Frequency Effects in Compound Processing

R. Harald Baayen
University of Alberta, Edmonton

Victor Kuperman
Stanford University

Raymond Bertram
Turku University

Mailing address for proofs:
baayen@ualberta.ca
R. H. Baayen
Department of Linguistics
4-55 Assiniboia Hall
University of Alberta
Edmonton T6G 2E5
CANADA

Abstract

This chapter discusses the role of compound token frequency, head and modifier token frequency, and head and modifier compound family sizes (type frequencies) in the reading of English and Dutch compounds, using data from word naming, visual lexical decision, and eye-tracking studies. Using generalized additive regression modeling, it is shown that these frequency measures enter into many complex interactions. These interactions argue against current staged models, favoring morphological processing as part of a complex dynamic system.

1 Introduction

The entry of the Collins English Dictionary for the word *index* offers as its first subentry the meaning ('an alphabetical list of persons, places, subjects, etc., mentioned in a printed work . . . '). The second subentry is a pointer to a compound: 'see thumb index'. Ever since Taft and Forster (1976), this set-up of dictionary entries has served as a model for decompositional theories of lexical processing. According to these 'prelexical' theories, lexical access to the meaning of a compound such as *thumb index* is crucially mediated by a representation of its modifier (*thumb*).

The Collins English Dictionary also provides separate main entries for many compounds. A main entry for *house*, for instance, is supplemented by a main entry for *houseboat*. This set-up has also served as a model for rival theories of lexical processing. According to these theories, access to the meaning of *houseboat* might just as well be guided by a separate orthographic access representation for the compound word itself. In the parallel dual route model of Baayen et al. (1997), the meaning of *houseboat* can be retrieved either directly through the access representation of *houseboat*, or indirectly through the access representations of *house* and *boat*.

What these rival theories have in common is their assumption that a frequency effect observed in an experiment for modifier, head, or compound reflects the activation of a lexical representation for that lexical unit. What they disagree about is the temporal order in which these representations become activated. In prelexical theories, the compound representation is activated only after its constituents have been accessed. In 'supralexical' theories (see,e.g., Giraudo and Grainger, 2001), the constituents are accessed following the activation of the whole word. In the parallel dual route framework, access representations for compound, head and modifier are accessed in parallel, and which interpretational route is the first to complete depends on the distributional properties of all three forms, including their frequencies.

The goal of the present study is to introduce evidence that all these theories underestimate the complexities involved in understanding compound words. We first discuss evidence from large-scale visual lexical decision and visual word naming studies carried out with English compounds. We then proceed with a reanalysis of an eye-tracking study of compound processing in Dutch.

In what follows, we investigate the predictivity for lexical processing of the following variables: (1) compound frequency (the summed frequencies of

		r
ModFreq	ModFreqResid	0.81
HeadFreq	HeadFreqResid	0.76
ModLength	ModLengthResid	0.94
HeadLength	HeadLengthResid	0.96

Table 1: Correlations of the original frequency and length measures and their residualized counterparts.

the compound and its inflectional variants), (2) head frequency and modifier frequency (also summing over inflectional variants), (3) head and modifier compound family size (the number of different compounds in which the head (or modifier) occurs as head (or modifier), and as a final control variable (4) the length (in letters) of head, modifier, and the compound.

2 Lexical Decision and Naming Latencies

From the set of compounds with nonzero frequency in the CELEX lexical database (Baayen et al., 1995) consisting of two monomorphemic nouns, those compounds were selected for which lexical decision and naming latencies were available in the English Lexicon Project (Balota et al., 2007), resulting in a set of 1252 words. On the basis of the information in CELEX, the following information was extracted or calculated for each compound: the frequency of the compound, the lemma frequency of the head, the lemma frequency of the modifier, the constituent family size of modifier and head and the lengths of both constituents in characters. Frequencies and family sizes were log-transformed in order to reduce artefactual outlier effects.

Given the high collinearity of this set of predictors ($\kappa = 29.44$), the constituent frequency measures were residualized on the corresponding family sizes, and the length measures were residualized on the family size and constituent frequency measures. The residualized measures were all highly correlated with their original counterparts (all $r > 0.75$, see Table 1). After decorrelation, the collinearity was no longer harmful, as indicated by the low value of the condition number 8.49. Thus, the *ModLengthResid* measure can be understood as a measure of the length of the modifier from which that part of the length that can be predicted from other measures has been removed.

The response latencies in both naming and lexical decision deviated sub-

stantially from normality. This deviance was corrected by a reciprocal transformation. We used the transformation $-1000/RT$ so that larger values in the transformed dependent variable correspond to larger values of the original variable.

2.1   Lexical Decision

Figure 1 and Table 2 summarize a regression model fitted to the lexical decision latencies, given the assumption that the effects of the predictors are linear in nature (the adjusted R^2 of this model is 0.33). The frequency of the compound itself emerges as the predictor with the largest effect size: The greater the frequency of the compound, the shorter the response latencies. Intermediate effect sizes come with the lengths of the constituents: Greater lengths for head and modifier lead to longer response latencies. Slightly smaller, facilitatory effect sizes can be seen for the compound family sizes. The smallest effect sizes are those for the head and modifier frequencies. All partial effects graphed in Figure 1 were highly significant (all $p < 0.002$), with the exception of the frequency of the head ($p = 0.31$).

	Estimate	Std. Error	t value	p
Intercept	-0.98	0.02	-60.87	0.0000
CompFreq	-0.06	0.00	-17.64	0.0000
ModFreqResid	-0.01	0.00	-3.22	0.0013
ModLengthResid	0.03	0.00	6.97	0.0000
ModCompFam	-0.04	0.01	-8.71	0.0000
HeadFreqResid	-0.00	0.00	-1.01	0.3118
HeadLengthResid	0.03	0.01	6.42	0.0000
HeadCompFam	-0.03	0.00	-7.33	0.0000

Table 2: A regression model fitted to the lexical decision latencies.

Previous research has shown that the assumption of the effects being linear may be too restrictive (see, e.g. Baayen et al., 2006). Nonlinearities can be modeled efficiently using restricted cubic splines (see, e.g., Harrell, 2001; Wood, 2006). A second implicit assumption underlying the present regression model is that there are no interactions between the different — numerical — predictions. In other words, the model shown in Figure 1 assumes that the effect of one predictor is the same irrespective of the value
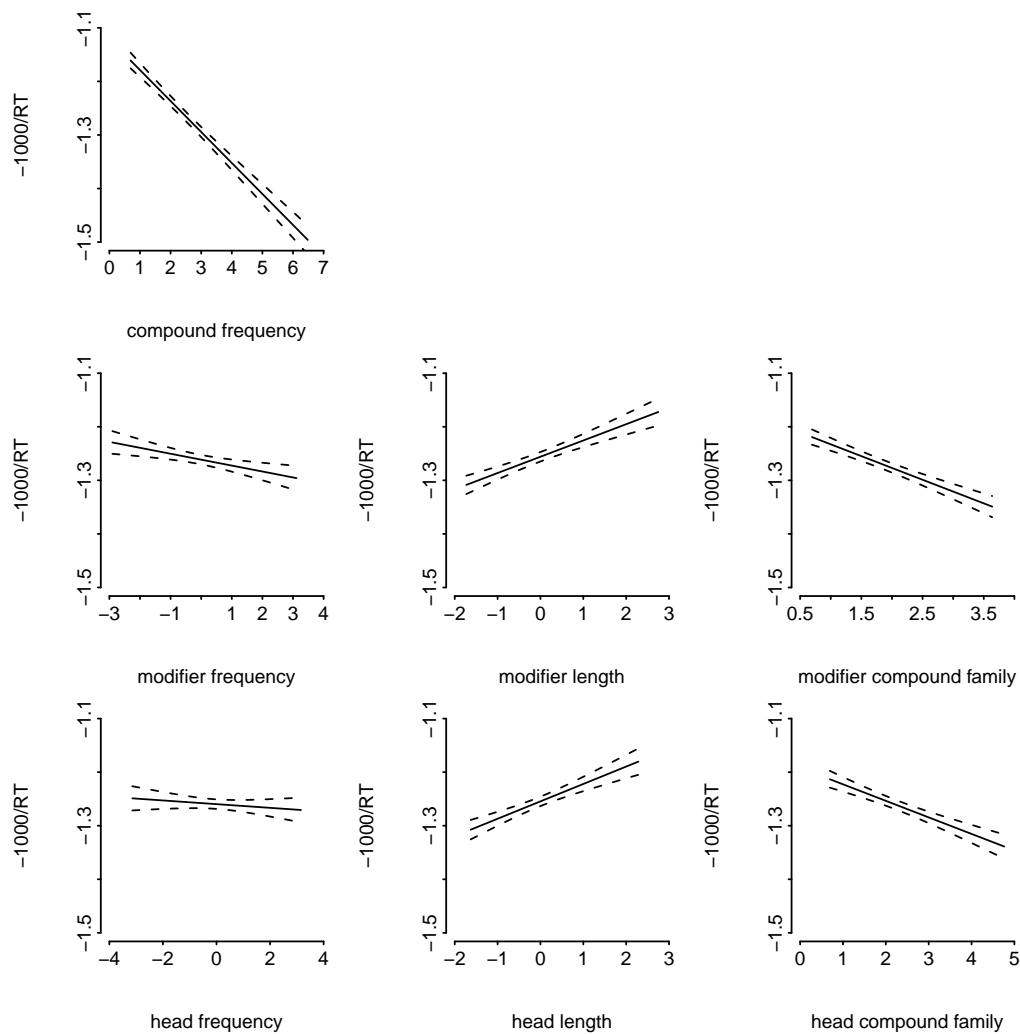
Figure 1: Partial effects of compound frequency, modifier and head frequency, length, and compound family size, in a simple regression model fitted to the (transformed) lexical decision latencies of 1252 two-constituent compounds as available in the English Lexicon Project.

of any of the other predictors. Although interactions between factors and between a factor and a numeric predictor are routinely explored in analysis of variance and analysis of covariance, interactions between two (or more) numeric predictors are seldom investigated. There are two reasons for this

neglect. First, when we consider interactions between numerical variables, we are effectively searching for a solution in a potentially nonlinear, highly complex multidimensional space. Harrell (2001) gives as practical advice not to explore interactions between numerical variables, unless they have already been reported to be relevant in the literature. This paradoxical advice — someone has to take the first step to explore and report such interactions — illustrates the conservatism that characterizes much research using multiple regression.

Second, modeling interactions between two or more numerical predictors is not a trivial enterprise. Within the standard linear regression framework, one can include a so-called multiplicative interaction, which effectively amounts to including as a new variable the product of the two predictors for which an interaction is required. For an interaction of compound frequency and modifier frequency, for instance, a new variable is created that, for each word, multiplies the compound frequency of that word with the frequency of its modifier. This new variable is then entered into the regression equation along with the other predictors.

The problem with this multiplicative interaction is that it imposes a very specific functional form on the interaction, as can be seen in Figure 2. The left panels show simulated data where the dependent variable z is a straightforward linear function of the predictors x and y, without an interaction. The top left panel shows the resulting plane in 3-D, the bottom left panel presents the corresponding contour plot, with darker shades of grey representing more negative values. Note that the contour lines run in parallel. The top right panel illustrates a multiplicative interaction. For small x, the effect of y is facilitatory, for large x, it is inhibitory. The bottom right panel presents the corresponding contour plot. Note that the two corners on the main diagonal are dark (large negative values), that the other two corners are light (representing negative values closer to zero), and that the center of the graph is intermediate between these extremes. Depending on the range of values of the predictors and the strength of the interaction, the interaction plot can show full reversals as illustrated here, or only part of this twisted but otherwise quite regular surface.

Given that the multiplicative interaction captures only a small subset of possible regression surfaces, we need to explore more flexible ways of understanding interactions between distributional variables in lexical processing. Sometimes it is possible to replace a multiplicative interaction by a theoretically motivated alternative. Kuperman et al. (2009a), for instance, observed
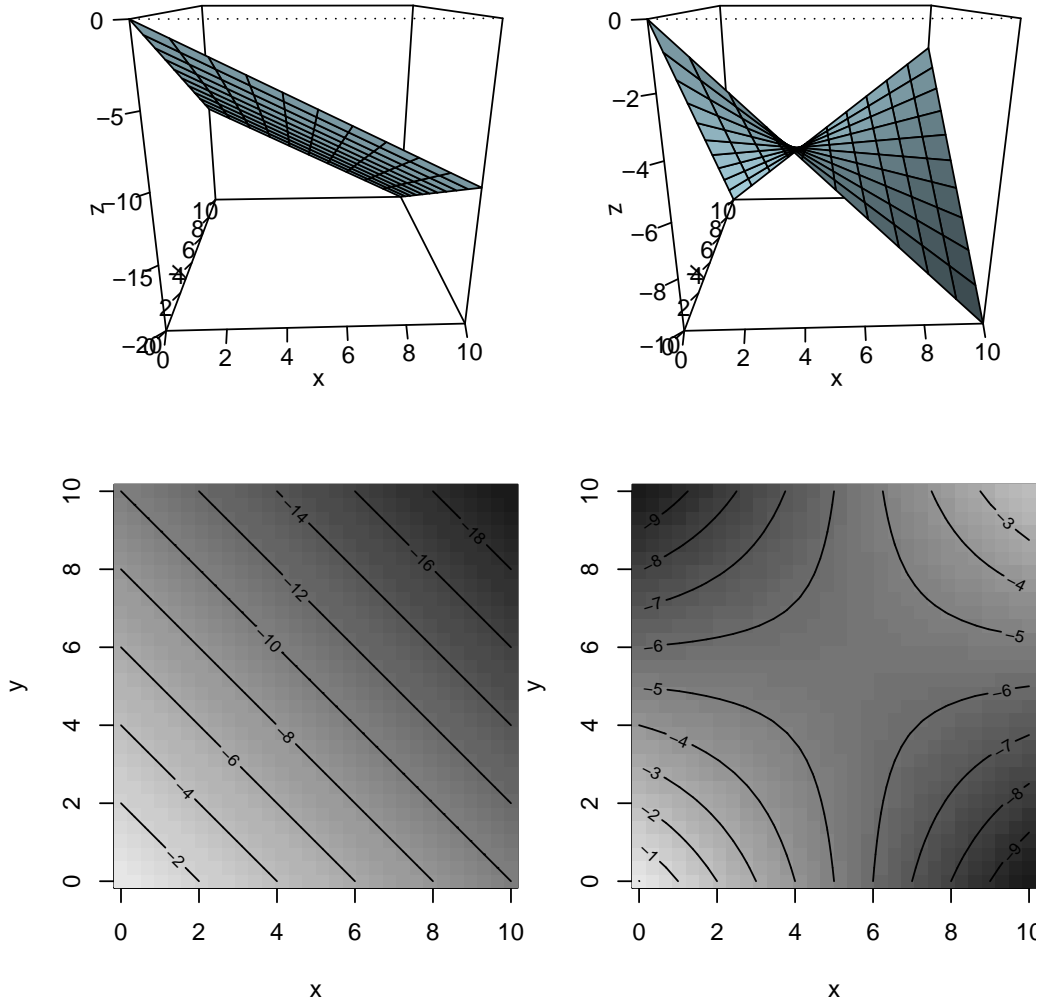
Figure 2: Interactions in the standard linear model. Left panels: no inter-action, right panels: a multiplicative interaction. The top panels present 3-D perspective plots, the lower panels the corresponding contour plots, with darker shades of grey representing larger (less negative) values. In the right panels, the dependent variable z varies substantially depending on the specific combination of values of x and y.

that an interaction of suffix productivity by base family size, emerging in the eye-movement record for Dutch derived words, could be more economically

modeled by considering the entropy of the families of suffix and base word. In the present study, we explore a more general and more flexible alternative, namely, Generalized Additive Models, henceforth GAMs.

A generalized additive model consists of two parts, a parametric part identical to that of standard linear models, and a non-parametric part that provides non-parametric functions for modeling wiggly surfaces in two or higher dimensions. In what follows, we make use of so-called tensor products to model wiggly surfaces, for details, the reader is referred to Wood (2006).

The tensor product functions are non-parametric in the sense that we will not be interested in the parameters that these smoothing functions use internally, but only in how well these mathematical black boxes succeed in capturing the wiggliness of a given surface. When fitting tensor smoothers to the data, it is crucial to avoid both undersmoothing and oversmoothing. We have used the Restricted Maximum Likelihood option of the MGCV package of Wood (2006) (version 1.5-5) to obtain the optimal balance between undersmoothing and oversmoothing. To evaluate whether a smooth function is contributing significantly to the model, we inspected the Bayesian p-values that the MGCV package documentation provides. Table 3 lists the significant tensor smooths and their associated statistics. The greater the estimated degrees of freedom (edf), the more the smooth invests in wiggliness.

	edf	Ref.df	F	p-value
te(CompFreq,ModFreqResid)	4.90	4.90	72.97	0.0000
te(HeadLengthResid,ModLengthResid)	8.62	8.62	15.05	0.0000
te(HeadCompFam,ModCompFam)	11.08	11.08	11.11	0.0000

Table 3: Tensor smooths (te) fitted to the lexical decision latencies.

The upper panels of Figure 3 summarize the tensor smooths of the GAM fitted to the lexical decision data. All smooths shown are highly significant $p < 0.0001$. The Head Frequency measure never reached significance. Other interactions were also considered (e.g., Compound Frequency by Modifier Family Size, and Head Frequency by Modifier Compound Family Size), , but turned out not to be robust predictors and were removed from the model specification. The adjusted R^2 for this model is 0.37. In Figure 3, darker shades of grey represent more negative transformed response latencies, and hence shorter response latencies in the original untransformed ms scale. The small dots represent the individual data points.
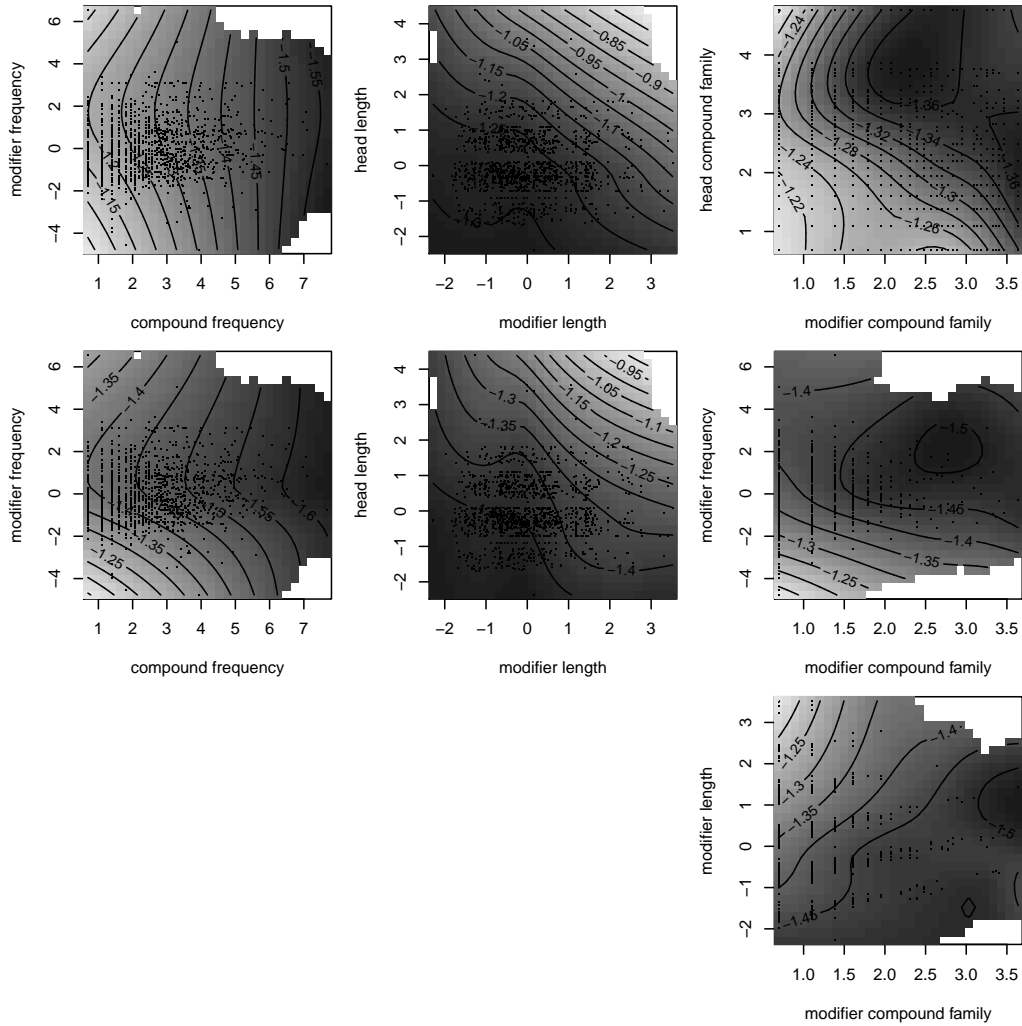
Figure 3: Nonlinear interactions in the (transformed) *lexical decision latencies* (top panels) and *naming latencies* (lower panels) to 1252 two-constituent noun-noun compounds as available in the English Lexicon Project. Lighter shades of gray indicate longer latencies. Dots represent data points. Grid nodes that are too far away from actual data points are excluded from the plot and show up in white.

First consider the top left panel. As Compound Frequency (horizontal axis) increases, we see darker shades of grey, indicating shorter response la-

tencies, as expected: Frequency effects are usually facilitatory. The effect of Compound Frequency is modulated by Modifier Frequency, but only for lower-frequency compounds. Here we see that latencies decrease to some extent for compounds with higher-frequency modifiers. Note that this facilitation tends to level off for higher modifier frequencies.

The top right panel visualizes the interaction of the orthographic lengths of head and modifier. Most observations are located in a fairly flat and somewhat irregular area between contour lines -1.3 and -1.2. The relatively small number of compounds with both larger left and right constituents elicited longer response latencies. Clearly, constituent length comes with increased processing costs.

The lower left panel depicts the interaction of head and modifier family size. The effect size of this interaction is small: Contour lines are 0.02 units apart, instead of 0.05 units as in the two upper panels. Facilitation for both family size measures is visible, but what strikes the eye (under magnification) is that facilitation is most pronounced for high values of one family size given intermediate values of the other family size.

2.2 Word Naming

The English Lexicon Project also provides the naming latencies for the present data set. The bottom row of panels in Figure 3 summarize the generalized additive model fitted to the naming data. As for the lexical decision data, all tensor smooths were highly significant ($p < 0.0001$, see Table 4). The adjusted R^2 of this model was 0.36.

The pattern of results obtained for the interaction of Compound Frequency and Modifier Frequency, and for modifier and head orthographic

	edf	Ref.df	F	p-value
s(HeadCompFam)	2.65	2.65	22.86	0.0000
te(CompFreq,ModFreqResid)	4.70	4.70	28.75	0.0000
te(HeadLengthResid,ModLengthResid)	6.95	6.95	9.03	0.0000
te(ModCompFam,ModLengthResid, ModFreqResid)	23.88	23.88	7.77	0.0000

Table 4: A restricted cubic spline (s) and tensor smooths (te) fitted to the naming latencies.

length, resembles that obtained for the lexical decision latencies. The effect of the frequency of the modifier seems to be somewhat stronger and more persistent for higher compound frequencies. It also seems to level off more substantially for higher modifier frequencies.

The effect of the family size measures emerges as qualitatively different from that observed for lexical decision. First, and not shown in Figure 3, a significant facilitatory effect of head family size was present (modeled with a restricted cubic spline), that was not modulated by the family size of the modifier as in lexical decision. This effect was linear for all but the highest head family sizes, for which it leveled off.

Second, the effect of the family size of the modifier was modulated by both modifier frequency and modifier length. This modulation was captured by a three-way tensor product, which outperformed a model with two pairwise tensor products. Figure 3 visualizes this complex interaction by means of the two pairwise interactions in the central and lower right panels. In the central right panel, the contour lines are more horizontally oriented than vertically oriented, indicating that most of the effect in this panel is carried by Modifier Frequency. Modifiers with higher frequencies elicited somewhat shorter naming latencies, but this facilitation leveled off for higher modifier frequencies. The evidence for facilitation from the modifier family size is much stronger in the lower right panel, especially so for longer modifier lengths.

3   The Eye-Tracking Record

The data from the English Lexicon Project, although informative, remain silent about the time course of information uptake during reading. As a consequence, it remains unclear whether the effects of compound frequency and effects linked to distributional properties of the head and modifier occur simultaneously or whether they are staggered in time. Early studies argued that constituent effects preceded whole word effects (Taft and Forster, 1976). As mentioned in the introduction, the reverse position, with whole-word effects preceding constituent effects in time, has also been argued for (Giraudo and Grainger, 2001). Eye-movement studies offer more fine-grained insight into the moments in time at which the different kinds of information carried by a compound are absorbed, and may help resolving this theoretical conflict.

Classic eye-tracking studies replicated results from lexical decision and naming studies suggesting a role for both constituents and whole words in

lexical processing (see, e.g. Andrews et al., 2004; Bertram et al., 2004). Pollatsek et al. (2000) observed that the frequency of the head and the frequency of the compound emerged in later measures of compound processing, after the first fixation on the word. The study by Hyönä et al. (2004) also suggested that for long compounds such as *dishwasher* there is early activation of the left constituent (*dish*) and only later activation of the right constituent (*washer*).

In what follows, we reanalyse the data of a recent study (Kuperman et al., 2009b) with generalized additive models. As GAMs currently do not integrate well with mixed-effects models, our analyses are carried out on by-compound mean durations, averaged over participants. In our reanalysis, we focus on compound frequency, head and modifier frequency, head and modifier compound family size, and head and modifier length, all defined in the same way as in the study of the lexical decision and naming latencies in the English Lexicon Project. For details on the decorrelation of these predictors, the reader is referred to (Kuperman et al., 2009b). The materials comprise 1250 Dutch compounds, written without hyphens or intervening spaces, for which at least one of the immediate constituents was itself morphologically complex. The length in letters of these compounds ranged between 8 and 12 characters.

Figure 4 summarizes the effects observed at the first fixation (left panels, all $p < 0.0001$), the subgaze duration on the modifier (central panels, all $p < 0.0001$), and the total gaze durations (all $p < 0.003$). Table 5 presents the corresponding tensor smooths and associated statistics. In Figure 4, contour lines are all equidistant. Going from left to right, the overall shades of grey become lighter, indicating longer durations (as expected for increasingly cumulative measures). Contour lines are all 0.1 unit apart, so that the size of the different effects can be gauged.

First consider the top three panels, which visualize the development of the effects of the orthographic lengths of head and modifier. The effects of the lengths are very small during the initial fixation. Notably the effect of modifier length becomes highly prominent in the modifier subgaze measure, to ebb away in the gaze durations with the exception of compounds with long heads. Note, furthermore, that the effect of modifier length is facilitatory during the initial fixation, inhibitory during subsequent fixations on the modifier as witnessed by the subgaze duration measure, and emerges as facilitatory in the overall gaze durations. Apparently, an initial short fixation on the modifier is followed by more refixations, balancing a first superficial scan

11

Tensor smooths fitted to the first fixation durations	edf	Ref.df	F	p-value
te(CompFreq,ModFreq)	4.38	4.38	6.68	0.0000
te(ModLength,HeadLength)	3.00	3.00	44.57	0.0000
te(CompLength,ModCompFam)	12.93	12.93	9.36	0.0000
Tensor smooths fitted to the modifier subgaze durations				
te(ModLength,HeadLength)	7.45	7.45	43.79	0.0000
te(ModCompFam,CompLength)	4.28	4.28	14.40	0.0000
te(CompFreq,ModFreq)	6.57	6.57	11.58	0.0000
Tensor smooths fitted to the gaze durations				
te(ModLength,HeadLength)	6.82	6.82	6.95	0.0000
te(HeadCompFam,FinTrigram)	4.09	4.09	5.40	0.0002
te(HeadFreq,ModCompFam)	5.57	5.57	10.68	0.0000
te(CompLength,ModCompFam)	3.99	3.99	4.13	0.0025
te(CompFreq,ModFreq)	5.69	5.69	36.42	0.0000

Table 5: Tensor smooths fitted to the first fixation durations, the modifier subgaze durations, and the gaze durations.

by more thorough visual processing. From the very beginning, this process seems to be modulated by the length of the head. The longer the head, the less superficial the initial scan is. The same holds, albeit to a lesser degree, for subsequent subgazes on the modifier. This suggests that knowing more substantial processing is going to be required for the right part of the visual input induces the system to invest more in understanding the left part of the word. In the overall gaze durations, the earlier investment in longer subgazes on longer modifiers pays off in the form of a modest overall processing advantage that manifests itself most clearly for compounds with more than average head lengths. Within the area for which the present experiment offers data, we see slightly elongated gaze durations for compounds with short modifiers and long heads, and perhaps also for compounds with short heads and long modifiers. Apparently, at this late processing stage, the system has optimized processing for heads and compounds of average length. Finally, it is noteworthy that the effect of head length is already apparent at the very first fixation. This effect should probably be interpreted as an approximate indicator (based largely on parafoveal information) of the amount of visual information that remains to be scanned.

The effect of orthographic length that emerges from the eye-tracking record of this Dutch experiment is remarkably different from the effects observed above for the English lexical decision and naming latencies. For the English experiments, latencies increased as both head and modifier increased in length. Two possible explanations may help understand this discrepancy. First, English writes many of its compounds with intervening spaces, and does so without exception for more complex compounds. Dutch, like German or Danish, consistently write compounds, however complex, without spaces. This language difference may lead to different reading strategies. Second, the English and Dutch data sample the space of lengths differently. The Dutch data are tightly constrained by the length of the compound, which varied between 8 and 12 letters. The English compounds vary in length between 6 and 14, and offer more compounds for which both head and modifier are relatively long. As a consequence, the Dutch data are more restricted in their exploration of the consequences of constituent length in compound processing. Further research is clearly required.

The second row of panels of Figure 4 summarize the joint effects of compound frequency and modifier frequency. Moving from left to right, the number of contour lines increases, notably for compound frequency, indicating that especially this whole-word frequency effect increases in magnitude as we move from early to late measures of information uptake. As the visual information about the full compound becomes available mostly after the first fixation, this increase in the magnitude of the compound frequency effect is as expected. A comparison with the English lexical decision and naming data shows roughly similar patterns in both sets of latencies and in the gaze durations. This suggests that the frequency effects that emerge in the response latencies are more or less fully developed at the point that the eye leaves the word.

An important question is, however, why the compound frequency effect emerges as early as it does, already at the first fixation duration. One potential explanation is to assume that compounds have some full-form orthographic access code that can be activated on the basis of partial matching information for the initial constituent, combined with parafoveal information about the compound's length, following, e.g., Pollatsek and Rayner (1982). This explanation fits well with parallel dual route models of lexical access in which orthographic representations for constituents and whole words are activated in parallel. Interestingly, the present evidence for an early effect of compound frequency, also observed for Finnish compounds in the regression

Figure 4: Nonlinear interactions in the eye-movement record.

14

study of Kuperman et al. (2008), is incompatible with 'postlexical' staged models of lexical access according to which properties of a complex word as a whole would become accessible only after access to their constituents has been completed.

There are several problems with the assumption of whole-word orthographic access representations mediating the activation of word meaning. First, whole-word frequency effects are much more robust than constituent frequency effects, and have been observed even for very low-frequency complex words (Baayen et al., 2007). The presence of whole word frequency effects in the absence of constituent frequency effects for low-frequency complex words would, under the traditional interpretation of frequency effects in psycholinguistic theories, deny a role for constituents in lexical access to many very low-frequency words, a conclusion that does not make sense from both linguistic and computational perspectives. It makes more sense to understand full-form frequency effects as tapping into the reader's knowledge about the joint probability of modifier and head, possibly in the form of memory traces for previous experience in parsing the compound into its constituents. Given that only partial information is available at the earliest stages of visual information uptake, the effect of compound frequency is, under this alternative, expected to increase over time and to produce its greatest effect in the gaze duration measure.

The third row of panels of Figure 4 brings together the contour plots for the family size of the modifier, which is also already present at the first fixation. Here, an inhibitory family size effect for short compounds levels off for longer compounds. Subsequent fixations on the modifier reverse the effect into facilitation independently of compound length. In the gaze durations, a modest facilitatory effect of modifier family size remains visible. For small-family modifiers it appears to be somewhat less facilitatory for medium-length compounds. The facilitation in the gaze durations mirrors the facilitation observed in the naming and lexical decision latencies.

The early presence of a family size effect in the eye-movement record is interesting in the light of the suggested semantic nature of this effect (De Jong et al., 2000). Although an early family size effect can be interpreted as a form effect relating to familiarity with parsing a word out as a modifier in a compound, as argued by Kuperman et al. (2009b), one would expect facilitation rather than the inhibition actually observed at the initial fixation, the prime moment for visual parsing.

It is therefore worth considering the possibility that with the modifier

15

family size effect we are actually observing an index of early semantic processing, which would harmonize well with the very early stages at which semantic effects have been demonstrated to arise (Feldman et al., 2009). Instead of assuming a form-then-meaning architecture for morphological processing and strictly serialized processing of constituents, we may need to consider more flexible architectures in which the link between form and meaning in reading is much tighter than has often been assumed.

Possibly, the constituent and compound frequency effects documented in the present study should also be interpreted as — at least partial — indices of meaning activation. The development over time of the interaction of compound frequency by modifier frequency, with an initially stronger effect of modifier frequency that later gives way to a compound frequency effect, would then represent an initial understanding of the modifier by itself that then gives way to an integrated understanding of the compound as the unification of the meanings of modifier and head.

There are two further interactions at work in the gaze durations. The one but last panel of Figure 4 graphs the joint effect of modifier family size and head frequency. For compounds with low-frequency heads, the effect of the modifier family size is slightly inhibitory. For compounds with higher-frequency heads, the majority of the compounds in this experiment, the effect reverses. The effect of head frequency shows a similar reversal, with facilitation for higher modifier families, but inhibition for smaller modifier families. The gaze durations, as a global measure of information uptake, therefore suggest a trade-off, perhaps even some form of competition between the head and the modifier. Possibly, modifiers with large families are stronger competitors for the functional interpretation as head of the compound.

The final interaction, shown in the lower right panel of Figure 4, concerns the modulation of the head family size by the token frequency of the word-final trigram. Word beginnings and word ends are visually the most salient part of the word. Kuperman et al. (2009b) therefore included the initial and final trigrams as control variables in their study. The pattern that emerges in the gaze durations suggests that the final trigram affords shorter gazes for compounds with intermediately-sized head families. Conversely, compounds with highly salient final trigrams are the ones for which facilitation from the head compound family is present. This suggests that in long compounds, a salient constituent ending is essential for boosting the head family size effect.

We conclude this reanalysis of the data of Kuperman et al. (2009b) with a methodological comment. The GAM models fitted to the durational mea-

sures from the eye-tracking record confirm the main patterns uncovered with the help of standard linear models with multiplicative interactions. Across models, we find, however, that the GAMs offer slightly more precise models that explain significantly more variance (6% in the case of the gaze duration measure, for instance). Furthermore, the GAMs are able to detect interactions that are not visible to standard models with linear interactions (in the case of the gaze duration analysis, this is the case for the interactions of the frequency of the head by modifier family size, and head family size by final trigram).

4   General Discussion

If one thing is clear from the present analyses, it is that the processing system is much more complex than previously thought. Psycholinguistic theories of morphological processing have been dominated by staged models, with different frequency effects reflecting different processing stages.

The independent effects illustrated in Figure 1 for a straightforward main effects linear model would seem to offer support for such an approach. For instance, the effect of Compound Frequency in Figure 1 seems independent of all other effects, and thus can be argued either to be accessed prior to, or posterior to the access to its constituents, following either Giraudo and Grainger (2001) or Taft (1979).

However, the enhanced precision obtained by carefully considering interactions suggests a fingerprint not of a a simple sequence of processing stages, but rather the hallmark of a complex dynamic system. Even for straightforward lexical decision and naming latencies, Figure 3 illustrates a much more intricate interplay of factors. The effect of Compound Frequency, for instance, is modulated by the frequency of the modifier. This is not what either prelexical or postlexical theories of morphological processing predict. The decrease in facilitation from constituent frequencies, or even its reversal into inhibition as seen in the eye-tracking record may reflect the increased independence of high-frequency complex words with respect to their constituents. For an experimental investigation of specifically the role of semantic transparency in compound processing and how transparency may modulate the effects of compound and constituent frequencies, see Kuperman and Baayen (2009).

The eye-movement record adds to this by clarifying that this interaction

17

can arise early, already at the initial fixation duration (see Figure 4). Even at this early stage of information uptake, when the modifier is fixated, the processing system is already anticipating the possibility that this modifier may be part of a larger compound structure.

Our data suggest that understanding the compound frequency effect as simply reflecting the activation of some unstructured, holistic lexical representation may be misguided. As explained by Kuperman et al. (2008), it is worth considering this effect as tapping into dynamic knowledge about the joint probability of head and modifier and their combinatorial semantics.

Our data further suggest that multiple sources of linguistic information are used dynamically as soon as they become available. These multiple sources include not only information from the fixated modifiers, but also information from the embedding compounds, and, also present from the initial fixation duration, information coming from the modifiers' morphological families, which are now known to play important roles for compound interpretation (Gagné and Shoben, 1997), the selection of interfixes in Dutch (Krott et al., 2001) and compound stress assignment in English (Plag et al., 2008).

References

Andrews, S., Miller, B., and Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, 16(1):285–311.

Baayen, R. H., Dijkstra, T., and Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, 36:94–117.

Baayen, R. H., Feldman, L., and Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, 53:496–512.

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

Baayen, R. H., Wurm, L. H., and Aycock, J. (2007). Lexical dynamics for low-frequency complex words. a regression study across tasks and modalities. *The Mental Lexicon*, 2:419–463.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., and Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39(3):445–459.

Bertram, R., Hyönä, J., and Pollatsek, A. (2004). Morphological parsing and the use of segmentation cues in reading Finnish compounds. *Journal of Memory and Language*, 51:325–345.

De Jong, N. H., Schreuder, R., and Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes*, 15:329–365.

Feldman, L. B., O'Connor, P. A., and Moscoso del Prado Martín, F. (2009). Early morphological processing is morpho-semantic and not simply morpho-orthographic: An exception to form-then-meaning accounts of word recognition. *Psychonomic Bulletin and Review*, in press.

Gagné, C. and Shoben, E. J. (1997). The influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23:71–87.

Giraudo, H. and Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin and Review*, 8:127–131.

Harrell, F. (2001). *Regression modeling strategies*. Springer, Berlin.

Hyönä, J., Bertram, R., and Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & cognition*, 32(4):523–532.

Krott, A., Baayen, R. H., and Schreuder, R. (2001). Analogy in morphology: modeling the choice of linking morphemes in Dutch. *Linguistics*, 39(1):51–93.

Kuperman, V. and Baayen, R. H. (2009). Semantic transparency revisited. *Manuscript Submitted For Publication*, 0:1–86.

Kuperman, V., Bertram, R., and Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes*, 23:1089–1132.

Kuperman, V., Bertram, R., and Baayen, R. H. (2009a). Processing Trade-Offs in the Reading of Dutch Derived Words. *manuscript*, Stanford University:1–63.

Kuperman, V., Schreuder, R., Bertram, R., and Baayen, R. H. (2009b). Reading Polymorphemic Dutch compounds: Towards a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, 35:876–895.

Plag, I., Kunter, G., Lappe, S., and Braun, M. (2008). The role of semantics, argument structure, and lexicalization in compound stress assignment in english. *Language*, 84:760–794.

Pollatsek, A., Hyönä, J., and Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human, Perception and Performance*, 26:820–833.

Pollatsek, A. and Rayner, K. (1982). Eye Movement Control in Reading: The Role of Word Boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, 8(6):817–33.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, 7:263–272.

Taft, M. and Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, 15:607–620.

Wood, S. N. (2006). *Generalized Additive Models*. Chapman & Hall/CRC, New York.