

Roots, concepts, and word structure: On the atoms of lexical semantics

Paolo Acquaviva, University College Dublin - paolo.acquaviva@ucd.ie

Abstract

This chapter examines the relation between the structure of words as linguistic objects and their conceptual content. It addresses two questions: what are the primitives of lexical semantic interpretation, and how they are expressed in the grammatical and morphological representation of a lexical item. The answer involves a characterization of roots as theoretical objects, followed by an argument to the effect that it is not roots, but larger structures of variable size which relate to lexical concepts. An in-depth discussion of nouns leads to the claim that the conceptual content of a lexical item does not reflect its grammatical structure, because a concept is not the meaning of a linguistically defined unit, but a language-external cognitive content, globally associated with the lexical word as a whole.

Keywords: Concepts, lexical semantics, morphology, roots.

1. Introduction

It is rather uncontroversial that the meaning of ‘substantive’, open-class lexical items like *table* or *say* comprehends a part that is grammatically relevant and a part that is not. When it comes to delineating the encyclopaedic and the linguistic components, however, positions differ, and theoretical choices must be made. The aim of this chapter is to propose a view of the relation between grammatically encapsulated content and the overall conceptual content of lexical words, based on a syntactic approach to lexical structure. After considering how this approach relates to other models of word-internal structure and its morphological interpretation, I will focus on the notion of lexical root as the putative locus of purely conceptual, non-grammatical meaning. In the face of evidence which makes this position problematic, I will outline an alternative, which distinguishes roots as a syntactic and as a morphological notion; this paves the way for the proposal that a word’s conceptual content is associated not with its root or any other linguistically defined piece, but with the word as a global map of linguistic and non-linguistic knowledge, where grammatically encapsulated content *constrains*, but is not a *part* of, lexical meaning.

2. Word meaning and word structure

published in: . F. Rainer, W. U. Dressler, F. Gardani, and H.C. Luschützky (eds). 2014. *Morphology and Meaning. Selected Papers from the 15th International Morphology Meeting*. Amsterdam: John Benjamins, 49-70.

2.1 Structuring word meaning

It is the recurring regularities discernible across lexical items that make it possible to study lexical competence as a linguistic and cognitive phenomenon. Several important contributions have attempted to isolate the conceptual primitives underlying the expression of lexical meaning, not just full-fledged words, but also lexical stems and lexeme-forming affixes (see the overview in Lieber 2004: 6-10). A major strand of this research programme investigates not just what lexical meaning consists of, but also how it is structured. Analyses of the internal constituency of lexical meaning differ, among other things, in how they treat non-linguistic knowledge. Some proposals explicitly posit distinct components. Lieber (2004) distinguishes between a ‘Skeleton’, which represents primitive conceptual properties in function-argument structure, and a ‘Body’, which contains what is encyclopaedic, holistic, non-decompositional, and culture-dependent. Examples of Skeletons are “[+material ([])]” for the noun *chair*, and “[+dynamic ([], [])]” for the verb *kiss* (Lieber 2004: 25), where the open places encode valency, and the features express the conceptual categories grouping together, respectively, descriptions of concrete objects and of non-stative events.

The influential model of verb meanings developed by Levin & Rappaport Hovav (1995, 2005) differs in that it lays a greater emphasis on the structural analysis of verbs than on the conceptual primitives of lexical meaning. Levin & Rappaport Hovav set out to identify invariant structural templates across lexical items, which define verb classes generalizing over many words. In simple cases, such lexical semantic templates involve a primitive predicate and a ‘constant’ or ‘root’, respectively corresponding to BECOME and BROKEN in the following representation:

- (1) noncausative *break*: [y BECOME *BROKEN*]

Constants / roots distinguish one verb from another within the same template, but they do not by themselves identify lexical items; in fact, the same constant / root may appear in different templates, as in the pair provided by *to shovel* as a verb of removal (*shovel the snow from the path*) or as a verb of putting (*shovel the sand on the truck*).

Another family of approaches treats lexical structure as syntactic. For Hale & Keyser (2002), syntactic structure defines argument structure and derives lexicalization patterns. For example, an interpretation corresponding to ‘they put salt in the box’ can be associated with the structure *they boxed the salt*, but not **they salted the box*, because use of the nouns *salt* and *box* in verbal capacity proceeds from the syntactic structure schematized in (2):

(2) [VP V [PP N(*salt*) P N(*box*)]]]

Selection holds between V and P, and P and N(*box*), but not between V and N(*salt*). It is this syntactically defined difference that motivates the different patterns of lexicalization, as the selected *box*, but not the unselected *salt*, can spell out the higher node V. Hale & Keyser's approach recapitulates most word-formation into syntax; other approaches instead envisage separate operations to morphologically interpret a syntactic structure. Such is the case of Distributed Morphology (Halle & Marantz 1993, Embick & Marantz 2008, among many others). Syntax is also the only structure-building component for Borer's model (2005a,b), which differs from Distributed Morphology on several counts (notably in its view of the relation between morphological exponence and syntactic constituency), and argues that *all* lexical properties that are grammatically relevant (like countability, event structure, or argument structure) are properties of syntactic structure; what she calls lexical 'listemes' neither contain grammatical properties nor determine those of their environment.¹

2.2 *The locus of idiosyncratic non-grammatical meaning*

It is instructive to briefly consider how other models that distinguish grammatically relevant from non-grammatically-relevant information relate this opposition to the structure they posit. Lieber (2004) consigns the encyclopaedic, culturally dependent, idiolectally variable components of lexical meaning to the Body, which is a property of specific words and word-forming morphemes, opposed to the hierarchically arranged function-argument structure defined in the Skeleton, which is a template applying to whole classes of lexical items. The distinction between Skeleton and Body does not correspond to that between root-external material and innermost root. This is clear, for example, in the following passage:

The skeletons of which compounds are formed will generally have accompanying bodies, but derivational affixes will often have little or nothing in the way of semantic bodies. Both derived words and compounds may, however, over time, develop substantial and distinctive bodies as a function of their lexicalization. (Lieber 2004: 10-11)

The view emerging from Distributed Morphology is different. Making allowance for the fact that a variety of analyses have been developed within this framework over 20 years, and that this work has consistently focused on the grammatical aspects of lexical meaning rather than on the non-grammatical ones, it is still clear that in Distributed Morphology roots determine the idiosyncratic

and encyclopaedic interpretation of the syntactic structure surrounding them (cf. the detailed discussion in Arad 2005). On this question, Borer's position is not substantially different; however, her approach gives much greater attention to the syntactic construction of meaning and takes a clear position on the relation between grammar-internal and grammar-external aspects of lexical knowledge, namely, that what does not reduce to linguistic competence is encapsulated in the 'listemes' around which the syntactic structure is built:

Within an XS-[exoskeletal] model, then, the particular final meaning associated with any phrase is a combination of, on the one hand, its syntactic structure and the interpretation returned for that structure by the formal semantic component, and, on the other hand, by whatever value is assigned by the conceptual system and world knowledge to the particular listemes embedded within that structure. These listemes, I suggest, function as modifiers of that structure. (Borer 2005b: 9)

I will argue that, while grammatical knowledge underlies much of lexical semantics, non-grammatical conceptual knowledge is not encapsulated in root-like listemes, but is embodied in linguistic structures of various complexity. The argument hinges precisely on those listed cores of lexical words which, even in an interpretation like Borer's, cannot be analyzed as parts of a syntactic shell produced by the vocabulary and principles of grammar.

2.3 *Conceptions of roots*

Aronoff's (1994: 40) definition of a root as "what is left after all morphological structure has been wrung out of a form" refers unambiguously to *form*, and identifies a root as an unanalyzable residue defined in morphological terms as a morphological object. All approaches to lexical decomposition necessarily include a corresponding notion for the core elements of lexical structure, but, as the type of structure hypothesized changes, so does the conception of root.

Levin & Rappaport's 'constants' or 'roots' are not purely morphological objects as in Aronoff's definition, since they are defined as irreducible components of an abstract *semantic* template; this does not coincide with a phonological form (called a verb's 'name' in Levin & Rappaport Hovav 1995: 28), nor with a tag that identifies each lexical item (cf. *shovel*₁ ≠ *shovel*₂). In fact, in class lectures Levin (2009: 2) explicitly states that her notion of root is "not to be confused with the notion of root used in morphology; e.g., Aronoff 1993.).". Coherently, roots in this sense are categorized as nouns, verbs, or adjectives, and have denotations which can be made

fully explicit, as in these examples from Koontz-Garboden (2011), where *red* is an adjective denoting a state holding of entities, and *break* is a verb denoting a caused change of state:

- (3) $\llbracket \text{red}_A \rrbracket = \lambda x \lambda s [red' (s, x)]$
 $\llbracket \text{break}_V \rrbracket = \lambda x \lambda y \lambda e \exists s [CAUSE (y, e) \wedge BECOME (e, s) \wedge non-whole' (s, x)]$

These morphemes are called roots because the denotations provided apply to morphologically unanalyzable lexical words, but they remain distinct from roots as morphologically defined forms. By contrast, the roots of Distributed Morphology are primarily defined in syntactic and morphological terms. More precisely, they are terminals filling category-free heads, and are identified with the cores of open-class words, in a theory that treats all such words as consisting of a core (the root) plus a distinct categorizing head, often null (Arad 2005). To illustrate, Embick & Marantz (2008: 4) represent the noun *cat* as (4):

- (4)
- $$\begin{array}{c} n \\ \swarrow \quad \searrow \\ \sqrt{\text{CAT}} \quad [n, \emptyset] \end{array}$$

The most distinctive aspect of this proposal is probably the dissociation between root and category as distinct syntactic heads, which raises the question of how to properly describe the content of the former in isolation from the latter; that is, how to state what *cat* means as a root abstracting away from its being a noun. Arad (2005), in the most in-depth treatment of this issue, qualified such roots as *signs*, which map a (possibly underspecified) phonological form with an underspecified semantic representation. For Arad, some roots identify a rather precise meaning, while others are associated with a much more impoverished content which underlies several distinct word meanings, as in the Hebrew series *šavar* “to break”, *mašber* “crisis”, *šever* “fraction”, *šavir* “fragile”, all built on the root *šBR* (Arad 2005: 193). These cases illustrate what Arad calls Multiple Contextual Meaning, defined (p. 90) as “the interpretation assigned to roots in different environments”: a root syntactically selects a number of local contexts, and each combination corresponds to a specific (word-) interpretation listed in the so-called Encyclopaedia.

A pairing of sound and meaning also features in Borer’s (2005b: 25) definition of the unanalyzable elements that appear in a syntactic environment but are themselves void of grammatical information : “By listemes we refer to a pairing of a conceptual feature bundle with a phonological index”. Like the roots of Distributed Morphology, these listemes lack a lexical category, which is inferred contextually from the syntactic context (without dedicated categorizing

heads). The question arising is how this “conceptual feature bundle” is linguistically encapsulated, if it is not the content of a noun or a verb. As we will see, an answer is made difficult by the fact that different approaches to lexical decomposition construe the notion of root in different ways, not directly comparable. Therefore, our discussion will proceed from an empirical claim about roots as morphological objects, as the least theory-dependent construal of this notion.

3. *Morphological roots do not encapsulate core lexical meaning*

Whatever relation exists between the semantic and the morphological structure of a lexical item, there is empirical evidence that roots, in the Aronovian sense, are not the formal side of a sign whose content is the core lexical meaning. To summarize the discussion in Acquaviva (forthcoming), words built on the same root often lack a common semantic nucleus, or are semantically related but do not share a coherent conceptual content, or share a conceptual content but still differ in lexical, non-grammatical properties.

3.1 *No common semantic denominator between corradical lexemes*

Arad herself (2005: 102) notes that many Hebrew roots fully determine the content of one and only one lexical concept for words referring to natural kinds, food, or kinship relations (which of course can be further derived). Even within the same Semitic language, then, there are whole areas of the vocabulary where the idea of meaningful but underspecified roots, compatible with one of several context-driven alternative meanings, fails to apply. When we consider languages in which roots do not have the cross-lexeme distribution they have in Semitic, the empirical inadequacy of this characterization of roots as morphosemantic atoms of lexical content stands out all the more clearly. It is not just that in a language like English most roots identify one lexical concept (further modifiable, as in *book*, *book-let*, *book-ish*, *book-ish-ness*, *book-worm*) rather than a constellation of loosely related concepts. The fact is that, when such a series of words exists, the root they share does not have any plausible content which could relate the lexemes with each other, no matter how loosely. In English, this pattern appears most clearly with Latinate roots like -CEIVE or in -MIT in *re-ceive*, *con-ceive*, *per-ceive*, or *ad-mit*, *com-mit*, *per-mit*, which Aronoff (1976: 11-14) identified as morphological entities, and not just pre-packed phonological strings, on the basis of their regular alternations <-ceive, -cept> <-mit, -mission> (contrast Borer 2005b: 353, who takes them as “phonological indices”). Revealingly, the same phenomenon obtains in Italian, where the numerous verbs sharing the root -METT- are not a learned class marked off from the rest of the native lexicon

(*mett-ere* “put”, *am-mett-ere* “admit”, *per-mett-ere* “permit”, *pro-mett-ere* “promise”, *s-com-mett-ere* “wager”, *s-mett-ere* “quit”, and several others). More generally, it suffices to think of cases like *stand* and *under-stand*, *with-stand* in English, or *hören* “to hear” and *auf-hören* “to quit” in German, to realize how often corradical words are semantically unrelated (see Aronoff 2007).

3.2 Semantic relatedness without a shared concept

Being semantically related does not mean the same as having some semantic content in common, in particular not semantic content that is the meaning of a sub-expression. The claim is coherent, and plausible for cases like *dog* – *dogbed* (less so for *dog* – *dogged*), but it remains a claim. When we question the justifications for reifying semantic relatedness as the meaning of a root, we quickly discover how problematic the putative shared nucleus can be. An apple and an apple-tree, for instance, correspond to concepts that are definitely related; one thing is to recognize this relatedness, though, and quite another to posit a concept for something that is neither a fruit nor a tree, but underlies both. In English, it is fair to say that the notion of apple tree is derivative from those of apple (the fruit) and tree. But when they are expressed by two nouns with the same root, like the Latin *mālum* (neuter) “apple” and *mālus* (fem.) “apple tree”, the idea of semantically contentful underspecified roots would lead us to see this relatedness as the meaning of the root. Yet there are no more reasons to do so in Latin than in English; in both cases, it is equally implausible to posit a totally abstract apple-concept with *none* of the properties specific to fruits or to trees; in fact, it is hard to see what other properties might substantiate this concept, and indeed how it could be learned at all.

The problem generalizes to many pairs that are not so obviously taxonomically related; consider just the Swahili *u-siku* (class 11/4) “night” – *siku* (class 9/10) “day” (Polomé 196: 95, 103), or the Russian *tsvet* “colour” – *tsvet-ók* “flower”, or the Italian *man-ic-o* “handle” – *man-ic-a* “sleeve”, which not only shares the root *man-* of *mano* “hand”, but also the suffix *-ic-*. In fact, establishing an umbrella meaning for related word meanings proves problematic even in case of a single word: as Fradin & Kerleroux (2009: 91-92) show, the senses “daughter” and “female child” of the French *fille* cannot be analyzed as sub-specifications of a shared underlying property, because the only shared component is so generic that it would encompass all lexemes referring to human females, failing to single out *fille*. These considerations, which are not novel in any way, show that it takes more than mere semantic relatedness to justify a semantically contentful root; and they cast doubt on Marantz’ and Arad’s view of roots as signs underlying the semantic relatedness of words like *mašber* “crisis” and *šavir* “fragile”.

3.3 Roots underdetermine lexical properties

The existence of different lexeme-specific properties for corradical words poses another problem for approaches that associate all lexical properties with the root and all grammatical properties with its syntactic environment. Consider how English phrasal verbs differ from the corresponding simplex verbs in selectional restrictions:

- (5) a the grandmother sewed a dress / # a deal (Basilico 2008)
b the grandmother sewed up a deal with the yarn company
- (6) a the criminals cooked a meal / #an evil scheme (ibid.)
b the criminals cooked up an evil scheme

In Basilico's analysis, this type of phrasal verbs consist of a complex root, in which both the verb stem and the particle occur below a verbalizing head [v]. This structural interpretation accounts for the ability of phrasal verbs to specify their own selectional restrictions, but it comes at a price: what counts as root can no longer be defined on purely morphological criteria. What determines selectional restrictions is not just a root, but a theoretical object corresponding to *cook* in (5a) but to *cook up* in (5b). Indeed, morphological roots and such theoretical objects may have conflicting lexical properties: the roots *wise* and *cosy* positively resist use as verbs (**to wise*, **to cosy*), but admit it in the phrasal expressions *wise up* and *cosy up*.

Event nouns illustrate a related point. Consider the two readings of *collection*, respectively as a nominalization of the predicate *collect*, which preserves the verb's argument structure and allows reference to the event's distribution in time (as in *the frequent collection of mushrooms by Nina*), and as a designation for an object, or 'result nominal' (as in *let me show you my collection of stamps*). According to Borer (2003), transparent nominalizations like *collection* in the first reading are best analyzed as full-fledged verbal structures, which determine event and argument structure, capped by a nominalizing morpheme. Notice that the root is the same in both construals, which are clearly related, witness the label 'result nominal' for the second. The root, then, fails to determine both its syntactic environment (verbal or nominal) and the ontological sort of its denotation (event or object).² This type of ambiguity may further determine differences at the morphological level, as shown by the contrast between *argument* as a logical category and as a nominalization from the verb *argue*: here the same root, suffixed by the same nominalizing morpheme *-ment*, allows two readings distinguished not only semantically, but also by their ability to feed further derivation in *-al* (cf. Acquaviva 2009, forthcoming, for a discussion of Aronoff's 1976 original analysis):

- (7) *argu-ment*¹ “logical category”
*argu-ment*² “event of arguing”
argu-ment-al “relative to *argument*¹ / # *argument*²”

A final piece of evidence is provided by semantically distinct inflectional forms; for instance, Italian plural doublets like *membri* “members” – *membra* “limbs” (from *membro* “member”), or the special mass reading of English plurals like *brains*, along the regular count reading based on the singular (Acquaviva 2008). Again, what sets apart the two readings cannot be a consequence of the root, which remains constant.

4. *Syntactic vs. morphological roots*

Before addressing the main issue of the relation between lexical content and grammatically expressed word structure, let us briefly consider an alternative conception of roots within a syntactic decomposition approach. In joint work with P. Panagiotidis (Panagiotidis 2011, Acquaviva forthcoming), we distinguish between roots as morphological objects and roots as elements of the syntactic computation. Syntactic roots, notated by capitals, are the innermost elements in the syntactic structure defining a lexical item, corresponding in this function to the roots of Distributed Morphology and to Borer’s lexical listemes, but with a crucial qualification: they are abstract elements of the syntactic computation, distinct from the morphological objects that realize them. This allows us to reconcile the theoretical claim that roots are maximally underspecified (implicit in all syntactic decomposition approaches, where a root is what is left after all *syntactic* structure has been abstracted from a word) with the evidence that morphological roots may be marked for a lexical category, or other syntactically relevant information. In such cases, a morphological root would spell out a syntactic root node fused with the featural information that originates from a distinct node, like [Tense] for the suppletive root allomorph of *go*:

- (8) a $\sqrt{\text{CAT}} \leftrightarrow \text{cat}$
 b $\sqrt{\text{GO}} \leftrightarrow \text{go}$
 $\sqrt{\text{GO}}, [\text{Tense: Past}] \leftrightarrow \text{went}$

Our view not only accommodates, but predicts the existence of lexical roots used to spell out a grammatical terminal, resulting in open-class elements with closed-class function, like the Irish *caitfidh* (future of *cait* “to cast, spend”) as modal auxiliary: *caitfidh mé imeacht* “I must leave”.

In addition, morphological roots can also display constraints on their form, such as the well-known three-consonant skeleton characteristic of Semitic roots. But such constraints only affect roots as morphological objects, not as syntactic terminals. In our view, the latter are abstract symbols, purely formal objects internal to the faculty of language in the narrow sense, where they act as formal indices marking lexical relatedness in a precise, syntax-internal sense. One and the same syntactic root may well correspond to different morphological roots, as the multiple ‘radicals’ or ‘stems’ of Latin inflection and derivation (Aronoff 1994). Apart from this purely differential function, a syntactic root for us has no meaning by itself. It is not a sign, but determines an interpretation for the syntactic structure that embeds it (possibly a structure where nothing else is spelled out at word level, as in *cat*). As a syntax-internal criterion of lexical identity, it marks formal, not conceptual relatedness; by itself, a root is no instruction to ‘fetch’ or activate a concept (contrast Pietroski 2008). While this negative conclusion accords with the evidence, as we have seen, it reopens the question of how a word’s conceptual content relates to its syntactic structure.

5. *Complex words, simplex concepts*

5.1 *Word structure and concept structure*

It seems self-evident that decomposing lexical items into a structure of meaningful elements should correspond (perhaps indirectly) to a decomposition of their conceptual structure. But asking how grammatical structure relates to conceptual *content* is not the same as asking how it relates to conceptual *structure*; for the second formulation of the question makes a substantial assumption, namely that the structure hypothesized by linguistic analysis is *constitutive* of conceptual content. The assumption leads us to expect that what remains after stripping away the grammatical pieces is the non-grammatical core of a word; and the expectation proves incorrect, because this linguistically defined residue, as we have seen, does not suffice to identify anything that may be called a lexical concept.

In fact, modelling the conceptual content of lexical items as complex structures of sub-word elements, more or less directly related to linguistic structures, is not at all as obvious as linguistic

tradition makes it appear. Laurence & Margolis (1999) identify a number of difficulties faced by theories of lexical concepts based on lexical decomposition:

- decomposition into primitives faces a regress problem: what do primitives mean?
- if lexical meaning was analyzable into constituent parts and their relations, we would expect definitions to accurately describe word meaning: but this typically fails
- proficient speakers often don't know important aspects (supposedly constitutive) of the meaning of words

To cite just two empirical problems, consider the difficulty of specifying what properties single out the concept DOG from the broader concept of ANIMAL, apart from the property of being a dog; or consider that even kinship terms like *grandmother*, which apparently consist entirely in a relational specification between humans, display prototype effects, so that some grandmothers are more typical than others, suggesting that the concept GRANDMOTHER does not in fact reduce to the relation it may be defined as. As Laurence & Margolis note (1999: 55-56), linguists disregard these problems, and represent lexical concepts as complexes of semantic primitives (like CAUSE, GO, BECOME, or categories like *Thing* or *Event*), mainly “because they aren't interested in giving a theory of concepts per se. They are interested, instead, in grammatically relevant aspects of word meaning”. They cite in this respect Grimshaw (1993, also cited in Jackendoff 2002: 338): “Linguistically speaking, pairs like [*break* and *shatter*] are synonyms, because they have the same structure. The differences between them are not visible to the language”. Such a reduction, while coherent, would in effect mean that almost all lexical semantic facts are not visible to the language (homonymy, hyperonymy, selection of ‘senses’ ...); not just the difference between *break* and *shatter*, but also between *dog* and *cat* (in Swahili, *day* and *night*), or *dog* and *animal*.

Among the alternatives to the ‘classical’ view of concepts as consisting of structured representations, Jerry Fodor's thesis of Conceptual Atomism takes the opposite stance and claims that basic lexical concepts have no internal structure. This does not mean that all concepts are atomic, for instance not a concept of RED SQUARE expressed by the corresponding phrase; the claim is rather that *simplex lexical* concepts have no decomposition into parts, features, or components. A key question, which remains unanswered in Fodor's work and more generally in all work outside linguistics, is what it means for a lexical concept to be simplex. We can get an idea of what this means, however, on the basis of Fodor's (1998: 121, 122, note 3) qualified statement of Conceptual Atomism as holding that “Most lexical concepts have no internal structure”, and especially of the following remark: “actually, of course, DOORKNOB isn't a very good example, since it's plausibly a compound composed of the constituent concepts DOOR and KNOB”. This approach, then, leaves room for the type of multi-concept word illustrated by (transparent) compounds, but otherwise holds that the conceptual content of lexical words is atomic.

Unsurprisingly, this is generally seen as the opposite of what lexical decomposition claims, in particular grammatical decomposition, like that which posits distinct predicates CAUSE and DIE in the linguistic representation of *kill*. I suggest instead that Conceptual Atomism is right, and that it is compatible with lexical decomposition (specifically, syntactic), once we reject the assumption that, at word level, linguistic structure determines conceptual structure.

Lieber (2004: 5) captured with lucidity the importance of a linguistically justified characterization of lexical items in order to choose how to represent conceptual content:

Fodor is right to question the nature of primitives. But in doing so, he declares that we have no grounds for preferring one set of primitives to another, and that the default set of primitives is “the lexicon of English”, that is, the set of words of which the lexicon is constituted. But surely we must consider carefully what constitutes the lexicon — what its parts are, what makes up words — before we decide that the *word* is the correct grain size for conceptual primitives.

Of course Conceptual Atomism does not elect English as the universal language of thought; but aside from this polemical barb, it is hard to disagree with Lieber’s point that a theory of lexical concepts cannot ignore a theory of lexical items, but must rather relate in a principled way to a theory of the morphological and semantic constituency of lexical knowledge. From this perspective, the evidence that roots do not encapsulate lexical concepts means that the pieces of grammatical structure are not also pieces of the conceptual structure. We can then claim that words are complex, but the concepts associated with them are simplex. There may be, perhaps, an abstract causative verb in the representation of *kill*, but it does not correspond to a concept; only *kill* as a whole does.

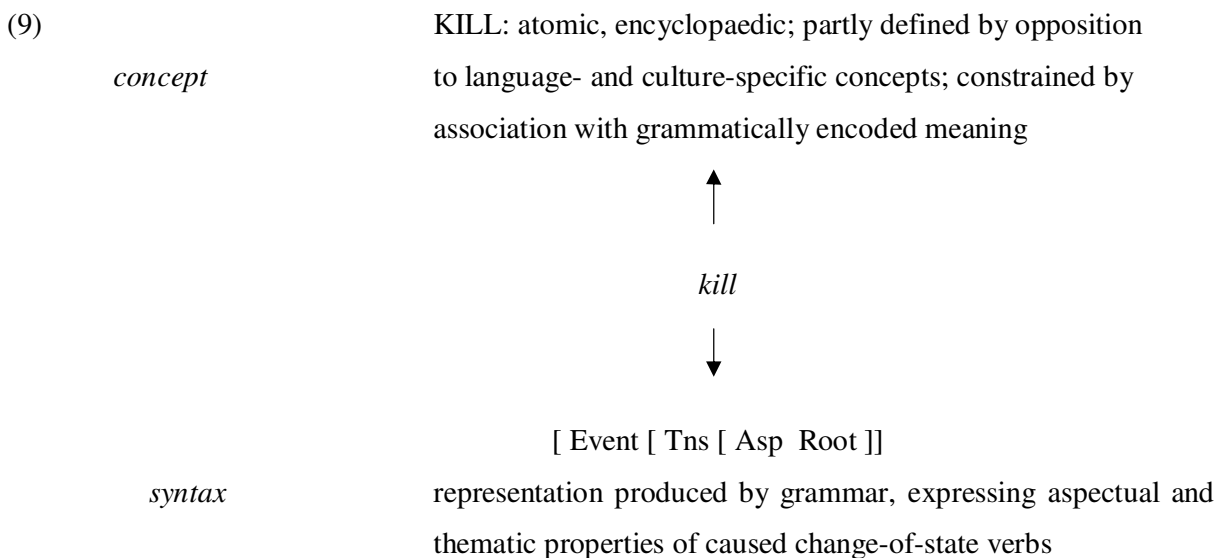
5.2 *Conceptual atomism and the grammatical decomposition of lexical meaning*

It is important to distinguish concepts from properties. To identify a concept with a structured bundle of properties is only one analytic option, and it has several empirical shortcomings. On the other hand, we can think that inferences like *bachelor* → *unmarried* are not constitutive of concepts, but follow from beliefs associated with them. BACHELOR does not include UNMARRIED as an integral part, since concepts have no parts; inferences of this kind are not a consequence of structural inclusion, neither between concepts nor (and this is the key point) between grammatical representations expressing them. As often noted, this accounts for the fact that even the inferences which seem most deeply constitutive of concepts can be defeated: the pope

doesn't seem to be a bachelor, a robot cat is a cat but not an animal, and a plastic flower does not instantiate a natural kind (Katz & Pitt 2000).

In addition, inferences can be *necessary* though not *constitutive*: even if *water contains hydrogen* is necessary, it is possible to have the concept WATER without having the concept HYDROGEN, as Fodor argues (1998: 74): “It’s perfectly consistent to claim that concepts are individuated by the properties they denote, and that *the properties* are individuated by their necessary relations to one another, but to deny that knowing about the necessary relations between the properties is a condition for having the concept”.

In sum, if lexical roots do not encapsulate conceptual content we can reject the assumption that the (syntactic) pieces of lexical structure are also pieces of the conceptual structure; in turn, this makes it possible to accept both Conceptual Atomism and syntactic decomposition, which would be desirable, since the former captures fundamental characters of lexical concepts, and the latter captures fundamental characters of lexical competence. To achieve that goal, we must explain what it means to say that the elements in the grammatical lexical representation have semantic content, but this content is not a ‘part’ of the overall lexical concept. The intuition I would like to pursue, with no claims to originality, is that these grammar-internal elements provide a semantic template, and a (grammar-external) lexical concept cannot conflict with the interpretation fixed by this template if it is to be expressed by the corresponding grammatical structure. For an example, the lexical item *kill* identifies a concept KILL, which is linguistically embodied as a formal representation constructed by the principles of grammar, and incorporating those elements and relations that define a caused change-of-state verb.



Grammatically encoded meaning is maximally abstract, culture-independent, and possibly organized in featural oppositions. As examples of what kind of grammatical elements may plausibly enter into the grammatical determination of specific lexemes, we may mention GO as a general motion verb (distinct from the lexical item *go*), perhaps a determinant of activity DO, and classifier-like elements which entails an individualized reading for a nominal. To give an idea of what this approach looks like in practice, it is precisely on nouns that I will focus on, as a case study in the relation between conceptual content and linguistic representation.

6. *Nouns and concepts*

6.1 *The conceptual bases of nominality*

As is natural for an approach which foregrounds the structural aspects of word meaning, work on grammatical lexical decomposition deals mostly with verbs and with deverbal nominalizations like *destruction*, concentrating on the representation of argument- and event structure. Consequently, not much can be said in this perspective about lexical items like *dog*, which lack an articulated argument- and event structure. Yet such basic terms are fundamental *explananda* in order to understand the place of conceptual content in the representation of words.

The situation is different in a syntactic approach that defines lexical categories constructionally, because of the hypothesis that even simplex nouns like *dog* are grammatical constructs. A statement ‘*dog* maps to DOG’ is insufficient, because the symbol *dog* conflates several lexeme-defining properties, like being a noun (in particular a common noun), being count, and denoting bounded entities extended in space and persisting through time, all of which taken together largely determine the type of concept notated DOG. Especially in Borer’s (2005a,b) radical version of this programme, which I follow here, the task lies precisely in teasing apart the components of lexical meaning that are and are not linguistically relevant, on the operational assumption that the former are represented through lexeme-independent grammatical means. Instead of concentrating on the grammatically encoded component, here I would like to discuss the conceptual component of nouns taking it in its own right, and not just as a residue of linguistic analysis.

To sum up proposals elaborated elsewhere (Acquaviva 2009, forthcoming), we should first distinguish the nominal style of reference, which is a property of DP, from the type of content associated with nouns as a lexical category. Quantification, anaphora, deixis, arise at the DP level. Nouns, as a lexical category, are the category for *naming*: they allow reference (via DP) to something viewed as an entity. In their core use, nouns name kinds, and the resulting nomenclature

constitutes a speaker's conceptual ontology. There are important parallels between this view and the conception of the philosopher Ruth Millikan, who groups together individuals (*mama*, *the Empire State Building*), substances (*water*), and natural kinds (*mouse*, but also *McDonald's* or *the number 3 bus*) as basic *substance concepts*, which speakers recognize as categories of entities which remain stable in time and guide inductive learning (Millikan 2000). Nouns, then, name the concepts which define entities; verbs have instead a different set of primitives, based on the notions of change and persistence through time.

This view of the basic conceptual function of nouns contrasts with attempts to define the category by reference to ontological categories like 'Thing', 'Event', 'Instrument', 'Property', 'Place' (Levin 2009, Jackendoff 1999), where lexical categories prototypically express one of more of these sorts. Apart from the fact that nouns quite often denote events (cf. *event*, *war*), or events, places, and things all at the same time (*end*; cf. Jackendoff 1999: 322), such a direct link between sorts and lexical categories is very dubious, if only because the labels chosen generally do not partition the ontological or conceptual space appropriately (in particular, it is unclear how an instrument or a place is not also a thing).

Naming substance concepts is the core and distinctive function of nouns, which sets them apart from other word classes; but it cannot *define* them, because not all nouns do so. Apart from non-nouns embedded in nominal syntax (*the why's*, *das Ich*), there are cases like *contents*, *beginnings*, *heights*, which denote entities sharing a contingent property without naming what kind they instantiate (Wierzbicka 1988, Acquaviva 2008); more importantly, transparent nominalizations like *destruction* or *collection* (in the eventive reading) also belong to this class, since their nominal morphology and syntax merely allows speakers to refer as an entity to a structured concept entirely determined by the corresponding verb (cf. Borer 2003). Therefore, I distinguish between nominality as a grammatical property, which encompass concept nouns as well as transparent nominalizations, from the distinctively nominal conceptual property of identifying kind-concepts. The latter are understood, along Millikan's guidelines, as reified categories of entities (of type <*e*>) which make up speakers' ontology at the kind-level, and underlie object-level reference in all cases where entities are categorized (excluding, then, pure ostension and cases like *contents* or *that "oh"*).

When construed as abstract categories, these hypothesized kind-level entities are all atoms; *water* is as atomic as *drop* at this level, and both are as abstract as *love*, as countability properties and object-level reference both arise in a grammatical context. This follows Borer's (2005a) view that it is not lexical items in isolation that determine countability, but grammar, and that a mass reading arises by default when no stable granularity is specified. The nominal structure is what turns these kind-names into predicates, and in so doing specifies the mereology of their extensions, that is, the granularity of their domain of reference, by specifying what entities a variable ranges

over in order to be true of the predicate. Even relational nouns like *sister* can be modelled at the same time as atomic kind concepts and as two-place relations; as the former, they name a concept associated with world-knowledge properties, which give rise to prototype effects (two biological sisters may conform to prototype sisterhood to different degrees); as the latter, such nouns are true of pairs of individuals standing in a certain biological relation. As for the plausibility of teasing apart grammatically constructed and conceptually given lexical meaning, where countability is a property of the former only, consider that the two are simultaneously accessible as sharply distinct readings in expressions like *there's too much Falstaff in this play*, where a name expresses a conceptually atomic individual while the DP is true of a mass, whose composition (Falstaff-events, -stages) is left vague. Essentially the same analysis applies to *there's too much water*, by taking *water* to name an atomic kind-level individual much as *Falstaff* does, with the difference that *water* and other common nouns receive this interpretation routinely (not in their kind reading, as in *the formula of water*), while *Falstaff* generally, if not here, denotes an individual both as a kind and as an individual object.

6.2 Constraining nominal semantic space

Crucially, this proposal allows us to make verifiable hypotheses about the conceptual content which can and cannot be associated with a simplex noun. To see how, consider first Fodor's (1998:164-165) argument that there are constraints on the content of an atomic concept. Given the atomic concepts notated RED and SQUARE, we can envisage a non-atomic concept RED + SQUARE (this would be the content of a phrase *red square*). Now, we can coherently think of acquiring the concept of a red square as a whole, without isolating its being red and its being square as distinct properties; for instance, by ostension to red square objects. However, countenancing both RED+SQUARE and REDSQUARE does not commit us to accepting just any concept formed in the same way. In particular, Fodor argues, not one made up of the contradictory ROUND and SQUARE: because in this case, there can be no object instantiating the supposed atomic concept ROUNDSQUARE, as a matter of logical necessity. Unlike, for instance, concepts for fictional entities, like UNICORN, a putative ROUNDSQUARE would not identify any entity in any possible world. A predicate can obviously have an empty denotation in all possible worlds, if it is made up of two contradictory properties; but a concept whose instantiations cannot even be imagined in alternative worlds, and of which no property would be true, seems unthinkable. In Fodor's words, "there can be no *primitive* concept without a corresponding property for it to lock to". In sum, one can have REDSQUARE as a primitive concept, but not the necessarily uninstantiated ROUNDSQUARE.

This argument concerns concepts, not linguistic entities; but in effect it constrains the possible content of nouns, because the notion of primitive, non-complex concept is understood in linguistic terms as a lexical concept, rather than one associated with a phrase or with a compound (cf. the comments about DOORKNOB in 5.1 above). This means that ROUNDSQUARE cannot be the content of a simplex noun, as opposed to a compound whose distinct parts are semantically accessible. In a similar vein, we may hypothesize that no simplex noun encapsulates the concept of NOT-BLUE THING, because the property of not being blue seems too uninformative to categorize a type of entities in such a way to support inductive learning and reidentification.³ What is more, we may be in a position to rule out a simplex noun meaning ‘undetached rabbit part’ (Quine 1960: 28, 52), which would apply felicitously to all and only the situations where the noun *rabbit* does, without having the same content. If kinds are conceptualized as entities which guide and support inductive learning (see 6.1), then we would have here a substance concept that could not be learned empirically as distinct from RABBIT. Moreover, the notion of an arbitrary part of an entity is as vague and functionally unmotivated as that of a thing that is not blue; in this it differs from concepts that identify *one* particular part of an entity (like *hood* as part of a car), identifiable from the rest of the entity.

7. *Concepts, morphemes and syntactic structure*

7.1 *Noun-specific features and the structure of DP*

Keeping the focus on nouns, let us now turn to the role of linguistic structure in shaping conceptual content. The constructional approach here followed locates on syntactic heads several components of ‘lexical’ meaning which lexicalist approaches view instead as determinations of listed lexical items. For nouns, two notions take centre stage in this connection: individuation and classification. These pretheoretical terms refer, respectively, to the degree to which a noun singles out a referent as an identifiable discrete whole, and to the ability to conceptualize referents as belonging to distinct types through linguistic paradigmatic oppositions, typically expressed by gender or noun class. On this point, Picallo (2005), De Belder (2008), and Alexiadou (2011) present arguments for a distinct head hosting lexeme-inherent features, located between the lexical root and the head hosting number features. This concerns not only gender or noun class, but also number, where the choice of a value determines a distinct reading for the lexical predicate, as in the mass reading of *brains* as “brain matter, intelligence” or in the spatially extended interpretation of plurals of abundance like *waters* or *snows*, or in the time-extended eventive reading exemplified by *rains*. In particular, an

‘inner’ encoding for number when it affects the conceptual value of the noun seems plausible when it is linked to a value for gender or noun class, as in the Italian plurals in *-a* which combine feminine gender, an irregular ending, and an often idiosyncratic reading: *membro* “member” (masc. sg.), regular plural *membri*, irregular *membra* “limbs” (fem. pl.) (see Acquaviva 2008 and, for a more detailed analysis, Kramer 2009).

We can then divide the DP structure into three functionally distinct areas: the outermost projections, at DP level, express a discourse referent via anaphora, quantification, or deixis; below this region there are grammatical formatives that define the granularity of the noun’s domain of reference, and determine the grammatical properties related to the count-mass distinction; in innermost position, but still as part of syntactic structure and not a mere determination of a listed root, grammar expresses information that identify a noun as a lexeme.

(10) [Discourse reference [Division of reference [Categorization ROOT]]]]

More precisely, I will assume the following structure, based particularly on De Belder’s (2008) analysis of affixes with individualizing value as realizations of a [Size] node inside [Division], which following Borer (2005a,b) is the locus of number marking; in a similar fashion, Déchaine et al. (2012) posit a [Sort] node inside the expression of number.

(11)	[Determiner	[Quantity	[Division	[Size ROOT]]]]
			inflectional plural	lexical plural
			mensural classifiers	sortal classifiers
				gender
				singulative affixes

7.2 Individuation

A syntactic decomposition of the type just outlined proposes an answer to Lieber’s question as to “what constitutes the lexicon — what its parts are, what makes up words”. In the case at hand, it proposes the functions of reference division and categorization as building blocks for nominal lexical semantics, structured over an articulated syntactic representation. It is not morphemes per se which determine an interpretation, but morphemes as realizations of a syntactic head with a certain value. This has two consequences: first, when there are reasons to associate several morphemes to the same head, a range of shared semantic values is predicted, rather than having to be posited separately for all morphemes; second, the same morpheme may in principle lexicalize different

heads, displaying a semantic variability whose boundaries can again be predicted, from the general values associated with the heads in question. To conclude, we will now consider how this analytical perspective contributes to a deeper understanding of the nominal semantic function loosely labelled ‘individuation’, which plays a central role in forming entity concepts.

In her study of the fundamental semantic categories of English derivational morphology, Lieber (2004: 37) identifies a common value for the suffixes *-ee*, *-er*, *-ist*, *-ent*, namely bringing about an interpretation of the referent as a concrete entity related to a function (in her feature system, [+material, +dynamic]). However, nothing in the characterization as ‘concrete entity’ makes it expected that these derived nouns are systematically count. In principle, such nominalizers may form nouns denoting concrete substances like fluids; instead, mass nouns are typically derived by other suffixes like *-age*, *-ery*, *-ure*, which Lieber rightly ascribes to a different semantic class (bounded collectives), as they systematically double up to name abstract notions and locations (Lieber 2004: 41-43, 148-150).⁴ An alternative analysis consists in treating these nominalizers as dividers of reference that express individual entities with a stable granularity; essentially, as classifier-like bound morphemes constructing singulative nouns, lexicalizing the inner division head labelled ‘Size’ in (11) (below the higher division node and therefore capable of being pluralized). This characterization is more precise than ‘concrete entity’, foregrounding the conceptualization of the referents as bounded and discrete, and captures the important fact that such entities may be extended in space but also in time, witness eventive formations like *all-nighter*.

Borer’s (2005a) analysis of the mass reading as deriving from a *lack* of stable partition provides a plausible reason for the rarity of mass nominalizers, since the lack of division seems an unlikely function for a dedicated morpheme. Massifiers exist, but they generally contribute additional information, like the Serbo-Croatian *-ina* (Carstairs-McCarthy 1992: 185), which suffixed to a noun denoting an animal denotes a substance obtained from that animal: *svinjat* ‘pig’, *slon* ‘elephant’ – *svinj-et-ina* ‘pork’, *slon-ov-ina* ‘ivory’ (not for instance ‘masslike collection of pigs / elephants’, like *cattle*). In this respect they differ from plural when used as a generic massifier, as in *brain* – *brains*, *fund* – *funds*, *work* – *works*.

The Italian suffix *-ista* finally, illustrates a slightly different scenario: this morpheme too can have the value of a classifier-like individualizer, but it also occurs as a generic nominalizer with a broader distribution. To summarize the detailed account of Lo Duca (2004: 206-208) and Rainer (2004a: 14, 2004b: 256-258), words derived by this morpheme give rise to two constellations of properties, exemplified by *socialista* ‘socialist’ and *linguista* ‘linguist’. The two display different semantic properties, which Lo Duca traces back to the opposition between ‘characterizing’ and ‘classifying’ agent formations (Grossmann 1998), where both characterize human individuals as habitual agents in an activity, but only the latter uses this to classify individuals by their role in

society. Strikingly, these interpretive differences correlate with the ability to occur as noun alone or as a noun or adjective. Even exceptions to the pattern $N \rightarrow Nista$ are not randomly distributed, since truncated bases (smaller than N) are a subset of noun-only derivatives. The two following subtypes emerge:

(12) subtype 1

- a possible use as adjectives, often but not necessarily relational
l'invasione fascista “the Fascist invasion”, *un atteggiamento fascista* “a Fascist attitude”
- b semantic derivation either from a base *X* or from *X-ismo*
camorra - *camorrista* “C. - member of / belonging to the C.” (not from **camorristismo*)
socialismo - *socialista* “Socialism - Socialist” (not from *sociale* “social”)⁵

(13) subtype 2

- a no use as adjectives; reference to human individuals only
*un'osservazione *linguista* “a linguistic observation”
- b some truncated bases
panegir-ic-o “eulogy” → *panegir-ista*; *fonet-ic-o* “phonetic” → *fonet-ista* “phonetician”
- c some opaque bases
base “basis, base” → *bas-ista* “criminal mastermind”
- d productive derivations N [instrument] → N [instrument operator], with a connotation of modern-day specialist
softwarista, *autostoppista*, ...

Why do these properties cluster precisely as they do? The hypothesis of a syntactic encoding for the individual reading offers a possible answer. Suppose that *-ista* is a unitary affix expressing morphosyntactic nominal features, which are in principle compatible both with the distribution of a noun and with that of a noun-modifying adjective. On its own, *-ista* simply derives a nominal word denoting a property, which is ‘dynamic’ in the sense that it applies to dispositions, attitudes, or activities, and this involves reference to sentient agents even without being directly predicated of them (*un atteggiamento pressapochista* “a sloppy attitude”). As noted, the property is derivative on a concept, but this semantic base is not necessarily expressed by the morphological base *X* in *Xista*: in particular, there is a systematic subregularity linking abstract nouns of the form *Xismo* to corresponding forms *Xista*, which express the property of having (following, supporting ...) *Xismo* (not just *X*, cf. again *socialista*).

The subtype in (13) arises as a special case of this generic characterization, when *-ista* lexicalizes the inner division node with the function of an individualizing noun-forming classifier. In this capacity, *-ista* forms nouns denoting human individuals (not just properties that relate to human agency), identified as doers of an activity; typically, but not necessarily, an activity made conceptually salient by the base. Nouns for entities conceptualized as instruments, like *liuto* “lute”, form *-ista* derivations denoting the agents of the activity identified by that instrument: *liutista* “lute player”, *scambista* “exchange operator”, but *arabista* “Arabic scholar”, and not “speaker”, where the language name identifies a field of professional activity, because languages are not conceptualized as instruments (cf. Rainer 2004: 14). This explains why, among the nouns in *-ismo* denoting medical conditions, a counterpart in *-ista* exists only for *tabagismo* → *tabagista* “tobacco user” and *alcolismo*, *etilismo* → *alcolista*, *etilista* “alcoholic”, where the condition results in an activity (contrast *rachitismo* → **rachitista* “affected by rachitis/rickets”). Morphologically, the suffix may attach to already derived words, but also to a bare root, giving the impression that the base has been truncated in cases like *fonet-ista*.

Returning to the general issue of word structure and conceptualization, it should not come as a surprise that the suffix *-ista* in subtype 2 contributes its own encyclopaedic content, namely the modern specialist connotation, when it lexicalizes a head expressing an individualizing nominalizer. This is not a property of the suffix in isolation, or of the syntactic structure alone, but of the suffix realizing a particular structure. The conclusion generalizes: all words are complex, as syntactic constructions if not also morphologically, and their conceptual content is determined globally by their linguistic structure and by the specific properties of the elements that interpret it.

8. Conclusion

The question of how lexical semantics relates to word structure is most naturally approached from a certain view about word structure. I have considered in some detail the implications of a model that represents word structure as a syntactic construction, as a convenient analytical framework to shed light on the role of grammatical and extragrammatical knowledge in lexical competence. The main goal has not been to defend this approach, but to reach a view of the constitutive elements of lexical meaning whose validity could be verified empirically, especially in connection with lexical roots. In the analysis here proposed, the grammatical meaning expressed in the structure making up a word creates an interpretation space, which constrains the conceptual content associated with the word; importantly, this content is a property of the whole lexeme, not of a syntactic or morphological element. This conclusion, which was reached on linguistic grounds, supports a view of lexical

concepts as unstructured atoms, as it decouples the structure of lexical items as grammar-internal linguistic objects from the internal structure of lexical concepts. Word meaning is cognitively complex, but not as a reflex of linguistic complexity.

References

- Acquaviva, Paolo. 2008. *Lexical Plurals*. Oxford: Oxford University Press.
- Acquaviva, Paolo. 2009. "Roots, categories, and nominal concepts". *Lingue e linguaggio* 8: 25-51.
- Acquaviva, Paolo. Forthcoming. The roots of nominality, the nominality of roots. *The Syntax of Roots and the Roots of Syntax*. ed. by Artemis Alexiadou et al. Oxford: Oxford University Press.
- Alexiadou, Artemis. 2011. "Plural Mass Nouns and the Morpho-syntax of Number". *Proceedings of the 28th West Coast Conference on Formal Linguistics* ed. by Mary Byram Washburn et al., 33-41. Somerville, Mass.: Cascadilla Proceedings Project. www.lingref.com, document #2433.
- Arad, Maya. 2005. *Roots and Patterns: Hebrew Morpho-Syntax*. Berlin: Springer.
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.
- Aronoff, Mark. 1994. *Morphology by Itself*. Cambridge, Mass.: MIT Press.
- Aronoff, Mark. 2007. "In the beginning was the word". *Language* 83: 803-830.
- Basilico, David. 2008. "Particle verbs and benefactive double objects in English: High and low attachments". *Natural Language and Linguistic Theory* 26: 731-729.
- Borer, Hagit. 2003. "Exo-skeletal vs. endo-skeletal explanations: syntactic projections and the lexicon". *The Nature of Explanation in Linguistic Theory* ed. by J. Moore & M. Polinsky. Stanford: CSLI.
- Borer, Hagit. 2005a. *In Name Only*. Oxford: Oxford University Press.
- Borer, Hagit. 2005b. *The Normal Course of Events*. Oxford: Oxford University Press.
- Carstairs-McCarthy, Andrew, 1992. *Current Morphology*. London: Routledge.
- DeBelder, Marijke. 2008. "Size Matters. Towards a Syntactic Decomposition of Countability". *Proceedings of the 27th West Coast Conference on Formal Linguistics* ed. by Natasha Abner & Jason Bishop, 116-122. Somerville, Mass.: Cascadilla Proceedings Project.
- Déchaine, Rose-Marie, Raphaël Girard, Calisto Mudzingwa, Martina Wiltschko. 2012. "The Internal Syntax of Shona Class Prefixes". Paper presented at the 15th International Morphology Meeting, Vienna, February 2012.
- Embick, David & Alec Marantz. 2008. "Architecture and Blocking". *Linguistic Inquiry* 39(1), 1-53.
- Fodor, Jerry. 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

- Fradin, Bernard & Françoise Kerleroux. 2009. "L'identité lexémique". *Aperçus de morphologie du français* ed. by Bernard Fradin, Kerleroux, & Marc Plénat. 83-102. Paris: Presses universitaires de Vincennes.
- Grossmann, Maria. 1998. "Formazione dei nomi di agente, strumento e luogo in catalano". *Atti del XXI congresso internazionale di linguistica e filologia romanza*, vol. 2, ed. by G. Ruffino, 383-392. Tübingen: Niemeyer.
- Halle, Morris & Alec Marantz. 1993. "Distributed Morphology and the pieces of inflection". *The View from Building 20* ed. by Ken Hale & Samuel J. Keyser, 111-176. Cambridge, Mass.: MIT Press.
- Hale, Ken & S.J. Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge, Mass.: MIT Press.
- Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 1999. "What is a concept, that a person may grasp it?" *Concepts: Core readings* ed. by Stephen Laurence & Eric Margolis, 305-333. Cambridge, MA: MIT Press. Originally published 1989, *Mind and Language*.
- Jackendoff, Ray. 2002. *Foundations of Language*. Oxford: Oxford University Press.
- Katz, Jerrold, & David Pitt. 2000. "Compositional idioms". *Language* 76:409-432.
- Koontz-Garboden, Andrew. 2011. The roots of change of state verbs. Paper presented at the workshop Approaches to the Lexicon (Roots III), Jerusalem, June 2011. Available at <http://personalpages.manchester.ac.uk/staff/andrewkg/cos-roots-jerusalem-final.pdf>
- Kramer, Ruth. 2009. Definite markers, phi-features, and agreement: a morphosyntactic investigation of the Amharic DP. PhD dissertation, University of California, Santa Cruz.
- Laurence, Stephen & Eric Margolis. 1999. "Concepts and cognitive science". *Concepts: Core readings* ed. by Stephen Laurence & Eric Margolis, 3-81. Cambridge, MA: MIT Press.
- Levin, Beth. 2009. "Lexical semantics of verbs II: The structure of event structure". Lecture handout. Available at <http://www.stanford.edu/~bclewin/lsa09event.pdf>
- Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity*. Cambridge, Mass.: MIT Press.
- Levin, Beth & Malka Rappaport Hovav. 2005. *Argument Realization*. Cambridge: Cambridge University Press.
- Lieber, Rochelle. 2004. *Morphology and Lexical Semantics*, Cambridge: Cambridge University Press.
- Lo Duca, Maria G. 2004. "Nomi di agente". *La formazione delle parole in italiano* ed. by Maria Grossmann & Franz Rainer, 191-218. Tübingen: Niemeyer.
- Millikan, Ruth. 2000. *On Clear and Confused Ideas*. Cambridge: Cambridge University Press.

- Panagiotidis, Phoevos. 2011. "Root 'content' and lexical semantics". Paper given at the workshop Approaches to the Lexicon (Roots III), Jerusalem, June 2011.
- Picallo, Carme. 2005. "Some notes on grammatical gender and *l*-pronouns". *Konstanzer Arbeitspapiere zur Sprachwissenschaft* 119, ed. by Klaus von Heusinger, Georg Kaiser & Elisabeth Stark, 107-121.
- Pietroski, Paul. 2008. "Minimalist meaning, internalist interpretation". *Biolinguistics* 2: 317-340
Available at <http://www.biolinguistics.eu/index.php/biolinguistics/article/view/70/87>.
- Polomé, Edgar. 1967. *Swahili Language Handbook*. Washington, D.C.: Center for Applied Linguistics.
- Quine, William van Orman. 1960. *Word and Object*. Cambridge, Mass.: MIT Press.
- Rainer, Franz. 2004a. "Il significato delle regole di formazione di parole". *La formazione delle parole in italiano* ed. by Maria Grossmann & Franz Rainer, 13-16. Tübingen: Niemeyer.
- Rainer, Franz. 2004b. "I suffissi *-ismo* e *-esimo*". *La formazione delle parole in italiano* ed. by Maria Grossmann & Franz Rainer, 256-260. Tübingen: Niemeyer.
- Wierzbicka, Anna. 1988. *The semantics of grammar*. Amsterdam: John Benjamins

¹ Borer (2005b: 354) takes idioms to map a listeme to a syntactic head, possibly with a pre-fixed value; however, this local exception to the general pattern whereby listemes do not determine syntax does not weaken the key claim that all linguistically visible properties of a word are ultimately syntactically encoded.

² If the eventive reading of the noun does not correspond to a verbal syntactic structure, then the root is purely nominal throughout; but the main point stands, and it is that the same root determines sharply distinct 'lexical' properties.

³ It may be objected that nouns like *thing*, *object*, *entity* have a denotation that is even less restricted than 'not-blue thing'. But these and similar nouns have a function as expressions of maximal generality, while 'not-blue thing' restricts the denotation in an arbitrary and functionally unmotivated way.

⁴ An exception is represented by the mass-noun-forming *-ware*, which however Lieber treats as a 'compounding stem' rather than a suffix proper, with good reason.

⁵ Although *-ista* forms which admit adjectival reading have a privileged relation with bases in *-ismo*, not all *Xista* based on *Xismo* belong to this second class; *alpinista* "mountaineer" or *turista* "tourist", for instance, have all the properties of the first.