things

Gillian Ramchand, UiT the Arctic University of Norway

Draft: January 2022

1 Introduction

Natural language as spoken by humans embodies a number of remarkable systemic properties which are arguably unique, co-evolved with human beings´ unique ability to organize and control their natural environment, and to store and pass on culturally accumulated knowledge over generations. One of natural language´s remarkable properties is the creative generation of multiplex meaning-bearing symbolic structures that are hierarchically organized. This hierarchy is evident at the level of the sentence, but also apparently, within units that we are used to calling 'words´ in common parlance.

The hierarchical organization of complex symbols is fascinating, all the more so because it seems so far to be unique to the human species. It is not surprising therefore that linguists in the modern era (at least since Chomsky 1957) have devoted time and energy to understanding the building blocks, joints, and bolts of the complex symbolic entity that is a human language sentence. What is the nature of these structured representations? What are the logical primes of the system and their modes of combination? Morphosyntacticians in particular do not confine themselves to superficial or apparent boundaries, but seek to understand hierarchical structuring whereever they find abstract evidence for it, both inside and outside the word. It has been controversial whether words contain the same kind of hierarchical structure as sentences do (Baker 1985, Ackema and Neeleman 2007, Julien 2004), what their status is as a unit, and whether the answers to these broad questions vary across languages (Baker 2001). Hagit Borer's work over the past few decades has been at the forefront of research seeking to understand the hierarchical nature of human language, within and across words, and her work always combines detailed empirical argumentation with high level theorizing

about the nature of the language faculty based on different morphosyntactic observable facts.

In this short paper, I offer a discussion of the common sense notion of 'word' that has been at the heart of much theoretical discussion in morphosyntax, taking as my starting point the Borer 2013 discussion and relating it to the DM (Distributed Morphology, Halle and Marantz 1993, Embick and Noyer 2001, Harley 2014) position. Adding to the debate, I will attempt to reconcile the theorizing of the formal linguists with the theorizing of the psycholinguists when it comes to the evidence for the basic primes of hierarchical composition. I would like to emphasize that the evidence that each group of linguists uses is in practice distinct: formal linguists make generalizations based on the internal logical patterning of forms, and what it would take to construct an explicit implementation that gets the right distributional results; psycholinguists theorize based on human behaviour in controlled tasks related to comprehension and production, positing primitives of linguistic competence that would make sense of these processing behaviours. Thus, the formal morphologist takes the weight of evidence from uncontroversial distributional and paradigmatic facts, but also secondarily from speaker judgements about grammaticality. The psycholinguist on the other hand grounds their data in processing behaviours that are controlled and manipulated under laboratory settings, from which they try to abductively infer the system of competence that might give rise to them (although a certain amount of corpus evidence from things like speech errors, or acoustic measures of attested speech can also be used). It seems to me that all of these kinds of evidence are important when building theories of morphosyntactic competence that are nevertheless psychologically 'real', a desideratum that has been on the agenda since Chomsky put it there in 1957. Synthesizing this evidence however is rarely attempted (but see e.g. Gwilliams 2020 as a prominent exception to this).

Before I embark on this synthesis of evidence, it is perhaps necessary to spend some time spelling out why such a synthesis is neither irrelevant nor premature. One response to lack of integration between the two fields, is to point to the fact that formal linguists' theories are computational theories of competence, whereas the psycholinguist studies data related to performance. Under this way of thinking, processing is rightly contrasted with the body of knowledge that underpins it, but the objection appears to rest on the idea that psycholinguists are studying performance while the formal linguist is the only one genuinely studying competence. However, this dichotomy is

a false one. To the extent that formal linguists also consult judgements of well-formedness, they are also thereby relying, on data from 'performance' to gain access to information about competence. Arguably this is *always* the case. The Chomsky (1965) emphasis on competence was an argument about what the *object of inquiry* should be, not to deny performance(s) as evidence for it (an underappreciated point). Performance on various tasks is currently our only evidence for competence (in the absence of a credible Vulcan Mind Meld technique), but we should be clear that the purpose of our investigation is not just to describe those behaviours themselves, but to reason from them to the state of the competence that gives rise to them. If grammaticality judgements have proved to be the methodology of choice for many formal syntacticians (and morphologists) in the modern era, it is only because they are argued to be a fairly *direct* reflection of competence, and also because the method is flexible and controllable in terms of the complexity of the phenomena being investigated. This gives the methodology a great advantage over corpus methods which rarely provide the linguist with the precise conditions for falsifying particular formal models or hypotheses.

But particular hypotheses and computational implementations have to come from somewhere, and it seems sensible to use independently known information about the mind/brain to put boundary conditions on those hypotheses, especially if our ultimate goal is more and more algorithmically realistic theories (Marr 1982) that begin to make inroads on desiderata of explanatory adequacy.

The goal of coming up with a psychologically real theory of morphosyntax that is consistent with what we actually know about the cognition of symbol memory and symbol processing is also not premature because there is a lot we do now know about both those things. Arguably, the minimalist agenda of Chomsky (1993) which seeks to situate the language system in the context of the interface conditions with the other facets of the human mind/brain requires us to take account of the discoveries and puzzles that are being uncovered in our sister disciplines within cognitive neuroscience. But in addition, the data that psycholinguists gather and reason about is also data that, by abduction, also tells us something about competence, and so we ignore it at our peril.

My view will be, in laying out the major architectural controversies, that there are a number of robust results from psycholinguistics that I believe are relevant to the discussion, and have potential bearing on their resolution.

2 Do Words Exist?

The notion of 'word' seems intuitive from a folk understanding point of view, but any attempt to define it formally from the point of view of the primitives of a particular module of grammar gives results that (i) do not exactly conform to our folk intuition and (ii) are not consistent with each other. Modern formal linguistics partitions the architecture of knowledge into modules corresponding to phonology/phonetics, semantics/pragmatics, morphology and syntax. The primitive building blocks of phonology are probably either abstract phonological features or segments, which need to be learned for each individual language as basic inventory that underlies all the morphemes in the lexicon of that language. But morphemes are not just strung together like beads on a necklace. In between the affix and the sentence, there are well-defined intermediate units defined by their internal phonological integrity (Kiparsky 1982).

The phonological word is a formally coherent unit defined by certain sound interactions/rules that exist within that domain but not outside of it. It includes 'words' such as *I'll* and *I'm* in addition to *cat* and *cats* (Dixon 1977, Matthews 1991, Bauer et al. 2013).

On the other hand, if we are looking for the building blocks of semantic composition, for the units that need to be memorized because they are sequences of sound that are associated with idiosyncratic (non-composed) meaning, then we agree with the phonologists on the primitive status of units like *cat*, but we are also stuck with units like *kick the bucket*, but not *cats* or *running* which are semantically regularly composed (Matthews 1991).

Finally, neither of these primitive notions corresponds to what we get if we pick out the elements classical Chomskian linguists would like to insert under the terminal node of a syntactic phrase structure tree, which once again includes *cat*, but also [+PAST], and maybe even *-ing* .

Linguists reaction to this state of affairs has been to say that the notion of 'word' is a folk notion, which has a perfectly reasonable and understandable use in the context of normal language, but which is not fit for purpose in linguistic theorizing. Instead it has to be replaced by other more precise terminological notions useable in the different modules.

Borer in *Taking Form* (Borer 2013) mounts a compelling argument against the reality of words as a linguistic concept (of course putting aside folk uses of the word 'word', and concentrating on its utility for the scientist of language).

> "Words", here, are perceived as units which are morphologically constructed, but are nonetheless simultaneously phonological, semantic, and syntactic objects.
> Borer (2013) pg 11

The argument against Word as a unit here is that it is not clear what it is a unit *of* , since it seems to be that the minimal units of phonology, semantics and syntax do not actually coincide, and none of the three definitions separately works to pick out a coherent set of units that conforms to the folk intuition.

Borer (2013) acknowledges the legitimacy of a distinguished phonological unit, what she calls the Phonological Rule Application Domain, or P-RaD, which is defined by a certain class of phonological properties and well formedness rules that apply to it in a particular language. Her point is that while it is perfectly possible to make a distinction between those structure building operations that apply within that unit vs. those that apply external to those units (as in Ackema and Neeleman 2007), such a distinction could not be made in a non-circular fashion outside of the P-RaD diacritic itself.

> What is, however, rather striking is that none of these accounts offers a definition for what a syntactic "word" or even a morphological "word" is, such that it is independent of P-RaD; i.e. independent of whatever domain is defined by the assignments of primary stress.

There is no doubt then in some sense that this phonologically defined domain is psychologically 'real', and a true artefact of grammar, *especially* since it cannot be derived from any other primitive properties of concepts or even consistently across languages. If we choose to use the word 'word' for this coherent unit, then we must admit that within a language 'words' exist, and knowing about them is an important part of knowledge of a language. The question that occupies Borer, then, is the secondary question of whether the construction of this unit also corresponds to a distinct set of combinatoric principles and rules (morphology?) from the ones that operate over those units (syntax?). Borer's position in her work has been that there *is* no such justification for assuming two distinct combinatoric systems, as in Distributed Morphology (DM) she subscribes to the view that it is 'syntax all the way down' (Marantz 1997b).

It is worth pointing out that while the primitives of hierarchical composition within and outside phonological word units have a great degree of overlap within and across languages (see e.g. Baker 1985, Julien 2004), there is at least one respect in which the formal system specific to a particular language is different inside and outside the word, and that is in the domain of linearization. Specifically, functional material that is hierarchically higher than the root is often linearized to the right of it (as a suffix), while the opposite linearization is observed outside of the word with higher functional material to the left. This tendency in many head initial languages, led to head-movement and left-adjunction analyses of complex word formation, and to the Mirror Theory of word internal agglutinative morphology in Brody (2000). It is quite common for languages to have distinct linearization principles for word internal bound morphemes than for the ordering of syntactically movable units. In fact, if we see linearization as a core property of phonologization, we can add the following phonological criterion another potential defining property of the word.

(1) **The Phonological Word as a Linearization Domain**
The phonological word is the unit of the external linearization algorithm; it is a product of the internal linearization algorithm. Internal and external linearization are different mapping systems.

Note that the above statement does not attempt to relate the phonological word to atoms of the syntactic or semantic representation, but claims that the phenomenon of "word" in this sense emerges whenever a language has two distinct linearization cycles.

Even if we give up on the Word and replace it with three separate notions (a) phonological word, (b) memorized meaning-bearing chunk, and (c) syntactic atom respectively, we still need to integrate those different notions into an understanding of grammar and how the system works architecturally. In other words, how do these different more specific 'units' relate to each other derivationally and ontologically.

For the purposes of the argument, let us agree to give the phonological, linearization domain the privileged label of "word". How are words in this sense related to the atoms of declarative memory that all language users must operate with? And how do those memorized units then relate to the primitives of the syntactic representations that give the best characterization of the formal system? The answer to the latter question is going to

depend on one's syntactic and morphological assumptions. The answer to the former question is going to require some input from psycholinguistic and neurolinguistic study.

3 Separation

The first main plank in the architecture that has gained a lot of prominence in recent decades is the idea of 'late insertion', which is really the idea of 'separation', once we divorce the idea from the standard derivational metaphor used to implement it computationally. 'Separation' says that the atoms and features of the syntactic representation are independent of, and are autonomous with respect to, the phonological material that spells it out. 'Separation' is also strictly enforced between rich conceptual content and the syntax, in the form of the segregation of roots. In this popular recent view, embraced by both DM (Distributed Morphology) and in Hagit Borer's own work, the idea is that there is a special class of memorized units called ROOTS (like $\sqrt{}$*cat*) which have no syntactic information associated with them, but which are associated with rich semantic content. They are then interpreted and pronounced in the context of morphosyntactic (allosemy) and phonological context (allomorphy) respectively.

Within this family of morphosyntactic approaches, Borer has always steered a slightly different path with respect to DM in her understanding of the status of roots. While early DM approaches gave roots a conceptual content and late inserted only functional information (Marantz 1997a, Marantz 1997b), later work reduces roots to just indices which then point to different allosemic and allomorphic interpretations in context (Myler 2016). Borer (2013) on the other hand argues that roots are basically phonological, and that this is what anchors their representation, while essentially embracing allosemy. Borer's argument against a core conceptual core for roots comes from the radically different meanings roots have in contexts, even while phonological abstract identity seems clear (Borer 2013 pg 24; cf. also Arad 2003, Arad 2005).

Omer Preminger in a series of recent presentations and blogposts argues explicitly against any phonological, syntactic or semantic identity for symbols (see also Rasin et al. 2021 for a reevaluation of Arad 2003). For him, the thing that is ontologically prime is the syntactic atom or feature. The form/materiality of the language, and the meanings that are invoked are each

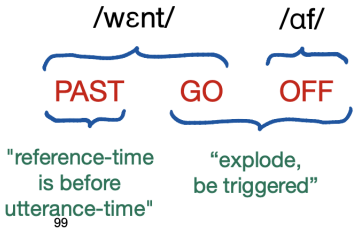
mapped separately and in non organically integrated ways from the syntactic spine. The following extract is from the slide show Preminger presented at the 14th annual conference on Syntax, Phonology and Language Analysis (SinFonIJA 14, Novi Sad).

> Can adequately abstract notions of "FORM" and of "MEANING" salvage a semiotic view of linguistic atoms as < FORM, (SYNTAX), MEANING > mappings?
> CLAIM:The answer is no. ...
> In other words, the proposed architecture of listed (a.k.a."lexical") knowledge:
> (A) fully abstract syntactic atoms (e.g. DOG, PAST, RUN, IN, etc.)
> (B) many-to-one rules from sets of nodes in (A) to units of FORM
> (C) many-to-one rules from sets of nodes in (A) to units of MEANING
> (https://omer.lingsite.org/files/Preminger-2021-Natural-language-without-semiosis-slides.pdf)

Thus, the ontologically privileged syntactic terminals are associated with a context sensitive spellout mechanism as well as a context sensitive intepretation mechanism. In this system, there is also no direct link between a FORM or set of FORMS on the one hand and a set of MEANINGS on the other—any linkage is mediated by overlaps in the syntactic input for the two kinds of rule. Since Preminger also embraces non terminal spell out, his system is ' informally characterized as "spanning with dissociated PF and LFspans."'.

To illustrate the position, it is instructive to look at Preminger's poster-child example, the past tense of the particle verb *go off*, meaning to explode. Given that *go* has an unpredictable past tense version *went*, the mapping between syntactic terminals and forms and meanings in the case of the phrase *went off*, would look as in (2).

(2)



Preminger claims that his account is not semiotic, since it does not embody a pairing of sound and meaning, but rather (a) < FORM, STRUCTURE > pairs and, separately, (b) < STRUCTURE, MEANING > pairs. In the above example, the (a) pairing contains a different structure from the (b) pairing. But crucially, the two structures have the syntactic atom GO in common.

I have two problems with the claims being made here. Firstly, if the syntactic atoms appear in the domain of many different mappings, either alone, or together with other syntactic atoms, then there is a sense in which each syntactic atom is part of a whole network of associations. Some of these associations are to "form" and some are to "meaning", but I fail to see what kind of implementation in human cognition would make that different from a semiotic system, since the GO hub would have connectivity to both form(s) and meaning(s), albeit a less simplistic mapping than that originally envisaged by Saussure. I conclude therefore that using semiotics over narrowly as a bogey man does not really advance matters much. Indeed, as we will see in the next section, there seem to be some interesting psycholinguistic evidence in favour of these 'hubs' for declarative memory within the linguistic system, except that one needs to make a distinction between open class and functional items.

Secondly, what *are* these syntactic atoms within the STRUCTURE domain? They include GO and OFF in addition to PAST, and also by Preminger's own admission DANCE and RUN. The abstract syntactic identity of these units labeled in small caps is "whatever is necessary to distinguish RUN from DANCE". But what is that, if it is not some kind of abstract semantics, especially considering that this toy model has to scale up to include all the verbs known by a particular speaker? Unfortunately, writing something in small caps does not make it syntax. I suspect that Preminger here is actually arguing for an abstract *semantics* as his ontologically prior start set from which all the mappings flow. (At the very least, Preminger and I interpret the word "semantics" in incompatible ways).

From what we know generally about the human brain, its plasticity and

learning capacity, it is highly implausible that the ontological primes Preminger invokes here are cognitively innate, given how highly specific to the linguistic system they are, and the differences found from language to language. In the view promoted by Preminger, the ontological primes are syntactic, but they have to be inferred from the knowledge of individual words and morphemes. Since these in Preminger's system are not even integrated or organic with respect to the ontological skeleton, this turns the *forms* of the language into highly unsuitable items to learn that system from, which seems unfortunate. Preminger's position would make perfect sense if there were a highly articulated innate set of syntactic featural primitives that the child relies on as the basic scaffolding for her complex words and sentences. But this would require us to believe that DOG and RUN are innate primitives of the system. The more reasonable alternative is to say that the primes of the fully mature system get established by inference from exposure to words. This indeed seems to be what Preminger believes. It is through processes of abstraction and generalization (cognitive capacities which are highly developed in our species) that the learner manages to create reusable symbols for her meaning-generating engine. So the memory hubs the language user has (whatever they turn out to be, and that is an empirical issue), must also support inferences to the underlying system of linguistic primitives in the speaker's language, as well as being operationally deployed in performance as psychologically real Things. So, just as phonological words were Things (as I argued above), so memory hubs are Things. They just happen to be different Things. The position I reject is the idea that *only* the syntactic primitives of the abstract representations built are Things.

Note that Borer (2013) assumes a different anchor point for characterizing the memory hubs we need. She claims that the hub for mappings to meaning in syntactic context is characterized (abstractly) phonologically, but that no abstract characterization for the semantics of /dog/ is possible. We will see in the discussion of polysemy vs. homonymy in the psycholinguistic literature, that semantic hubs appear to be necessary, and that the challenge instead is to overturn the classical view of what that semantic identity consists in.

4 The View from Psycholinguistics

Borer (2013) argues that there is a coherent definition of word that is phonological. The minimal units on the semantic side or the syntactic side do

not necessarily coincide with this well defined unit however. We have seen that Preminger rejects any ontologically primitive status for the primes of memory. But the primes of memory are the central characters in any psycholinguistic story, and are at the very least epistemologically primary, and the evidence is that speakers track statistical regularities of co-occurrence to pick out morphemes and contentful stems. The research programme on the psycholinguistics side has revolved around the nature and structure of the information stored by the user who demonstrates natural language competence in a particular language.

In general, there is strong evidence for interconnected networks in memory. The cohort model of Marslen-Wilson and Tyler (1980) assumes a straightforward forward feeding model of recognition from bottom up phonetic information as it unfolds in time. These cohorts gradually get winnowed down to a unique most highly activated target as the incoming information becomes fully discriminatory. Later models, such as the TRACE model of McClelland and Elman (1986) build in top down information flow from higher levels of representation such as at the word level, allowing many different competitors to be activated and to compete for recognition based on similarity and frequency. The current consensus appears to be that recognition involves both bottom up and top down processes of this kind.

Through priming studies of lexical access in comprehension, we can gain evidence for which representations co-activate others. Briefly, the speed of word recognition is reduced if that word has been previously activated in memory, but it is also, interestingly, affected by the prior activation of phonologically and semantically similar items. The strength and latency of these effects varies (with identity priming being the strongest, and semantic priming being the weakest and also with a longer latency), giving additional evidence for the architecture of the links within this kind of network.

We also know from production studies in picture naming that lexical access proceeds via the semantic or conceptual representation, thereby accessing an abstract lexical entry or 'lemma level', which in turn activates the abstract phonological representations and articulatory gestural programmes required to pronounce the word. This network and its latencies are now fairly well understood, giving rise to interactive models of competition and frequency effects that mirror the behavioural evidence found under experimental conditions, and also account for patterns in speech errors (Levelt 1999).

4.1 The Lemma Level

Over the course of the last five decades or so, we have accumulated a lot of evidence concerning the location of lexical recognition activation and its time course. It appears that the mid temporal gyrus (MTG) is involved in semantic lexical access independent of whether the sensory input is visual or auditory (Indefrey and Levelt 2004, Hickok and Poeppel 2007. Friederici 2012). Activation in this area can also be tracked using MEG and fMRI. Based on both neurolinguistic and behavioural evidence, we have strong support for the existence of the LEMMA which is the lexeme family underlying a symbol *and all of its inflectional forms*. Specifically, we know that lemma frequency as a whole (not the frequency of individual FORMS) modulates effects in the 300/450 ms time window in the MTG (Solomyak and Marantz 2010).

This literature is important because it shows that there is a lemma hub for all inflectional forms of the 'same' lexeme. But what constitutes 'sameness' in this sense? How does a learner decide to group FORMS heard under the same umbrella lemma, the 'same lexical entry' if you will. We know from the production literature in picture naming that the presence of semantically related distractors slows down lexical access in production tasks in predictable ways (Glaser and Düngelhoff 1984). Interestingly, it was shown early on that ambiguous words seemed to show a processing *advantage* when it came to naming (Kellas et al. 1988, Borowsky and M.E.J.Masson 1996, Azuma and Orden 1997). However, this effect is now known to discriminate sharply between genuine homonymy and the existence of polysemic variants. In a lexical decision study, Rodd et al. (2002) manipulated ambiguity vs. polysemy directly and found that the processing advantage was entirely due to the polysemous items, and that the ambiguous items in fact had an inhibitory effect, in line with the behaviour of semantically related distractors noted in the production literature. Moreover, in an MEG followup study from Beretta et al. (2005), the polyseme advantage over the monoseme baseline was shown to exist at the earliest level of lexical access, while the ambiguous items were delayed at this point.

While in practice, it is not always easy to decide whether a pair of meanings associated with a form are homonyms or polysemic variants, or what leads learners/speakers to classify them as such, the evidence now seems clear that we can distinguish between cases where there must be two 'lexical entries' versus cases where there must be one. The cases where we have clear evidence for one lexical entry involve lemmas which characteristically

embrace a large number of polysemic variants. Thus, polysemy is sharply distinguished in terms of cognitive consequences from homonyny, or genuine ambiguity, in which two distinct lemmas happen to share the same form. Polysemous readings are bunched together for the purposes of priming. Polysemous meanings are facilitory in word recognition, while genuine homonyms are inhibitory and cause slow downs in processing because of more alternatives remaining active. This is in line with one strand of theoretical work such as Nunberg (1979) and Pustejovsky (1995) who emphasize the cognitive naturalness of polysemy and the generalizations that exist concerning the types of relations and extensions that are tolerated in such meaning families.

To my mind this shows that the primes of memory, the thing I will call the 'symbol' must be a Thing which has a coherent semantic network associated with it. If we identify the term 'semantics' with the idea of a fully specified single denotation, then there is no single denotation associated with a symbol. But if we look at it in a more holistically it is clear that symbols have a unity at the level of meaning contribution which is key to their individuation as primes in the linguistic system. The lesson we learn from the homonymy vs. polysemy literature is that cognitively speaking, meaning is a criterion for individuation in a symbolic sense. The secondary lesson we learn is that we need to find ways of elucidating what those Meanings are in a way that goes beyond the classical formal semantics toolbox (see also Pietroski 2018).

4.2 Inflection vs. Derivation

A 'word' like *formed*, consists clearly of the 'form' part that tells you what kind of eventuality is being described, but it also contains an 'ed' part that tells you that the eventuality was instantiated at a time period before now. In many of the world's languages, we have evidence for temporal information being attached to symbols describing event types. In languages where temporal information is expressed with a separate free morpheme, the evidence from the syntax shows that is is hierarchically higher than the lower verb (Julien 2003); in languages where one of the morphemes is free and the other is bound, it is the tense morpheme that is bound and attached to the free symbol describing the nature of the eventuality. The above example is a case of what has been called 'inflection'. If one form is free and another form is bound to it, it is never the case that basic symbols corresponding to PAST and PRESENT are then modulated with suffixes that give more specific content to those eventualities. Pervasive asymmetries like this have given rise

to the distinction made in most theories of grammar between roots, or open class lexemes, and functional items (like tense marking, in this case). These asymmetries have implications for both word and sentence structure, leading famously to strong claims about the relationship between syntactic hierarchies and internal word structure (Baker 1985, Brody 2000). We have seen from the above discussion that the lemma level in lexical item corresponds to these open class items, and that the family of inflectional forms that is associated with the lemma are connected to it in a network-like fashion.

The other source of internal word structure is 'derivation'. In one intuitive articulation of the difference between inflection and derivation, inflection modifies or modulates a particular basic lexeme depending on syntactic context; derivation creates a new lexeme from another one (Matthews 1991, Bauer et al. 2013). Thus a word like *formation*, consists of the 'form' part that tells you about a particular kind of eventuality, and an 'ation' part that converts the word into a noun that describes the action of, or outcome of, that eventuality.[1] Crucially for the way the distinction is being made, *form* and *formation* are two completely different lexemes. In the above example, *formed* is simply a different inflected version of *form*.

Borer in her work makes an important distinction between inflection and derivation. For inflection, she accepts late insertion (separation) and sees inflected forms as the spell out of morphosyntactic features on roots. Derivational morphemes are distinguished from inflection in being represented by structural nodes in the syntactic representation. Classical DM on the other hand tends to adopt the same kind of architecture for both derivation and inflection.

What does the psycholinguistic evidence tell us in this case? Do speakers morphologically de-compose complex word forms? Is there a difference between productive and non-productive affixation? And is there a difference between inflectiom and derivation in this sense?

On the issue of decomposition, just because a linguist can decompose a form into sub pieces (with or without identifiable semantics), does not mean that that particular piece *is* an independent unit of memory. And just because a linguist can write down a rule or subrule which encodes a generalization over larger or smaller domains, does not mean that human brains

[1]Many derived nouns in *-ation* are systematically ambiguous between a noun denoting the action, and a noun describing the result of that action. This is interesting, but irrelevant for the point at hand. To understand the issues involved, I refer the reader to Borer 2013).

reliably treat those bits of information differently from memorized words like *cat*. For this reason, we cannot simply inspect the morphological forms extant in a language inventory, more detailed behavioural and neurological evidence is required. Fortunately, this work is now being done in a number of different domains.

One of the questions that has occupied psycholinguists is the distinction between declarative and procedural memory, corresponding to the intuitive difference between memorized word forms and online 'rules' for creating forms. Much of the work in EEG, MEG and fMRI has been focused on looking for a distinction between productive and nonproductive affixation, with the additional factor of semantic transparency also taken into account (Gwilliams 2020. Leminen et al. 2018).

Interestingly, with respect to the question of whether one can reliably distinguish between inflection and derivation, there are fewer studies that start with this question explicitly. Productivity, transparency and the issue of whether speakers decompose cuts across this distinction. As far as decomposition is concerned, our most recent evidence suggests that speakers do aggressively decompose and extract affixes at an early stage of processing, even when the base stem is unique to that form and decomposition, and regardless of semantic transparency (Rastle et al. 2004, Gwilliams and Marantz 2018). At the same time, there is also evidence of whole word form access, especially for semantically opaque or highly frequent complex words. This is true of both inflection and derivation it seems. And while in general there seems to be support for a dual route model, the evidence is not completely conclusive and the research to date has been disproportionately skewed in the direction of English. Work on other languages has if anything shown quite different kinds of patterns to those found for English (Leminen et al. 2018).

My own impression of this issue is that a simplistic cut off between productive (rule based) and unproductive (memory) routes is not tenable, and that even productive, semantically transparent forms show some whole word effects, especially in derivation. Individual morphemes also clearly seem to be tracked and individuated, and although their role in meaning composition is still not clear, there is evidence that this kind of complexity adds to processing time at the comprehension level. However, with respect to productivity, decomposition, and the different types of memory that might be involved, it is not clear from the literature that a sharp distinction can be made between inflection and derivation. Early evidence suggests that

memory and processing differences among languages with different kinds of inflectional and derivational systems are likely to be substantial.

Putting the notion of productivity aside, if we look more closely at the neuroimaging data, and the details of lexical access and priming, clear differences do emerge between the two traditional morphological types.

Inflection directly engages the LIFG (Left Inferior Frontal Gyrus) (Marslen-Wilson and Tyler 2007, Whiting et al. 2014). This is also true even when morphology is covert (Sahin et al. 2009). The evidence for engagement of the LIFG on the other hand is much weaker (inconclusive) for derivation (Leminen et al. 2018).

Finally, as we have seen, processes that are sensitive to lemma frequency (for which there is robust evidence), lump together the open class item *with its inflectional instantiations* and not with all its derivatives, indicating that when it comes to individuation, derivational forms do seem to have more autonomous lexical entries than inflected ones. General priming seems to distinguish inflection from derivation. Recall that forms that successfully prime each other (like *form* and *formation*) do not need to be listed under the same lemma to interact in a network of activation spreading. The phonological similarity between the forms would be enough to give rise to priming (although Frost et al. 2000 demonstrate for Hebrew that morphological relatedness primes even over and above the form relatedness that often goes along with that relationship). Importantly then, in addition to derivational priming being somewhat weaker than inflectional priming and identity priming, there is also evidence that derivational suffixes *can prime each other* (Marslen-Wilson et al. 1996), whereas this has not been found for inflection. This suggests that derivational morphemes have their own lives as Things, in a way that inflectional forms do not.

My conclusion therefore is that Borer is right in this case, and that inflection should be distinguished in the system from derivation. Adopting a realizational strategy for inflection and a strategy for derivation which involves the analysis of derivational morphemes as heads in the phrase structural decomposition is consistent with the evidence from priming.

Further psycholinguistic work is required to address this question directly.

5 Reversing the Burden of Proof

Looking at the last few decades of neuro and psycholinguistic experimentation on 'words', I have argued that a coherent picture emerges concerning the architecture and organization of lexical access. We find mappings on the one hand from a coherent identifying network of semantic/conceptual representations associated with an abstract hub or lemma. At the same time, phonological information concerning inflectionally modulated variants of that lemma also seem to be accessed directly from that lemma hub.

This is not so different from the figure that Preminger proposes in (2), with his two sets of mappings: (i) from DOG to all its possible allosemes in different contexts and (ii) from DOG to all its possible allophones in different contexts. However, in the Preminger set up, DOG is a syntactic prime, and the unity of the network is not emphasized.[2] In what sense is the hub for the lemma corresponding to DOG or RUN a piece of syntax? Standard DM would deny that the root in this sense has any syntactic information associated with it at all. Preminger on the other hand, lists RUN as a *syntactic* prime. I have argued that the truth is probably not either of these things. Granted, to the extent that a symbol hub is not a required part of a mind that does not have language, the lemma is as abstract and linguistically specific a construct as you are likely to find. If this means that it is syntactic, then Preminger and I are not using the word 'syntactic' in an equivalent way, although I agree that the lemma is abstract and not derivable from facts about real world information. On the other hand, it is clearly the hub on which various different listed properties hang, both conceptual, syntactic and realizational. To my mind, this makes it part of the listed or declarative part of linguistic knowledge, which should be kept distinct from a description of combinatoric system and its properties per se. Moreover, the defining properties of the lemma's individuation as an element in the linguistic system seem to be conceptual/semantic. This goes against Borer's (2013) position, which denies that such unities form an organizational hub for lexical access.

However, as we have seen from the priming and lexical access literature, inflection seems to be treated differently from derivation. While both inflected and derivational forms of a given stem prime the bare stem strongly

[2]The mappings he considers to be real are not strictly speaking from DOG alone, but from a complex syntactic object containing DOG. However, the presence of DOG in both sets of rules is an important unifying factor, and arguably precisely what the psycholinguistic evidence suggests.

in experimental paradigms, inflectional priming is stronger. In addition, derivational morphemes may have their own abstract hubs, unlike inflectional morphemes which potentially only emerge as the output of a mapping from a hub to an abstract phonological representation in a syntactic context. On this point, the DM position of treating inflectional and derivational morphology as equivalently realizational based on abstract syntactic features, seems less convincing than Borer's own implementation.

As is clear from this short paper, I take the view that what we learn about neurocognition and psycholinguistic behaviour are important and central sources of evidence for theory construction. If abstract theorizing tells us that the memorized form - meaning units of a language are illusory, but neuro and psycholinguistic data backs up the cognitive reality of those units, then the solution has to be to reexamine and modify the basic assumptions of the theory.

To clarify, I am not advocating a surfacey, construction based, or frequentist, interpretation of grammar here (see e.g. Goldberg 1995). I think there is also good evidence for abstract mental representations over and above the constructions and forms that give rise to them. It seems to me that human cognition is extremely good at pattern finding, and generalizations. We are able to store abstract symbolic units which can then be deployed in various contexts because our representations of them are set up to be so modulated. Mental representations must be abstractions over exemplars in some sense.

In my view, part of the problem with the formalizations we have inherited in linguistic theory is that on the semantic side they are explicated in terms of an externalist tracking of effects rather than an internalist conception of what the speaker knows or has represented. Formal semantics in particular concentrates on describing the truth conditions of utterances in specific contexts and then reverse engineers 'meanings' of those symbolic items inside the utterance to derive those truth conditions. But this ends up giving us as many different formal denotations as there are truth condiitionally different polysemic variants. As we have seen, the evidence from processing shows that polysemy is a real and ubiquitous part of symbolic meaning, and that polysemy exists within a unified symbol representation.

The challenge at this point is to understand the nature of these abstract semantic hubs, and how they can be subsequently modulated in contexts (even those contexts distinct from the ones the learner was exposed to in acquiring the abstraction). To summarize the major positions in the literature, we have seen that Borer has consistently advocated for phonological

abstract unity being the anchor for these memory units. On the other hand, it seems as though DM is moving towards a version of the architecture in which the hub is a highly abstract index purely, and where mappings must be listed to both meaning and form (a version of the Preminger position). In my own view, simply listing allosemic subcategorization frames (Marantz and Wood lume, Myler 2016) lets us off the explanatory hook, and allows us in principle to list allosemes that could never be acquired or integrated within a human knowledge system, essentially begging the most interesting questions of unification via generalization and abstraction.[3]

	Memory Unit	**Variants**
Borer (2013)	Phonologically Anchored Hub	Allosemic
New DM[4]	Abstract Index	Allosemic *and* Allophonic
Ramchand	Conceptually Anchored Hub	Inflectional/Realization

Lemmas are abstract and polysemous hubs of semantic association, associated with abstract phonological features, and also potentially with some syntactic (categorizing) information, Arguably, it *must* be so in order to bootstrap the whole process of acquiring the mature system. Work such as Pustejovsky (1995) goes in the direction necessary to directly engage with these hard questions, the point being that truth conditional denotations are simply not appropriate at the level of symbolic content (cf. also Pietroski 2018), and give rise to non-explanatory disjunctive allosemic lists.

As we have seen, we have more and more subtle evidence that human parsers decompose automatically, at the same time as having entries for full word forms. And what about the "word" iself? As I have argued, we do not need to define it as the minimal unit of semantic association, or the thing associated with a terminal node. I essentially agree with Borer that the only coherent use for "word" is as a unit of phonology/linearization. There is a separate linearization algorithm that operates within it, as opposed to over it. Two recursive domains of linearization give rise to words. Words by definition are the primes of the linearization algorithm within syntax. while morphemes are linearized within words. As the primes of the linearization system, they do not need to conform to memory hubs, i.e. the thing I am calling the symbol.

[3]Recall that not all learning contexts lead to unified polysemous lemmas, sometimes the learner acquires homonyns, which as we know, behave differently within the user's linguistic system.

In this short paper, I have taken Borer (2013) as a starting point for a general architectural discussion of the primes of the linguistic system of a particular natural language. Borer's work in this domain has been profound and influential, and while we have often disagreed on details, we have agreed on the nature of the questions and the importance of the enterprise. In this sense, we have been fellow travellers.

I have ended up advocating the unfashionable view that while not being in any sense atomic at the core levels of phonology syntax or semantics respectively, that both "words" (phonological) and semiotic "symbols" (memory hubs) are indeed Things. Despite not being the same, or atomic, they are the items that allow the user to infer the atoms of the abstract system and importantly, provide the reusable hubs for comprehension and production lodged in declarative memory.

References

Ackema, P. and A. Neeleman (2007). Morphology ≠ syntax. In G. Ramchand and C. Reiss (Eds.), *The Oxford Handbook of Linguistic Interfaces*, pp. 325–352. Oxford and New York: Oxford University Press.

Arad, M. (2003). Locality constraints on the interpretation of roots: the case of hebrew denominal verbs. *Natural Language and Linguistic Theory 21(4)*, 737–778.

Arad, M. (2005). *Roots and patterns: Hebrew morpho-syntax*. Dordrecht: Springer. Studies in Natural Language and Linguistic Theory 63.

Azuma, T. and G. V. Orden (1997). Why safe is better than fast: the relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language 36*, 484–504.

Baker, M. (1985). The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry 16*(3), 373–415.

Baker, M. C. (2001). Phrase structure as a representation of "primitive" grammatical relations. In W. D. Davies and S. Dubinsky (Eds.), *Objects and other Subjects: Grammatical Functions, Functional Categories and Configurationality*, pp. 21–51. Dordrecht: Kluwer.

Bauer, L., R. Lieber, and I. Plag (2013). *The Oxford Reference Guide to English Morphology*. Oxford, UK: Oxford University Press.

Beretta, A., R. Fiorentino, and D. Poeppel (2005). The effects of homonymy and polysemy on lexical access: an MEG study. *Cognitive Brain Research 24*, 57–65.

Borer, H. (2013). *Taking Form: Structuring Sense Volume III*. Oxford: Oxford University Press.

Borowsky, R. and M.E.J.Masson (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology. Learning 22*, 63–85.

Brody, M. (2000). Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry 31*(1), 29–56.

Chomsky, N. (1957). *Syntactic Structures*. 's-Gravenhage: Mouton.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Ma.: MIT Press.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale and S. J. Keyser (Eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, pp. 1–52. Cambridge, Ma.: MIT Press.

Dixon, R. (1977). *A Grammar of Yidin*. Cambridge, UK: Cambridge University Press.

Embick, D. and R. Noyer (2001). Movement operations after syntax. *Linguistic Inquiry 32*(4), 555–595.

Friederici, A. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences 16*(5), 262–268.

Frost, R., A. Deutsch, O. Gilboa, M. Tannenbaum, and W. Marslen-Wilson (2000). Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition 20*(8), 277–288.

Glaser, M. O. and F.-J. Düngelhoff (1984). The timecourse of picture-word interference. *Journal of Experimental Psychology. Human Perception and Performance. 7*, 1247–1257.

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Gwilliams, L. (2020). How the brain composes morphemes into meaning. *Philosophical Transactions of the Royal Society B*, 375.20190311120190311.

Gwilliams, L. and A. Marantz (2018). Morphological representations are extrapolated from morpho-syntactic rules. *Neuropsychologia 114*, 77–87.

Halle, M. and A. Marantz (1993). Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyse (Eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, pp. 111–176. Cambridge, Ma.: MIT Press.

Harley, H. (2014). On the identity of roots. *Theoretical Linguistics 40(3/4)*, 225–276.

Hickok, G. and D. Poeppel (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience 8(5)*, 393–402.

Indefrey, P. and W. J. Levelt (2004). The spatial and temporal signatures of word production components. *Cognition 92(1-2)*, 101–144.

Julien, M. (2003). On the negated past in Finnic and Saami. In D. Nelson and S. Manninen (Eds.), *Generative approaches to Finnic and Saami linguistics*, pp. 419–446. CSLI Publications.

Julien, M. (2004). The syntax of Scandinavian nominal phrases.

Kellas, G., F. R. Ferraro, and G. B. Simpson (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology. Human Perception and Performance. 14*, 601–609.

Kiparsky, P. (1982). Lexical morphology and phonology. In T. L. S. of Korea (Ed.), *Linguistics in the Morning Calm*, pp. 1–91. Seoul: Hanshin.

Leminen, A., E. Smolka, J. D. nabeitia, and C. Pliatsikas (2018). Morphological processing in the brain: the good (inflection), the bad (derivation) and the ugly (compounding). *Cortex 116*, 4–44.

Levelt, W. (1999). Models of word production. *Trends in Cognitive Sciences 3*(6), 223–232.

Marantz, A. (1997a). 'cat' as a phrasal idiom: Consequences of late insertion in Distributed Morphology.

Marantz, A. (1997b). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In A. Dimitriadis and L. Siegel (Eds.), *Proceedings of the 21st Annual Penn Linguistics Colloquium*, University of Pennsylvania Working Papers in Linguistics, pp. 201–225. Philadelphia: University of Pennsylvania.

Marantz, A. and J. Wood (this volume). The interpretation of external arguments. In R. D'Alessandro and I. Franco (Eds.), *The Verbal Domain*. Oxford University Press.

Marr, D. (1982). *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W. H. Freeman and Company.

Marslen-Wilson, W. and L. K. Tyler (1980). The temporal structure of spoken language understanding. *Cognition 8*, 1–71.

Marslen-Wilson, W. T., M. Ford, L. Older, and X. Zhou (1996). The combinatorial lexicon: Priming derivational affixes. *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society 18*, 223–227.

Marslen-Wilson, W. T. and Tyler (2007). Morphology, language and the brain: the decompositional substrate for language comprehension. *Transactions of the Royal Society of London. Biological Sciences 1481*(362), 823–836.

Matthews, P. H. (1991). *Morphology* (2nd ed.). Cambridge: Cambridge University Press.

McClelland, J. L. and J. L. Elman (1986). The trace model of speech perception. *Cognitive Psychology 18(!)*, 1–86.

Myler, N. (2016). *Building and Interpreting Possession Sentences*. MIT Press.

Nunberg, G. (1979). The non-uniqueness of semantic solutions: polysemy. *Linguistic Philosophy 3*, 143–184.

Pietroski, P. (2018). *Conjoining Meanings*. Oxford University Press.

Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, Ma.: MIT Press.

Rasin, E., O. Preminger, and D. Pesetsky (2021). A re-evaluation of arad's argument for roots. *Proceedings of WCCFL 39*.

Rastle, K., M. H. Davis, and B. New (2004). The broth in my brother's brothel: morpho-orthographic segmentation in visual word recognition. *Psychon. Bull. Rev 11*, 1090–1098.

Rodd, J., G. Gaskell, and W. Marslen-Wilson (2002). Making sense of semantic ambiguity: semantic competition in lexical access. *Journal of Memory and Language 46*, 245–266.

Sahin, N. T., S. Pinker, S. Cash, D. Schomer, and E. Halgren (2009). Sequential processing of lexical, grammatical, and phonological information within Brocas area. *Science 5951*(326), 445–449.

Solomyak, O. and A. Marantz (2010). Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience 22(9)*, 2042–2057.

Whiting, C., Y. Shtyrov, and W. Marslen-Wilson (2014). Real-time functional architecture of visual word recognition. *Journal of Cognitive Neuroscience 2*(27), 246–265.

Wood, J. and A. Marantz (2017). The interpretation of external arguments. In I. F. Roberta D'Alessandro and A. Gallego (Eds.), *The Verbal Domain*, pp. 255–278. Oxford University Press.