

LEXICAL DECOMPOSITION MEETS CONCEPTUAL ATOMISM

PAOLO ACQUAVIVA
PHOEVS PANAGIOTIDIS

ABSTRACT: Asking what can be a substantive word in natural language is closely related to asking what can be a lexical concept. However, studies on lexical concepts in cognitive psychology and philosophy and studies on the constitution of lexical items in linguistics have little contact with each other. We argue that linguistic analyses of lexical items as grammatical structures do not map naturally to plausible models of the corresponding concepts. In particular, roots cannot encapsulate the conceptual content of a lexical item. Instead, we delineate a notion of syntactic root, distinct from that of morphological root; syntactic roots are name-tags establishing lexical identity for grammatical structures. This makes it possible to view basic lexical items as mappings between syntactically complex structures, identified by their root, with simplex concepts, where the constructional meaning of the former constrains the content of the latter. This can lead to predictive hypotheses about the possible content of lexical items in natural language.

KEYWORDS: concepts, roots, lexical decomposition

1. CONCEPTS AND WORD STRUCTURE*

Taking DOG to represent the concept associated with the word *dog* seems a straightforward choice, but it presupposes a clear notion of what is a word. To see that there is an issue, and that the issue is linguistic in nature, it suffices to ask whether the two word forms present in *put up* map to one or to two concepts; or whether *break* corresponds to a single concept BREAK

* We wish to thank the audience and organizers of the first NetWordS workshop. P. Acquaviva's research was supported by an Alexander von Humboldt fellowship, which is gratefully acknowledged. Faults and omissions are the authors' responsibility.

in all its uses, including idiomaticized ones like *break wind* (contrast *wind-breaker*, in the sense of a garment). Far from being a terminological quibble, this is a substantive issue about the discrimination between simple and complex concepts. If the notion of wordhood relevant for conceptualization is morphological, *break up* is not only a complex expression but also a complex concept, assembled out of two simple concepts like BROWN COW. If instead a simple concept corresponds to a ‘semantically simple’ word, then we must decide what counts as semantically simple—a task that seems identical to the task of deciding what counts as a concept, resulting in circularity: *a b* jointly form a simple concept because they are a concept.

Researchers in linguistics and in cognitive psychology are not generally overly worried about the linguistic bases of concept individuation. Beginning their overview of the extensive literature on concepts, Laurence and Margolis (1999:4) note that ‘For a variety of reasons, most discussions of concepts have centered around *lexical concepts*. Lexical concepts are concepts like BACHELOR, BIRD, and BITE—roughly, ones that correspond to lexical items in natural languages.’ Correspondingly, Fodor (1998:122) makes it clear that consisting of a single word is indeed crucial in defining lexical concepts: ‘actually, of course, DOORKNOB isn’t a very good example, since it’s plausibly a compound composed of the constituent concepts DOOR and KNOB.’ Expressions like ‘concepts (/lexical meanings)’ (Fodor 2008:26) are symptomatic of this perspective. Laurence and Margolis (1999:4) acknowledge that defining simple concepts as those associated to lexical words is not straightforward, and add that ‘the concepts in question are ones that are usually encoded by single morphemes. In particular, we don’t worry about the possibility the one language may use a phrase where another uses a word, and we won’t worry about what exactly a word is’. Still, being monomorphemic is not the same as being semantically basic; consider morphologically but not semantically complex lexemes like *con-ceive*, or lexicalized compounds like *home run*, or indeed *cranberry*.

On the linguistic side, research into the primitives and the constituent structure of lexical meaning represents a long and richly diverse tradition of studies, but typically without much dialogue with psychological research into the representation of concepts. Most analyses decompose the content of lexical words into representations differing in the primitives and in the type of structure envisaged (cf. Pinker 1989, Jackendoff 1990, Pustejovsky 1995, among many others). Some approaches distinguish separate structural representations in the meaning of a lexical item, like Jackendoff’s (2002:334-339) Conceptual Structure and Spatial Structure, or Lieber’s (2004) encyclopaedic ‘Body’ and semantically regimented ‘Skeleton’ (with function-argument structure). In contrast, Levin and Rappaport Hovav

(1995, 2005) envisage a single representation, expressing the argument- and event structure of a verb by means of primitive predicates (like BECOME) and constants/roots (like BREAK), forming a lexical semantic template:

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

Finally, among the approaches that decompose lexical meaning into a grammatical structure, a family of analyses explicitly take this structure to be generated according to the same principles that underlie sentence construction (Hale and Keyser 1993, 2002, Arad 2005, Halle and Marantz 1993, Embick and Marantz 2008, Borer 2005a,b, Ramchand 2008). We will focus on analyses of this type, questioning the way they deal with non-structural, idiosyncratic aspects of lexical meaning that are essential to a word's conceptual content but apparently lack any grammatical relevance.

As Laurence and Margolis (1999) note, representing the content of a lexical item as a structured arrangement of primitives results problematic for the view it presupposes of lexical concepts. In particular, the idea that linguistic word-internal structure may explain the content of lexical concepts and their mutual relations inherits the problems associated with 'classical' theories of concepts as structures articulated into smaller components:

- decomposition into semantic primitives faces a regress problem: what do primitives like CAUSE or THING mean, if they are not the same as the corresponding lexical words?
- if lexical meaning was analyzable into constituent parts and their relations, we would expect definitions reflecting the structural decomposition of a concept to accurately describe its content: but this typically fails, since word meaning systematically cannot be give a unique and precise definition or paraphrase;
- if lexical concepts were constituted of linguistic constructs, possession of these concepts would require being aware of their content; yet competent speakers often don't seem to know certain aspects of the meaning of the words they use, even supposedly constitutive ones;
- prototype effects, like the fact that a certain representation of grandmother exemplifies the concept better than others, are unexpected if the content of GRANDMOTHER consists in a hierarchical arrangement of semantic primitives, defining in this case a biological relation.

Such empirical issues do not seem to have had an impact on linguistic analyses of the structure of lexical items. In part this is due to a widespread perception that such matters do not concern what speakers know about

lexical items as *linguistic* representations; the content of lexical concepts certainly includes a fair deal of non-linguistic knowledge, but, it may be argued, this is irrelevant for an account of what speakers know when they know a word as a product of the language faculty. Grimshaw 1993 (cited in Laurence and Margolis 1999: 55, Jackendoff 2002: 338) has articulated this position in a particularly strong form: ‘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.’

Most syntactic decompositional analyses to lexical structure share this view, in practice if not in principle. This is a problem, however. If the semantic relations between concepts like DOG and CAT lie outside the purview of linguistic theory, as a theory of the computational capacity of the mind to represent linguistic knowledge in a way that explains the productivity and compositionality of linguistic expressions, then it is hard to see why the relation between DOG and ANIMAL should not be likewise ‘not visible to the language.’ And if such a canonical example of hyponymy falls outside the scope of linguistics, much of what speakers know about the relations between word meanings becomes inaccessible to linguistic explanation. Thus, the semantic deviance of comparisons like # *a dog is smaller than an animal*, involving two nouns in hyponymy relation, can only receive a non-linguistic explanation, or none at all. However, these facts are part and parcel of what speakers know about words and their combinatorial possibilities in sentences. Current syntactic approaches to lexical semantics are forced to ignore this empirical domain, which was an important part of earlier work in generative grammar (the example comes from Bever and Rosenbaum 1970). The result has been a near-exclusive attention to argument- and event structure in verbs, which is a significant limitation and blocks the way towards a linguistically informed theory of a possible word.

By contrast, we hold that a theory of UG should have something to say about the way lexical items relate to their conceptual content. We will take as our point of departure a specific syntactic approach which most clearly dissociates the grammatical components of a word from a non-grammatical core, and focus on the properties which can and cannot be attributed to this root element as a key locus for the relation between syntactic representation and conceptual content.

2. ROOTS IN LEXICAL DECOMPOSITION

Work in Distributed Morphology and Borer’s (2005a,b) Exoskeletal approach both envisage maximally underspecified root terminals embedded inside a number of syntactic shells, which collectively define syntactic

constructions that define lexical categories; a noun, adjective, or verb is for a construct, in whose innermost core lies a category-neutral root. There are many important differences between the two approaches, and indeed between the two conceptions of roots, the most apparent being that Distributed Morphology, but not Borer, mandates the presence of categorizing heads, [n], [v], or [a], immediately governing the root and categorizing it (with possible complications for complex roots). For present purposes, however, what counts is the role of the root in determining lexical semantic properties, understood as lexeme-related properties which remain constant across grammatical contexts. Both models assume that all roots are non-categorized, so even the unique categorial determination of monomorphemic words like *fun*, *tall*, or *idiot* is inferred from the context; categorial underspecification, however, does not directly imply that roots lack the kind of semantic information which makes a difference between a noun and a verb. Analyses within Distributed Morphology, when they address the topic, typically treat the root as a meaningful element, whose content selects a suitable syntactic context. Importantly, however, work in this framework stresses that a root's meaning is emergent in a context. In the most comprehensive treatment of the issue in this framework, Arad (2005) defends a view of roots as radically underspecified but still meaningful elements which give rise to distinct interpretations depending on their immediate context. More precisely, Arad distinguishes roots with a relatively stable and well-defined meaning, from a more theoretically interesting type of roots whose semantic content cannot be stated in isolation, but emerges as a cluster of conceptually related words, giving rise to what Arad calls Multiple Contextual Meaning. Roots of the first type tend to form one or very few words only (as Hebrew nouns for animal, plants, food, or kinship terms, like *kelev* 'dog', *sukar* 'sugar', *□ax* 'brother', *□axot* 'sister'); roots of the second type give rise to larger word families, with a more or less recognizable semantic relatedness which can be very faint indeed; for example, XŠB in *xašav* 'think', *xišev* 'calculate', *hexšiv* 'consider', (Arad 2005:82), or QLT in *miqlat* 'shelter', *maqlet* 'receiver', *qaletet* 'cassette', *qalat* 'absorb, receive' (Arad 2005:97). While roots of this second type do not define a lexical sense without a context, they are unambiguously qualified as semantically contentful signs.

In contrast, the category-free heads which correspond to roots in Borer's (2005a,b) framework lack any kind of grammatically legible information (with the exception of idioms; cf. Borer 2005b:354). In a framework that consigns to syntax *all* grammatically relevant information of lexical words, these elements are the non-grammatical residue, which appear as listed phonological forms, or 'listemes': 'By listemes we refer to a pairing

of a conceptual feature bundle with a phonological index' (Borer 2005b:25). Borer's listemes thus encapsulate the non-syntactic information which defines a lexical item. In different ways, then, Distributed Morphology and Borer's Exoskeletal model posit contentful root elements at the core of their syntactic decompositions of substantive lexical items, which determine lexeme-specific and encyclopaedic aspects of lexical semantics either by themselves or as a function of their context. Our claim, now, is that roots in a syntactic decomposition sense cannot have this sort of content.

3. ROOTS ARE NOT MEANINGFUL SIGNS

In this section we will review some empirical evidence that roots do not carry any meaning and/or semantic content that could be identifiable outside of a grammatical structure, not just because they need a local context to determine a specific interpretation, but more radically because they are not signs. In fact, the evidence suggests that any sort of lexical meaning is a property of roots embedded in a grammatical structure, which can be of a rich and complex nature. The conclusion that will emerge is that there is no such thing as non-structural meaning, even at the level of 'word'.

Let us begin with some remarkable cases. It is received wisdom within the Distributed Morphology research on the systematic idiomacity of the structure below the first categorizing shell (e.g. *nP* or *vP*) that the categorizer projection acts as a sort of limit, below which interpretation is / can be / must be non-compositional (Marantz 2000; see also Marantz 2006, where inner versus outer morphology phenomena are explained in this way). In this perspective, the opposition between event nominalization and result nominal of *collection* in (2) must be due to different grammatical structures corresponding to the two readings (see Borer 2003). But since the root is the same, neither the difference in syntactic structure nor that in ontological typing (event vs. object) can be even indirectly a function of the root:

- (2) *collection*₁ 'the frequent collection of mushrooms by Eric'
*collection*₂ 'let me show you my collection of stamps' (result nominal)

Still, it can be argued that the two structures, while different, share a semantic core because they only differ in terms of outer morphology, above the first categorizing shell. However, as discussed in Panagiotidis (2011), we can have radically different meanings *across* the first categorizing shell. A telling example is the one below from Greek:

- (3) a. [VoiceP *nom-iz-*] 'think'

- b. [_{nP} [_{VoiceP} *nom-iz-*] *ma*] ‘coin, currency’
- c. [_{aP} *ne-* [_{VoiceP} *nom-iz-*] *men-*] ‘legally prescribed’

A large number of words relating to law, regulations and the like is derived from the root *nom-*. However, when the root is verbalized, yielding the verbal stem *nom-iz-* in (i) above, the meaning assigned is ‘think, believe’. So far, this is just what Marantz (2000; 2006) and Arad (2005) predict, namely that roots are not assigned meaning until they are categorised.

See however what happens when we take the verbal stem, a *vP* by hypothesis, and nominalize it, using the run-of-the-mill nominalizer *-ma* in (3b). Unlike the explicit predictions in Arad (2005), and as Borer (2009) points out with similar examples, the already categorized element *nomiz-* does not keep its meaning. What happens instead is that the whole [_{nP} *n vP*] structure is (re-)assigned a new, unrelated and completely arbitrary meaning, that of ‘coin, currency’. Perhaps equally interestingly, the participle derived in (3c) from the selfsame verbal stem carries a meaning as if *nomiz-* meant ‘legislate, prescribe by law’. In other words, in (3c), the *vP* embedded within an adjectival shell also fails to keep its “fixed” meaning of ‘think, believe’ and the whole *aP* participle means ‘legally prescribed’.

The question raised by such and similar examples concerns the semantic malleability of roots. Assuming that they are very underspecified semantically, one might ask how underspecified they can be before they become semantically vacuous. The most obvious example is provided by Latinate roots like *-ceive*, *-mit*, or *-verse*, which in English underlie a variety of semantically unrelated lexemes like *con-ceive* and *re-ceive*, *ad-mit* and *per-mit*, *con-verse* and *per-verse*. Their likes can be found in a number of languages, like the Greek *esth-*:

- (4) *esth-an-o-me* ‘feel’
- esth-is-i* ‘sense’
- esth-i-ma* ‘feeling’, ‘love affair’, ‘boyfriend / girlfriend’
- esth-an-tik-os* ‘sensitive, emotional’
- esth-it-os* ‘perceptible’, ‘tangible’,
- an-esth-it-os* ‘unconscious’, ‘insensitive’
- esth-it-ir-ios* ‘sensory’
- esth-it-ik-os* ‘esthetic’, ‘beautician’

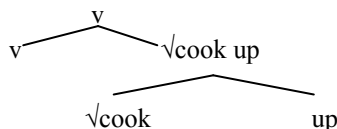
Despite the illusory affinities suggested by the Latinate English glosses (G. Longobardi, p.c.), the concepts of words derived from *esth-* is so broad that it is impossible associate the root *itself* with any cognitively coherent concept, no matter how underspecified, even to the exclusion of ‘beautician’.

The problem is not just that all too often a single root lacks a single identifiable content. In some cases there is evidence that the different interpretations are visible for grammatical processes. This happens when the same root yields interpretations of different ontological types (like (2) above), which differ for the purposes of further morphological derivations, *after* the root has been categorized, as in the following Greek example:

- (5) *paradosi*₁ ‘tradition’ (result / *process nominal)
*paradosi*₂ ‘delivery’, ‘surrender’ (result / process nominal)
paradosiakos relative to *paradosi*₁ (i.e. ‘traditional’), # *paradosi*₂

Even clearer examples where the same root under-determines lexical properties are the ones studied in Basilico (2008), where the same (atomic) root is compatible with different selectional restrictions, according to the grammatical environment in which it is embedded:

- (6) *the criminals cooked a meal* / #*an evil scheme* (Basilico 2008)
the criminals cooked up an evil scheme



This type of examples is particularly instructive, as it brings out an ambiguity in the notion of root: atomic element individuated morphologically (here, *cook*), or innermost category-free element, defined syntactically and possibly complex (here, *cook up*). This will play an important part on our discussion.

Finally, we can push further the empirical point that lexical meaning is not fixed within the first categorizing shell; in fact, we also find cases where the basic lexical predicate is determined only by the choice of inflectional morphemes, after a significant amount of structure has been built. Consider Russian, where the root *tsvet* in different noun declensions derives both the word for ‘colour’ and the word for ‘flower’:

(7)	SINGULAR		PLURAL
	<i>tsvet</i>	‘colour’	<i>tsvet-á</i> ‘colours’
	<i>tsvet-ók</i>	‘flower’	<i>tsvet-’y</i> ‘flowers’

Even though FLOWER is a basic-level concept, the noun lexicalizing this concept is derived in the singular by the addition of the diminutive suffix *-ok* with individualizing function. There are, to be sure, an archaic form *tsvet* with the meaning ‘flower, blossom’, and a regular plural *tsvet-kí* from *tsvet-ók*; but in so far as the paradigm in (7) reflects a stable synchronic pattern, it shows that what individuates the concept FLOWER is neither the root by itself (also appearing in *tsvestí* ‘to blossom’) nor, crucially, the root with a nominal suffix, which is absent in the plural, but the choice of one among two alternative inflectional classes, which emerge in the nominative / accusative plural. Further evidence that lexical meaning can be fully established at the inflectional level comes from the numerous idiosyncratic (specialized) interpretations for morphologically regular inflectional plurals (cf. Acquaviva 2008), like the English *brain* (count) - *brains* (count / mass), or the Cypriot Greek *nero* (‘water’), plural *nera* (‘heavy rain’).

4. TWO TYPES OF ROOTS

For Distributed Morphology, roots are syntactically active elements (but see De Belder 2011 for an interesting alternative). Moreover, they are:

- (8) i. category-neutral and categorized in the course of the derivation;
- ii. meaningful, although there is no consensus on how much content they have;
- iii. phonologically identified as *forms*.

We have a number of objections on these (see also Acquaviva 2009b, Borer 2009; Harley 2012). The first is of a conceptual nature: if roots are indeed meaningful, then they are equivalent to verbs, nouns and adjectives except for a categorial label. This in turn raises serious concerns on the nature, purpose and necessity of categorization in natural language (see Panagiotidis 2011 for discussion). The second objection concerns two interlinked facts: on the one hand, there exists unconstrained variation between roots that appear to be very specified (e.g. *sugar*), extremely impoverished (e.g. *mett-* in Italian or *mit-* in English) and all the in-between shades. Moreover, even if we argue for impoverished and semantically underspecified roots, we are

still with left with the empirical problems adumbrated in the previous section, namely that roots too often do not capture a coherent meaning (what connects, for instance, the noun *book* to the verb *to book*? what logical or ontological type should the root *book* have?). This renders unlearnable the purported ‘common semantic denominator’ roots are supposed to express.

It seems, then, that roots in the technical sense this term has in Distributed Morphology cannot have all the three properties attributed to them. Taking into account also the recent contributions by Borer (2009) and Harley (2012), we propose an alternative which abandons (8ii) and crucially qualifies (8iii) (see also Acquaviva, forthcoming, Panagiotidis 2011, Acquaviva and Panagiotidis, in prep.).

First, we think that it is necessary to distinguish between roots as morphological objects and roots as elements of syntactic computation. In doing so, we embrace generalized Late Insertion, not just for non-root syntactic material, as in Galani (2005: Ch. 5-6); Siddiqi (2006: Ch. 3); Haugen (2009). Thus, syntactic roots will be associated with different morphological roots (Vocabulary Items, essentially: *forms*) in particular syntactic contexts, as sketched below:

- (9) $\sqrt{\text{CAT}} \longleftrightarrow \text{cat}$
 $\sqrt{\text{GO}} \longleftrightarrow \text{go}$
 $\sqrt{\text{GO}}, [\text{Tense: Past}] \longleftrightarrow \text{went}$

Given this dissociation, we can use the notion of morphological roots to account for the multiple ‘radicals’ or ‘stems’ that occur, for instance, in Latin inflection and derivation (Aronoff 1994). Thus conceived, morphological roots *may* display specifications like being exclusively nominal or verbal, and we expect there to exist constraints on their form (like the Semitic three-consonant skeleton). Moreover, the same Vocabulary Item (form) that spells out a root may also spell out functional terminals, as is the case of *will* (future marker or noun); see also De Belder (2011). So, a notion of morphological root distinct from that of syntactic root correctly predicts the existence of such ‘semilexical’ categories.

The consequence of the above dissociation is that we can now conceive syntactic roots, as distinct from morphological ones, as *abstract indices* (cf. Acquaviva 2009b, Harley 2009; 2012). By this we mean purely formal objects internal to the faculty of language in the narrow sense; that is to say, elements that are defined only as constituents of a formal syntactic representation, but have no grammar-external status—in particular, not definable, independently of a syntactic structure, as sound-meaning

mappings, or even as abstract instructions to ‘fetch’ or ‘activate’ concepts (contrast Pietroski 2008:319, 9 Boeckx 2010:28-29). What we are essentially claiming is that a syntactic root is a syntax-internal criterion of lexical identity, so that two otherwise identical syntactic constructions count as different formal objects if they differ in the root, and as identical (that is, tokens of the same type) if the root is the same. Given this characterization, there is no semantic variation to explain between root types, nor learnability problems raised by some elusive conceptual content independent of any one lexical item; because roots have no semantic content. Instead, we argue, they act as labels to identify (UG-internally) the structures which correspond to lexical words, and it is *these* which support conceptual content. The following section will make explicit the implications of our proposal for the relation between conceptual content and syntactic structure.

5. MAPPING CONCEPTS WITH WORD STRUCTURE

It seems a truism that if lexical items are grammatically complex, then the corresponding lexical concepts are also complex. If the hypothesis we put forward can be substantiated, however, the structural complexity of a word as a linguistic object does not necessarily correspond to complexity in its conceptual content. Recall that syntactic decompositional approaches aim at representing in syntactic terms the grammatically relevant information encapsulated in a lexical word, by means of a structure generated by the same principles underlying the productive construction of sentences. Now, lexical words also have a non-grammatical content, idiosyncratic and encyclopaedic, which cannot be associated to a grammatical shell. It seems natural to associate this irreducibly lexical residue to a root element. But if independent empirical and conceptual arguments make it problematic to associate with roots even this type of content, the question where idiosyncratic lexical meaning is represented must receive a different answer.

The answer we suggest is that a word’s conceptual content does not correspond to one piece of the syntactic construction, but corresponds to a construction as a whole. Syntactic heads express content regimented into grammatical features, and collectively determine a grammatical interpretation; say, count noun, or unaccusative change-of-state verb. A root at the core of such a construction merely *labels* it; for that purpose, it does not matter whether it is a single node, realized as an invariant phonological form, or a complex node like *cook up* in (6). Assuming that pairs like *break* and *shatter* or *dog* and *cat* have identical structural representations, what we claim is that they are differentiated, in the abstract syntactic representation before morphological spellout, by distinct syntactic roots. These do not

differ by virtue of semantic content, but by a differential marking, like subscripts. It is by virtue of having different subscripts that the structures corresponding to *dog* and *cat* count as different syntactic objects, independently of semantic interpretation. Conversely, when two distinct structures have the same root, they can correspond to different concepts, as in (4) and in the corradical ‘colour’ and ‘flower’ in (7); or, less frequently, they can map to the same concept, as in the singular and plural of ‘flower’ in (7). Syntactic roots, then, mark lexical across syntactic representations. Lexical concepts map to these representations, not directly to roots; what the latter do is provide a UG-internal signature for lexical concepts.

Consider Borer’s (2005b:9) statement that a lexical item consists of ‘its syntactic structure and the interpretation returned for that structure by the formal semantic component, and [...] whatever value is assigned by the conceptual system and world knowledge to the particular listemes embedded within that structure.’ Instead of claiming that the conceptual component is associated to grammatically inert listemes (which has no independent motivation, although it may appear natural as a null hypothesis), we claim that an empirically more satisfactory solution consists in taking the structural-grammatical meaning as a semantic template which *constrains* the conceptual content associated with the structure. If the syntax of a verb involves a causative *v* head, the lexical concept associated with it should be a causative verb (like *kill*); but a semantically causative verb does not have to decompose into a non-causative part and a CAUSE predicate definable independently of this concept. In essence, then, we argue that there exist morphological and syntactic roots, but that there are no semantic roots as distinct from basic lexical concepts; in particular, not as the semantic content of syntactic or morphological roots.

Of course, it is at best insufficient, and at worst circular, to say that a concept may map to ‘whatever’ grammatical construct defines a lexical word (N. Hornstein, p.c.); but the claim that concepts do not map to fixed-size syntactic pieces is coherent and compatible with the data. As cases like the Russian *tsvet-ók* show, a single concept can be expressed by a noun with different structures in the singular and plural; and especially a category like number may easily be an intrinsic component of the lexical concept. This appears clearly in ‘collective’ plurals like the Spanish *padres*, which shifts the meaning of *padre* / *padres* ‘father / fathers’) to that of ‘parents’, but only denoting mother-father pairs (so, a mother and her mother are both parents but are not *padres*). In addition, not just any structure can map to a lexical concept, for principled language-internal reasons. It seems plausible that the domain of conceptual lexicalization cannot extend beyond a nominal or verbal extended projection, probably definable as a syntactic Phase

corresponding to a DP or a vP; in fact, this is expected if we take seriously the notion of Phase as derivational cycle whose output is consigned to interpretation (Acquaviva and Panagiotodis, in prep.).

6. CONCLUSION: COMPLEX WORDS, SIMPLE CONCEPTS

Lexical decomposition, as a hypothesis on the constituency of words as linguistic representations, captures fundamental aspects of lexical competence. On the other hand, it is problematic as a hypothesis on the internal constituency of lexical concepts. Our main point is that decomposition becomes problematic even from a linguistic perspective, as soon as we ask where a lexical grammatical structure hosts non-grammatical conceptual content; resorting to roots, in particular, proves empirically inadequate. Our alternative hypothesis, linguistically motivated, is that a word can be linguistically complex but conceptually simplex.

Conceptual atomism, as defined by Fodor (1998:121), holds that ‘most lexical concepts have no internal structure’. Since we still claim that the grammatical structure of words comprehends meaningful elements, we do not take this thesis to mean that lexical words are semantically unanalyzable as linguistic objects (in particular, they are not semantic atoms in a Mentalese; contrast Fodor 2008). What we claim is rather that a word’s conceptual content is not on a par with grammatically encoded meaning, as the content of one syntactic piece, but belongs outside UG-generated representations and is mapped to them in such a way as to respect the semantic templates defined by grammar. Unstructured concepts, then, can map to complex syntactic structures.

The difference we envisage between lexical concepts (UG-external) and the content of syntactic representations (UG-internal) does not mean that the relation between them is arbitrary and unconstrained. On the contrary, a principled relation between the two can lead to predictive hypotheses on what can be a possible lexical word in a natural language. For instance, Fodor (1998:164-165) argues that while REDSQUARE is conceivable as a primitive concept, without having RED and SQUARE, there can be no *primitive*, atomic concept ROUNDSQUARE, as opposed to the complex ROUND SQUARE (as the conceptual content of the phrase *round square*). Such a basic concept could never identify anything at all; while contradictory properties can be entertained, it could not exist as a basic concept (‘there can be no *primitive* concept without a corresponding property for it to lock to’). But this is a prediction about language: a noun with that content is impossible in natural language. Further hypotheses about the conceptual bases of lexical nouns can rule out words meaning ‘number

of planets’ or ‘undetached rabbit part’ as simplex lexical concepts (Acquaviva, forthcoming).

It bears emphasizing that the thesis of conceptual atomism, and our contention that syntactic lexical decomposition is compatible with it, does not deny the cognitive complexity of concepts. The content of a word enters in a complex network of relations with the content of other words, as CAT and ANIMAL. But inferences can be necessary though not constitutive: taking *water contains hydrogen* to be necessarily true, it is still possible to have the concept WATER without having the concept HYDROGEN.

Word meaning, in conclusion, is indeed cognitively complex, but not as a reflex of grammatical complexity. We take it to be a strength of our analysis that it makes linguistically motivated decompositions of lexical items (more) compatible not only with conceptual atomism, but also with views that, without embracing conceptual atomism, emphasize the lack of one fixed structure for lexical concepts; cf. Murphy (2002:441): ‘Thus, it can be very difficult to know where to draw the line between what is part of the word meaning “per se” and what is background knowledge. It is not clear to me that drawing this line will be theoretically useful.’ A linguistic analysis of lexical content which could be related to a psychologically and philosophically plausible view of lexical concepts is certainly a desirable goal. Our proposal is a contribution towards that goal.

REFERENCES

- Acquaviva, P. (2008). *Lexical Plurals*. Oxford: Oxford University Press.
- Acquaviva, P. (2009a). Roots, categories, and nominal concepts. *Lingue e linguaggio* 8, 25–51.
- Acquaviva, P. (2009b). Roots and Lexicality in Distributed Morphology . In A. Galani, D. Redinger & N. Yeo (Eds.) *York-Essex Morphology Meeting 5*, 1–21.
- Acquaviva, P. Forthcoming. The roots of nominality, the nominality of roots. In A. Alexiadou et al. (Eds.), *The Syntax of Roots and the Roots of Syntax*. Oxford University Press.
- Acquaviva, P. & P. Panagiotidis. In preparation. Roots and lexical semantics. University College Dublin - University of Cyprus.
- Arad, M. (2005). *Roots and Patterns: Hebrew Morpho-Syntax*. Berlin: Springer.
- Aronoff, M. (1994). *Morphology by Itself*. Cambridge, MA: MIT Press.
- Basilico, D. (2008). Particle verbs and benefactive double objects in English: High and low attachments. *Natural Language and Linguistic Theory* 26, 731–729.

- Bever, T. & P. Rosenbaum. (1970). Some lexical structures and their empirical validity. In R. A. Jacobs and P. S. Rosenbaum (Eds), *Readings in English Transformational Grammar* (pp. 3–19). Waltham, MA: Ginn and Company.
- Boeckx, Cedric (2010b). Defeating Lexicocentrism. Ms., CLT/UAB. Available from <http://ling.auf.net/lingBuzz/001130>
- Borer, H. (2005a). *In Name Only*. Oxford: Oxford University Press.
- Borer, H. (2005b). *The Normal Course of Events*. Oxford: Oxford University Press.
- Borer, H. (2009). *Roots and Categories*. Paper presented at the 19th Colloquium on Generative Grammar. University of the Basque Country, Vitoria-Gasteiz.
- DeBelder, M. (2011). *Roots and affixes: eliminating lexical categories from syntax*. LOT: Utrecht.
- Embick, D. & A. Marantz. (2008). Architecture and Blocking. *Linguistic Inquiry* 39(1), 1–53.
- Fodor, J.A. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.
- Fodor, J.A. (2008). *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.
- Galani, A. (2005) *The morphosyntax of verbs in Modern Greek*. Unpublished PhD thesis, University of York
- Grimshaw, J. (1993). Semantic structure and semantic content in lexical representation. Ms, Rutgers University. Published in Grimshaw, J. (2005). *Words and Structure* (pp. 101–119). Chicago: CSLI.
- Hale, K. & S. J. Keyser. (1993). On argument structure and the lexical expression of syntactic relations. In K. Hale & S. J. Keyser (Eds.), *The View from Building 20* (pp. 11–41). Cambridge, MA: MIT Press.
- Hale, K. & S. J. Keyser. (2002). *Prolegomenon to a Theory of Argument Structure*. Cambridge: MA: MIT Press.
- Halle, M. & Alec Marantz. (1993). Distributed Morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *The View from Building 20* (pp. 111–176). Cambridge, MA: MIT Press.
- Harley, H. (2009). *Roots: Identity, Insertion, Idiosyncracies*. Paper presented at the Root Bound workshop, USC, February 21, 2009
- Harley, H. (2012). *On the Identity of Roots*. Ms., University of Arizona.
- Haugen, J. (2009). Hyponymous objects and Late Insertion. *Lingua* 119, 242–262
- Jackendoff, R. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
- Jackendoff, R. (2002). *Foundations of Language*. Oxford: Oxford University Press.
- Laurence, S. & E. Margolis. (1999). Concepts and cognitive science. In S. Laurence & E. Margolis (Eds.), *Concepts: Core readings*. Cambridge, MA: MIT Press.
- Levin, B. & M. Rappaport Hovav. (1995). *Unaccusativity*. Cambridge, MA: MIT Press.
- Levin, B. & M. Rappaport Hovav. (2005). *Argument Realization*. Cambridge: Cambridge University Press.
- Lieber, R. (2004). *Morphology and Lexical Semantics*. Cambridge: Cambridge University Press.
- Marantz, A. (2000). *Words*. Unpublished ms. MIT.

- Marantz, A. (2006). *Phases and words*. Unpublished ms. NYU.
- Murphy, G. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.
- Pietroski, P. (2008). Minimalist meaning, internalist interpretation. *Biolinguistics* 2, 317–340.
- Pinker, S. (1989). *Learnability and Cognition*. Cambridge, MA: MIT Press.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Ramchand, G. (2008). *Verb Meaning and the Lexicon*. Cambridge: Cambridge University Press.
- Siddiqi, D. (2006). *Minimize exponence: economy effects on a model of the morphosyntactic component of the grammar*. PhD thesis, University of Arizona.
-

Paolo Acquaviva
University College Dublin
Newman Building, Belfield, Dublin 4
Ireland
e-mail: paolo.acquaviva@ucd.ie

Phoevos Panagiotidis
University of Cyprus
Dept. of English Studies
Kallipoleos 75, 1678 Nicosia
Cyprus
e-mail: phoevos@ucy.ac.cy