Testing hypotheses about compound stress assignment in English: a corpus-based investigation

Ingo Plag, Gero Kunter, Sabine Lappe & Maria Braun

Universität Siegen

October 27, 2006

Corresponding author:

Ingo Plag
English Linguistics
Fachbereich 3
Universitaet Siegen
Adolf-Reichwein-Str. 2
D-57068 Siegen

http://www.uni-siegen.de/~engspra/
tel. +49-(0)271-740-2560
tel. +49-(0)271-740-2349 (secretary)
fax tel. +49-(0)271-740-3246
e-mail: plag@anglistik.uni-siegen.de

Testing hypotheses about compound stress assignment in English: a corpus-based investigation

Abstract

This paper tests three factors that have been held to be responsible for the variable stress behavior of noun-noun constructs in English: argument structure, semantics, and analogy. In a large-scale investigation of some 4500 compounds extracted from the CELEX lexical data base (Baayen et al. 1995), we show that traditional claims about noun-noun stress cannot be upheld. Argument structure plays a role only with synthetic compounds ending in the agentive suffix *–er*. The semantic categories and relations assumed in the literature to trigger rightward stress do not show the expected effects. As an alternative to the rule-based approaches, the data were modeled computationally and probabilistically using a memory-based analogical algorithm and logistic regression, respectively. It turns out that probabilistic models and analogical algorithms are more successful in predicting stress assignment correctly than any of the rules proposed in the literature. The behavior of the analogical model suggests that the left and right constituent are the most important factor in compound stress assignment. This is in line with recent findings on the semi-regular behavior of compounds in other languages.

1. Introduction[1]

The present paper deals with an area of grammar which is more variable than generally assumed: stress assignment in English noun-noun (NN) compounds. In general, it has often been claimed that compounds tend to have a stress pattern that is different from that of phrases. This is especially true for nominal compounds, the class of compounds that is most productive (e.g. Plag 2003: 145). While phrases tend to be stressed phrase-finally, compounds tend to be stressed on the first element. This systematic difference is captured in the so-called nuclear stress and compound stress rules (Chomsky and Halle 1968:17). Phonetic studies (e.g. Farnetani and Cosi 1988, Ingram et al. 2003) have shown in addition that segmentally identical phrases and compounds (such as *bláckboard* vs. *black bóard*) differ significantly not only in their stress pattern, but also in length, with phrases being generally longer than the corresponding compounds. While the compound stress rule apparently makes correct predictions for what seems to be the majority of nominal compounds, it has been pointed out, e.g. by Kingdon (1958), Fudge (1984), Liberman and Sproat (1992), Bauer (1998), Olsen (2000, 2001), and Giegerich (2004), that there are also numerous exceptions to the rule. Some of these forms are listed in (1). The most prominent syllable is marked by an acute accent on the vowel.

(1) geologist-astrónomer
apricot crúmble
Boston márathon
aluminum fóil
apple píe
Michigan hóspital
Penny Láne
May flówers
scholar-áctivist
Madison Ávenue
summer níght
silk tíe

In view of this situation, the obvious question is how we can account for the variability in stress assignment of noun-noun constructs. Basically, one finds three kinds of hypotheses that are spelled out in the literature to different degrees of explicitness. These hypotheses, which will be discussed in more detail shortly, refer to either structural, semantic, or analogical factors that are held responsible for the

[1] Acknowledgements to be added after review.

stress of NN-constructs. It has to be stated, however, that systematic empirical or experimental work on the variability of compound stress is scarce, and the hypotheses mentioned were develped the basis of data that had the following rather questionable properties. First, the provenance of the data remained obscure. Authors generally did not say where they took their data from. The selection of data does not seem in any way systematic but more designed to prove the point of the respective author. The second problem is that the amount of data is usually quite small, ranging from only a handful of pertinent examples to a few hundred forms. The third problem is that most of the studies do not discuss the details of their methodological decisions, such as the assignment of particular examples to a given analytical category.

In sum, there is a need for a large-scale empirical investigation of compound stress variability using an independently gathered set of data. The present paper will provide such a study. We will present the results of the investigation of all noun-noun compounds (some 4500 types) extracted from the CELEX lexical data base (Baayen et al. 1995). It will be shown that traditional claims about noun-noun stress cannot be upheld. As an alternative to the rule-based approaches, we will model the data computationally using a memory-based analogical algorithm, and probabilistically using regression analysis. Overall, it turns out that the probabilistic and analogical models are more successful in predicting stress assignment correctly than any of the rules proposed in the literature. The behavior of the analogical model suggests that the left and right constituent are the most important factor in compound stress assignment.

Before we turn to the discussion of the hypotheses to be tested, a word is in order with regard to the notorious problem of whether NN constructions should be analyzed as compounds or phrases. Since we will use the CELEX lexical database, this decision has already been taken by the compilers of that source. The constructs that we investigate were considered words, hence compounds, by the compilers, since CELEX only contains words, and not phrases.

The structure of the paper is as follows. Section 2 reviews the three hypotheses on the variability of compound stress mentioned above. In section 3 we describe the CELEX lexical data base and our data coding procedure, discussing the methodological problems involved. Section 4 presents the results for the structural hypothesis, section 5 for the semantic hypothesis and section 6 for the analogical hypothesis. This is followed by the final discussion and conclusion in section 7.

2. Three hypotheses on stress assignment to compounds

Three types of approaches have been taken to account for the puzzling facts of variable NN stress. The first is what Plag (2006) has called the 'structural hypothesis'. Proponents of this hypothesis (e.g. Bloomfield 1933, Lees 1963, Marchand 1969 or Payne/Huddleston 2002) maintain that compounds are regularly left-stressed, and that word combinations with rightward stress cannot be compounds, which raises the question of what else such structures could be. One natural possibility is to consider such forms to be phrases. However, such an approach would face the problem of explaining why not all forms that have the same superficial structure, i.e. NN, are phrases. Second, one would like to have independent criteria coinciding with stress in order to say whether something is a lexical entity (i.e. a compound) or a syntactic entity (i.e. a phrase). This is, however, often impossible: apart from stress itself, there seems to be no independent argument for claiming that *Mádison Street* should be a compound, whereas *Madison Ávenue* (or *Madison Róad*, for that matter) should be a phrase. Both kinds of constructs seem to have the same internal structure, both show the same meaning relationship between their respective constituents, both are right-headed, and it is only in their stress patterns that they differ. Spencer (2003) also argues that we find compounds with phrasal stress and phrases with compound stress, and hence that stress is more related to lexicalization patterns than to structural differences, a point taken up by Giegerich (2004, to be discussed in more detail shortly). A final problem for the phrasal analysis is the fact that the rightward stress pattern seems often triggered by analogy to other combinations with the same rightward element. This can only happen if the forms on which the analogy is based are stored in the mental lexicon. And storage in the mental lexicon is something we would typically expect from words (i.e. compounds), and only exceptionally from phrases (as in the case of *jack-in-the-box*).

Most recently, Giegerich (2004) has proposed a new variant of the structural hypothesis. On the basis of the fact that in English syntax complements follow the head, he argues that, due to the order of elements, complement-head structures like *trúck driver* cannot be syntactic phrases, hence must be compounds, hence are left-stressed. Modifier-head structures such as *steel brídge* display the same word order as corresponding modifier-head phrases (cf. *wooden brídge*), hence are syntactic structures and regularly right-stressed.[2]

This means, however, that many existing modifier-head structures are in fact not stressed in the predicted way, since they are left-stressed (e.g. *ópera glasses, táble cloth*). Such aberrant behavior, is, according to Giegerich, the result of lexicalization. The problem with this idea is that lexicalization is, first, not a categorical notion, but rather a gradual one, and, second, that it is not exactly clear how it can be decided whether a given item is lexicalized or not. For compounds, four criteria come to mind: frequency, spelling, semantic transparency, and phonological transparency. In this study we will use frequency and spelling as indicators of lexicalization. Higher frequency indicates a higher degree of lexicalization, and one-word spellings should also be most prevalent with lexicalized compounds, while less lexicalized compounds should prefer two-word spellings.

Lexicalization as an explanation for leftward stress makes interesting predictions. With regard to corpora data, we should expect that the amount of leftward-stressed compounds should vary according to frequency.[3] Thus, we should find more modifier-head structures with leftward stress among the more frequent items. In addition, we would expect a higher proportion of left-stressed compounds

[2] Giegerich characterizes modifier-head structures in terms of their lack of argument-predicate semantics. We prefer the term 'argument-head' instead of 'argument-predicate' in the context of this paper because of its parallelism with 'modifier-head'.

[3] Cf. Lipka's definition, according to which lexicalization "is defined as the process by which complex lexemes tend to become a single unit, with a specific content, *through frequent use*" (1994:2165, my emphasis). Bauer (1983:51) mentions irregular stress assignment in English derivatives and Danish compounds as prototypical cases of (phonological) lexicalization. See also Adams (1973:59), who writes that "in established NPs *which are used frequently* and over a period of time the nucleus tends to shift from the second to the first element although this does not always happen " (our emphasis).

among those spelled as one word than among those spelled as two words, with hyphenated compounds being somewhere in between.

Furthermore, the structural hypothesis predicts that we should never find rightward stress among those NN constructs that exhibit complement-head order. This is, however, not true, as pointed out by Giegerich himself, who cites *Tory léader* as a counterexample. The structural hypothesis also entails that compounds with the same rightward constituent exhibit different stress patterns, depending on whether the leftward constituent is an argument or a modifier. Pairs such as *yárd sale* vs. *bóok sale* (or *trúck driver* vs. *Súnday driver*) suggest that this prediction is probably not always in accordance with the data. In such cases lexicalization would have to kick in to explain leftward stress.

In any case, none of these predictions has ever been systematically tested against larger amounts of data. In a recent experimental study, Plag (2006) found the expected argument-structure effect, but no lexicalization effect (based on frequency). Novel, i.e. newly invented, modifier-head compounds showed the same type of variability in stress behavior as existing modifier-head compounds. Argument-head compounds thus behaved as expected, while for modifier-head compounds the hypothesis did not make the right predictions.

Before turning to the discussion of what we call the 'semantic hypothesis' we would like to point out that what has been labeled 'structural hypothesis' is the hypothesis that rests largely on the argument-modifier distinction. Although this distinction clearly has strong semantic implications, there are, as pointed out above, crucial structural facts that correlate with this distinction. This is our reason for calling the hypothesis structural, athough the underlying distinction might be semantic.

The second approach to variable compound stress is what can be called the semantic hypothesis. A number of scholars have argued that words with rightward stress such as those in (1) above are systematic exceptions to the compound stress rule (e.g. Sampson 1980, Fudge 1984, Ladd 1984, Liberman and Sproat 1992, Olsen 2000, 2001, Spencer 2003). Although these authors differ slightly in details of their respective approaches, they all argue that rightward prominence is restricted to only a limited number of more or less well-defined types of meaning categories and relationships. For example, compounds like *geologist-astrónomer* and *scholar-áctivist* are copulative compounds, and these are uncontroversially and regularly right-stressed.[4] Other meaning relationships that are often, if not typically, accompanied by rightward stress are temporal (e.g. *summer níght*), locative (e.g. *Boston márathon*), and causative, the latter of which is usually paraphrased as 'made of' (as in *aluminum fóil, silk tíe*), or 'created by' (as in *a Shakespeare sónnet, a Mahler sýmphony*). It is, however, unclear how accurate the membership in a given class can really predict the locus of stress. The leftward stress on *súmmer school, súmmer camp* or *dáy job*, for example, violates Fudge's (1984: 144ff.) generalization that NNs in which N_1 refers to a period or point of time are right-stressed. Furthermore, it is unclear how many, and which, semantic classes should be set up to account for all the putative exceptions to the compound stress rule (see also Bauer 1998:71 on this point). Finally, semantically very similar compounds can behave differently in terms of stress assignment (*Fífth Street* vs. *Fifth Ávenue*). And again, we have to state that, apart from the copulative compounds (Olsen 2001) and compounds expressing an authorship relation (Plag 2006), detailed and systematic empirical studies are lacking for the classes postulated to trigger rightward stress.

In the afore-mentioned experimental study by Plag (2006) it was tested whether the semantic hypothesis makes the right predictions for compounds with an authorship relation. Testing the authorship relation (as in *Kauffmann sonata*) against a relation that is not predicted to trigger righthand stress (as in *Twilight Sonata*), it turned out that the data show either no effect, or show an effect in the opposite direction of what the semantic hypothesis would have predicted.

Note that we use the label 'semantic hypothesis' in this paper to refer to approaches that set up semantic categories and correlate these with stress patterns. Although these approaches actually never refer explicitly to the modifier-argument distinction, the semantic categories that are alleged to produce rightward stress would all involve modifier-head compounds, but never argument-head compounds.

[4] Even this nice generalization has its (apparently very few) exceptions, for example *mán-servant*, which is left-stressed.

Thus, structural and semantic hypothesis converge on the point that they expect rightward stress to be largely restricted to modifier-head compounds.

Finally, a third approach can be taken which draws on the idea of analogy and hypothesizes that stress assignment is generally based on analogy to existing NN constructions in the mental lexicon. Plag (2003:139) mentions the textbook examples of *street* vs. *avenue* compounds as a clear case of analogy. All street names involving *street* as their right-hand constituent pattern alike in having leftward stress (e.g. *Óxford Street*, *Máin Street, Fóurth Street*), while all combinations with, for example, *avenue* as right-hand constituent pattern alike in having rightward stress (e.g. *Fifth Ávenue, Madison Ávenue*). Schmerling (1971: 56) provides more examples of this kind, arguing that many compounds choose their stress pattern in analogy to combinations that have the same head, i.e. rightward constituent. It is, however, unclear how far such an analogical approach can reach. Along similar lines, Spencer (2003: 331) proposes that "stress patterns are in many cases determined by (admittedly vague) semantic 'constructions' defined over collections of similar lexical entries." In a similar vein, Ladd (1984) proposes a destressing account of compound stress which would explain the analogical effects triggered by the same rightward constituents as basically semantico-pragmatic effects.

What is considered the effect of lexicalization in some approaches would emerge naturally in an analogical system, in which existing (i.e. lexicalized) compounds influence new (i.e. non-lexicalized) compounds to behave similarly. This raises the question on which basis similarity could be computed (cf. also Liberman and Sproat 1992: 176 on this point). In principle, any property could serve that purpose, for example, the number of syllables of the right constituent, the semantic properties of the left constituent, or, perhaps absurdly, the third segment of the left constituent, or a combination of these. One rather simple assumption to start out with is that it is the left or right constituent that is responsible for the choice of the stress pattern. Given, for example, a set of compounds with the same right constituent, we would first expect that the vast majority of items in that set are stressed in a certain way, e.g. leftward, and that any novel form with that right-hand constituent will also receive leftward stress. A more sophisticated analogical model would incorporate of course more, and different types of linguistic information (phonological, semantic, structural, frequential).

Plag (2006) found a robust effect of the right constituent on the stress of the novel compounds used in the experiment, irrespective of the semantic relation expressed by the compound. However, other potential factors playing a role in analogy were not investigated in that study.

At this point a note is in order on the notion of analogy as used in different traditions. The traditional notion of analogy has been rightly criticized by many because it is difficult to see how any falsifiable prediction might be obtained with it. For instance, in Goldberg and Jackendoff (2005: 475) we still find the statement that 'analogy is notoriously difficult to constrain'. Recent work in computational morphology has shown, however, that a formal, constrained, and computationally tractable notion is available that offers new ways of understanding the ways in which linguistic rules actually work. Such formal analogical models have been quite successful in predicting both regular and irregular morphology in general, and variable compound behavior in particular. For example, Krott and her collaborators (Krott et al. 2001, Krott et al. 2002, Krott et al. 2004) analyzed the semi-regular behavior of the linking morphemes in Dutch compounds in terms of analogy, using a memory-based analogical learning algorithm (TiMBL, Daelemans et al. 2000).[5] They compared the algorithm's performance with that of native speakers in an experiment with novel compounds and found that the variable occurrence of the three linking morphemes in Dutch compounds is much better accounted for by a dynamic analogical mechanism than by traditional symbolic rules. Although the analogical hypothesis has been evoked here and there (and quite informally) in treatments of English compounds, it has never been tested empirically or formally modelled (cf. Spencer's above-cited remark on the vagueness of possible analogical sets).

To summarize, there are three reasonable hypotheses available to account for the variability of NN stress, all of which are in some sense problematic and all of which are still in need of serious empirical testing. One way to do this is to carry out

[5] Analogical effects in compound interpretation have been shown to exist by Gagné and her collaborators (e.g. Gagné and Shoben 1997, Gagné 2001).

experimental studies such as Plag's (2006), in which the data can be carefully controlled for the different potential factors involved. The present investigation takes a different approach and uses the CELEX lexical database.

3. Methodology: general remarks

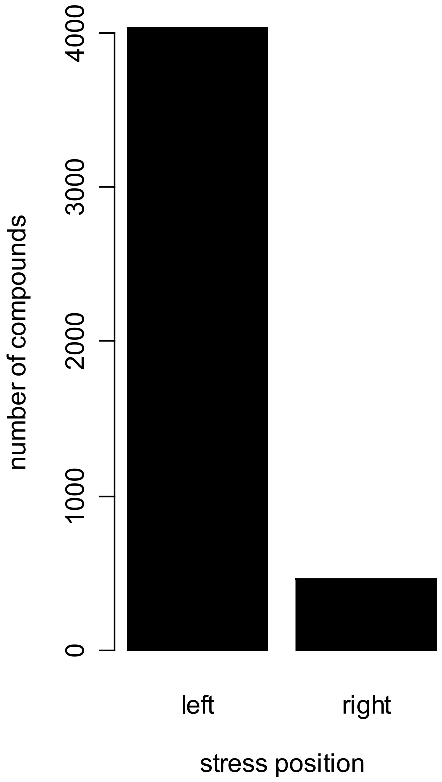
CELEX is a lexical database which contains data from German, (British) English and Dutch, and has been successfully employed in linguistic and psycholinguistic research, including compounds (e.g. Krott et al. 2001). Apart from orthographic features, the CELEX database comprises representations of the phonological, morphological, syntactic and frequency properties of lemmata. In this study we use the English part of CELEX, which has been compiled on the basis of dictionary data and text corpus data. The dictionary data come from the *Oxford Advanced Learner's Dictionary* (1974, 41,000 lemmata) and from the *Longman Dictionary of Contemporary English* (1978, 53,000 lemmata). The text corpus data come from the COBUILD corpus, which contains 17.9 million word tokens. 92% of the word types attested in COBUILD were incorporated into CELEX. The frequency information given in CELEX is based on the COBUILD frequencies. Overall, CELEX contains lexical information about 52,446 lemmata, which represent 160,594 word forms.

From the set of lemmata we selected all words that had two (and only two) nouns as morphological constituents. This gave us a set of 4491 NN compounds. For each of the compounds we extracted the information about stress and frequency. We also coded each compound according to the categories held to be responsible for stress assignment in the literature (and some more, to be discussed below). For those variables where categorization proved to be problematic due to the ill-defined nature of the categories mentioned in the literature, each compound was coded independently by two raters and we analyzed only that subset of the data where the two raters came up with the same categorization. Overall, three raters were enganged in the coding, all of them holding both an MA and a PhD in English linguistics.

To test the structural and semantic hypotheses we modeled the data statistically using logistic regression, and compared the predictive accuracy of our model with that of the two hypotheses. To test the analogical hypothesis we modeled the data using a memory-based learner (TiMBL 5.1, Daelemans et al. 2004). Further details of the methodologies employed will be discussed as we go along.

Before turning to the individual hypotheses let us take a first look at the data. The overall distribution of stresses for the CELEX NN compounds is given in the following bar graph.

Figure 1: Overall distribution of stress in NN compounds (N = 4491)



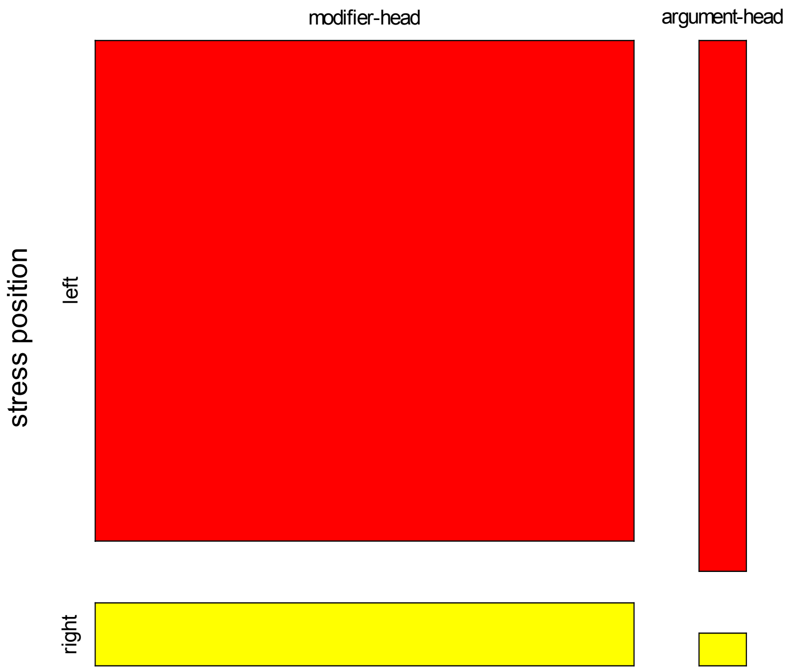
	left stress	**right stress**
number of observations	4029	462
percentage	89.7	10.3

We see that roughly 90 percent of the compounds in CELEX are given as left-stressed, and 10 percent as right-stressed. In the following sections we will investigate the nature of this variation in detail.

4. Testing the structural hypothesis

If we first take a look at the role of the argument structure distinction, we want to take into account only those compounds where both ratings agreed. Thus for a compound to be included here, both raters had to assign the same structural analysis, for example 'argument-head'. This reduces our data set to 4139 compounds. The mosaic plot in figure 2 shows the distribution of stress by structure:

Figure 2: Distribution of stress by structure (N = 4139)



stress	modifier-head	argument-head
left	3387	314
right	419	19

Mosaic plots represent the number of observations in each subset of the data as an area. We can thus see that the majority of compounds has a modifier-head structure, and that the proportion of right stresses is lower with argument-head compounds. A chi-square analysis reveals that the difference between modifier-head and argument-head compounds is statistically significant (χ^2 = 8.55, *df* = 1, p = 0.003457, φ = 0.05). Although the effect goes in the direction expected under the structural hypothesis, the hypothesis that modifier-head structures be right-stressed is clearly falsified, since of the 3806 modifier-head compounds, only 11 percent are right-stressed.

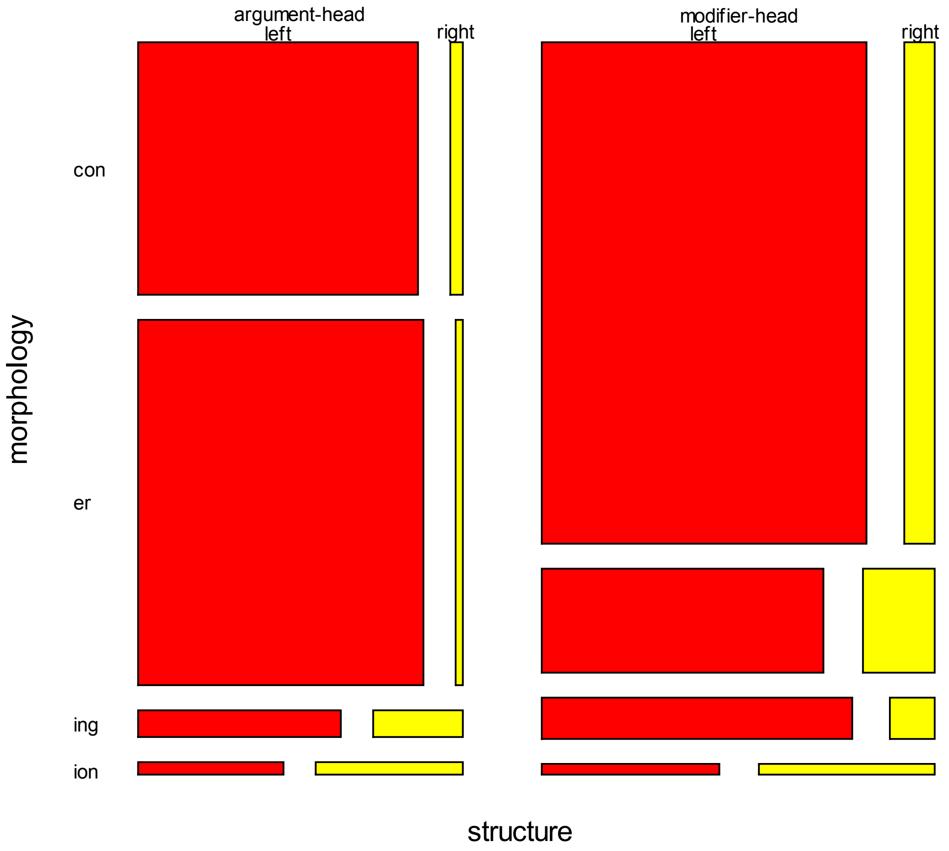
In order to take a closer look at what is going on here we coded the morphological makeup of the deverbal head nouns and investigated whether we would find an interaction of suffix and argument structure. Under the structural hypothesis we should expect that there be significant differences also between those argument-head compounds and modifier-head compounds that share the same head morpheme. The following table gives examples of the kinds of combinations we found in the data:

Table 1:

morphology of head	**argument-head**	**modifier-head**
conversion	*fish slice*	*side glance*
-er	*squadron leader*	*belly dancer*
-ing	*trend setting*	*sea-bathing*
-ion	*blood transfusion*	*Mercator projection*

We also found a few heads that ended in the deverbal suffixes *–age, –al,* and *–ance*, but these were too rare to be included in the statistical analysis. Overall, the heads of 683 items contained one of the suffixes shown in table 1 as the outermost suffix. The distribution of stresses for this set of data is given in figure 3:

Figure 3: Interaction of head morphology, argument structure and stress



An inspection of the plot already shows that for conversion, *-ing*, and *–ion*, the distribution of stress does not seem to differ according to the modifier-argument distinction. A logistic regression analysis of the interaction of argument structure and righthand-head morpheme indeed reveals a significant effect only for those compounds that have *–er* as their right-hand head morpheme (p = 0.0375, C = 0.723). This restriction of the argument structure effect to *–er* compounds is the same as the one recently found by Plag et al. (2006) in a study using an American speech corpus (Boston University Radio Speech Corpus, Ostendorf et al. 1996). These findings can be interpreted in such a way that the argument-structure effect hypothesized in the literature is in fact an effect of only one particular subgroup of synthetic compounds, those ending in *-er*. Not surprisingly, this is the subgroup that is almost exclusively discussed in the literature, while the other subgroups are being largely ignored.

Let us compare the performance of our logistic regression model with that of the stress assignment rule of the structural hypothesis. We used our logistic regression model to calculate the probability of right stress for each compound on the basis of the variables 'argument structure' and 'morphology of the head'. If the probablity of right stress was < 0.5 for a given item, we interpreted this item as left-stressed, and as right-stressed if otherwise. These probabilistic predictions were then compared to the stress positions found in CELEX – a match was counted as a correct prediction. We also estimated the accuracy of the structural hypothesis by categorically assigning left stress to argument-head compounds and right stress to modifier-head compounds, and then comparing these stresses with the CELEX stresses. The following table illustrates the methodology and gives the results for the logistic regression model.

Table 2: accuracy of predictions, logistic regression model

	Total	**correct predictions**	**incorrect predictions**	**prediction accuracy**
CELEX has left stress	623	617	6	99.0%
CELEX has right stress	58	6	52	10.3%
Total	681	623	58	91.5%

The following table 3 compares the accuracies of the model and the structural hypothesis:

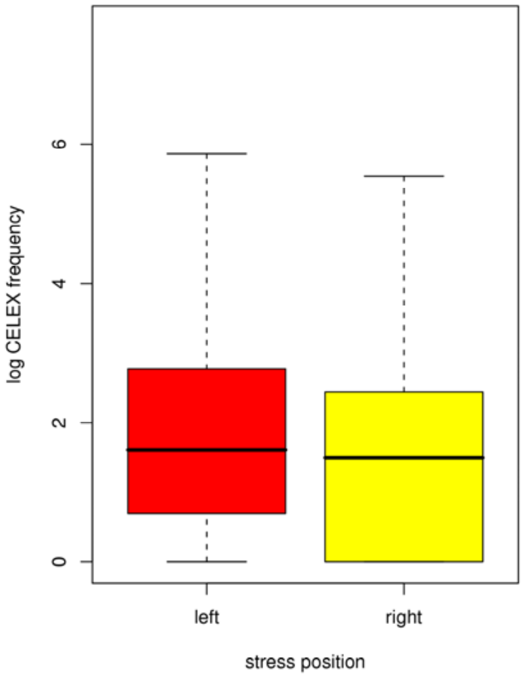
Table 3: Accuracy of prediction, structural hypothesis

prediction	**regression model accuracy**	**structural hypothesis accuracy**
of left stresses	99.0%	46.9%
of right stresses	10.3%	72.4%
total	91.5%	49.0%

The overall accuracy is far better for the logistic regression model (91.5% vs. 49.0%). The structural hypothesis captures exsisting right stresses to a relatively high degree of 72.4% (as against only 10.3% correct right stresses in regression), but it also assigns incorrectly right stresses to the large number of modifier-head compounds that are actually left-stressed. In view of this problem the obvious escape hatch for the structural hypothesis is lexicalization, to which we now turn.

We will first investigate lexicalization using frequency. The problem with the CELEX frequencies is that many compounds in CELEX are taken from the dictionaries and are not attested in the COBUILD corpus, so that they are listed with a frequency of zero. We took only those compounds whose frequency is larger than zero, which gives us still 2118 observations. For this subset we do not find the expected lexicalization effect. Compounds with leftward stress do not have a significantly higher CELEX frequency (*t* (2128) < 1) than compounds with rightward stress. This is illustrated in the boxplot in figure 4:

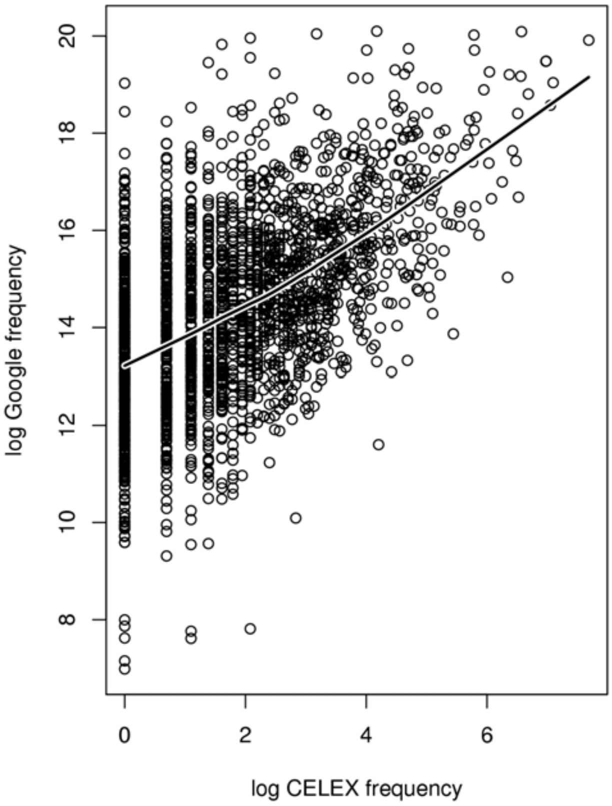
Figure 4: Stress position and CELEX frequency (all items with frequency > 0)



There is no interaction between stress position and argument structure (*F* (1, 2126) < 1), which means that neither the modifier-head compounds nor the argument head compounds show the expected effect. This means that, contra the structural hypothesis, there is no lexicalization effect observable (if we use the CELEX frequencies).

In order to overcome the difficulty of having lost half of the data points due to lacking CELEX frequencies, we also took a look at the log Google frequencies of all compounds. First we checked the reliability of the Google frequencies[6] by correlating them with the CELEX frequencies, which come, as mentioned above, from a controlled corpus. The reliability of the Google frequencies was confirmed by their strong correlation (Spearman's ρ = 0.511, $p < 0.05$) with the CELEX frequencies. Only for items with a log CELEX frequency lower than log (11) = 2.398 does the relation deviate from linearity, as shown by the scatterplot smoother in figure 5:[7]

Figure 5: Google log frequency by CELEX frequency

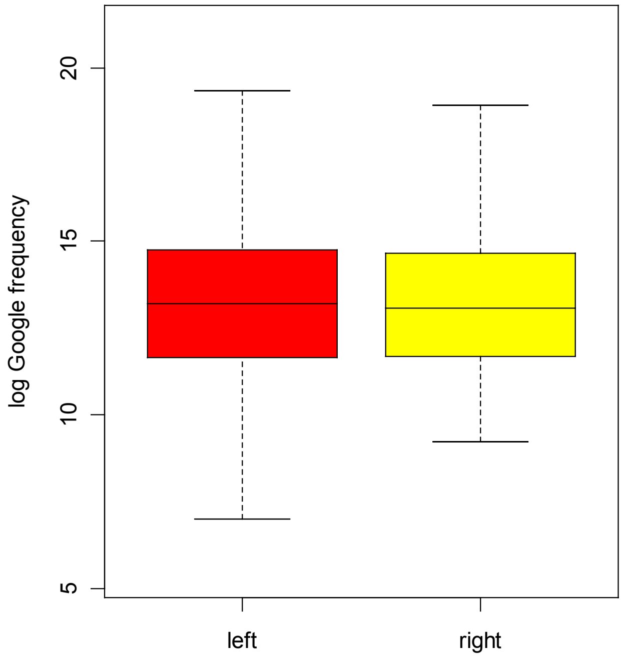


[6] See Plag (2006: 159) for a detailed discussion of the problems involved in using Google for investigating compound frequencies.

[7] Given the wide usage of *newsletter* in an Internet context, and the resulting extremely high Google frequency, this item was excluded from the comparison.

Having established the reliability of the Google frequencies, we repeated the frequency analysis from above with the whole set of compounds in CELEX, now with their Google frequencies. Using the Google frequencies does not improve the situation for the structural hypothesis, however. We still do not find the expected effect (*t* (4469) < 1). Consider figure 6 for illustration:

Figure 6: Stress position by Google frequency

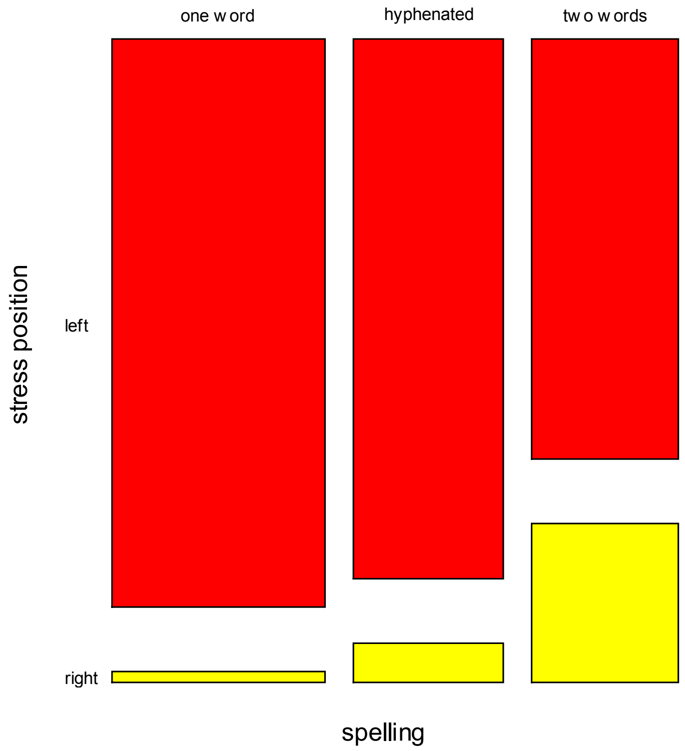


Again, there is no interaction between stress position and argument structure (*F* (1, 4466) = 2.801, p > 0.05), which means that neither the modifier-head compounds nor the argument head compounds show the expected effect. Contra the structural hypothesis, there is no lexicalization effect observable.

Finally, we investigate the idea that spelling may be an indicator of lexicalization, and can thus be used to test the structural hypothesis. It can be assumed that one-word spellings should be most prevalent with lexicalized compounds, while less lexicalized compounds should prefer two-word spellings. According to the structural hypothesis we would therefore expect the proportion of right stresses to be highest among the two-word compounds, lower among the hyphenated compounds, and lowest among the compounds spelled as one orthographic word. Indeed this effect can be found (χ^2 = 512.08, *df* = 2, *p* < 0.01).

Figure 7 nicely illustrates the trend: the tighter the orthography, the more likely becomes leftward stress.

Figure 7: Stress by spelling



However, when examining the difference between argument-head and modifier-head compounds, an analysis of deviance of a generalized linear model revealed no significant reduction of the residuals for the interaction between spelling and stress position (logit, df = 2, p = 0.1), so that, contra the hypothesis, we find a general lexicalization effect, but not one that is restricted to modifier-head compounds.

To summarize our results for the structural hypothesis, we can say that this hypothesis is not very successful in predicting compound stress. The argument structure effect is restricted to compounds ending in *–er*, and the predicted lexicalization effect is not measurable (when using frequency as a correlate) and is not restricted to modifier-head compounds (when using spelling). Let us turn to the semantic hypothesis and see whether it fares better.

5. Testing the semantic hypothesis

As mentioned in section 2 we often find claims concerning rightward stress assignment which are based on semantic considerations. In general these considerations refer either to the semantic relationship between the two compound constituents, or to the properties of individual compound constituents or of the compound as a whole. For ease of reference we will refer to the former set of semantic entities as '(semantic) relations', and to the latter set of semantic entities as '(semantic) categories'.

The literature predicts rightward stress explicitly for the following semantic categories (e.g. Fudge 1984: 144ff, Liberman & Sproat 1992, Zwicky 1986); 'N1' refers to the left constituent, 'N2' to the right constituent):

(2) N1 refers to a period or point in time (as in *night bird*)
N2 is a geographical term (*lee shore*),
N2 is a type of thoroughfare (*chain bridge*)
The compound is a proper noun (*Union Jack*)
N1 is a proper noun (*Achilles tendon*)

In addition, the literature claims that rigthward stress is triggered by the following semantic relations (e.g. Fudge 1984: 144ff, Liberman & Sproat 1992):

(3) N2 DURING N1 (*harvest festival*)
N2 IS LOCATED AT N1 (*promenade concert*)
N2 IS MADE OF N1 (*tin hat*)
N1 MAKES N2 (*worm hole*)

There are a number of methodological and theoretical problems with testing these claims. First of all, the semantic categories and semantic relations mentioned in the literature (such as ,N1 is a material', 'N2 is located at N1') seem generally ill-defined.

Second, items are often ambiguous, i.e. they show more than one relation.[8] Third, on a theoretical level it is unclear how many and what kinds of relations and categories would be expected to play a role. There may be many more (or less) than the eight categories and relations mentioned above that have an effect on stress assignment.

In order to deal with, if not solve, these problems we used a set of 18 semantic relations that are more or less established as useful in studies of compound interpretation. The bulk of these relations come from Levi (1978), a seminal work on compound semantics, whose relations have since been employed in many linguistic (e.g. Liberman & Sproat 1992) and, more recently, psycholinguistic studies of compound structure, stress and meaning (cf., for example, Gagné & Shoben 1997, Gagné 2001). Levi's catalogue contains fewer than our 18 relations, but we felt that some additions were necessary, especially to ensure the possibility of reciprocal relations. For example, Levi's list has a relation N2 USES N1, but no relation N1 USES N2. In such cases we added the missing relation to our set of relations to be coded. Furthermore, we added a few categories that we felt were missing from her set, such as N2 IS NAMED AFTER N1. In (4) we present the final list of our relations. The relations are expressed by supposedly language-independent predicates that link the concepts denoted by the two constituents (see Levi 1978 for discussion).

[8] Cf., for example, *worm hole*, which could also be interpreted in a locative sense.

(4) List of semantic relations coded, illustrated with one example each

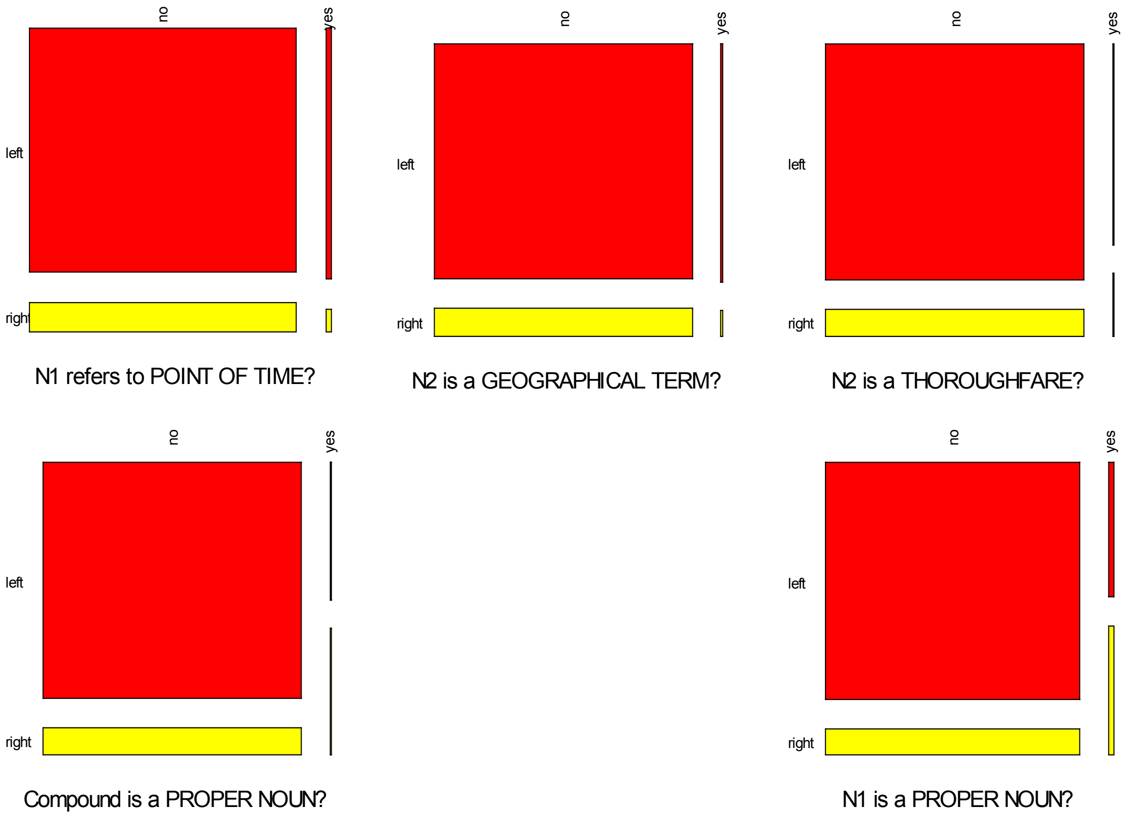
	Semantic relation			**example**
1.	N2	CAUSES	N1	*teargas*
2.	N1	CAUSES	N2	*heat rash*
3.	N2	HAS	N1	*stock market*
4.	N1	HAS	N2	*lung power*
5.	N2	MAKES	N1	*silkworm*
6.	N1	MAKES	N2	*firelight*
7.	N2	IS MADE OF	N1	*potato crisp*
8.	N2	USES	N1	*water mill*
9.	N1	USES	N2	*handbrake*
10.	N1	IS	N2	*child prodigy*
11.	N1	IS LIKE	N2	*kettle drum*
12.	N2	FOR	N1	*travel agency*
13.	N2	ABOUT	N1	*mortality table*
14.	N2	IS LOCATED AT/IN/...	N1	*garden party*
15.	N1	IS LOCATED AT/IN/...	N2	*taxi stand*
16.	N2	DURING	N1	*night watch*
17.	N2	IS NAMED AFTER	N1	*Wellington boot*
18.	OTHER			*schoolfellow*

Some of the categories proved especially difficult to code consistently, so that additional guidelines were developed. These concerned mainly the interpretation of the predicates CAUSE, MAKE, and IS. CAUSE was pertinent in cases where a cause (denoted by the one constituent) triggers an effect (denoted by the other constituent), while MAKE was coded in cases of purposeful creation or of production. IS subsumes three cases, the first being that the left constituent denotes a subset of the denotation of the right constituent (*poison gas*), the second being that left and right constituents are not in a subset-superset relation and IS works in both directions (*girl-friend*), the third being same-level appositional compounds (*owner-driver*).

Given that compounds in English are in principle ambiguous, a compound could be assigned multiple relationships. As mentioned above, each compound was coded by two independent raters and only those compounds were analyzed that were assigned to the same category by the two raters.

The data were subjected to two separate logistic regression analyses, the first using the five semantic categories referring to compound constituents or the compound as a whole, the second using the 18 relations as predictors. Let us first look at the categories referring to compound constituents or the compound as a whole. The five panels in figure 8 illustrate the distributions of stresses for these types of compound:

Figure 8: Stress position by semantic category



N1 refers to POINT OF TIME?

N2 is a GEOGRAPHICAL TERM?

N2 is a THOROUGHFARE?

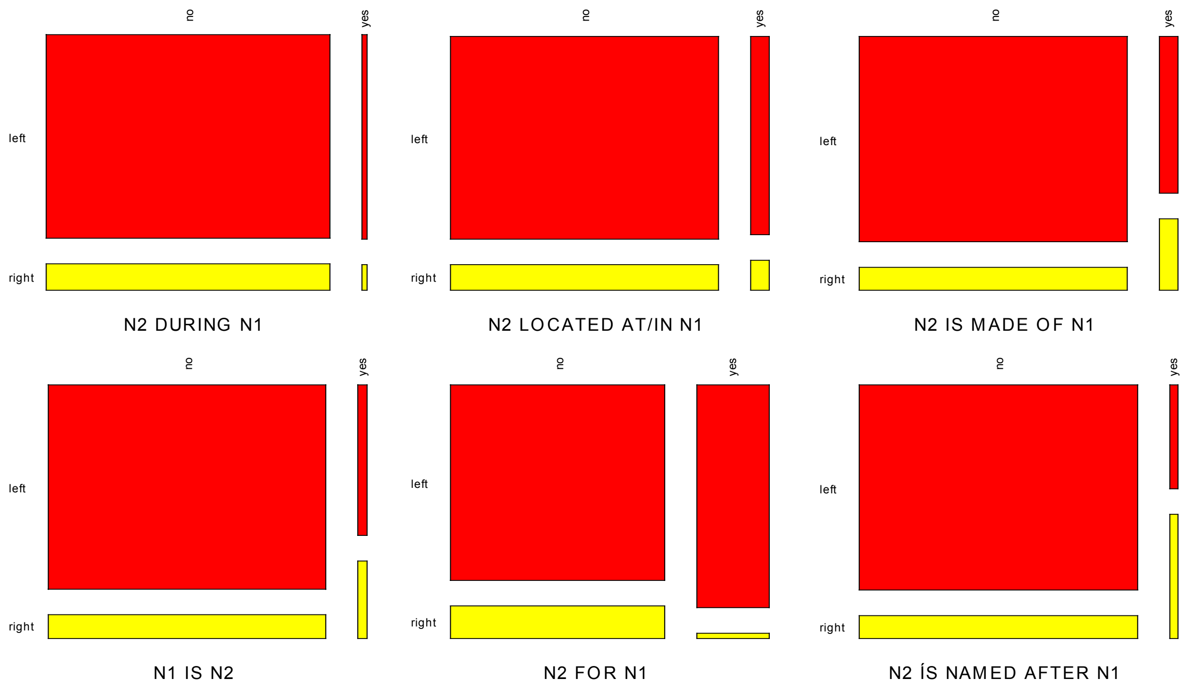
Compound is a PROPER NOUN?

N1 is a PROPER NOUN?

The logistic regression analysis with the five categories as predictors shows that for three of the five categories (upper panels: 'N1 refers to a point of time' and 'N2 is a geographical term', and 'N2 is a thoroughfare') we do not find the predicted effect. For the two other categories ('compound is a proper noun', and 'N1 is a proper noun') we do find the expected effect, i.e. compounds with these categories have a higher proportion of right-stresses than other compounds ('compound is a proper noun': p = 0.004, and 'N1 is a proper noun': p < 0.001). However, the majority of the pertinent compounds are still left-stressed, so that the overall predictive power of the model is extremely low (C = 0.551).

Let us turn to the regression analysis of the effect of semantic relations on stress assignment. Of the four relations predicted to favor rightward stress only three could be tested (N2 DURING N1, N2 LOCATED AT N1, N2 IS MADE OF N1), because there were only two pertinent cases for the relation N1 MAKES N2. In the analysis of deviance of a logistic regression model, two relations did not have the predicted effect (N2 LOCATED AT N1: p = 0.53, N2 DURING N1: p = 0.97), while the third, N2 IS MADE OF N1, did show a significant effect in the right direction (p < 0.001). In addition, our model shows significant rightward stress effects also for N1 IS N2 (p < 0.001) and N2 IS NAMED AFTER N1 (p < 0.001), and a significant leftward stress effect for N2 FOR N1 (p < 0.001). The overall power of the model is not too impressive (C = 0.772). Figure 9 shows the distributions:

Figure 9: Stress position by semantic relation



The findings for semantic relation thus closely resemble those for semantic categories discussed above, in that only a subset of the proposed variables show an effect in the expected direction of rightward stress, but that the majority of the pertinent compounds are still left-stressed.

If we combine all semantic relations and all semantic categories in one combined logistic regression model, the four relations N1 IS MADE OF N2, N1 IS N2, N2 FOR N1, N2 IS NAMED AFTER N1 remain significant, while of the semantic categories only 'The compound is a proper noun' remains in the model.[9] All significant predictors increase the likelihood of rightward stress, only N2 FOR N1 increases left-stress. A comparison of the performance of the final regression model and the semantic hypothesis (based on the categorical implementation of the categories and relations found in the literature) again reveals a better overall accuracy (the two models differ significantly: $\chi^2 = 179.984$, $df = 3$, $p < 0.01$) for the regression model, with the categorical rules (again) overpredicting right stresses and underpredicting left stresses. The regression model, on the other hand, has more trouble accounting for the occurrence of right stress in the data set. Table 4 gives the pertinent figures:

Table 4: Accuracy of prediction, semantic hypothesis

prediction	**regression model accuracy**	**semantic hypothesis accuracy**
of left stresses	98.3%	85.0%
of right stresses	15.2%	30.0%
total	89.0%	78.7%

We may now turn to the final hypothesis, analogy.

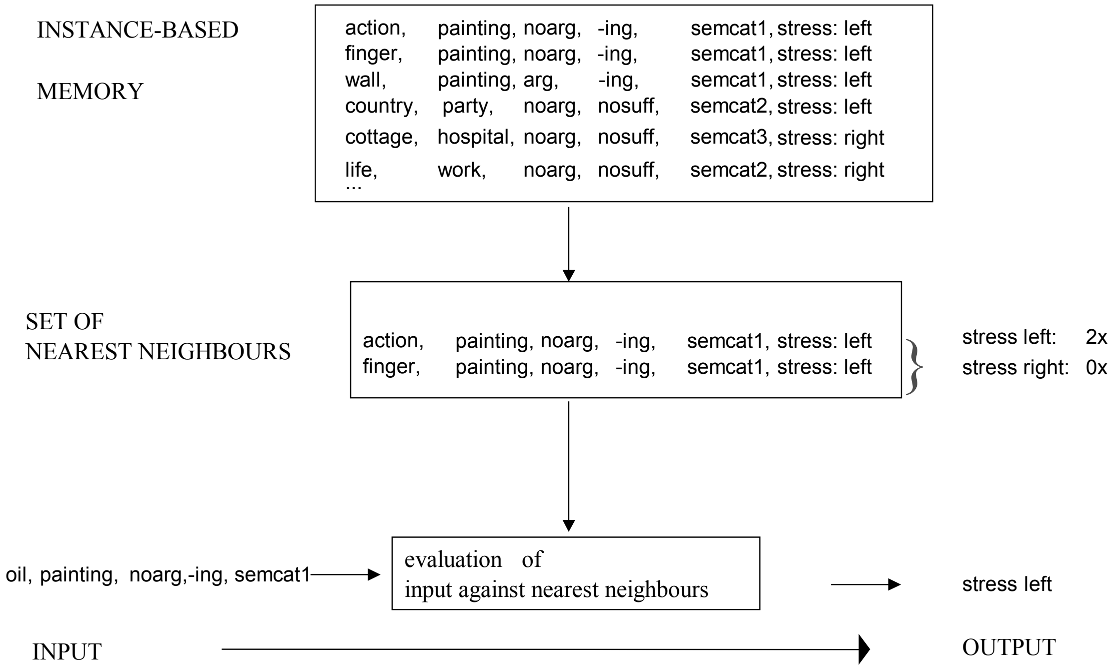
[9] The significances in the analysis of deviance for N1 IS MADE OF N2, N1 IS N2, N2 FOR N1, and N2 IS NAMED AFTER N1 are all p < 0.001, for 'The compound is a proper noun' we get p = 0.011.

6. Testing the analogical hypothesis

In general terms, the analogical hypothesis claims that stress in compounds is determined by the stress pattern of the majority of similar instances that are stored in memory. For example, a given word *carpet beater* would be assigned leftward stress because the putatively most similar exemplar stored in memory (say, e.g., *eggbeater*), has leftward stress. The crucial problem is of course how to measure the similarity between compounds. In the present study, we used TiMBL 5.1 (Daelemans et al. 2004) as a computational algorithm to test the analogical hypothesis.

TiMBL computes similarities by counting identical values of the variables that encode the properties of the items in the lexicon and the input. This works as follows. For every compound in the lexicon, we have coded its semantic and structural properties and its stress. When a new compound comes in and needs to be assigned stress, the new compound is compared in all its properties with all exemplars in the lexicon. The algorithm selects a set of so-called nearest neighbors which contains only those compounds that are most similar to the input. The algorithm then assigns the kind of stress that is most frequent among the nearest neighbors. Figure 10 illustrates the procedure using only five predictor categories (left constituent, right constituent, argument structure, morphology, and semantics). In the example, *oil painting* is assigned leftward stress since both compounds in the set of nearest neighbours are left-stressed.

Figure 10: Stress assignment by TiMBL


INSTANCE-BASED
MEMORY
action, painting, noarg, -ing, semcat1, stress: left
finger, painting, noarg, -ing, semcat1, stress: left
wall, painting, arg, -ing, semcat1, stress: left
country, party, noarg, nosuff, semcat2, stress: left
cottage, hospital, noarg, nosuff, semcat3, stress: right
life, work, noarg, nosuff, semcat2, stress: right
...
SET OF
NEAREST NEIGHBOURS
action, painting, noarg, -ing, semcat1, stress: left
finger, painting, noarg, -ing, semcat1, stress: left
stress left: 2x
stress right: 0x
oil, painting, noarg,-ing, semcat1
evaluation of
input against nearest neighbours
stress left
INPUT
OUTPUT

All claims about analogy that can be found in the literature make reference to the left and right constituents of compounds (see again section 2 for details and examples). In order to be able to specifically test these claims (in addition to testing the potential analogical effects of argument-structural, semantic and morphological factors) we used only those compounds whose left or right constituent occurred more than once in the corpus. In other words, if we wanted to test the effect of consituent families, we had to have only those compounds for which the pertinent information was available. While this procedure reduced our data set to 2643 items, it made it possible for us to test whether the absence vs. the presence of the constituent family information would make a difference in the accuracy of predictions.

Let us see how good the algorithm is at taking the right decisions for the 2643 compounds. Table 5 gives the correct and incorrect predictions for each type of stress. It also compares the predictive accuracy reached by TiMBL with that of a regression model based on this subset of the data.

Table 5: Accuracy of prediction, analogical hypothesis

prediction	regression model accuracy	TiMBL's Accuracy
of left stresses	100%	99.0%
of right stresses	0.6%	21.2%
total	94.1%	94.4%

We can see that TiMBL, like our regression models, overpredicts leftward stress and underpredicts rightward stress. If we compare TiMBL's performance with that of the rules and models discussed in the previous two section, we have to state that TiMBL has the best overall performance of all, but considerable problems with predicting rightward stresses. The logistic regression analysis, which, like TiMBL, makes use of all predictor variables, does not differ significantly in its overall accuracy from TiMBL. It detects 100% of the left stresses, but notably only 0.6% of the right stresses, which leads to an overall accuracy of 94.1%.

The important question is of course which kinds of features prove to be most important for the analogical algorithm. Given the claims in the literature, what is of particular interest is the contribution of the constituent family information. The above figure of 94.4% accurracy has been achieved by taking all kinds of feature into account. If we take each set of features (structural, morphological, semantic, and constituent family) separately, i.e. if we ignore all other features while testing one set of features, we, quite surprisingly, always arrive at around 94% accurarcy. Thus, any given set of features is as good a predictor as any other set. If we do the opposite and use all sets of features but one, basically the same picture emerges. The performance does not drop, with the significant exception of the factor 'constituent family'. If we leave out the information on the left or right constituent the accuracy rate drops significantly (left constituent: χ^2 = 7.425, p= 0.0064, right constituent: χ^2 = 4.834, p = 0.0279, left and right constituent: χ^2 = 4.834, p= 0.0279).

This result shows that left and right constituent indeed add valuable information to the classfication task in an analogical model. Our findings thus constitute robust evidence for the influence of constituent families on compound stress assignment. Importantly, this evidence is not based on small sets of handpicked forms that supposedly show the validity of that hypothesis, but emerges through the formal analysis of a large number of independently gathered data points.

7. Discussion and conclusion

In this paper we have presented the first large-scale investigation of compound stress in English based on independently available data. Overall the study has shown that existing hypotheses about compound stress are not able to explain the variability in the data. This is in line with recent experimental and speech corpus-based studies (such as Plag 2006 and Plag et al. 2006).

In particular, the idea that argument structure plays a role in compound stress assignment has been shown to only hold for compounds ending in the agentive suffix *–er*, and not for compounds featuring other right-hand head morphemes (such as *–ion, -ing,* or conversion). With regard to lexicalization effects, we found a clear interaction of spelling and stress, with one-word compounds exhibiting almost exclusively leftward stress. However, we did not find a lexicalization effect based on frequency data, which runs counter to the spelling results. How can this discrepancy be explained?

An answer to this question can be found if we look at the compounds spelled as two words. First, we should note that the two-word compounds are the smallest subset of compounds in CELEX (cf. again figure 7 above). Second, apart from two exceptions, all of the 1268 two-word compounds in CELEX have a frequency of zero, i.e. they all have been taken from the two dictionaries, and not from the COBUILD corpus. This means that the (presumably very many) two-word compounds that occur in the texts of the COBUILD corpus were simply not sampled for the CELEX data base. For practical reasons of data base creation, only orthographic words, i.e. continuous letter strings between two spaces, were sampled from the COBUILD corpus.[10] Since we can expect that the proportion of non-lexicalized compounds is highest among two-word compounds, we have to state that CELEX has in general a bias towards lexicalized compounds. This would also explain the rather high amount of 90% left stresses in the data.[11]

The fact that CELEX has this bias towards lexicalized compounds may seem disappointing, but we have to recall that – apart from Giegerich (2004) – none of the hypotheses found in the literature is explicit about the role of lexicalization. In fact, many of the examples cited in the pertinent literature to back up the respective claims, are indeed well-known, hence lexicalized, compounds. Therefore, even a database that has a lexicalization bias such as CELEX is an appropriate testing ground for these theories.

With regard to the semantic hypothesis we have shown that only few predictions are borne out by the facts, that many claims do not hold, and that it is possible to find new effects. Finally, the analogical modeling of the data showed that the constituent families play an important role in stress assignment. This supports pertinent claims in the literature, which so far have rested on a few pertinent examples only.

An overall comparison of models shows that analogical modeling is most successful with the data. Table 6 below combines tables 3 through 5 for convenience. We refer to the accuracy reached by the application of the categorical rules of the structural and semantic hypotheses as 'hypothesis-based accuracy' :

[10] That this is indeed the case has been confirmed by Harald Baayen (p.c.), one of the authors of CELEX. The two exceptional items *station wagon* and *India rubber* are probably not from the COBUILD corpus, but received their frequency of 3 due to an error.

[11] It is presently unclear in what proportions left and right stresses occur in English compounds. In a recent perception experiment using a random sample of compounds from the Boston University Radio Speech Corpus, Kunter & Plag (2006) found, roughly, two thirds left stresses and one third right stresses (type-wise). Plag et al. (2006) even find a majority of right-stressed compounds among their 4400 compounds, counting token-wise. It remains to be shown what kind of a distribution can be regarded as representative of the language as a whole.

Table 6: Accuracy of predictions across hypotheses and models

	structural factors		**semantic factors**		**all factors**	
prediction	**regression model accuracy**	**hypothesis-based accuracy**	**regression model accuracy**	**hypothesis-based accuracy**	**regression model accuracy**	**TiMBL's accuracy**
of left stresses	99.0%	46.9%	98.3%	85.0%	100,0%	99.0%
of right stresses	10.3%	72.4%	15.2%	30.0%	0.6%	21.2%
total	91.5%	49.0%	89.0%	78.7%	94.1	94.4%

What these comparisons show us is that rule-based models cannot adequately cope with the variability of the data. Probabilistic and analogical models have higher overall accuracy rates, even though they are not good at detecting rightward stress. Rule-based approaches account better for the -- in the CELEX data -- much less frequent rightward stresses, while at the same time overgeneralizing rightward stress incorrectly to many compounds with leftward stress. The overall results are in line with the findings of recent studies of compounds in other languages (e.g. Krott et al. 2001, 2002, 2004), which have also shown that variable compound behavior is best accounted for by probabilistic or analogical models, instead of rule-based ones. Future studies of compound stress will have to show whether this general assumption about the organization of compounds in the mental lexicon still holds if more non-lexicalized data are factored in.

Finally, the present study has found additional evidence for the importance of the constituent family in explaining compound behavior. Krott et al. (2001, 2002, 2004) have already shown that the constituent family has significant influence on the choice of the linking morpheme in Dutch compounds, and Gagné (2001) provided evidence that the constituent family has an effect on compound interpretation. Our study now demonstrates that such effects also pertain to the phonology of compounds.

References

Adams, V. (1973). *An introduction to Modern English word-formation*. London: Longman.

Baayen, H. R. H., R. Piepenbrock, and L. Gulikers (1995). *The CELEX lexical database* (CD-ROM). Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Bauer, L. (1983). *English word-formation*. Cambridge: CUP.

Bauer, L. (1998). When is a sequence of two nouns a compound in English? *English Language and Linguistics* 2: 65–86.

Bloomfield, L. (1933). *Language*. Chicago: Holt.

Chomsky, N. & M. Halle (1968). *The sound pattern of English*. New York: Harper and Row.

Daelemans, W., J. Zavrel, K. van der Sloot, and A. van den Bosch (2004). *TiMBL: Tilburg Memory Based Learner*, version 5.1, Reference Guide. ILK Technical Report 04-02, available from http://ilk.uvt.nl/downloads/pub/papers/ilk0402.pdf.

Farnetani, E. & P. Cosi (1988). English compound versus non-compound noun phrases in discourse: An acoustic and perceptual study. Language and Speech 31: 157-180.

Fudge, E. C. (1984). *English word-stress*. London: George Allen & Unwin.

Gagné, Ch. & E. Shoben (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory and Cognition* 23: 71-87.

Gagné, Ch. (2001). Relation and lexical priming during the interpretation of noun-noun combinations. *Journal of Experimental Psychology: Learning, Memory and Cognition* 27: 236-254.

Giegerich, H. (2004). Compound or phrase? English noun-plus-noun constructions and the stress criterion. *English Language and Linguistics* 8: 1–24.

Goldberg, A. & R. Jackendoff (2005). The end result(ative). *Language* 81.2: 474-477.

Ingram, J. , T. A. T. Nguyen & R. Pensalfini (2003). An acoustic analysis of compound and phrasal stress patterns in Australian English. Submitted for publication.

Kingdon, R. (1958). The groundwork of English stress. London, etc.: Longmans, Green & Co.

Krott, A. , H. Baayen & R. Schreuder (2001). Analogy in morphology: Modeling the choice of linking morphemes in Dutch. *Linguistics* 39: 51-93.

Krott, A. , P. Hagoort & H. Baayen (2004). Sublexical units and supralexical combinatorics: The case of Dutch interfixes in visual processing. *Language and Cognitive Processes* 19: 453-471.

Krott, A. , R. Schreuder & H. Baayen (2002). Analogical hierarchy: Exemplar-based modeling of linkers in Dutch noun-noun compounds. In Skousen, R. , D. Lonsdale & D. S. Parkinson (eds.) *Analogical modeling: An exemplar-based approach to language*. Amsterdam: Benjamins. 181-206.

Kunter, G. & I. Plag (2006). What is compound stress? Paper presented at the University of Edinburgh, May 2006.

Ladd, D. R. (1984). English compound stress. In Gibbon, Dafydd & Helmut Richter (eds.) *Intonation, accent and rhythm*. Berlin: Mouton de Gruyter, 253-266.

Lees, R. B. (1963). *The grammar of English nominalizations*. The Hague: Mouton.

Levi, J. N. (1978) *The syntax and semantics of complex nominals*. New York: Academic Press.

Liberman, M. & R. Sproat (1992). The stress and structure of modified noun phrases in English. In Sag, I. A. & A. Szabolcsi (eds.) *Lexical Matters*. Stanford: Center for the Study of Language and Information. 131-181.

Lipka, L. (1994). Lexicalization and institutionalization. In Adger, R. E. (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergamon Press. 2164-2167.

Marchand, H. (1969). *The categories and types of present-day English word-formation*. 2nd ed. München: Beck.

Olsen, S. (2000). Compounding and stress in English: A closer look at the boundary between morphology and syntax. *Linguistische Berichte* 181: 55-69.

Olsen, S. (2001). Copulative compounds: A closer look at the interface between syntax and morphology. In Booij, G. E. & J. van Marle (eds.) *Yearbook of Morphology 2000*. Dordrecht/Boston/London: Kluwer. 279-320.

Ostendorf, M., P. Price, & S. Shattuck-Hufnagel (1996). *Boston University Radio Speech Corpus*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.Payne, J. , and R. Huddleston (2002). Nouns and noun phrases. In Huddleston, R. & G. K. Pullum, *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press. 323–524.

Plag, I. (2003). *Word-formation in English*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Plag, I. (2006). The variability of compound stress in English: structural, semantic and analogical factors, *English Language and Linguistics* 10.1, 143-172.

Plag, I., G. Kunter, S. Lappe & M. Braun (2006) Variable stress assignment in noun-noun compounds: new evidence from corpora, *Directions in English Language Studies*, University of Manchester, 6-8 April 2006.

Sampson, R. (1980). Stress in English N+N phrases: A further complicating factor. *English Studies* 61: 264-270.

Schmerling, S. F. (1971). A stress mess. *Studies in the Linguistic Sciences* 1: 52-65.

Spencer, A. (2003). Does English have productive compounding? In Booij. G. E., J. DeCesaris, A. Ralli & S. Scalise (eds.). *Topics in morphology. Selected papers from the third Mediterranean morphology meeting (Barcelona, September 20 – 22, 2001).* Barcelona: Institut Universitari de Lingüística Applicada, Universtitat Pompeu Fabra. 329—341.

Zwicky, A. M. (1986) Forestress and afterstress. *Ohio State University Working Papers in Linguistics* 32, 46-62.