

DRAFT December 1996  
Alec Marantz, marantz@mit.edu

“Cat” as a phrasal idiom:  
Consequences of late insertion in Distributed Morphology

Theories of the Lexicon tend to equate sorts of linguistic units that are at least conceptually distinct. Of course, claiming that conceptually distinct entities are in fact the same can be an empirical claim about the way language works. But to understand what’s at stake in unifying conceptually distinct entities, we need to re-separate the parts. As prelude to this paper, then, I will distinguish three sorts of lists of items that, arguably, belong to our grammatical knowledge. These I will call the “(pure) Lexicon,” the “Vocabulary,” and the “Encyclopedia.”

In the spirit of Government-Binding theory, many syntactic theories suggest that the computational system of grammar (the “syntax,” broadly construed) operates on items chosen from a place called the “Lexicon.” Let us then preserve the title “Lexicon” for the list (perhaps a generated list) of the set of bundles of features that are input to the computational system of grammar and are relevant to this system. These bundles of features would be the “terminal nodes” of the computational system. For the conceptually clean or pure notion of Lexicon, we include in this list of bundles of features only features that actually play a role in the computation (in computing “merger” and “movement” in Chomsky’s Minimalist terms). Thus any features relevant only at the “interpretive” levels of PF and LF would not be part of Lexical items, narrowly defined. To the extent that phonological and semantic features are irrelevant to the computational system, these are not features of Lexical items as defined above. When it becomes necessary to distinguish this notion of Lexicon from others, e.g., the Lexicon of “lexicalist” theories, we’ll call this defined notion (bundles of computationally relevant features that serve as terminal nodes in the computational system) the “pure” Lexicon.

About Lexical items as defined above, it’s important to note that no one has ever claimed that all such items in a language should be stored. For example, in a Minimalist (and “checking”) theory of syntax, where the computational system deals with Ns and Vs that contain inflectional features (e.g., Case, Tense, and Agreement features), no one would suppose that the Lexicon explicitly lists all the case forms of a noun, all the inflected forms of a verb, all the underspecified forms listed multiply for all of their possible uses.... Thus “the Lexicon” as the input to the computational system is always considered generative in some sense. Even within Distributed Morphology, where Case and Tense and Agreement would be separate terminal nodes from N and V heads and all composition of terminal nodes is assumed to be syntactic rather than lexical, still the Lexicon as defined above would be generative. The possible combination of person, number and gender features in a terminal node, for example, is a matter of universal and perhaps language specific constraints — substantive constraints listing the possible features and combinatorial constraints about feature cooccurrences. If gender and number features can generally cooccur with person features in a language, then, one would not wish to list all the possible Lexical items consisting of combinations of person, number and gender features; rather, the creation of such combinations — as possible pronouns for the language, for example — should involve a “generative Lexicon.”

In general, the features relevant to the computational system, and any language particular constraints on these features (substantive or combinatorial), determine a “space” of Lexical items — i.e., paradigms. This notion of paradigm figures into every theory, even one in which all functional features are features only of functional heads. What features a

Tense node, for example, can bear determines a paradigm space of tense Lexical items and thus a space for organizing morphophonological expressions of tense. If you believe that some Lexical items are stored, then paradigms serve as an interface between the stored Lexical items and the bundles of features used in the computational system. (cf. Lieber's use of categorial signatures; Wunderlich's explicit system of paradigms generated from the features of affixes; informal assumptions about feature fill-in post-Lexicon; etc.)

In Distributed Morphology, where the morphology and phonology are purely interpretive rather than filtering, the computational system as a whole determines a vast paradigm space for morphophonological realization. Both the combination of features in terminal nodes and the merger and movement of these nodes map out the space of possible expressions in the language to which the morphophonology assigns phonological realizations. In terms of "paradigms," then, Distributed Morphology makes no distinction between inflectional and derivational morphology, nor between derivations internal and external to the (phonological) word.

Given the above definition of "Lexicon," a definition that speaks only of features relevant for the computational system of grammar, not for the phonological or semantic interfaces, Lexical items would not contain phonological features (i.e., any phonological features beyond those relevant for the syntax). We therefore may define the "Vocabulary" as the set of connections between Lexical items and phonological features. Although the term "morpheme" has been used in a variety of ways through the history of linguistics, within DM Vocabulary items come very close to matching usual notions of "morphemes." Since Vocabulary items are the minimal stored chunks of phonological information and since they are inserted one per terminal node from the syntax, where word-internal structure is maximally decomposed into terminal nodes, Vocabulary Items are the minimal units of sound - (syntactic/semantic) feature connections in a language.

Recall that the space of Lexical items is in some sense generated by the features of Lexical items. The Vocabulary items are underspecified with respect to this space. Every theory recognizes this split. No theory, for example, would list 5 different lexical items for present tense "walk" each with the same pronunciation (first singular present, first plural present, second singular present....) — nor would a theory like Distributed Morphology list 5 different present tense suffix Vocabulary items with zero pronunciation. Every theory recognizes that present tense "walk" (or the zero present tense suffix) is underspecified as to the person and number of the subject. This underspecification of Vocabulary items gives us a use/mention distinction with respect to the features of Lexical items. The features of Lexical items are \*used\* by the computational system. But they are \*mentioned\* by the Vocabulary items that relate these features to phonological form. Again, the computational system works in the paradigmatic space defined by the features of Lexical items. Vocabulary items connect this space to phonological form by mentioning some of the lexical features and describing how they relate to phonological features.

Standard generative theories and, more recently, Anderson's A-morphous morphology, make a cut between roots and derivational morphemes on the one hand and inflectional morphology on the other. The roots and derivational morphemes are standardly assumed to be a combination of pure Lexical items and Vocabulary items, i.e., bundles of features that include both the features relevant to the computational system and the features relevant to phonological interpretation (as well as, of course, the features relevant to semantic interpretation). Inflectional morphology, on the other hand, is often taken to separate its pure Lexical and Vocabulary aspects, with "abstract" feature bundles playing a role in the syntax and phonological realization accomplished by rules similar to phonological rules (see Anderson for one way to play out this sort of theory).

The Lexical items were defined in way to exclude as lexical features any semantic features that play no role in the computational system (of syntax). For example, the semantic difference between “dog” and “cat” drives no syntactic or compositional semantic principle, rule, or constraint; thus, whatever feature of “dog” that distinguishes “dog” from “cat” would not be a feature of Lexical items. The Vocabulary was the list that related the Lexical items to phonological features that were irrelevant to the computational system of syntax. We will call the list of connections between nodes in the syntactic derivation and meanings that are irrelevant for the computational system the “Encyclopedia.”

“Lexicalist” theories claim that the actual units manipulated by the syntax combine information from the Lexicon, Vocabulary, and Encyclopedia, as I have defined them. Since, as we have seen, the “pure Lexical items” can’t all be listed, all lexicalist theories endorse a computational Lexicon that actively makes the connections between Vocabulary, Encyclopedic, and pure lexical information.

Distributed Morphology claims, in contrast, that the only system to make the connection between Pure Lexical, Vocabulary, and Encyclopedic information is the grammar as a whole. In particular, the computational system works only with Lexical items, which consist of atomic bundles of Lexical features. This means both that the terminal nodes of the syntax proper contain no phonological and no syntactically irrelevant semantic features and that the syntax “starts with” only simple, nonbranching constituents. All internal structure of words is a product of the syntax and of (morpho- and morphophonological) operations in the PF component. The phonological features for Lexical items is provided via Vocabulary insertion at PF, i.e., after the computational system. Distributed Morphology thus endorses “late insertion” of Vocabulary items.

This paper is concerned primarily with connection between Lexical and Vocabulary items on the one hand and Encyclopedic knowledge on the other. While it’s true that PF must know whether a node contains “dog” or whether it contains “cat” and while it’s also true that semantic interpretation of a sentence containing “dog” or “cat” must also know which one appears at a given node, it is not clear that the computational system of language, down to whatever principles, filters or constraints operate at LF, must know whether a node contains “dog” or “cat.” Distributed Morphology takes the position that this difference between “dog” and “cat” is a matter of Encyclopedic knowledge and that such knowledge is used in semantic interpretation of LF, but not in grammatical computations over LF or involving LF. A node at LF might be such that at PF, either “dog” or “cat” might equally well have been inserted. Thus the use of the Encyclopedia to interpret LF semantically must involve knowledge of Vocabulary insertion at PF, where a choice between “dog” and “cat” may have been made. We claim that **any** (unforced) choice made in the course of a derivation is input to semantic interpretation. Thus, for example, any choice of Lexical items for computation, any choice to combine X and Y rather than X and Z, any movement that is not forced by grammatical principle, any choice of Vocabulary item that is not forced by the “Elsewhere” principle (choose the most highly specified item that fits) — any of these choices may be interpreted. The choice of “cat” as opposed to “dog” (and as opposed to any of the numerous other Vocabulary items that might be inserted into something like an “animate count noun” node from the syntax) must be registered as part of the derivation subject to semantic interpretation with the help of the Encyclopedia. On this view, of course, LF alone isn’t the input to semantic interpretation. Rather, the derivation as a whole is interpreted. But the specific aspects of the derivation that get interpreted are precisely any and all unforced choices made (cf. for historical purposes the notion of a T-marker in old transformational analyses).

We take the strong position, to be explicated below, that the syntax (and thus LF) provides the only way the mind has to represent compositional meanings. Thus the Encyclopedia may list only the noncompositional meanings of items — compositional meanings are constructed by the computational system. Noncompositional meanings of a node may be conditioned by the syntactic environment of the node, with idioms presenting the most extreme case of contextually conditioned noncompositional meanings. On one extreme, a root like “-ceive” has a noncompositional interpretation entirely dependent on its context (cf. conceive, deceive, receive — where the compositional relation between an aspectual prefix and a root is of course preserved in all these words, since compositional meanings are inherent in the syntactic configuration (inside and outside words) and not even “idioms” can mess with compositional meanings — see below). On the other, almost anywhere “cat” appears, it means about the same thing (but see below).

The main purpose of this paper is to counter an intuition that gives some force to “Lexicalist” theories, which otherwise fare very poorly from the standpoint of any real linguistic arguments. This intuition identifies the word as the locus of special meaning and thus supports a notion that there is a place for listing specialized meanings, a place that lists words, i.e., the Lexicon. Once one acknowledges a storage for “words,” with phonological, lexical, and Encyclopedic information all bundled together, one is led to a computational Lexicon independent from the computational system of syntax. Assuming a computational Lexicon that yields words leads to an “early insertion” theory of Vocabulary and Encyclopedic knowledge and to all sorts of ordering paradoxes surrounding the interaction of phonological and syntactic/morphological form. Contrary to our powerful intuitions, careful examination shows that the phonological word has no special status with respect to idiosyncratic meanings. The environment within which a particular Vocabulary choice may find a trigger for a special noncompositional meaning (the domain of idioms, one might call it) is restricted, but this local domain for contextual determination of special meaning doesn’t correspond to any notion of “word” or “zero level category” but is, rather, syntactically defined in a natural way. These local domains are related to the domains in which the morphological categories of N, V and Adj are realized, and thus there is a connection between the categories “verb” and “noun” and special meanings, but I show below that these connections emerge coherently only in a non-Lexicalist theory. Once we dismiss the notion that the phonological word has special Encyclopedic status and identify the syntactic domain of special meanings, the intuitive force behind early insertion and lexicalist theories dissolves.

**Lexicon:** The (generated) list of bundles of features that enter the computational system; contains only those features relevant to the principles of syntax. What features appear in these bundles is determined by general principles and perhaps language-particular inventories and co-occurrence restrictions.

**Vocabulary:** Vocabulary items are connections between lexical features and phonologically relevant features. All the Vocabulary items compete for insertion at each terminal node, but in the phonological component (i.e., after “spell-out”).

**Encyclopedia:** Encyclopedia entries connect (pieces) of the output of the grammar — derivations of PF and LF connections — to noncompositional meanings. These entries are used in the interpretation of linguistic structures, where we take the relevant input to semantic interpretation to be the whole derivation, not simply LF, which is the syntactic representation of compositional meanings.

This paper is divided into three parts. In the first part we “kick the bucket” with Ray Jackendoff, exploring the connection between the non-compositional semantics of words

and of phrasal idioms. In the second, we look at the late insertion of Vocabulary Items in Distributed Morphology and why the atomic units of grammar aren't the kind of things you know non-compositional meanings about. Finally, we take a revisionist look at "Remarks on Nominalization" and conclude that nouns and verbs are always phrasal, not atomic. These sections connect to my title in the following way: The noncompositional and computationally irrelevant meaning of "cat" — furry domestic feline.... — makes it an idiom (i.e., an Encyclopedic entry). "Cat" the noun with the noncompositional meaning isn't a unit in the computational system of language, since Encyclopedic knowledge is used only in semantic interpretation of the output of this system. And the noun "cat," about which you know noncompositional meaning things and which isn't a unit in the grammar but is an idiom, is phrasal, rather than atomic. Thus: "cat" is a phrasal idiom.

I. The Encyclopedia: Where non-compositional meanings of linguistic representations are stored.

In a couple recent papers, Jackendoff discusses the importance and relevance of the fact that non-compositional meanings are not limited to word-size units. We all know about idioms — phrasal combinations of morphemes with unpredictable non-compositional meanings — but we tend to think that the bulk of noncompositional meanings rests in single words (until we begin to learn a foreign language at any rate).

Jackendoff points out that knowledge of special meanings and usages of linguistic material is pervasive at all sizes of constituents. He's collected — or rather, has had his daughter collect — what he calls the "Wheel of Fortune Corpus." The puzzles on this game show include a high percentage of idiomatic material. This TV show has run, daily, for years and continues to appear on TV without repeating clues or showing any signs of running out of material. Idiomatic material from English, then, seems virtually inexhaustible.

(1) From Beth Jackendoff's Wheel of Fortune Corpus

Sitting Pretty  
You Scared the Daylights Out of Me  
Eat Humble Pie  
Keep your Shirt On  
Pie in the Sky  
Stick in the Mud  
Off the Top of My Head  
Never Throw in the Towel  
Not Playing with a Full Deck  
You're Whistling in the Dark

Jackendoff rightly remarks on how the Wheel of Fortune puzzles include a range of material from proper names to song titles to cliches, catch phrases, etc. He suggests that there is no particular reason to draw a line anywhere among these pieces of language that speakers know, separating, for example, what we know about the meaning of "cat" (other than its strictly "pure" lexical features, say singular animate count noun) and what we know about "you're whistling in the dark." If the "Lexicon" is that storage house of idiosyncratic information about linguistic units, then the "Lexicon" must include the items in (1) as well as "cat." We'll label Jackendoff's Lexicon, "Lexicon-j" to keep it distinct from the pure Lexicon defined above.

Jackendoff uses the prevalence of larger than word-sized idioms to argue for a particular view of lexical-j items and their storage. In his conception of the grammar, there

are essentially three levels of analysis, PS (phonological structure), SS (Syntax), and CS (Conceptual Structure) — there's no LF, but this absence is irrelevant for the present discussion, and we should take any LF that we might motivate as being part of SS for Jackendoff. Jackendoff views these levels as independently constructed and related via correspondence rules. Each lexical-j entry becomes, then, a miniature correspondence rule, one that relates PS features to SS features to CS features. Morphemes do this generally for a single node at each sort of representation while idioms relate whole chunks of structure at each level. So in (2) you see the phonological features of “cat” corresponding to a set of syntactic features and these in turn corresponding to some sort of conceptual structure features.

(2) Word                      ←-->    N            ←-->    [<sub>Thing</sub> TYPE: CAT]  
      Syl                                  Count  
  Sing  
      /kæt/

Phrasal idioms link larger chunks of structure together, as shown in (3), Jackendoff's representation for the lexical-j entry of the idiom “kick the bucket”:

(3) LCS:                    [DIE ([ ])<sub>h</sub>]  
      Syntax:                [VP<sub>a</sub>[V] [NP<sub>b</sub>[Det]<sub>c</sub>[N]]]<sub>h</sub>  
      Phonology:            <sub>a</sub>/kIk/ <sub>b</sub>/ðV <sub>c</sub>/bvkʌt/

The representation in (3) is from a separate section of the paper containing (2), hence the slight difference in notation (and I have reproduced (3) more or less precisely from Jackendoff — I don't know why there's an open parenthesis in the LCS). The important claim that Jackendoff makes is that there is no difference in the lexical-j representation of noncompositional meanings between words as in (2) and phrasal idioms as in (3). “Compositional meanings” would be captured by regular, general correspondence rules that link the structure of Syntax to the structure of Conceptual Structure representations. The entry for an atomic word like “cat” makes a link between noncompositional aspects of meaning and the syntactic and phonological features of a word. A phrasal idiom actively suppresses the usual correspondences between syntactic and conceptual structure, providing instead a particular noncompositional conceptual structure for a patently compositional syntactic structure.

This idea of “lexical” entries as mappings between different representations was actually an idea I endorsed in my book (Marantz 1984). As you recall, a central reason to adopt such a position is to allow lexical entries to stipulate unusual mappings. At the time, I wanted to allow for unusual connections between semantic argument structure and syntactic structure, for example. The mapping theory gave me a way to encode such things.

Jackendoff, too, is employing entries as correspondence rules precisely to allow lexical-j items, including now phrasal idioms, to override general correspondence rules relating, for example, CS and SS. However, I now believe that we know that such special correspondences don't exist. Jackendoff himself does not argue in the papers I am citing for unusual compositional meanings; he argues for special non-compositional meanings. That is, idioms in general don't yield special compositional meanings; they seem to suppress compositional meanings and yield non-compositional meanings. An idiom won't, for example, make the subject a patient and the object an agent. An idiom might, as in “kick

the bucket,” suppress the referentiality of a phrase. But an idiom won’t create a novel compositionality among the syntactic elements it contains.

Jackendoff denies this conclusion. He claims that “Construction Grammar” (championed by Fillmore and Kay) has provided evidence of specially compositional structures, where novel and particular compositional relations between syntax and conceptual structure must be stipulated. I give some of Jackendoff’s examples below:

- (4) a. Babe Ruth homered his way into the hearts of America.
- b. The more he likes her, the more she dislikes him.
- c. Bill hates chemistry class, doesn’t he?
- d. One more beer and I’m leaving.
- e. Fred doesn’t even like Schubert, let alone Mahler.

I’ve written about the “way” construction in (4a), showing that its compositionality is perfectly regular in English. Fillmore and Kay make a lot of “the more...the more” construction in (4b), but I don’t see their point at all (apparently, the intuition here is that the structures in (4b-e) all involve phrases acting like sentences but without the usual sentential structure, e.g., “One more beer” lacks a verb and, probably, the subject). How would you argue that a construction shows unusual compositionality? “Unusual” should imply at least two things: doesn’t follow from universal principles (interacting with the specific parameter settings of individual languages) and must be learned through example by a child. The second is particularly important, since the claim must be that some construction actually violates a child’s expectations.

To show that a construction involves a special semantic/syntax correspondence, one would want to show not just that the construction is peculiar to a language but that similar constructions in other languages have different compositional interpretations. And then one would want to show both that a child has sufficient exposure to the construction and its special meaning to acquire it and that children make mistakes on the construction (and that they acquire it at different ages, some children never acquire it, etc.). What Jackendoff and the construction grammarians do instead of making this kind of argument is rather to argue that on their particular view of the mapping between syntax and semantics for English, these constructions are somewhat exceptional. The answer to them is, of course, that these constructions simply show that their view about the mapping between syntax and semantics is wrong. That is, these constructions are counterexamples to their theories about the syntax/semantics interface, not motivation to adopt theories that allow for exceptional mappings.

However, what’s really behind construction grammar is precisely the point that compositional meanings don’t reside in Lexical items — they reside in structures. The work in construction grammar, then, argues against a “Lexicon” as the locus of information about compositionality, if the “Lexicon” is where special information about the connection between phonological form and syntactic features is stored. Where we differ from construction grammar here is not in the acknowledgement of structural meanings divorced from particular Lexical items (agreement here) but in the denial that such constructional meanings are special and should be listed construction by construction.

Jackendoff’s survey and analysis of specialized meanings and semi-fixed expressions doesn’t then argue for a single Lexicon-j of special mapping rules that relate syntactic structure to compositional and non-compositional meanings for morpheme as well as for phrasal units. Rather, Jackendoff’s work argues for a single storage for morpheme-sized and phrasal units that relates the non-compositional meanings of morphemes to their syntactic environments — i.e., the Encyclopedia. Note a redundancy in Jackendoff’s

representation of the phrasal idiom, “kick the bucket.” The structure of a transitive verb phrase is listed in this lexical entry, but presumably this structure is licensed by the syntax independent of the construction. Almost 100% of phrasal idioms conform to the syntax of the language. Clearly we don’t want the lexical entry for “kick the bucket” itself to license a transitive verb phrase. Rather, what Jackendoff’s entry is doing is saying, “given a structure of a certain sort built by the grammar, we can interpret the pieces of the structure in a particular way” (the meaning of the structure itself is fixed by the structure, as explained below). This is precisely Encyclopedic knowledge about structures independently licensed. Or, to put it another way, grammar is a theory about the mental representations of linguistic objects. We would need a similar structural theory about visual representations. So-called “lexical(-j) knowledge” of idioms is knowledge about these linguistic representations. Such knowledge doesn’t enter into the licensing or construction of the representations, only into their interpretation.

But what we say about Jackendoff’s entry for “kick the bucket” applies equally well to his entry for “cat.” As far as the licensing structure of grammar is concerned, all we need to know about “cat” is that a particular bunch of phonology can be associated with a syntactic node with the features, singular count noun (perhaps also, animate). It’s certainly nothing special about “cat” that singular count noun maps onto “THING.” Moreover, nothing we know about cats is relevant to the syntax or more generally to the representation of the structure of sentences with the noun “cat” in them. Given the presence in the computational system of grammar of a relationship between the features singular count noun and the phonology /kæt/, our knowledge of “cat” might best be expressed as Encyclopedic knowledge about linguistic representations containing that relationship.

As Jackendoff urges, when it comes to the usual informal notion of the meanings of “words,” there is no reason to make a decisive split between units the size of phonological words and bigger, phrase-size units. The special meaning of “cat” — what separates cats from dogs, for example — is meaning relevant only to the semantic interpretation of nodes in independently licensed syntactic contexts. In an important sense, all such meanings are “idiomatic” for here we are precisely dealing with meanings that follow from no generalizations about the relationship between syntactic structure and semantic interpretation.

The morals we draw from Jackendoff’s observations are:

First, “There are too many idioms and other fixed expressions for us to simply disregard them as phenomena ‘on the margin of language’ “ p 22

Second, noncompositional meanings of idioms involve knowledge of linguistic representations independently licensed by the computational system of grammar — there is no reason to suppose that listed or lexical knowledge messes with the compositional system in special or idiosyncratic ways.

Third, noncompositional meanings of simple Lexical items like “cat” can be treated on par with the noncompositional meanings of phrasal idioms — all such meanings involve Encyclopedic knowledge about linguistic representations independently licensed.

Two footnotes to this discussion of Encyclopedic entries:

First, it’s important in the present context to distinguish among different notions of “special” meanings with respect to composition and to clearly separate the possibility of special “non-compositional” meanings from the (im)possibility of “special



compositional” meaning. In semantic interpretation, the Encyclopedia assigns to atoms of syntactic composition noncompositional (atomic) meanings based on the choice of Vocabulary item for that atom and based on the syntactic context. When a specialized context is listed in the Encyclopedia that “bleaches” or negates the (canonical/default) semantic effect of the choice of a Vocabulary item, this is a case of special noncompositional meaning. For example, in “kick the bucket,” the choice of “bucket” might lose its canonical noncompositional meaning referring to buckets in the syntactic context of “kick.” Although contextually determined special noncompositional meanings imply composition — syntactic context implies compositional semantics — the actual Encyclopedic information is about the noncompositional meanings of atoms in a syntactic context. The claim is that information about noncompositional meanings of atoms in context is all the Encyclopedia ever contains; the Encyclopedia \*couldn’t\* contain special compositional meanings.

I take it to be the case that the only way one would have to represent to oneself compositional meanings that are encoded through syntactic relations is through bits of syntactic structure itself. That is, there is not, as Jackendoff assumes, a translation from a syntactic representation of compositional semantics to a similar Conceptual Structure representation of compositional semantics. The one syntactic representation is all you get, and this will be interpreted semantically, whatever that comes to, but semantic interpretation won’t be simply a translation into another syntactic-like representation. The implication of this assumption is that the only way that the Encyclopedia could list a “special compositional” semantics for a syntactic construction would be to associate the construction with another construction. For example, to take a particular transitive clause and make the subject the patient and the object an agent would just be to transform the syntactic structure into a different syntactic structure, e.g., one in which subject and object are switched. Our claim is that the Encyclopedia does not map syntactic representations onto syntactic representations but only provides a key to the noncompositional meanings of syntactic atoms in syntactic environments.

So, when we claim that the Encyclopedia lists special non-compositional meanings, we mean precisely that — and not that the Encyclopedia lists special compositional meanings. The Encyclopedia doesn’t relate linguistic representations to other linguistic representations but precisely provides the kind of information that isn’t encoded in linguistic representations at all.

Two very careful studies of “idioms,” broadly construed, support the claim that the Encyclopedia can’t provide a new compositional semantics for a licensed syntactic structure with a predictable semantic interpretation. Both Nunberg et al. 1994 and Ruwet 1991 claim that idioms generally if not uniformly preserve the compositional meanings of their syntactic structures. What’s special about idioms, other than connotations and figurative associations and allusions, is the special, contextual noncompositional meanings of their atomic elements. For example, the noncompositional meaning of “bucket” in “kick the bucket” differs from its canonical noncompositional meaning referring to a container. This sort of special contextual noncompositional meaning must be contrasted with the notion of special noncompositional meaning as the suppression of predicted compositional meaning. The picture of the Encyclopedia as the locus of noncompositional meanings in the sense of meanings that can’t be structurally represented denies the possibility for the Encyclopedia to suppress compositional meanings in idioms. My only understanding of what it would mean to suppress a compositional meaning would be to replace one compositional meaning with a “smaller” one, one that required fewer pieces and/or relations. But this again would involve mapping one syntactic structure onto another, since only syntax represents the sort of compositional semantics that can be represented syntactically, and the mapping of syntactic structure onto another is precisely beyond the bounds of the Encyclopedia.

When “kick the bucket” is given its special, noncompositional meaning in the Encyclopedia, then, this special meaning can’t be the compositional meaning of a verb like

“die” (and here I follow the discussion in Nunberg and Ruwet of “kick the bucket” and related matters). To associate “kick the bucket” with “die” would be to transform one bit of syntactic structure into another. Rather, the idiomatic/Encyclopedic knowledge can only say special noncompositional things about a structure. In the case of “kick the bucket,” the punctual, completive aspect reading of a transitive clause with a definite direct object is preserved in the idiom — to say that someone “kicked the bucket” is to imply, perhaps in jest and certainly metaphorically, that the person accomplished something in dying. Someone can be said to have been dying for months, but not “kicking the bucket for months.” The strange thing about “kick the bucket” is the pure non-referentiality of “the bucket.” Idiomatic objects often have a special, perhaps vague sense, but usually can be placed into some metaphoric reading with respect to the idiom. When you keep “tabs” on someone, we don’t know exactly what “tabs” are, but they stand in for “an eye,” or “watch.” Thus “kick the bucket” patterns with “shoot the breeze” and a very few other idioms in which it is difficult to assign \*any\* noncompositional meaning to some of the atoms.

So “kick the bucket” has the punctual completive aspect of a transitive verb with a definite direct object — as such, its compositional meaning is normal, if metaphoric. Idioms may, in fact, only mess with the noncompositional meanings of their parts, with the effects of, e.g., the suppression of referentiality accounting for some of the behavior of idioms without need to appeal to a suppression of compositional meanings as well. If this line of reasoning is correct, then the syntactic atom has special status with respect to meaning and in contrast to syntactic structures; syntactically complex elements always carry with them their associated compositional meanings, even through idioms, while syntactic atoms have their meanings determined contextually. This conclusion would, though, give special semantic status to the “word” only to the extent that words are the syntactic atoms — and we’ll argue against any such notion below. Let me repeat, though, that there is an essential difference between the syntactic atom and the syntactically complex object — the latter will have an inalterable compositional meaning (to the extent it consists of semantically interpretable elements; it could be that the syntactic combination of functional elements involves no compositional semantics). However, elements of all levels of complexity have noncompositional meanings, and this is Jackendoff’s main point, suitably translated.

Our second footnote to this section concerns the possible domain of the context for Encyclopedic determination of (special) noncompositional meanings. What’s most important in this paper is that the domain of special meaning is not the (phonological) word and that there is no particular difference between contextually dependent meanings within words and within larger structures. Ruwet and others have emphasized the point that what might be called “light-verb” constructions show the semantic properties of “lexical” verbs. Constructions in English with “do” “take” “give” and other light verbs have the semantics of single verbs and call into question the notion that the phonological word is the distinguished locus of idiosyncratic meaning.

- (5) a. Take a leap
- b. Take a leak
- c. Take a piss
- d. Take a break.....

Although light verb constructions and idioms show that the domain of specialized meanings is not the phonological word, there do seem to be locality constraints on the contextual determination of specialized meaning. Note that in light verb constructions/idioms with “make,” for example, a lower verb cannot be agentive.

- (6) a. Make X ready  
 b. Make X over  
 c. Make ends meet  
 d. \*Make X swim/fly a kite/etc. (only pure causative meaning on top of independent reading of lower VP)
- e. Marie a laissé tomber Luc. ‘Marie dropped Luc like a hot potato’  
 f. On lui fera passer le goût du pain. ‘They’ll kill him (lit. make the tast of bread pass him)’  
 g. \*Marie a laissé/fait V (NP) (à) NP\*, with special meaning of “V” not available outside the causative construction where NP\* is an agent

In fact, a number of observations in the literature about the domain of idiomization follow from the hypothesis that the node that syntactically projects an agent serves as a boundary for special meanings — a boundary that limits the domain of any contextual information contained in an Encyclopedic entry; see Harley (1995) for some discussion of the syntactic projection of agents and the role of the agent in setting the boundary for contextual determination of meanings. This observation about the domain of special meanings allows us to show a complete double disassociation between the phonological word and special meanings. We showed that the domain of special meanings may be \*bigger\* than the phonological word; when a single phonological word includes the head that projects agents, the domain of idioms is \*smaller\* than the word.

The restriction to non-agentive lower VPs for special meanings in causative constructions was noted already by Ruwet for French constructions with “faire” and “laisser” (see (6e-g)). Work by Miyagawa, extended by Harley and Kuroda, shows a similar effect in Japanese, where the Japanese equivalent to “make” or “faire” is the suffix -sase. An example from Japanese clearly illustrates that the locality domain for idiomatic interpretation is independent of the phonological packaging into words. Indirect causatives with -sase, where the causee is a persuaded/coerced actor, never have a special idiomatic reading that is not associated with the stem to which they attach. Direct causatives in -sase, where the causee is not an actor/agent, may have such a reading. Direct and indirect causatives may be morphophonologically identical (e.g., when a direct causative employs -sase) and certainly both types illustrated in (7) are single phonological words.

- (7) tob-ase “fly-make” = demote someone to a remote post — direct causative (non-agentive lower VP) with idiomatic reading  
 suw-ase “smoke-make” = make someone smoke — indirect causative (agentive lower VP) and no possible idiomatic reading that isn’t present when the root is used independently of -sase)

Despite its packaging into a single word with a root, the causative affix -sase cannot serve as the environment for special meanings of the root if the agent-projecting head falls between the root and -sase, i.e., if the verbal root has an agentive interpretation.

In discussing the domain for idiomatic/special noncompositional meanings, we must be careful not to forget that all the structures under consideration, whether idiomatic or not, retain their \*compositional\* meanings. All the light verb constructions with English “make” or French “faire” and all the Japanese causative constructions with -sase have the causative reading that is structurally determined. The special meanings are noncompositional readings associated with the choice of Lexical items within the domain of the causatives. The context of a causative head cannot reach down below an agent to effect the noncompositional reading of a lower element.

The notion that the syntactically projected agent marks a boundary for contextually determined special meanings gives us an interesting angle on the noted difference between

“eventive” and “stative” readings of passive morphology and the possibility of passive idioms. Ruwet, using examples from French and English, clarifies the claim that there cannot be any “passive idioms” (idioms in passive form where the corresponding active lacks the idiomatic reading). He shows that clauses passive in form can be idiomatic, but only on a stative, non-agentive, non-eventive reading. So passive morphology can serve as the context for special meanings, but only when the construction is stative, i.e., when no agent is syntactically projected (and we assume that true “implicit” agents must be syntactically projected). Note again that what is at stake here are the noncompositional meanings — the “passive in form” stative idioms all have stative interpretations associated with their syntactic structure; only the noncompositional interpretation of choices of Vocabulary items below the passive morphology is special.

- (8) a. Chaque chose à sa place, et les vaches seront bien gardées.  
 ‘Each thing in its place and everything will be OK.’  
 b. Cet argument est tiré par les cheveux.  
 ‘This argument is far-fetched (lit. pulled by the hairs).’

This difference between statives, where agents are not projected syntactically, and passives, where they are, is brought out dramatically in a recent paper by Dubinsky and Simango on Chichewa (although they use the evidence in an argument that statives are lexical while passives are syntactic). Both stative and passive morphology in Chichewa are verbal suffixes that pattern morphologically with such suffixes as applicative and causative morphemes and go inside normal tense and agreement affixes. The stative suffix, by a variety of tests, does not project a syntactic agent and yields a stative, noneventive, nonagentive interpretation. The passive, on the other hand, does involve the syntactic agent, which can be expressed in an oblique phrase. Dubinsky & Simango note that there are stative idioms but no passive idioms in Chichewa, just like in French and English (although they don’t point out this parallel and although in Chichewa, unlike French and English, the morphology for passives and statives is phonologically distinct). Again, the stative idioms are indeed stative; they just involve an interpretation of the stem not available when the stative affix is not present. Like the Japanese example given above, the Chichewa evidence clearly shows that the domain of idioms can be smaller, as well as larger, than the phonological word. The domain is syntactically/semantically marked by the projection of an agent.

- (9) a. Chimanga chi- ku- gul -idwa ku-msika.  
 corn AGR-PROG-buy-PASS at-market  
 ‘Corn is being bought at the market.’  
 [no idiomatic reading, and none possible with passive]
- b. Chimanga chi- ku- gul -ika ku-msika.  
 corn AGR-PROG-buy-STAT at-market  
 ‘Corn is cheap at the market.’  
 [idiomatic reading of ‘buy’ in the context of STAT]
- c. Chaka chatha chimanga chi- na- lim -idwa.  
 year last corn AGR-PROG-cultivate-PASS  
 ‘Last year corn was cultivated.’
- d. Chaka chatha chimanga chi- na- lim -ika.  
 year last corn AGR-PROG-cultivate-STAT  
 ‘Last year corn was bountiful.’

## II. Distributed Morphology: Vocabulary Items ain't "signs," but neither are the atoms of the computational system.

But why would you want a theory of linguistic structure that separates the noncompositional meaning of "cat" from the connection between the phonology /kæt/ and such syntactically relevant features as singular count noun? Here we turn to Distributed Morphology, a way of looking at morphology that Morris Halle, Rolf Noyer, and I have been developing for a few years. Recall that Distributed Morphology is an aggressively Item and Arrangement theory — a piece theory of morphology. We have published what we consider to be knock-down, dead-on-arrival arguments against A-morphous and paradigmatic approaches to morphology. However, we needed to take something from Item and Process views of morphology — a basic insight that has kept such theories in business for centuries. The insight is this, stated in terms that Robert Beard has used: the pieces that the computational system of syntax and morphology combine are not "signs" in the Saussurian sense — the computational system does not deal with elements that are inextricable combinations of sound and meaning. The separation between the bundles of features operative in the computational system and the pieces of phonology we call stems, affixes and words is most clear in the domain of inflectional morphology. Inflectional affixes, when identified by their phonological form, are underspecified with respect to syntactic features. For example, in English the suffix /-d/ has simply something like the feature "past" — it is underspecified with respect to the feature that distinguishes past tense "he saved" from past participle "he had saved" or "he was saved." One major appeal of paradigm-based morphological theories is their ability to have underspecified forms spread to fill cells in paradigms defined by syntactic features. DM turns the whole computation system into a system for producing a paradigm space to be filled via competition among Vocabulary items.

The structure of grammar assumed within DM is shown in (10), with some updates from previous displays of this structure in the literature. We now find it important to adopt the basic architecture of Chomsky's "Bare Phrase Structure" theory. In place of a syntax connecting D-structure and S-structure, we have a bare computational system that puts pieces together and moves features, sans X-bar theory, sans "Greed," sans just about everything. In departure from other diagrams of DM, I have left out any box or level for Morphology (or Morphological Structure) per se. It has always been a basic claim of DM that Morphology, like Phonology, is interpretive rather than filtering. Morphological well-formedness and phonological well-formedness are both conditions that are met by local insertions, deletions, and minor rearrangements of features or structure. There are no principles that need to refer to or that define a level of Morphological Structure. Therefore, we might well look at morphology as part of the Phonology, i.e., the interpretive component that relates an output of the computational system to PF (and here we'll assume that PF isn't Chomsky's "phonetic form" but some hierarchically organized prosodic structure).

(10) "computational system" ("merge" and "move")

Phonology:	fusion	LF
	morphological merger	
	insertion of nodes like N or V (e.g., "do" insertion)	
	impoverishment	
	Vocabulary insertion	
	fission	

Some comment on the apparent wealth of morphological operations in the Phonology is in order here. I have encountered many cranky complaints that DM is vacuous since the

combinations of insertions, deletions, movements, and feature changing allowed by various versions of the theory could describe anything. Since no alternative system that covers the same sorts of facts (i.e., the entire morphophonology, with the necessity of getting the syntax and semantics correct as well) has ever claimed to be more “restrictive” in any relevant sense, it’s hard to know why the descriptive power of DM should be at issue. But perhaps the complaints arise because in narrow sub-areas of rules and principles and in certain restricted domains, I or my colleagues may have claimed that DM is explanatory because it does rule out possible systems. In the spirit of cooperation with linguists, including our students, that wish to pursue various possible competing paths within the general assumptions of DM (and the general assumptions are those of a piece theory of morphology, with underspecified Vocabulary items and late insertion), we have not chosen specific, radically restrictive versions of the various mechanisms we propose with the aim of momentary “explanatory” power. At this stage in the game, we prefer simply to avoid taking advantage of the possible descriptive excessives of the theoretical machinery, leaving aside for the moment the question about whether what we avoid should be ruled out in principle. For example, we suppose that “fission” of feature bundles into multiple terminal nodes for Vocabulary insertion is purely local and unable to interact with anything else (e.g., fissioned features can’t “merge,” undergo further movement or fuse with other nodes). This locality of fission could be directly insured if we adopt Noyer’s vision of fission as motivated by the contents of the Vocabulary — fission happens in the context of Vocabulary insertion into a particular node. On the other hand, McGinnis argues that we might take fission to be feature movement in the sense of Chomsky’s recent proposals. Mechanically, in my own work I will, for the moment, use fission in the strictly local, non-interactive way that is consistent with Noyer’s vision, although I am not ready to adopt his major claim of how the structure of the Vocabulary drives the morphophonology. Those interested in restrictiveness are welcome to assume that DM, as an orthodoxy, makes the most restrictive theoretical choice here, although individual DMers may go their own ways.

We have here assumed that no feature changing operations exist in the morphophonology. Apparent feature changing might result from feature deletion (impoverishment) followed by a fill-in of the unmarked value of a feature. Note, however, that simply disallowing feature-changing doesn’t make the data that apparently motivate feature changing go away. There is class of phenomena that Zwicky handles through “rules of referral” and that might motivate feature-changing in the morphology. Careful and exhaustive evaluation of these phenomena would be necessary for anyone that wished to claim that feature-changing is impossible. Here, of course, the issues are almost completely independent of DM; they arise in any theoretical framework.

In the case of the addition of features and nodes in the phonology, in general we suppose that the sorts of features and nodes inserted are precisely those that don’t exist in the computational system. For example, we have endorsed a system in which Case and Agreement morphology are purely morphophonological, so Case and Agreement nodes are added in the phonology. In this assumption, I am trying to be responsible to my own (syntactic) work on case and agreement. If the arguments suggested below are correct, then even “N” and “V” nodes are morphological, not syntactic, and thus must be inserted in the “phonology” broadly construed. In the desire for restrictiveness, then, one might suppose that the only features explicitly added in the “phonology” (other than, perhaps, default features added after deletions — see above) are those that are completely absent from the syntax. This move makes the evaluation of restrictiveness depend on the substantive theory of these non-syntactic features and their deployment rather than on the potentially devious re-manipulation of features previously distributed by a constrained syntax.

Fusion as an operation between sisters is already highly constrained. One can explore limiting fusion to sisters whose relation is created via head movement. I have been examining the possibility of eliminating fusion entirely, although this move requires more work out of fission. Again, however, if one is simply worried about descriptive power, fusion need not present an embarrassment of riches.

In general, once one looks at the typical deployment of morphological operations in DM — how they're used and for what sorts of phenomena — one can postulate restrictions on these operations that might, cumulatively, restrict the descriptive power of the theory. However, our understanding of morphology is not likely to be quickly advanced by theory comparison that pretends to rely on evaluations of descriptive power. It is quite possible to make strong arguments on real issues in morphology even though no one has a theory that actually predicts anything, in the strongest sense of predict. The connection between theoretical constructs and particular ranges of data is clear enough to make strong arguments about essential features of morphology. For example, I used to endorse the Lexicalist assumption that individual “morphemes” (what DM would now call Vocabulary items) include subcategorization frames that determine their distribution, both hierarchically and linearly, inside words. I am now convinced that there is overwhelming evidence to the contrary — that Vocabulary items may specify whether they are prefixes, infixes, or suffixes and may have contextual restrictions that put constraints on their distribution but that the hierarchical positioning of a Vocabulary item is entirely determined by the features of the nodes into which it may be inserted. Inkelas (1993) clarifies the issue and shows what kind of predictions might be made if the older subcategorization feature theory was correct. The facts of Nimboran simply falsify her analysis and show what things should look like if morphemes can't have subcategorization frames. One can make this argument against subcategorization frames even though neither competing theory can really make actual predictions in any strict sense.

It should be obvious why underspecified Vocabulary items aren't appropriate atoms for the computational system of grammar. To use underspecified "walk" in the syntax, one needs to know that it cannot be used with third person singular subjects in the present tense. This information is clearly the result of competition with "walks" and not of information directly associated with the form "walk" itself. In Marantz and Halle (1994) we discuss the various approaches to this problem. There seems to be agreement that the proper solution to underspecification of Vocabulary items is the independent generation of paradigm space, with competition among Vocabulary items for filling out this space. The two major alternatives here are to have this competition take place prior to the syntax or after the syntax. The former type of theory would have paradigms as lexical spaces, with competition to fill out paradigm space a lexical phenomenon. DM of course claims that only the computational system itself generates paradigm space, with competition among Vocabulary items at PF, after the syntax.

In Marantz & Halle (1994) we show that since Vocabulary items may have contextual features determining their insertion, paradigms must include not only the substantive features of the node being realized via insertion but also contextual features detailing word-internal and, to some extent, word-external contexts. Since DM generates the paradigm space in the syntax, this contextual information is of course available at PF. Lexical paradigm theories, we argue, must repeat the computations of the syntax within the Lexicon in order to provide paradigms with the appropriate features for the proper spread of underspecified Vocabulary items.

Situations of Impoverishment, we argue, highlight the problem of trying to restrict the competition among Vocabulary items to lexical paradigms. A canonical example of Impoverishment is provided by Harris's analysis of the spurious "se" in Spanish. In

construction with third person ACC clitics, third person DAT clitics show the form "se," which Harris argues is the least marked of the clitic Vocabulary items in Spanish. Contextual deletion of features at PF prior to Vocabulary insertion -- Impoverishment -- can properly account for these cases, where the traditional analyses actually change "le" and "les" to "se" in the environment of the ACC clitics via a morphophonological rule after the syntax. The point should be clear: the Vocabulary item "se" is not the kind of element that one wants the computational system of the syntax to operate with, and providing underspecified "se" with the necessary features to become a possible computational atom would involve recreating the syntax within the Lexicon in order to place essentially syntactic contextual features on par with substantive features as definers of lexical paradigm space.

(11) Impoverishment — the emergence of the unmarked

Spanish: le/les lo di --> se lo di “I gave it to him/them”

The implication of impoverishment is clear: the items that are pronounced — the Vocabulary items — are not appropriate elements for the computational system that feeds LF. Vocabulary items are underspecified as to syntactic features, but the elements that operate in the computational system must be fully (or at least more fully) specified. The late insertion of the Vocabulary items insures that the paradigm space being filled out by the phonological realization of combinatorial atoms is generated by the computational system of grammar — the syntax.

Now most of the arguments for late insertion in the literature discuss inflectional or derivational affixes and clitics. Why should “cat” also be inserted late? Why shouldn’t stems come into the computational system with their phonological features and special Encyclopedic noncompositional information?

There are two basic reasons to treat “cat” and all so-called lexical roots as we treat inflectional affixes, and insert them late.

first, it’s extremely difficult to argue that roots behave any differently from affixes with respect to the computational system. No phonological properties of roots interact with the principles or computations of syntax, nor do idiosyncratic Encyclopedic facts about roots show any such interactions. Moreover, there are no categorical differences between roots and affixes in the phonological component. So why do early insertion of roots? What principled reason would there be to separate the treatment of roots from other Vocabulary items with respect to the locus of insertion?

second, where stem suppletion and stem allomorphy occur, they behave just like affix suppletion and affix allomorphy (as Anderson is reluctantly forced to admit in A-Morphous morphology — and so he adopts two principles of disjunction for choice of stem allomorph that mirror his principles of disjunction for the application of morphophonological rules). To the extent that contextual allomorphy, Impoverishment, and the emergence of the unmarked are diagnostic of late insertion, then at least some roots must be inserted late in any case.

I have provided an extended discussion of “late insertion” for stems elsewhere (see Marantz 1994). From the morphophonological point of view, the most important empirical domain connected with this issues is that of stem suppletion and allomorphy. It is possible to force DM into making very strong claims about the possibility of stem suppletion (recall that stem allomorphy may be the result of morphophonological rules). For example, one



might propose that DM rules out stem suppletion for a Vocabulary item like “dog” that competes with “cat” (and “ties” with it) for insertion in a node. If “dog,” for example, had a suppletive plural allomorph, of the sort shown in (12), this would block out the insertion of “cat” in the environment of “plural,” ruling out “cats” (and all regular plurals of nouns in competition with “dog”). Only “functional” stems, stems whose insertion is completely determined by features and thus whose stems aren’t involved in “ties” for insertion, could have suppletive allomorphs on this view. This approach would require a particular analysis of apparently “light” lexical stems that show suppletive allomorphy, including “go/went” in English as well as the words for “person” “child” “have” “be” etc. in various languages.

- (12) [animate, count, N] <--> /hawnd/ in env. [+pl]  
 suppletive plural stem for “dog” would block out “cats”

Late insertion involves making a specific claim about the connection between LF and semantic interpretation. LF can’t by itself be the input to semantic interpretation. If “cat” is inserted in the phonology at a node at which “dog” could just as well have been inserted — and if, as we assume, the difference between “cat” and “dog” makes a difference in semantic interpretation — then the phonological representation, specifically the choice of Vocabulary items, must also be input to semantic interpretation. The conclusion from our discussion of Jackendoff is even broader — what gets interpreted semantically is, in a sense, the whole derivation, in particular the choices made in the course of a derivation. I suggested that we might look at any unforced choice in the derivation as being input to semantic interpretation. Where “dog” and “cat” share all features relevant to Vocabulary insertion at LF, they tie for availability for insertion at a node and the decision to pick one over the other is open for semantic interpretation, using the Encyclopedia. Encyclopedic knowledge is knowledge about complete derivations and representations, with phonology and syntax included.

### III. Nouns and Verbs are Syntactically Complex

So we have seen why we might want to treat “cat” as a non-compositional idiom. The late insertion of Vocabulary items combined with the claim that Vocabulary items are not signs implies that semantic interpretation needs to look both at the phonology where Vocabulary items are inserted and at LF. An Encyclopedic entry provides the connection between a piece of an entire linguistic representation — PF and LF included — and special meaning. The meaning of “cat” beyond “singular, animate, object” at least, is special in precisely the sense of an idiom — this meaning is irrelevant to the computational system of grammar, irrelevant to anything except semantic interpretation.

Although I’ve explained why “cat” should be an Encyclopedic idiom, I haven’t yet justified calling “cat” a \*phrasal\* idiom. The idea here is that our words — the things we actually pronounce and the things the Encyclopedia helps interpret — our words are not simple combinatorial units in the computational system but are rather always embedded in some syntactic structure. Words like sentences are the interpreted output of the computational system of grammar. In fact our usual syntactic categories of nouns, verbs, and adjectives aren’t syntactic categories at all but morphological categories, categories that emerge during a derivation only in the context of certain functional projections. Something becomes nominal, then, when it merges with D(eterminer). Speakers of highly inflected languages won’t find this notion too foreign — it is often the case that to pronounce a noun, you have to put it in some particular case or other. Or to pronounce a verb, you need to use some inflected form — there is no way to pronounce the bare stem or root. Such inflected

forms in a sense carry a syntactic tree with them. My claim is that this is generally true, even of languages like English without much overt case or tense/aspect morphology. So “cat” is a word that implies a syntactic environment that produces nouns. “Catty” as in, he’s a catty conversationalist, involves presumably the same root in an environment that produces adjectives. The word “catty” involves special noncompositional meaning for the stem “cat,” but so does the word “cat.” Every word is the result of a syntactic derivation interpreted with the use of the Encyclopedia.

To make my point clear here, I’d like to turn back to Chomsky’s “Remarks on Nominalization.” This is often taken to be the locus of the birth of the Lexicon as a place for nonsyntactic derivation, but a close reading of the paper leads to exactly the opposite conclusion. Not only is there no argument for a special Lexicon in the paper, there are specific arguments against the notion that there should be any structures derived in the Lexicon. This is not the appropriate context to develop and support a complete revisionist reading of “Remarks”; instead, I’d like to illustrate my point with an updated account of the paradigms in (13-14).

- (13) a. that John destroyed the city  
 b. \*that the city destroyed  
 c. John’s destruction of the city  
 d. John’s destroying the city
- (14) a. that John grows tomatoes  
 b. that tomatoes grow  
 c. \*John’s growth of tomatoes  
 d. the tomatoes’ growth  
 e. John’s growing tomatoes  
 f. tomatoes’ growing (there would surprise me)

Chomsky’s concern in “Remarks” is the possibility of a transformational derivation of nominalizations like (13c) from sentences like (13a). In retrospect, among the arguments and considerations Chomsky provides for his conclusion, the most compelling and still pertinent are his observations about the differences in syntactic possibilities between sentences and certain types of nominalization. In particular, the “-ing” aspectual nominalization of “destroy” in (13d) or “grow” in (14e-f) parallel tensed sentences in the array of arguments that may appear and in their marking. However, when non-aspectual nominalizations occur as in (13c) or (14c,d), “destroy” and “grow” part company. “Destroy” in “destruction” seems to retain its agent while “grow” loses the possibility of expressing the agentive causer with “growth.”

While Chomsky’s solution to this pattern of data is called “lexical,” it in fact is purely syntactic. Simply put, the lexical root behind both the verb “destroy” and the noun “destruction” includes causative meaning. On the other hand, the root behind “grow” and “growth” lacks causative meaning. Sentential environments allow the introduction of a causative head to project an agent for the root “grow,” but in nominal environments, this causative head is excluded, leading to the lack of an agent for “growth.”

The observations central to current analyses of nominalizations are that certain types of constructions allowed in sentential environments are excluded in nominalizations. These include raising, ECM, double object, and certain experiencer as object constructions (e.g., \*the clown’s amusement of the children). Chomsky’s original proposal points the way to treat all these differences between sentential and nominal environments as syntactic, not lexical. Included among these differences between the sentential and nominal environments is the possible presence of a syntactically projected agent; sentential — verbal — environments allow these agents while nominal environments do not, thus the distinction between “destruction” and “growth.” Pesetsky (1995) can be seen as taking up

Chomsky’s proposal here, attributing many of the constructions that fail to appear in nominalizations to (mostly phonologically null) \*syntactic\* heads whose failure to appear in nominalizations results from generalizations about possibilities of (syntactic) affixation. The “cause” head projecting agents is among these phonologically null syntactic heads excluded for morphological reasons from within nominalizations. I believe we can turn Chomsky’s arguments from nominalizations into a more radical and possibly more explanatory claim about the nature of the so-called “lexical categories.”

Let me update Chomsky’s arguments in a “Bare Phrase Structure” type of theory. In such a theory, there is no X-bar structure. There is a single general way of putting things together: asymmetrical merger. In asymmetrical merger, two elements, possibly already derived via merger, are put together — merged — with one of the elements identified as the head. As far as evaluating whether a given merger is well-formed, one waits to LF to see if the resulting structure finds an interpretation. Thus so-called c- or constituent selection is reduced to s- or semantic selection. Subcategorization of argument taking elements is reduced to the potential of an item for semantic combination.

Recall again that in arguing against a powerful transformation that would take a sentence like (13a) into a nominalization like (13c), Chomsky did not argue for the formation of nominalizations in the Lexicon. Rather, he argued that, behind both ‘destroy’ and ‘destruction’ is a single root that combines with a direct object. The combination of root and complement is nominal in a nominal environment and verbal in a verbal environment. We might provide the structural analyses in (13) to illustrate this, using Pesetsky’s notation for roots.

- (15) a.                    D
- D            √DESTROY
- √DESTROY    the city
- b.                    aspect                    where this “aspect” may merge with an agent
- aspect            √DESTROY
- √DESTROY    the city

Hold on for the moment the question of whether the morphology inserts an N node adjoined to the root DESTROY in (15a) and a V node adjoined to the root in (15b). We will return to this issue briefly below. The claim would be that the possessive agent in (13c) “John's destruction of the city” is not a projected agent. Rather, the root DESTROY implies an agent much the way it implies a time of destruction. Thus, “John's destruction of the city” is to be treated on par with “Yesterday's destruction of the city.” The range of possessive subjects for non-aspectual nominalizations like (15a) is well-known to be different from the range of subjects projected in a verbal/aspectual environment.

On the other hand, the aspect node in (15b) syntactically may project an agent. We will put aside the important (and perhaps deadly to the present analysis) question of the relation between the agent inherent in the root DESTROY, the one that appears in the non-aspectual nominalization, and the syntactically projected agent of (15b).

In addition to the non-aspectual nominalizations -- “fake” nominalizations since they embed no verbal context -- there are also true nominalizations in the language. These arise when one places a root in a verbal context then places the verbal context in a nominal context, as in (16).

(16) D  
 D aspect [may merge with an agent]  
 aspect √DESTROY  
 -ING  
 √DESTROY the city

In the structure in (16), the aspectual -ing may project an agent. Placing the aspectual phrase in a nominal environment, i.e., merging aspect with D, now creates a nominalization with a syntactically projected agent, which appears as the possessive in "John's destroying the city."

The crucial aspect of Chomsky's analysis is the observation that the root GROW, unlike the root DESTROY, is non-agentive. As a consequence, when GROW is placed in the nominal environment as in (17a), there is no agentive argument for the possessive phrase, and we get only "the growth of tomatoes" or "the tomatoes' growth." However, in a verbal environment such as (17b), a syntactically projected agent may appear, yielding "John grows tomatoes."

(17) a. D  
 D √GROW [no implicit agent]  
 √GROW the tomatoes  
 b. aspect [may merge with an agent]  
 aspect √GROW  
 √GROW the tomatoes

Another crucial issue that we will put aside here is the nature and number of the "verbal environment" heads that are involved in "verbalizing" roots like GROW and DESTROY. The most central question here is whether there are two "aspect" heads involved in verbalizing GROW, one that projects an agent and the other, call it aspect-2, that does not. Aspect-2, then, would be involved in verbalizing GROW in "The tomatoes are growing." Chomsky's original proposal was something more like (18), in which a "cause" head is involved in syntactically projecting the agent of inherently non-causative roots like GROW. Regardless of one's approach here, the difficulty in the analysis concerns why roots with inherent agents like DESTROY can't co-occur with aspect-2, i.e., why certain roots resist the causative/inchoative alternation — see (19). To put it differently, what does having an inherent agent amount to?

(18) aspect  
 aspect V?  
 V? √GROW  
 (cause) √GROW tomatoes

(19) \*The city was destroying.

Whatever the ultimate solution to these problems are, they do not seem to argue for lexical composition of nominalizations -- or of verbs, for that matter. What's crucial is the semantic category of roots, from which their syntactic distribution seems to follow. Nor was Chomsky arguing that the Lexicon is the place for arbitrary classifications of elements that interact with syntactic principles to yield the distribution of morphemes within words and sentences. It's not an arbitrary fact about DESTROY that it involves an inherent agent nor an arbitrary fact about GROW that it does not include causative meaning. Chomsky is claiming that there in fact could not be a root, GROW\*, with the meaning [cause GROW]. The causative agent of transitive "grow" \*is\* the syntactically projected agent.

NOTE: The discussion here is complicated by the possibility that "-tion" or "-ment" nominalizations might potentially be aspectual. That is, "John's destruction of the city" might be ambiguous between a structure that includes "aspect" and thus a verbalizing environment and one that does not. Potentially, this distinction might make the same cut as Grimshaw's event vs. non-event readings of nominalizations. On such a view, aspectual "destruction" should be compared with "growing" as in "John's growing of tomatoes." Roots potentially in the same semantic class as GROW, then, might be expected to allow something parallel to "John's growth of tomatoes" with -tion nominalization, as long as this nominalization includes "aspect." So, "John's inflation of the balloon" (cf. "John inflated the balloon," "The balloon inflated," "The balloon's inflation") would not be a counter-example to anything said here on this view.

The evidence here supports Chomsky's proposal, and in a way surprising to any theory that does not accept late insertion. For ease of exposition, let's assume that the verbalizing heads come in at least two varieties, the agent projecting aspect-1 and the non-agent projecting aspect-2, as hinted above. Now it's possible that the realization of a root may differ in the environment of aspect-1 and aspect-2. A root in the same semantic category as GROW, RISE, shows allomorphy dependent on this contextual difference:

- (20) a. aspect-1 projects agents (is the head that defines the boundary of the contextual domain of special meanings)  
 b. aspect-2 does not project agents

- (21) a. John raised his glass. [RISE in environment aspect-1]  
 b. The sun rose in the east. [RISE in its elsewhere form]

Despite the fact that RISE is pronounced differently as a causative and as an inchoative, the causative form does not appear in non-aspectual nominalizations, since without aspect-1, no agent is projected.

- (22) a. \*John's raise of the glass [no aspect]  
 b. The elevator's rise to the top floor [no aspect]  
 c. John's raising of the glass [aspect-1]

Now roots of course can take on special non-compositional meanings in particular environments. RISE does take on special meaning in the context of aspect-1, meaning not present in the context of aspect-2.

- (23) a. John raised a pig for bacon. [special meaning for RISE in context of aspect-1]  
 b. \*The pig raised for bacon. [special meaning absent without aspect-1]

Even though there is a specialized meaning for transitive "raise," this form may not appear in nominalizations any more than the non-special "raise of glass" can in (22a) — see (24).

- (24) \*John's raise of the pig for bacon. [no aspect, thus no "raise"]

This is completely predicted by the present analysis: the special meaning of "raise" is created by the syntactic environment of aspect-1, which is of course absent from the non-aspectual nominalizations.

If nominalization were a lexical process and the Lexicon was the storage-place for specialized meanings, the facts in (21-24) would be completely unexpected. Since transitive "raise" has a special meaning, the verb must be listed in the Lexicon. But if causative, transitive "raise" with the special meaning is listed in the Lexicon, why doesn't it have a nominalization? The nominal form exists in the language — "John's raise was 10% more than expected" — as a resultive nominalization. A lexicalist theory would attribute the ungrammaticality of (24) to some sort of lexical gap, missing the point that \*no\* non-aspectual nominalizations exist with what can independently be identified as (purely) syntactically projected agents.

The behavior of RISE brings to mind an argument from Levin and Rappaport-Hovav that roots like RISE and GROW should be treated as inherently causative, contrary to Chomsky's arguments. They suggest that the fact that the causative uses of such roots have a wider range of meanings implies that the causative meaning is "basic" while the intransitive use is "derived." I take the same data to yield the opposite conclusion: the Encyclopedia lists special contextually determined non-compositional meanings and the agent-introducing aspect-1 will serve as the context for special non-compositional meanings of roots. L & R-H point out that we "break contracts" but that contracts don't break. However, aspect-2 as well can produce special contextual meanings: storms break, but we don't break storms (and since the bare nominalization is aspect free, we lose the special meaning in the nominalization here as well — \*the break of the storm (OK, the break in the storm)). We see these contextual meanings in light verb constructions, where the syntactic complexity of the forms is apparent. We see them also in Japanese derived causative constructions, as long as the root verb is non-agentive. Cross-linguistically, the causative environment is seen to yield special meanings. Nonetheless, the evidence from English nominalizations indicates that these special meanings are syntactic, not lexical, since they are prohibited from syntactic environments that prohibit the syntactically projected causatives.

Chomsky's conclusion about "destruction" is surprisingly close to that of the previous transformational analyses of nominalizations, the analyses he was criticizing. In the transformational story, "N" is a distributional category defined by the PS rules in which it appears. Thus "destruction" is not a noun — it doesn't have the distribution of nouns, which don't take complements. Rather, "destruction" is a verb that becomes morphologically noun-like — a nominalization, but not of the category "N." Chomsky's story is the same — "destruction" is not inherently a noun — its lexical entry is unspecified as to syntactic category. It's morphologically nominal in a nominal environment, morphologically verbal in a verbal environment.

Hale and Keyser in recent work on lexical decomposition make a similar distinction between morphological and syntactic categories. "have", for example, is really a preposition used in a verbal environment. I would question whether syntactic and morphological categories are really drawn from the same pool, however. It could well be that notions like "verb" and "noun" are only morphological notions, defined in terms of the merger of roots and functional categories. Words belong to morphological categories, but are always syntactically derived, and syntactically complex on this story.

Given that the so-called "lexical categories" of N, V, and Adj are morphological categories created in syntactic environments at PF, there is still the issue of whether independent nodes for Vocabulary insertion should be added at PF for the morphological category features. Should "cat" the noun, for example, look like (25a) or (25b) prior to Vocabulary insertion (of course the root here shouldn't be CAT but rather some indication of the syntactic category that includes both CAT and DOG)?



- Halle, M. and A. Marantz. 1994. Some Key Features of Distributed Morphology. MITWPL 21, 275-288.
- Harley, H. 1995. Subjects, Events and Licensing. Unpublished MIT PhD dissertation, soon to be available from MITWPL.
- Inkelas, S. 1993. Nimboran Position Class Morphology. NLLT, 11.559-624.
- Jackendoff, R. 1994. The Boundaries of the Lexicon. Brandeis ms.
- Jackendoff, R. 1994. Lexical Insertion in a Post-Minimalist Theory of Grammar. Brandeis ms.
- Kuroda, S.Y., 1993. Lexical and Productive Causatives in Japanese: An Examination of the Theory of Paradigmatic Structure. Journal of Japanese Linguistics, 15.1-81.
- Levin, B and M. Rappoport Hovav 1995. Unaccusativity. Cambridge, MIT Press.
- Marantz, A. 1984. On the Nature of Grammatical Relations. Cambridge, MIT Press.
- Marantz, A. 1992. The Way-Construction and the Semantics of Direct Arguments in English: A Reply to Jackendoff. In T. Stowell and E. Wehrli, eds., Syntax and Semantics vol. 26, New York, Academic Press, 179-188.
- Marantz, A. A Late Note on Late Insertion. In Y-S Kim et al., eds., Explorations in Generative Grammar: A Festschrift for Dong-Whee Yang, Seoul, Hankuk Publishing Co., 396-413.
- McGinnis, M. 1995. Fission as Feature-Movement. To appear in MITWPL 27.
- Miyagawa, S. 1995 (S)ase as an Elsewhere Causative and the Syntactic Nature of Words. To appear in the Journal of Japanese Linguistics.
- Noyer, R. 1992. Features, Positions, and Affixes in Autonomous Morphological Structure. MIT PhD dissertation available from MITWPL.
- Nunberg, G., I. Sag, and T. Wasow, 1994. Idioms, Language, 70.491-538.
- Pesetsky, D. 1995. Zero Syntax. Cambridge, MIT Press.
- Ruwet, N. 1991. On the Use and Abuse of Idioms. In Syntax and Human Experience, J. Goldsmith trans., Chicago, the University of Chicago Press, pp. 171-251.



