

Compounding in English

Rochelle Lieber

This paper discusses the phenomenon of compounding in English. It starts with an overview of the data of root and synthetic compounding in English, and a consideration of the various criteria that have been proposed for compoundhood in English. Issues such as the productivity and headedness of various sorts of compounds are then covered, as well as structural issues concerning the placement of inflectional and derivational affixes with respect to compounds. Finally, two special issues are considered: the thematic interpretation of compounds and the existence of phrasal compounds.

0. *Introduction.*

The purpose of this paper is to survey some of the basic data of compounding in English and to point out several interesting and still open issues in their analysis. The subject of compounding in English has given rise to a large literature in the last two decades or so, including Marchand (1969), Lees (1961), Levi (1977), Bauer (1978), Allen (1978), Roeper & Siegel (1978), Selkirk (1982), Lieber (1983, 1992), Botha (1984), Sprout (1985), Roeper (1988), and Grimshaw (1990), among others. I will make no effort here to review the literature systematically, but will merely try to give an overview of the facts of compounding in English, and touch on some of the major issues concerning their analysis.

1. *Data and Description.*

Compounding is of course one of the main ways by which new words are formed in English. Native compounds of two major types are usually distinguished. Synthetic compounds (also called verbal, deverbal, or verbal nexus compounds) are those in which the second stem is derived from a

verb; ¹ root compounds (also called primary compounds) are all others. We will begin with a survey of root compounding in English and then move on to synthetic compounding.

1.1. Root Compounds.

(1) contains examples of root compounds derived from various combinations of categories:

- (1) N N file cabinet, towel rack, catfood, steelmill
- N A sky blue, leaf green, stone cold, rock hard
- A A icy cold, red hot, green-blue, wide awake
- A N hard hat, bluebird, blackboard, poorhouse
- A V dryfarm, wetsand, doublecoat, sweettalk
- N V handmake, babysit, spoonfeed, machine wash
- V N drawbridge, cutpurse, pickpocket, pulltoy
- V V stir-fry, blow dry, jump shoot, jump start

The examples in (1) indicate that just about any combination of the open class categories N, V, and A can be found in compounds. However, compound types vary with respect to interpretation and productivity. Compounds composed of two stems of the same category may be *dvandva* compounds: this is almost always true of VV and AA compounds, as the examples in (1) indicate. There are also NN *dvandva* compounds in English: *producer-director*, *fighter-bomber*, *prince consort*, etc. More often than not, however, NN compounds are semantically, and arguably syntactically, rightheaded. That is, as has been frequently pointed out, a *file cabinet* is a type of *cabinet*, *catfood* is a type of *food*, and so on. All other types in (1) are rightheaded as well; *sky blue* is a shade of *blue* rather than a type of *sky*, a *hardhat* is a kind of *hat*, to *dryfarm* or *spoonfeed* is to *farm* or *feed* something in a specific way. Of course, all of these compounds are syntactically as well as semantically rightheaded.

The compounds in (1) also vary in terms of productivity. NN compounds are highly productive. Nearly any two nouns can be concatenated to form a new compound — for example, *armadillo dog*, which can in turn be compounded with another noun to form a still longer compound, *armadillo dog symposium*. NA, AN, and AA compounds are also quite productive in English; for example, *mudbrown*, *greenlady*, and *slippery-runny* are possible new compounds. Of the examples in (1), the least productive are compounds containing verbs. For a number of these types new forms are hard to coin: ?*to *loudtalk*, ?*to *child feed*, ?*to *slip slide*, ?*a *stealcar* (i.e. someone who steals cars). Why it should be so hard to create new root compounds containing verb stems is a subject that is raised in Lieber (1983), although to date I know of no really satisfying answer to this question.

¹ Note that stem and word are usually indistinguishable in English.

As Selkirk (1982) points out, prepositions also seem to take part in root compounding. The examples in (2) are from Selkirk (1982: 14-15):

- (2) P N overdose, underdog, uptown, inland
- P A overwide, underripe, overabundant
- P V outlive, overdo, offset, uproot

These compounds, like the majority of compounds in (1) are both syntactically and semantically rightheaded. All three types seem at least somewhat productive; compounds like the noun *overwig*, the adjective *overhot*, and the verb *overwish* seem possible. Belonging to a closed class, however, prepositions cannot head compounds.

1.2. Synthetic Compounds.

Synthetic (verbal, deverbal, verbal nexus) compounds are those in which the second stem is derived from a verb. Examples are given in (3):

- (3) a. truckdriver, meat eater, cat lover, flower grower, giftgiver
- b. truckdriving, meat eating, cat loving, flower growing, gift giving
- c. powerdriven, motheaten, well loved, home grown, law given

Synthetic compounds, like root compounds, are uniformly rightheaded, and typically the first stem is interpreted as an internal argument of the deverbal second stem. This is a principle which is referred to by Roeper & Siegel (1978) as the First Sister Principle, and by Selkirk (1982) as the First Order Projection Condition. What Roeper & Siegel and Selkirk show is that in compounds of this sort the first or lefthand stem often corresponds to the direct internal argument of the verb in the righthand stem, if it has one. Moreover, if the verb has more than one obligatory internal argument, no synthetic compound is possible. (4a) gives examples of compounds in which the verb's external argument is the lefthand stem of the synthetic compound. (4b) gives examples of synthetic compounds from verbs which have more than one obligatory internal argument. All such compounds are unacceptable.

- (4) a. *girl-swimming, *child-sleeping, *man-snooring
- b. *shelf-putter, *toy-hander, *book-placer

Lieber (1983) points out that when a verb lacks an internal argument, whether because of its normal argument structure, or because it is in its passive participle form, a synthetic compound can be formed from that verb in which the first stem is interpretable as an adjunct:

- (5) lake-swimming, home-dining, stream-wading,
home grown, hand made

In fact, when internal arguments are satisfiable outside the synthetic compound, the first stem is also interpretable as an adjunct:

- (6) the home-growing of tomatoes
the machine-picking of strawberries

Note further that although examples of synthetic compounds are typically forms in which the second stem is an agentive nominal (*truckdriver*), and *-ing* nominal (*gift-giving*), or a passive participle (*motheaten*), these being the most clearly productive among the synthetic compounds, synthetic compounds are formed from other sorts of nominals as well:

- (7) cost containment, treaty enforcement
waste disposal, truth denial
word coinage, tire slippage
load tolerance, stress endurance

That is, the examples in (7) suggest that the same sort of relationship that is exhibited between first and second stem in synthetic compounds based on *-er*, *-ing*, and passive participle forms shows up between first and second stems where the second stem is among the less productive forms of nominalization (*-ment*, *-al*, *-age*, *-ance*, etc.). In fact, just such a relationship shows up between first and second stem when the second stem is converted from verb to noun with no overt affix:²

- (8) air traffic control, bug fix, clam bake, wienie roast, barium swallow

In fact, as Laurie Bauer (p.c.) has pointed out, these examples shade off into examples in which the second stem is arguably not derived from a corresponding verb at all:

- (9) speech synthesis, data analysis, tea merchant

Yet the relationship between first and second stem is the same as in the cases where the second stem is transparently deverbal. We will return to the issue that this raises in section 4.

1.3. *Endocentric vs. Exocentric Compounds.*

English compounds may be further classified into endocentric and exocentric compounds. Endocentric compounds are those in which both the semantic and syntactic category of the compound corresponds to that of the head. In contrast, exocentric compounds are those in which either semantically or syntactically (or both), the compound as a whole does not bear the characteristics of its head. Among the exocentric compounds are the so-called *bahuvrihi* compounds like *redhead*, *paleface*, and *birdbrain*, all of which denote types of people rather than types of head, face, or brain, respectively. Note that while the usually cited bahuvrihi compounds are root compounds, there are synthetic bahuvrihi compounds as well. For

² The compound noun *bugfix* was cited in the February/March 1991 issue of *Technology Review*, and the compound noun *barium swallow* is used several times in the *Merck Manual* entry on hiatus hernia.

example, *posy-sniffers* and *tree-buggers* do not necessarily sniff posies or hug trees. The former is a disparaging term for a pacifist, and the latter for an environmentalist.

Note also that both the root and the synthetic bahuvrihi compounds are semantically, but not syntactically exocentric. In all cases, the compound as a whole bears the category of the second stem, as is typically the case in English compounds. There are, however, some examples of English compounds which appear to be syntactically exocentric, that is, where the compound as a whole does not bear the category of either of its members. These are forms like *show off*, *runaway*, *gobetween*, and *throwback*. All are nouns, yet they seem to consist of a verb plus a particle or a preposition. In terms of derivation, however, these are arguably not created by compounding at all. Rather, they might be analyzed as cases of conversion (or zero-affixation) from the corresponding verb plus particle construction.

1.4. *Conclusions.*

Sections 1.1-1.3 give an overview of many of the sorts of compounding found in English. Rather than discuss the semantics of these data in great detail here, I refer the reader to the extensive discussion and classification of compounds in Marchand (1969), and to the other sources cited above. Next I will turn to a consideration of some of the issues that surround the analysis of compounding in English.

2. *Criteria for Compoundhood in English.*

Works such as Marchand (1969), Levi (1977), and Bauer (1978) draw attention to the fact that it is difficult to find a single criterion which unequivocally distinguishes compounds from noncompounds in English. The most frequently cited method for distinguishing compounds from phrases is stress. English compounds are typically stressed on the left-hand stem: *blackbird*, *stir-fry*, *truckdriver*, *pasta-eating*, etc. However, all three of these authors, among others, point out that not all compounds display left-hand stress. AA root compounds typically do not: *icy cold*, *blue-green*, and the like have level stress. Similarly, among the synthetic compounds, those formed on the passive participle typically exhibit level stress: *hand-made*, *machine-stitched*, etc. Some NN compounds in fact have the stress pattern typical of phrases, that is, stress on the rightmost stem. Compare, for example, the stress pattern of *apple pie* to that of *apple cake*, or of *stone wall* to *stone mason*. Stress then cannot be taken as a necessary condition for compoundhood.

Nor can factors like spelling, lexicalized meaning, or lack of inflection on the first stem be taken as necessary conditions. It is well known that compounds in English are spelled variously as one word (*blackbird*), as hyphenated words (*pasta-eating*), or as two separate words (*apple cake*). And although compounds frequently have lexicalized meanings (e.g. a *blackboard* does not necessarily have to be black), some forms of compounding are extremely productive and consequently their outputs are perfectly compositional in meaning. It is also frequently pointed out that the nonhead member of an English compound typically may not be inflected (**towelrack*, **leavesgreen*, **babies sit*). But inability to inflect on the first stem cannot be a necessary condition either; there are some compounds in English in which the first stem is inflected as plural, possessive, or both: *children's hour*, *girls' club*, etc.

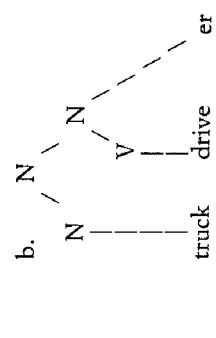
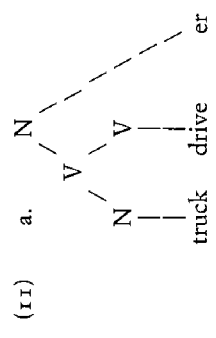
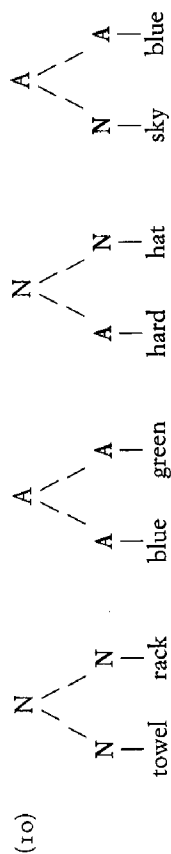
There is one test, however, which seems to work consistently for all sorts of compounds, and which therefore might be proposed as a necessary criterion for compoundhood. That test is inseparability. The elements of a compound in English may not be separated by an intervening modifier of any sort, at least not without eliminating the compound's meaning: **bläck heavy board*, **to stir quickly fry*, **blue-light green*, **truck careful driver*. In any case, it is usually the case that one or more of the possible criteria for compoundhood are present for any given item which is intuitively felt to be a compound in English.

3. The Internal Structure of Compounds.

Several issues surrounding the analysis of compounds in English have to do with their internal structures, specifically with the internal structure of synthetic compounds, the placement of inflection in all types of compounds, and with the availability of compounds for further derivation. We will take up each of these topics in turn.

3.1. The Structure of Synthetic Compounds.

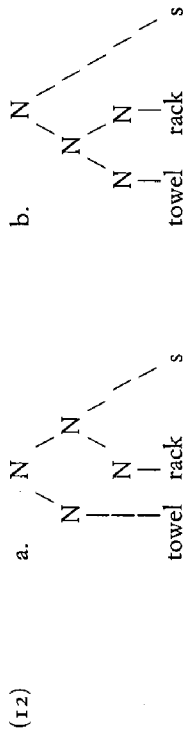
Whereas the internal structure of a root compound is relatively uncontroversial (see the examples in (10)), there are two *a priori* plausible structures for synthetic compounds, as illustrated in (11):



In (11a), the affix *-er* is sister to a verbal compound *truckdrive*, whereas in (11b) it is a sister of *drive* which is then compounded with the noun *truck*. The structure in (11a) has the advantage that it allows us to say that an obligatory complement of a verb is satisfied in the first projection above the verb in synthetic compounds, just as it is in phrasal syntax. Such a structure has in fact been proposed by Lieber (1983). As has been pointed out by Booij (1988), however, this structural analysis suffers from a rather serious drawback in that it claims, at least implicitly, that there should be verbal compounds like *truckdrive* in English. Yet this sort of compound is not formed productively in English. The few verbal compounds that do exist (*babysit*, *aircondition*, and the like) are generally acknowledged to be the result of backformation from synthetic compounds (*babysitter*, *airconditioner*, etc.). In other words, a grammar which generated (11a) as the structure for synthetic compounds would have to overgenerate large numbers of nonoccurring (and impossible) verbal compounds. Works such as Selkirk (1982) and Lieber (1992) therefore choose (11b) as the internal structure for synthetic compounds. Of course, if (11b) is the correct structure for English synthetic compounds, it must still be explained how the first stem in what is essentially a NN compound comes to be interpreted as an argument of the verb. This is an issue we will return to in section 4.

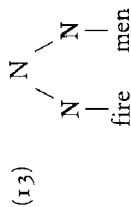
3.2. The Position of Inflection.

Another structural question concerns the position of inflection in compounds. In a plural compound noun such as *towel racks*, is the inflectional morpheme sister to the second stem, as illustrated in (12a), or to the compound as a whole, as in (12b)?

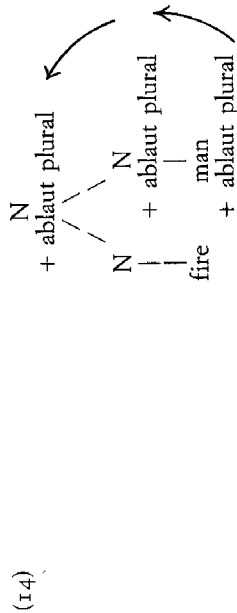


It turns out not to be easy to choose between the two structures. A possible argument in favor of structure (12a) might seem to come from cases in which the second stem has an irregular plural, for example,

fireman ~ *firemen*. Such cases might at first glance suggest that it is the second stem that is pluralized rather than the compound as a whole. But this argument would not go through given most current theories of word formation. In a theory like that of Lieber (1980), irregular inflectional stems are listed in the lexicon; a stem like *men* would be directly available for compounding, and the structure of the resulting compound would merely be that in (13):



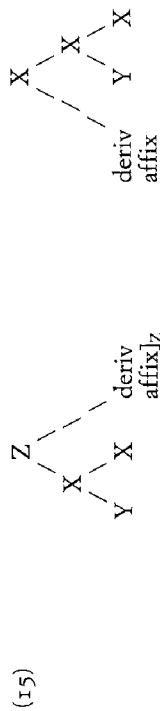
Even in a theory in which irregular inflection is effected by readjustment rules, cases of irregular inflection shed no light on the problem at hand. Assuming that the righthand stem is the head of the compound, all features of this stem would percolate to the node dominating the compound. If the inflectional rule is triggered by such a feature, it could be triggered by the feature on *man* or on the compound *fireman* equally well:



Again, irregular plurals shed no light on the issue raised here. In fact, I know of no solid argument to decide between (12a) and (12b) as the correct structure for inflected compounds, and therefore leave the issue open.

3.3. Derivation Outside Compounding.

Perhaps an equally vexed issue is whether there exist cases in which derivational affixes attach outside of compounds, as illustrated in (15). The schemata in (15) presuppose that suffixes change category and prefixes do not, as is generally the case in English.



where X, Y and Z are lexical categories, and X and Y dominate stems

As has been known at least since Allen (1978), certain derivational affixes do seem to attach outside of compounds. The examples in (16a) are derivations based on synthetic compounds, and those in (16b) on root compounds:

- (16) a. nonhabit-forming, nonmeateater, nonfleabitten, truckdriverish, truckdriverhood, truckdriverlike, habitformingness, extruckdriver
 b. non-icy cold, hounddogfish, hounddoglike, exfireman, icycoldness

The primary reason for believing that the affix is attached outside the compound is that in all cases the only plausible interpretation is one in which the affix has scope over the compound as a whole. Something which is *nonhabit-forming* is not habit-forming; it cannot concern the forming of non-habits. Similarly, an *extruckdriver* is someone who no longer drives trucks, rather than someone who drives vehicles which used to be trucks. In other words, semantic interpretation supports the suggestion that the items in (16) have the structure in (15).

The problem with the assumption that affixes attach outside of compounds, and that interpretation stands as evidence for this comes within a strict theory of Lexical Phonology (for example, Mohanan 1982, Kiparsky 1982, Kaisse & Shaw 1985). If compounding is strictly ordered after both Level I and Level II morphology and phonology, where derivation is added, then examples like (16) should not exist. Moreover, it does not seem possible to solve the 'ordering' problem by allowing a loop which permits compounds to be fed back into Level I, Level II, or both. As the examples in (17) show, while some Level I affixes seem able to attach outside compounds, others do not. Similarly, (18) shows that some Level II affixes may attach outside compounds, while others do not.

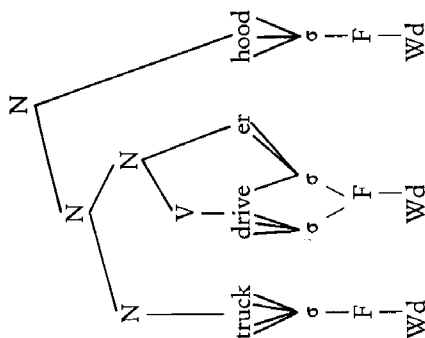
- (17) Level I *inhabit-forming, *inicycold, *icycoldity, set theoretic, transformational grammarian, nuclear physicist

- (18) Level II *unhabit-forming, *unicycold, *dechewinggum,³ *deforemast truckdriverhood, nonicycold, hounddogfish, exfireman

Any looping solution to this problem would inevitably lead to the generation of the unacceptable examples in (17) and (18), as well as to the acceptable ones. But if derivational affixes are not permitted to attach to the outside of compounds, then examples such as these are Bracketing Paradoxes. In terms of phonology, the affix must be attached before compounding takes place, but in terms of semantics, the affix has scope over and thus is in some sense "outside" the compound.

A copious literature has grown up around the problem of Bracketing Paradoxes in English and in other languages. I will not review this literature here (see Booij & Rubach 1984, Pesetsky 1985, Sproat 1985, 1988,

³ This example and the next are from Allen (1978).



The difference between a suffix like *-ian* classed as Level I and one like *-hood* classed as Level II lies in whether the affix forms part of the phonological word with the preceding prosodic constituent or not. Level I affixes do form part of the phonological word with the preceding prosodic constituent, and therefore part of that domain for phonological rules. Level II affixes like *-hood* constitute separate prosodic domains and therefore do not affect stress or internal phonology of preceding prosodic constituents. Given assumptions such as these, we can then answer in the affirmative the question with which we began this section: derivational affixes do attach outside compounds in English, at least with respect to morphosyntactic structure.

4. Compounding and Thematic Relations.

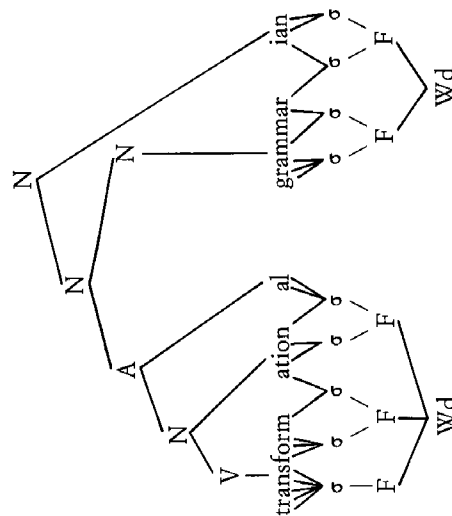
It has long been recognized that the semantic relationships exhibited within synthetic compounds like *truckdriver* or *pasta-eating* are closely related to the semantic relationships that obtain in sentences containing the base verb from which the compound's second stem is derived: *X drives a truck*, *X eats pasta*. Put in a somewhat different way, the first stem in a synthetic compound typically acts as an argument of, or bears a thematic relation (θ -role) to the second stem. A long literature has accumulated which attempts to show how this insight can best be captured within the framework of generative grammar. I will first briefly review the literature and then raise a further issue: whether thematic relationships between the stems of compounds are in fact confined to synthetic compounds, or whether the same sort of relationships arise in root compounds as well.

The earliest treatments of synthetic compounding within generative grammar in fact pose full sentences as the source of synthetic compounds. Lees (1961), for example, creates a compound like *oil-yielding* from an

Hoeksema 1987, Marantz 1988, Cohn 1989, among others), but rather will merely sketch one possible solution to Bracketing Paradoxes concerning compounding and derivation, namely the solution proposed in Booij & Lieber (1990), which is a refinement of a proposal made in Booij & Rubach (1984), and also closely related to proposals made by Sproat (1985, 1988) and Marantz (1988).

The solution proposed in Booij & Lieber (1990) is to assume that items like *transformational grammarian* or *nuclear physicist* have two simultaneous structures: a morphosyntactic structure and a prosodic structure. Morphosyntactic structure is the familiar hierarchically arranged word tree. Prosodic structure is also hierarchically structured, but consists of prosodic units such as syllable (σ), foot (F), and prosodic word (Wd) (see Selkirk 1984, Nespor & Vogel 1986). A word like *transformational grammarian* will have two coexisting representations which are not isomorphic, as illustrated in (19):

(19)



The morphosyntactic structure will have the affix *-ian* attached to the compound *transformational grammar*; from this structure the semantic representation of the compound can be recovered. Phonological rules, however, including the word stress rule in English, will have the phonological word rather than the morphosyntactic word as their domain. Since the suffix *-ian* forms part of the phonological word with *grammar*, stress rules will apply within this domain, as well as within the separate phonological word *transformational*; for the purposes of the phonology, that is, the suffix *-ian* will be treated as most closely associated with the word *grammar*, and not with the compound as a whole.

Compounds like *truckdriverhood* or *bounddoggish* which have Level II affixes outside of compounds will also have two simultaneous structures:

underlying sentence *the well yields oil*. Transformations of various sorts rearrange elements in the base sentence, deleting some along the way. Such an analysis is implicit in Marchand's (1969) treatment of synthetic compounds as well, and refinements of this sort of analysis can be found in Levi (1977) and Bauer (1978). Presumably whatever accounts for the thematic relationships between a verb and its arguments in a sentence will also account for the relationship obtaining in the compound. Traditional such as these, since the sentence is the source of the compound. Traditional transformational derivation of synthetic compounds has serious drawbacks, however, in that the number and complexity of the transformations needed to derive compounds from underlying sentences may be very great.

Consequently, most analyses of synthetic compounds since the late seventies have chosen to express the thematic relationship of the stems of synthetic compounds in other ways. Roeper & Siegel (1978) argue for an analysis in which a *Lexical Transformation*, repeated in (21), rearranges the lexical subcategorization frame of the deverbal stem to derive the synthetic compound:

(21) Compound Rule, Roeper & Siegel (1978:209)

$$[[\text{empty}] + \text{verb} + \text{affix}]_{[X, N + \text{word}]} W_1 \rightarrow$$

1	2	3	4	5
			[[word] + verb + affix] W	
			4	2
			3	5

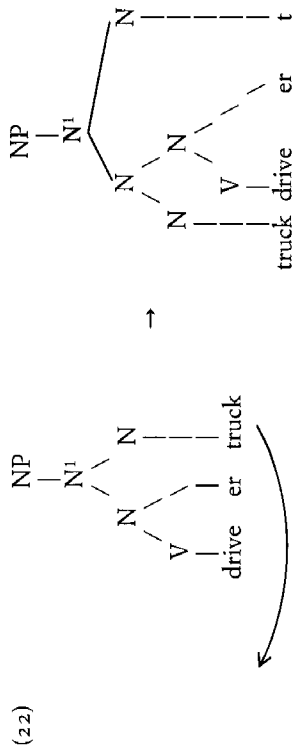
where W ranges over subcategorization frames and X_{+N} stands for lexical categories N, A, Adv

(21) takes a compound with an empty first stem position and fills that position with the word in the position following the compound; this is intended to capture what Roeper & Siegel call *The First Sister Principle*. Lieber (1983) points out several problems with this sort of analysis, chief among them that it requires the addition of a rather complex rule type — the lexical transformation — to the grammar.

Selkirk (1982) proposes to account for the relationship between the first and second stems of a synthetic compound through an interpretive principle that allows the direct internal argument of the verb to be satisfied by the first stem of the compound; this is the *First Order Projection Condition* (Selkirk 1982:37). The FOPC accounts for the interpretation of compounds such as *truckdriver* and *pasta-eating*, as well as for the impossibility of such compounds as **girl-swimming* and **shelf-putting*, and it does so with a simple interpretive principle.

Lieber (1983) proposes a principle of Argument-Linking which has the effect for synthetic compounds that the second stem must be able to link all its internal arguments inside the compound; in effect, her Argument-Linking Principle does roughly the work of Selkirk's FOPC, although it is meant to account for both the relationships exhibited in synthetic compounds, and relationships that obtain (or fail to obtain) in root compounds.

More recent analyses of synthetic compounds such as Roeper (1988) and Lieber (1992) have returned to the idea that synthetic compounds in English have a syntactic derivation. Roeper (1988) likens the formation of synthetic compounds in English to Noun Incorporation in languages like Mohawk (Baker 1988), and treats their formation as an instance of Head Movement. Lieber (1992) forms synthetic compounds in English by Head Movement as well, but for rather different reasons. In her theory, the position of complements (internal arguments) with respect to their heads is set by general X-bar parameters just once per language. Direction of complement placement is expected to hold not only within the syntax, that is, above X^0 , but also within words, that is, below X^0 . See Lieber (1992) for a full explanation of this restriction. Given this restriction and the fact that in English complements generally follow their head, synthetic compounds must be derived as shown in (22):



The N *truck* in this analysis is generated as a complement of *driver* under N^1 , where complements are regularly generated, and receives its θ -role from *driver* in exactly the same way that all complements do. No special principle like the FOPC or Argument-Linking is needed to explain the interpretation or assignment of θ -roles in synthetic compounds. However, since *truck* is an N^0 , and not an NP, it cannot receive case (case is assigned only to a maximal projection), and it is therefore forced to move.⁴ It adjoins to *driver*, where as part of a lexical item it does not require case. Lieber (1992) gives a number of arguments based on language acquisition, the history of English, and on internal facts of English to support this analysis. One such argument can be found in section 5.⁵

Of course, since no special principle of interpretation or assignment of θ -roles is necessary for synthetic compounds within the theory of Lieber (1992), this raises the question whether general principles of θ -assignment play any role in the interpretation and derivation of root compounds as

⁴ Note that if *truck* is generated under a maximal projection (NP), it will receive case through the insertion of the preposition *of* and will not move. In this way, the phrase *driver of the truck* will be generated.

⁵ See Lieber (1992) for a discussion of differences between this analysis and that of Roeper (1988).

well. That is, whereas synthetic compounds have typically been viewed as thematic and root compounds as nonthematic, there is no reason to believe that this has to be the case. This could have several ramifications. First, it leaves open the possibility that the generation of such nonexistent root compound as **putself* or **handbaby* (*hand* a verb) could be ruled out on a thematic basis, since verbs like *put* and *hand* fail to satisfy their argument structure in compounds. Second, it allows Lieber's (1983) explanation of the productivity of NN, NA, AN, and AA compounds to stand; because nouns and adjectives typically don't have argument structures to be satisfied (unless they are deverbal), they may be compounded quite freely. Finally, if general principles of θ -role assignment play a role in the interpretation of root as well as synthetic compounds, an analysis is available for items such as *speech synthesis* and *data analysis* mentioned above in section 1.2. Assuming that the second stems in these compounds are unusual in that they have argument structures (that is, that they take complements even if they are not deverbal), the compounds will be derived by movement, as in (22). They require no special treatment and need not be regarded as oddities.

5. Phrasal Compounds.

We come finally to a last issue concerning compounding in English, namely the analysis of so-called phrasal compounds such as the items in (23):

- (23) the [[Charles and Di]]*[[syndrome]]*
 a [[floor of a birdcage]]*[[taste]]*
 an [[late too much]]*[[headache]]*
 a [[who's the boss]]*[[wink]]*
 his [[God is dead]]*[[theology]]*

Compounds of this sort in various languages have been discussed in Botha (1980), Savini (1983), Toman (1983), Fabb (1984), Hoeksema (1985, 1988), Sproat (1985), Lieber (1992), and elsewhere, and consensus seems to be that they are indeed compounds. In English the first element of such compounds may be any maximal phrase belonging to an open class: NP, VP, AP, PP, and CP. Note that a functional class such as DP cannot occur as first element of a phrasal compound: **a [[the floor of a birdcage]]*[[taste]]*.⁶ Items such as those in (23) satisfy the criteria for compoundhood in English as well as any ordinary compounds do. That is, although not all such compounds exhibit left-hand stress for all speakers, all items like those in (23) satisfy the criterion of inseparability. It is impossible to insert a modifier between the phrasal element and the righthand stem in phrasal compounds — **a [[floor of a birdcage] salty [taste]]*. Note also that phrasal compounding*

⁶ I am assuming here the DP analysis of Abney (1987), in which NP is a complement to D.

is productive in English; the first element in such compounds is not confined to a fixed set of lexicalized expressions, but rather may be any maximal phrase of the categories NP, VP, AP, PP, or CP.

Phrasal compounds are of theoretical interest because they raise profound questions about the organization of the grammar. In particular, they present a problem for any theory of word formation that adheres strictly to the Lexicalist Hypothesis of Lapointe (1980) and others, in that phrases — that is, elements created as part of the syntactic component — seem to occur inside words. Since the phrases in question are typically not lexicalized, the morphology must have access to at least part of the output of the syntax. Lieber (1992) in fact uses the existence of such phrasal compounds to justify a radical departure from lexicalism. Rather than advocating the simple but unexplanatory device of a loop feeding phrases back into a separate morphological component, I suggest that words are not generated in a separate component by separate principles, but instead are created in the syntax by precisely the same mechanisms that sentences are, namely (within GB theory), X-bar theory, theta theory, case theory, and so on. While it would go beyond the scope of this short article to propose and justify such a theory (see Lieber 1992), I will merely outline here what such a theory says about phrasal compounds.

Lieber (1992) argues that one of the word order parameters that needs to be set in each language is a parameter fixing the position of restrictive modifiers with respect to the head. I argue further that this parameter is set for English so that modifiers precede their heads. Such a parameter setting accounts not only for the ordering in phrases like *green bananas*, but also for the ordering in root compounds like *blackboard* or *file cabinet*. X-bar principles allow either a maximal phrase or an X⁰ to occur on either side of a head, and recursion of X⁰ is permitted as well, allowing the correct structural configurations for either phrases or complex words to be generated. Moreover, since there is no separation at all between word syntax and sentence syntax, phrases can appear in the "modifier" position within words; the result is the structure needed for phrasal compounds.

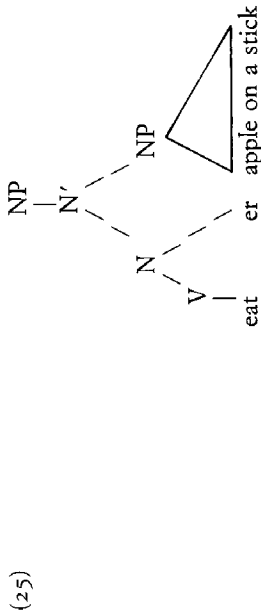
The analysis of phrasal compounds within Lieber (1992) also provides an explanation for some interesting and relatively subtle facts about phrasal compounds. Note that the compounds in (23) are all like root compounds, in that the righthand stem is a simple, not a deverbal noun. Hence we might refer to these as *root phrasal compounds*. This raises the possibility that there might then be phrasal counterparts to synthetic compounds, that is, items with a phrase as a first element and a deverbal noun or adjective as the second element. Yet, as the examples in (24) show, these *synthetic phrasal compounds* are impossible:⁷

⁷ Roeper (1988) also observes the impossibility of synthetic phrasal compounds.

- (24) *an [[apple on a stick][eater]]
 *a [[God is dead][believer]]
 *no [[who's the boss][asking]]
 *some [[ate too much][complaining]]
 *two [[pajama and slipper][packers]]

We must then ask why root phrasal compounds should be possible in English, but not synthetic phrasal compounds.

Recall from section 4 that Lieber (1992) generates ordinary synthetic compounds via movement. Since complements ordinarily follow their heads in English, complements must follow their heads in synthetic compounds as well, at least in d-structure. Since the complement in an ordinary synthetic compound is an X⁰ (see (22)), it cannot be assigned case, and Head Movement is therefore forced. Consider, however, the derivation we would have to assume for synthetic phrasal compounds:



The items in (24) would have to start out with d-structures like that in (25) — that is, since the phrase *apple on a stick* is a complement of *eater*, it must start out following *eater*. However, since the complement in this instance is a maximal phrase, it can receive case. If we assume, with Chomsky (1989), that movement cannot occur unless it is forced, the theory in question thus correctly provides no derivation for synthetic phrasal compounds. The phrase *apple on a stick* stays where it is, and the preposition *of* must be inserted as a case marker. Synthetic phrasal compounds are therefore predicted not to exist.

6. Conclusions.

There are no doubt other interesting issues that will arise with respect to compounding in English. Nevertheless, it seems safe to say that a great deal is known at this point about this very productive form of word formation in English.

Address of the Author:
 English Dept., Hamilton Smith Hall
 University of New Hampshire,
 Durham, NH 03824

References

- Abney, S. (1987), *The English Noun Phrase in its Sentential Aspect*, PhD dissertation, MIT, Cambridge.
- Allen, M. (1978), *Morphological Investigations*, PhD dissertation, University of Connecticut, Storrs.
- Baker, M. C. (1988), *Incorporation: a theory of grammatical function changing*, Chicago, Chicago University Press.
- Bauer, L. (1978), *The Grammar of Nominal Compounding with special reference to Danish, English, and French*, Odense, Odense University Press.
- Booij, G. (1988), "The Relation Between Inheritance and Argument Linking: Deverbal Nouns in Dutch", in M. Everaert, A. Evers, R. Huybregts, and M. Trommelen (eds.), *Morphology and Modularity*, Foris, Dordrecht: 57-74.
- Booij, G. & R. Lieber (1990), "On the Simultaneity of Morphological and Prosodic Structure", ms., Free University of Amsterdam and University of New Hampshire.
- Booij, G. & J. Rubach (1984), "Morphological and Prosodic Domains in Lexical Phonology", *Phonology Yearbook*, 1: 1-28.
- Botha, R. (1980), "Word-Based Morphology and Synthetic Compounding", *Stellenbosch Papers in Linguistics*, No. 5, University of Stellenbosch.
- Botha, R. (1984), *Morphological Mechanisms*, Oxford, Pergamon Press.
- Chomsky, N. (1989), "Some Notes on Economy of Derivation and Representation", *MIT Working Papers in Linguistics*, 10: 43-77.
- Cohn, A. (1989), "Stress in Indonesian and Bracketing Paradoxes", *Natural Language and Linguistic Theory*, 7: 167-216.
- Fabb, N. (1984), *Syntactic Affixation*, PhD dissertation, MIT, Cambridge.
- Grimshaw, J. (1990), *Argument Structure*, MIT Press, Cambridge.
- Hoeksema, J. (1985), *Categorical Morphology*, Garland Press, New York.
- Hoeksema, J. (1987), "Relating Word Structure and Logical Form", *Linguistic Inquiry*, 18: 119-125.
- Hoeksema, J. (1988), "Head Types in Morpho-Syntax", *Yearbook of Morphology*, 1: 123-138.
- Kaisse, E. & P. Shaw (1985), "On the theory of Lexical Phonology", *Phonology Yearbook*, 2: 1-30.

- Kiparsky, P. (1982), "From Cyclic Phonology to Lexical Phonology", in H. van der Hulst & N. Smith (eds.), *The Structure of Phonological Representations*, Foris, Dordrecht.
- Lapointe, S. (1980), *A Theory of Grammatical Agreement*, PhD dissertation, University of Massachusetts, Amherst.
- Lees, R. (1961), *The Grammar of English Nominalizations*, Indiana University, Bloomington, and Mouton, The Hague.
- Levi, J. (1977), *The Syntax and Semantics of Complex Nominals*, New York, Academic Press.
- Lieber, R. (1980), *On the Organization of the Lexicon*, PhD dissertation, MIT, Cambridge [reprinted 1981 by Indiana University Linguistics Club, and 1990 by Garland Publications].
- Lieber, R. (1983), "Argument-Linking and Compounds in English", *Linguistic Inquiry*, 14: 251-286.
- Lieber, R. (1992), *Deconstructing Morphology: Word Formation in Syntactic Theory*, University of Chicago Press, Chicago.
- Marantz, A. (1988), "Apparent Exceptions to the Projection Principle", in M. Everaert, A. Evers, R. Huybregts, and M. Trommelen (eds.), *Morphology and Modularity*, Foris, Dordrecht.
- Marchand, H. (1969), *The Categories and Types of Present-Day English Word Formation*, C. H. Beck'sche Verlagsbuchhandlung, München.
- Mohanan, K. P. (1982), *Lexical Phonology*, PhD dissertation, MIT, Cambridge.
- Nespor, M. & I. Vogel (1986), *Prosodic Phonology*, Foris, Dordrecht.
- Pesetsky, D. (1985), "Morphology and Logical Form", *Linguistic Inquiry*, 16: 193-248.
- Roeper, T. (1988), "Compound Syntax and Head Movement", *Yearbook of Morphology*, 1: 187-228.
- Roeper, T. & M. Siegel (1978), "A Lexical Transformation for Verbal Compounds", *Linguistic Inquiry*, 9: 197-260.
- Savini, M. (1983), "Phrasal Compounds in Afrikaans", MA Thesis, University of Stellenbosch.
- Selkirk, E. (1982), *The Syntax of Words*, MIT Press, Cambridge.
- Selkirk, E. (1984), *Phonology and Syntax*, MIT Press, Cambridge.
- Sproat, R. (1985), *On Deriving the Lexicon*, PhD dissertation, MIT, Cambridge.
- Sproat, R. (1988), "Bracketing Paradoxes, Cliticization and Other Topics: the Mapping Between Syntactic and Phonological Structure", in M. Everaert, A. Evers, R. Huybregts, and M. Trommelen (eds.), *Morphology and Modularity*, Foris, Dordrecht.
- Toman, J. (1983), *Wortsyntax*, Max Niemeyer Verlag, Tübingen.