Morphological embedding and phonetic reduction: The case of triconstituent compounds

Gero Kunter Ingo Plag

February 11, 2016

Abstract

In this paper we propose that the internal bracketing of a word with more than two morphemes is reflected in the phonetic implementation. We hypothesize that embedded forms show more phonetic reduction than forms at higher structural levels ('Embedded Reduction Hypothesis'). This paper tests the prediction of the Embedded Reduction Hypothesis with triconstituent compounds. The analysis of the durational properties of almost 500 compound tokens shows that there is a lengthening effect on the non-embedded constituent, and a shortening effect on the adjacent embedded constituent. Yet, this predicted effect of embedding interacts with other lexical factors, above all the bigram frequency of the embedded compound. At a theoretical level, these effects mean that the durational properties of the cross-boundary constituents are indicative of the hierarchical structure and of the strength of the internal boundary of triconstituent compounds. Hence, morphological structure is reflected in the speech signal.

1 Introduction

Theories of morphology and the lexicon often rest on the assumption that morphological boundaries vary in strength depending on the morphemes involved. For example, approaches as different as Lexical Phonology (e.g. Kiparsky 1982, Kaisse and Shaw 1985) and Natural Morphology (e.g. Dressler 1985) share the assumption that compounds have stronger word-internal boundaries than derived or inflected words. The assumed differences in boundary strength are taken to have phonological consequences (see, for example, Chomsky & Halle 1968's distinction between strong and weak boundaries). In English, for example, the prefix *in-* is said to have a weak boundary because, among other things, it shows assimiliation and degemination when attached to nasal-initial bases, whereas *un-* is said to have a strong boundary because it resists assimilation and degemination.

Studies investigating the order of derivational affixes in English (Hay and Plag 2004; Plag and Baayen 2009; Zirkel 2010) have provided evidence that affix boundaries differ in boundary strength: in a word of the form [[base-X]-Y], the outer boundary between [base-X] and affix Y is stronger than the inner boundary between the base and affix X. In contrast to earlier binary distinctions between weak and strong boundaries, these studies propose a gradient of boundary strengths. Hay (2003) and Plag and Baayen (2009) argue that affixes at weaker boundaries will show a higher degree of phonological integration than affixes at stronger boundaries.

In the framework of Lexical Phonology, different degrees of morphological boundary strength are not expected to have a subtle phonetic effect, as such sub-phonemic changes are handled at the post-lexical level, which does not have access to the internal morphological structure. Yet, acoustic studies such as Sproat (1993), Sproat and Fujimura (1993), Smith et al. (2012), and Hay (2007) provide evidence that weaker morphological boundaries show not only more phonological integration, but that the phonetic implementation can also depend on the strength of a morphological boundary in the vicinity.

Notably, all previous acoustic studies have been restricted to words with only one morphological boundary, and have looked at what happens at this boundary. The results of these studies raise, however, another interesting question, namely that of the relation between phonetic implementation and morphological embedding. Combining the results from the acoustic studies (different degrees of boundary strength have traces in the phonetic implementation) with those from the affix ordering studies (inner boundaries are weaker than outer boundaries), it may be expected that the phonetic implementation of a word with more than two morphemes will reflect its morphological structure, i.e. its internal bracketing. The degree of phonetic reduction would then correlate with the degree of embedding. Based on this reasoning, we propose what we call the 'Embedded Reduction Hypothesis'. Informally, it states that constituents that are embedded deeper in a morphological structure show more phonetic reduction than less deeply embedded constituents.

The Embedded Reduction Hypothesis poses a challenge for theoretical frameworks which argue that the internal structure of derived forms is not accessible anymore at the post-lexical stage, with which phonetic alternations like sub-phonemic durational differences or phonetic reduction are usually associated. If there is reliable evidence that the acoustic signal contains phonetic detail signaling the internal structure of morphologically complex words, and if this detail affects the processing of the acoustic signal by listeners, the strict division between lexical and post-lexical components needs to be revised. Testing the Embedded Reduction Hypothesis is therefore of great theoretical relevance.

In this paper we present a pilot study that investigates the Embedded Reduction Hypothesis by looking at the acoustic duration of embedded and non-embedded morphemes. As a test bed we will use triconstituent compounds such as those shown in (1).

(1) Left-branching [[*day care*] *center*] [[*science fiction*] *shocker*]
Right-branching [*business* [*credit card*]] [*state* [*health program*]]

Based on the analysis of about 500 compound tokens taken from an experiment by Kösling (2013) we show that the Embedded Reduction Hypothesis is generally supported by the data. The predicted effect of embedding is affected by other lexical properties, most notably by the bigram frequency of the complex constituent. This means that duration is indeed an acoustic correlate of hierarchical morphological structure, albeit in interaction with other factors.

The remainder of the paper is structured as follows. The next section discusses in more detail the relation between morphological structure and the phonetic implementation, as well as its theoretical implications. Sections 3 introduces our data set and methodology, while section 4 outlines the statistical modeling procedure. The results are presented in 5, followed by a discussion and conclusion.

2 Morphological structure and phonetic realization

The idea that different morphological processes may result in different types of word-internal boundaries has been present in the theoretical linguistic literature for a considerable time. In particular, the framework of Lexical Phonology (e.g. Mohanan 1986; Kiparsky 1982; Kaisse and Shaw 1985) assumes a stratal organization of morphological processes. The processes located at a given lexical stratum create a particular type of morphological boundary that is associated with a set of phonological alternations. In this view, affixes that belong to one stratum can be attached only to other affixes from the same or a lower stratum. Higher-stratum affixes cannot be attached to lower-stratum affixes. For example, in English it is assumed (e.g. Kiparsky 1982, 141) that suffixes that cause a stress shift (e.g. *-ion* in *hýphenate–hyphenátion*) are associated with the first lexical stratum, and are thus closest to the base. Stress-neutral suffixes such as *less* are assigned to the second stratum, and thus occur in a morphologically complex word after all suffixes from the first stratum (as in *propositionless*). A crucial feature of Lexical Phonology is the notion of Bracket Erasure (e.g. Kiparsky 1982, 11), i.e. the assumption that all morphological information is discarded at the end of a stratum. In consequence, a phonological process that refers to a morphological boundary can only be activated within that stratum. Phonological processes that are applied automatically to a word, and which are not affected by its internal structure, are activated at the post-lexical stage, which takes place after Bracket Erasure. This last feature of the Lexical Phonology framework, i.e. the absence of morphological information at the post-lexical stage, is also found in influential models of speech production such as that proposed in Levelt et al. (1999).

However, research into the order of derivational affixes in English (Hay and Plag 2004; Plag and Baayen 2009; Zirkel 2010) has provided evidence that speaks against some of the main claims of Lexical Phonology, and there is acoustic evidence (discussed below) that suggests that the morphological structure is, after all, encoded in the acoustic signal. In particular, affix ordering in English has been shown to be more fine-grained than a stratal organization would predict. Instead, Hay and Plag (2004) Plag and Baayen (2009), and Zirkel (2010) propose a view in which the boundary strengths between affixes and their bases can be arranged in a hierarchy according to their degree of decomposability, i.e. the ease of analyzing a complex word into its components in perception. Affixes with a low degree of decomposability (which thus form a weak boundary) are usually not found to attach to bases involving affixes with high decomposability (i.e. with a strong boundary), but may attach to bases involving an affix with even lower decomposability. Applying the hierarchy presented in Hay & Plag (2004) to a word such as *owlishness*, the inner boundary between the base *owl* and *-ish* is considered a weaker boundary than the outer boundary between the embedded constituent complex form *owlish* and *-ness*.

Hay (2003), Hay and Plag (2004), and Plag and Baayen (2009) argue that these boundary strength differences have implications for the acoustic representation of the complex word: *-ish*, the affix at the weaker boundary, is predicted to integrate more with the base to which it is attached than *-ness*, the affix at the stronger boundary. This integration also affects the acoustic realization of the complex word: *-ish* (or its derivative) is expected to be more prone to phonetic reduction, and to show a relatively shorter acoustic duration. Thus, this view assumes a correlation between the degree of morphological embeddedness and the degree of boundary strength, and claims that this correlation is evident in the extent of phonetic reduction of the involved affixes.

Several acoustic studies provide evidence for the effect of different types of boundaries on phonetic reduction. While an early study of word and segment duration (Lehiste, 1972) did not find acoustic evidence for a hierarchical organization of boundaries, later studies revealed that different boundary types can lead to different phonetic realizations of a given speech sound. For instance, Sproat (1993), Sproat and Fujimura (1993) investigate the phonetic realization of the English phoneme /l/ in different environments. This phoneme is described (e.g. in Trask 1996) to be realized by the alveolar lateral approximant 'clear l' in pre-vocalic position and by a variant with a strongly retracted dorsum 'dark l' in post-vocalic position. However, Sproat (1993) and Sproat and Fujimura (1993) show that the degree of velarization, as well as the segment's duration, is related to the boundary type at which it occurs. The phoneme is shortest and least dark (i.e. with least retracted dorsum) in word-internal position where no morphological boundary is present, longer and darker at a suffix boundary, longer and darker still at a compound boundary, and longest and darkest at a word boundary. Their finding is compatible with the conclusion of the affix order studies that weaker boundaries show more phonological integration. Hay (2007) finds a comparable difference in phonetic reduction of the phoneme sequence *un-* depending on whether the sequence is morphemic or non-morphemic, and in the latter case, on the decomposability of the complex words. Non-morphemic *un-* as in *until*, which is at the extreme low end of the decomposability scale, is significantly more reduced than morphemic *un-*, which in turn is more reduced in less decomposable words than in more decomposable words.

In another study of prefixed vs. non-prefixed words, Smith et al. (2012) investigated differences in duration and amplitude between the initial segments /mɪs/ and /dɪs/ in pairs such as *distasteful* vs. *distinctive*, and *mistimed* and *mistakes*. Durational properties consistently distinguish between the two types of word: voice onset time, the initial five segments, and [ɪ] are longer in prefixed words, while [s] is shorter. Amplitude differences were found for the /mɪs/ words.

In sum, these studies provide evidence for phonetic boundary effects (see, however, Hanique and Ernestus 2012 for a challenge of this view). However, they have focused on boundary effects in words with one morphological boundary only. It is still unknown how these effects shape the phonetic structure of words with two or more boundaries, which are very frequent in languages such as English.[1]

The present paper is devoted to this problem. In particular, we investigate the relation between morphological embedding and phonetic implementation in words with more than one morphological boundary. Combining insights on the phonetic realization of bimorphemic words and insights from studies of multiple affixation, we may expect to find that the internal bracketing of words with more than two morphemes is reflected in the phonetic implementation of the constituents. Specifically, it may be suggested that the degree of embedding somehow correlates with the degree of phonetic reduction. We put forward what we call the 'Embedded Reduction Hypothesis', as formulated in (2).

(2) **Embedded Reduction Hypothesis**
In a complex word with more than two constituents, embedded forms show more phonetic reduction than forms at higher derivational levels.

The following sections will be devoted to investigating this hypothesis with experimental data consisting of triconstituent noun-noun-noun (NNN) compounds such as *day care center* and *business credit card.*

3 Methodology

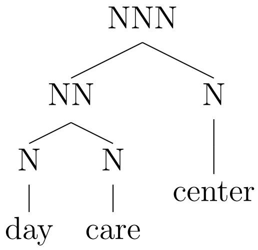
3.1 The structures under investigation: Triconstituent NNN compounds

English NNN compounds are a good testbed for investigating the Embedded Reduction Hypothesis: they show the necessary hierarchical morphological structure, are fairly frequent and allow us to control for sequencing effects since the embedded complex constituent can occur in the left and in the right position. They are usually analyzed as having a two-level binary structure. At the level of the immediate constituents, we find one complex constituent consisting of two nouns, i.e. an NN compound, and one other immediate constituent consisting of a single noun. The two kinds of immediate constituents may occur in either order, as illustrated in (3) and (4).

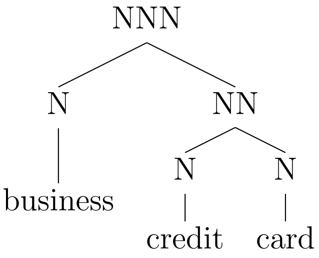
---

[1] Furthermore, it is unclear whether listeners take advantage of the morphological information that is encoded in the acoustic signal in morphological processing. This question is beyond the scope of the present paper.

(3)



(4)



In those cases where the embedded complex constituent is on the left-hand side we speak of 'left-branching' compounds, as in (3), and in those cases where the embedded complex constituent is on the right-hand side we speak of 'right-branching' compounds, as in (4). The non-embedded constituent will be referred to as 'free' while constituents that are part of the embedded compound will be referred to as 'embedded'. For example, in the compounds [[*day care*] *center*] and [*business* [*credit card*]], *center* and *business* are the 'free' constituents, while *day* and *care*, and *credit* and *card* are all 'embedded' constituents. Due to their location at the boundary next to the free constituents, *care* and *credit* will additionally be referred to as 'adjacent embedded constituents'.

Shortening is considered in the literature as a primary acoustic correlate of phonetic reduction (see, for example, the work by Ernestus and colleagues, such as Ernestus and Warner 2011; Pluymaekers et al. 2005, 2006). In order to test for a reduction effect of morphological embedding, we will compare the acoustic duration of the embedded constituents to that of the free constituents. We can then derive the predictions in (5) from the Embedded Reduction Hypothesis:

(5) **Predictions for NNN compounds based on the Embedded Reduction Hypothesis**

a. The embedded constituents should have a relatively short acoustic duration.

b. The free constituents should have a relatively long acoustic duration.

c. Predictions a. and b. should hold irrespective of branching direction (i.e. they should hold for both left-branching compounds and right-branching compounds)

These predicted duration differences in (5a) and (5b) should be observable on top of those caused by differences in the number of phonemes between the constituents. Since the Embedded Reduction Hypothesis is formulated without reference to the direction of branching, it follows that branching direction should not have an effect (unless there were some other as yet unknown kind of influence of embedding on phonetic reduction). This expectation is expressed in prediction (5c).

3.2 Data

The present study is the first to explore the phonetic correlates of morphological embedding and has the character of a pilot study. For this pilot study we decided to re-use a data set available to us from a previous investigation of compound prominence (Kösling (2013, ch. 5) and Kösling et al. (2013)), as a corpus for the present investigation of phonetic reduction in compounds.

The purpose of the original experiment was to compare the prominence pattern in triconstituent compounds with different internal branchings, and the data set is therefore not ideally suited for purposes of the present paper, as not all variables pertinent to the present study were experimentally controlled. However, the experimental data can be used as a corpus as it contains the right kinds of compound structures, and potentially confounding variables can be statistically controlled by introducing pertinent covariates. In the following we describe the stimuli and the experimental procedure.

The test stimuli consisted of 20 left-branching and 20 right-branching triconstituent compounds. As is well known, the branching direction of compounds with more than two constituents is not always straightforward. For example *kitchen towel rack* could be a rack for kitchen towels (and thus left-branching), or a towel rack for the kitchen (and thus right-branching). While cases exist where the branching direction is unclear, such compounds seem to be in the minority. For example, of the 505 triconstituent compounds randomly sampled from the Boston University Radio Speech Corpus by Kösling and Plag (2009), the authors considered only 11 percent to be neither clearly right-branching nor clearly left-branching. The qualification 'clearly' means that a speaker of English would naturally interpret the compound in a particular way in the context in which they appeared in the corpus, and would do so in spite of the fact that an alternative interpretation is in principle always available (due to independent principles of compound interpretation in English, see Bauer et al. 2013, ch. 20). For example, *ice cream parlor* is naturally interpreted as left-branching, although a right-branching interpretation would in principle be available (a cream parlor having to do with ice).

Not surprisingly, the speakers' branching intuition is usually supported by independent evidence. In the case of *ice cream parlor*, the left immediate constituent *ice cream* is an established compound. It can be found in dictionaries, and has a much higher frequency than the putative embedded compound *cream parlor*. In lexical access, listedness and high bigram frequency usually lead to the desired branching direction (and interpretation) of the compound as a whole.

In order to make sure that the participants in the experiment used the branching direction intended with a given item, Kösling chose compounds with an unambiguous branching direction (in the above sense). She made sure that only compounds were chosen as embedded compounds that were attested in at least one of the four dictionaries (the *Oxford Student's Dictionary of American English* (Hornby, 1983), the *Longman Dictionary of American English* (Bullon, 2002), the *Longman Advanced American Dictionary* (Summers, Della, 2000), and the *Oxford Advanced Learner's Dictionary* (Hornby, 1995). Additionally, the constituents were selected such that the bigram frequency for the two embedded constituents was much higher than the bigram frequency for the free constituent and the adjacent embedded constituent.

Participants were asked to read out aloud sentences while being recorded. The sentences either contained the constructed stimuli or some other structures that served as distractors. The compounds were embedded in carrier sentences that resembled a natural speech context. Thus, instead of using the same standard carrier sentence for all compounds, each compound came with its own individual carrier sentence that fitted the semantics of the triconstituent compound used in this sentence. Although the syntactic position of a compound does not seem to have a crucial effect on its prominence pattern (e.g. Plag 2006) this factor was controlled by placing each compound in object position of its carrier sentence. Furthermore, all compounds were presented as discourse-new information. Discourse-old information may lead to a deaccentuation of the head constituent of a given compound (e.g Hirschberg 2002). All compounds were followed by a two-word temporal adverbial starting with *this* in sentence-final position. Examples of carrier sentences with left-branching and right-branching compounds are given in (6) and (7), respectively. (see Appendix 2 for the full list of sentences used in the experiment).

(6) She worked at a [[*day care*] *center*] last year.
He started [[*hay fever*] *treatment*] last week.

(7) He signed up for a [*business* [*credit card*]] last month.
He missed the [*family* [*Christmas dinner*]] last night.

All utterances were digitally recorded in a sound-proof booth at the University of Toronto. If a sentence contained obvious disfluencies or pronunciation errors, the participant was asked to repeat the sentence at the end of the recording session. The data set contains recordings from 7 female and

6 male speakers of North American English, aged between 18 and 27 years. Further details of the methodology are given in Kösling (2013, ch. 5).

For the analysis in Kösling (2013), the recorded compounds were manually annotated by a phonetically trained linguist using the phonetic software Praat (Boersma and Weenink, 2013). Four compound tokens turned out to be either mispronounced or to be pronounced with discontinuities, and were therefore excluded from the data set. Also excluded were the recordings from one male speaker due to his highly idiosyncratic use of intonation (which might have been due to his profession as an actor).

4 Analysis

4.1 Overview

Previous studies of the acoustic duration of words have shown that duration is influenced by various acoustic and non-acoustic factors, for example speech rate, frequency, phonological length, accentuation, or phrase-final lengthening. In order to investigate the effect of embedding on the relative durations of the constituents in the target NNN compounds, it is therefore necessary to control these factors statistically, for example by using multivariate regression. We used linear mixed effects regression models for our study (for an introduction, see Baayen et al. 2008). The set of co-variates included in the present study is similar to that used in other studies on the duration of morphologically complex words (for example Pluymaekers et al. 2005, 2010; Hanique et al. 2013). We fitted a linear mixed-effects regression model to the data. This type of multivariate model allows us to look at the contribution of any predictor variable of interest while at the same time accounting for the contribution of the other predictor variables. It also offers a way of bringing under statistical control the various potentially confounding factors mentioned above (and discussed below in more detail). The variables included in our statistical analysis are listed in table 1. We will briefly discuss each variable, starting with the dependent variable and the variables of interest, and moving on to the covariates in the next subsection.

Dependent variable: Duration. Constituent duration (in seconds) was the dependent variable of the model. We measured the duration of each constituent as the time span between the starting point and the end point of each constituent in the remaining 477 compounds. With three constituents per compound, the total number of measurements amounted to 1431 observations. The mean duration of individual constituents was 0.365 s. The constituent with the shortest mean was the second constituent *ring* in *dia-*

Table 1: Variables included in the initial regression model

Dependent variable	
DURATION	Measured constituent duration (in seconds)
Independent variables	
CONSTITUENT	Position of constituent within compound (N1, N2, or N3)
BRANCHING	Branching direction of the compound (Left or Right)
FREQUENCY	Corpus frequency of constituent
BIGRAMFREQN1N2	Corpus frequency of bigram N1N2
BIGRAMFREQN2N3	Corpus frequency of bigram N2N3
ACCENT	Is a pitch accent expected on constituent? (Yes or No)
PITCH	Mean pitch in semitones of constituent
PHONOLENGTH	Phonological length (principal component of phoneme and syllable number)
Interactions	
CONSTITUENT × BRANCHING	
BRANCHING × BIGRAMFREQN1N2	
BRANCHING × BIGRAMFREQN2N3	
CONSTITUENT × BRANCHING × BIGRAMFREQN1N2	
CONSTITUENT × BRANCHING × BIGRAMFREQN2N3	
Random effect	
SPEAKER	Speaker identifier

mond ring exhibition ($M = 0.205$ s, $SD = 0.030$ s), while *manufacturer* in *silicon chip manufacturer* had the longest mean ($M = 0.724$ s, $SD = 0.034$ s).

Variables of interest: Constituent and branching. As we are interested in a comparison between the acoustic duration of the embedded and the free constituents, an interaction term between the categorical factors Constituent and Branching was included to incorporate the two different compound structures shown in the tree diagrams (3) and (4) above.[2]

4.2 Covariates

Frequency of constituent, **Bigram frequencies**. Frequency affects phonetic duration, such that more frequent words tend to show more reduction (e.g. Bybee 2001, 78, Jurafsky et al. 2001). In order to account for the shortened duration of words with high frequencies of occurrence, the corpus frequency of each constituent was included as the variable Frequency in the analysis. In addition, we also incorporated the corpus bigram frequencies of the first and second constituent and of the second and third constituent as variables BigramFreqN1N2 and BigramFreqN2N3, respectively. The primary reason for including the two bigram frequencies lies in the observation that the predictability of a word in its context has a notable effect on its acoustic duration (e.g. Jurafsky et al. 2001; Pluymaekers et al. 2005): words occurring in a highly predictable context are likely to have a shorter duration than in an unexpected context. Therefore, two nouns that co-occur very frequently are likely to be relatively short, and this acoustic reduction may occur independently of any reduction that is due to their position within a NNN compound.

However, this influence of the two bigram frequencies can be expected to be modified by the branching direction: in a left-branching compound, BigramFreqN1N2 corresponds to the frequency of the embedded constituent, whereas in a right-branching compound, this bigram frequency corresponds to the frequency of the combination of the free constituent and the adjacent embedded constituent. These potential interactions were accounted for by including two three-way interactions in our model: Constituent × Branching × BigramFreqN1N2 and Constituent × Branching ×

---

[2] At first glance, a simpler way of investigating the acoustic effect of embedding might be to code each constituent as either 'free' or 'embedded', and compare the estimated mean durations for these two types. However, this coding scheme would pool duration measurements from completely unrelated compound constituents: the free constituent in right-branching compounds is N1, but it is N3 in left-branching compounds. Coding both as 'free' obscures the structural differences between left-branching and right-branching compounds, and may therefore disguise potentially important duration differences.

BigramFreqN2N3. In effect, these interactions allow us to assess the influence of the two bigram frequencies on the three constituents separately for left-branching and for right-branching compounds.

All frequency measures were obtained from the DVD version of the Corpus of Contemporary American English (COCA, Davies 2008-) using the corpus query tool *Coquery* (Kunter, 2015). The bigram frequencies include all three spelling variants found in English noun-noun compounds (i.e. written as a single word, written as two hyphenated words, written as two words separated by a space). Due to the highly skewed distribution typically found with lexical frequencies, they were log-transformed before they were included in the analysis.

There is, however, a potential problem that originates from the nature of the data set chosen for the present study. As stated above, the triconstituent compounds were constructed in such a way that the two bigram frequencies were explicitly unbalanced: high bigram frequencies of the embedded constituents co-occurred with low bigram frequencies of the cross-boundary constituents. In other words, BigramFreqN1N2 was generally much higher than BigramFreqN2N3 in left-branching compounds, and vice versa in right-branching compounds. The results in Kösling (2013) and Kösling et al. (2013) indicate that this construction served its purpose as to disambiguate the internal branching direction both for left-branching and right-branching compounds. These acoustic studies were concerned primarily with the shapes of the pitch contours of the compounds, and these contours are probably not affected very much by the imbalance between the two bigram frequencies.

Yet, as it will be shown below, it can be difficult to disentangle the effect of these frequencies on the individual constituent durations in a statistical analysis. For example, while the bigram frequency of *care center* in the left-branching *day care center* is clearly lower than the bigram frequency of the embedded constituent *day care* (666 vs. 2148), it surpasses by far the bigram frequency of the embedded constituent of *internet page* in the right-branching compound *company internet page* (*internet page*: 5, *company internet*: 0). So far, no research is available on the effects of such bigram frequency differences, and we have to acknowledge that our data set is not very well-suited to identify effects potentially resulting from these uneven distributions.

Length There is a perhaps somewhat trivial relation between acoustic duration and the phonological length of a constituent: in the compound *hay fever treatment*, the first constituent should be shorter, and the last constituent should be longer not because to any effect of morphological embedding, but simply because the constituent itself is longer: when assuming a standard American pronunciation in citation form, *treatment* consists of eight, *fever* of five, and *hay* of two phonemes. In addition, the constituents in our data set differ not only in the number of phonemes, but also in the number of syllables: the shortest constituents consist of a single syllable like *hay*, but there are also constituents with four (*community*) or five (*manufacturer*) syllables, and they can be expected to have a longer acoustic dutation than words with fewer syllables. And of course, there will also be a relation between the number of phonemes and the number of syllables: in general, the number of phonemes will increase with increasing number of syllables.

In order to incorporate the relation between the phonological length of the constituents and their acoustic duration, we obtained the number of syllables and the number of phonemes as listed in the lexical database CELEX (Baayen et al., 1995). Yet, as expected, we found a high correlation between the two measures ($r_{Sp} = 0.884$, $p < 0.001$). To avoid the potentially harmful side-effects that the inclusion of highly correlated predictors can have in a multiple regression analysis, we subjected them to a principal component analysis, which revealed that the first component alone can already account for 96.7 percent of the variance expressed by the two measures. An inspection of the rotation matrix showed that this component is dominated by the phoneme number (the loading is 0.92), and to a much lesser degree by the syllable number (component loading: 0.40). We therefore used the first principal component as a predictor variable named PHONOLENGTH (for additional details on principal component analysis, and for an example of a similar procedure in multiple regression, see Baayen 2008, ch. 5.1.1 and ch. 6.22).

Accentuation It is well known that the presence or absence of a pitch accent on a syllable affects the duration of that syllable: other things being equal, accented syllables are notably longer than unaccented syllables (see, for instance, Sluijter and Heuven 1996; Turk and White 1999). This pattern also holds for English binominal compounds: in right-prominent noun-noun compounds, the second constituent is generally longer than in left-prominent compounds due to the presence of a pitch accent on that constituent (e.g. Farnetani et al. 1988, Kunter 2011, ch. 5). Thus, in order to separate the effect on acoustic duration of accentuation from that of morphological embedding in the present analysis, it is necessary to account for the prominence pattern of the triconstituent compounds.

The problem with triconstituent compounds is that it is far from trivial to determine their prominence pattern. It had long been assumed that the branching direction alone is responsible for the prominence pattern ('Lexical Category Prominence Rule', Liberman and Prince 1977). More recent work (e.g. Kösling and Plag 2009; Giegerich 2009) has shown, however, that such an account is empirically inadequate. In particular, Kösling and Plag (2009) using data from a radio speech corpus, and Kösling et al. (2013),

Table 2: Accent patterns by compound type

Type of compound	accent on N1	accent on N2	accent on N3
L/N1	yes	no	no
L/N2	yes	yes	no
R/N2	yes	yes	no
R/N3	yes	yes	yes

Kösling (2013) using the present data set have shown that the prominence pattern in these compounds depends on the branching direction and on the prominence pattern within the complex constituent. They demonstrate that there are different prominence patterns, each of which is realized by a particular constellation of the presence or absence of pitch accents on the three constituents.

Table 2 gives the three different accent patterns that emerged from their data analysis. These accent patterns are averaged over the respective subset of compounds in the data set, and are based on an analysis of the pitch contours measured for these compounds (see (Kösling et al., 2013, 547) for details). 'L/N1' refers to left-branching compounds whose embedded constituent is left-prominent, 'L/N2' refers to left-branching compounds whose embedded constituent is right-prominent,'R/N2' refers to right-branching compounds whose embedded constituent is left-prominent, 'R/N3' refers to right-branching compounds whose embedded constituent is right-prominent.

For example, based on the generalization over the respective subset of compounds in our data set, the left-branching compound *hay fever treatment* in our data set is expected to have a single pitch accent on the first constituent, while the right-branching compound *business credit card* in our data set can be expected to have one pitch accent on the first constituent and a second pitch accent on the second constituent. Given that these accentuation patterns were derived as generalizations from the very same data set that is used in the present analysis, we decided to encode the expected presence or absence of a pitch accent as shown in table 2 as the variable ACCENT in our analysis.

Yet, as these expected accents are based on averages over each of the four types of compound, and are therefore glossing over speaker- and type-dependent variations in the accentuation patterns, some tokens may not conform to the general trend in its category, which may reduce the predictive power of this variable. We therefore included actual pitch measurements as a second co-variate that addresses the accentuation pattern of the triconstituent compounds.[3] It has been shown in numerous publications that pitch is the most reliable cue to compound prominence (e.g. Kunter and Plag 2007; Kunter 2011; Kösling et al. 2013). The analyses in these publications used the mean pitch of each constituent as variables affecting prominence, but other ways of measuring pitch are also possible, such as the maximum pitch, or the pitch range in a given constituent. Which measure, then, is most appropriate in the present case to incorporate the pitch information?

We decided to determine the most appropriate model by the following approach. First, we fitted a linear mixed effects model that contained all other predictors described in this section, but which excluded any predictor pertaining to accentuation or pitch. We then fitted several models with additional predictors, and compared them to the reference model in order to see in how far the addition of the predictors improved the model fit. Model comparison was done by means of a log-likelihood test (with one degree of freedom): if the additional acoustic predictor in a model is significantly affecting the acoustic duration of the constituent, the likelihood of that model will be significantly higher than the likelihood of the reference model. We also investigated the changes in the Akaike Information Criterion (AIC, cf. Akaike 1974) as an index of the goodness-of-fit of the model. This index basically relates the likelihood to the number of estimated parameters: a model with a lower AIC is a model that yields a higher likelihood while investing as many or even fewer parameters than the reference model.

The following four acoustic parameters were tested in this way: MeanPitch (the average pitch in each constituent), RelativePitch (the difference between MeanPitch and the average pitch for the whole speaker), MaximumPitch (the highest pitch measurement in each constituent), PitchRange (the difference between the highest and the lowest pitch measurement in each constituent). The acoustic measures were obtained using *Praat* (Boersma and Weenink, 2013). Similar to the procedure in Kösling et al. (2013), the pitch contour was first extracted from the speech signal, then interpolated to fill any gap in the contour, and finally smoothed to reduce the effect of microprosodic variation and algorithmic imprecisions. All pitch

---

[3] In principle, it would also be possible to use listener ratings of the accentuation patterns for each token. However, such an approach has several disadvantages. First of all, it has been shown that identifying accentuation patterns in compounds is not a trivial task: in the perception study presented in Kunter (2011, ch. 4), only 17 out of 32 participants were able to provide highly reliable perceptual ratings for the prominence pattern in noun-noun compounds. Thus, in order to obtain reliable prominence ratings for the present data set, an additional large-scale perception study would be necessary. Given the pilot character of the present investigation, we did not carry out such an additional study, but used the generalizations from Kösling et al. 2013 and acoustic pitch measurements instead. As it turned out, these variables reliably showed the expected effects.

Table 3: Model comparisons for different pitch measurements

Pitch measurement	log likelihood	χ^2	p	*AIC*	Reduction
(none, i.e. the reference model)	2011.0	n.a.	n.a.	-3990.0	n.a.
Mean pitch	2011.1	0.273	0.601	-3988.2	-1.8
Relative pitch	2011.1	0.298	0.585	-3988.3	-1.7
Maximum pitch	2013.1	4.195	0.041	-3992.2	2.2
Pitch range	2017.2	12.483	≤ 0.001	-4000.4	10.4
Accent	2037.5	52.971	≤ 0.001	-4040.9	50.9
Accent and Pitch range	2070.9	60.286	≤ 0.001	-4093.8	103.8

measurements were taken separately for each constituent in semitones relative to 100 Hz. We also fitted a model that added ACCENT to the reference model, which allows us to compare the predictive power of this categorical variable to those of the different acoustic measurements. Finally, we fitted a model that included the predictors of the reference model, the predictor ACCENT, and that acoustic predictor that turned out to perform best when added to the reference model. These models allows us to test whether we need ACCENT, or the acoustic predictor, or both of them in the final model.

Table 3 shows that adding MEANPITCH and RELATIVEPITCH to the model do not lead to a significant improvement over the reference model. MAXIMUMPITCH and PITCHRANGE are both significant, but the AIC reduction of PITCHRANGE is notably greater than that of MAXPITCH. We therefore consider PITCHRANGE to be the most appropriate predictor out of the four pitch measurements. Yet, as the last column of the table shows, the improvement of the model that is gained by adding ACCENT is much larger than that stemming from either of the acoustic measurements. This suggests that the accent-induced lengthening found in the present compounds can be captured by the factor ACCENT quite efficiently already, much more so than by any of the conventionally used pitch measurements. Adding both the categorical predictor ACCENT and the best acoustic measurement PITCHRANGE to the reference model even further increases the AIC reduction, as the last row in the table shows. Therefore, we incorporated these two predictors in our model to accompany the effects of accentuation on constituent durations.

Speaker. Finally, the variable SPEAKER was included as a random factor in the model. This component accounts for speaker-dependent durational differences (for example, if speakers produce the constituents on average faster or slower than the other participants).

5 Results

The last model from table 3 served as the initial model for the analysis. In order to identify individual outliers that exerted an unduly high influence on the regression analysis, Cook's distances d for the observations were calculated using the `influence.ME` package[4]. All in all, the Cook's distances were found to be relatively low (overall median of d: $M = 0.0002$, $IQR = 0.0007$). A visual inspection of the distribution of d revealed that there were a few observations which had relatively high values of d that fell outside of the overall distribution. Accordingly, 13 observations with Cook's distances larger than 0.0055 were discarded from the regression analysis. An inspection of the variables included in the model suggested no need for further data trimming, so the initial model was refitted on the reduced data set of 1418 durations. The usual diagnostic plots for this refitted model revealed no violation of the model assumptions. In particular, there was no indication of heteroscedasticity, and the residuals of the model were considered to be sufficiently normally distributed.

As table 4 shows, all terms included in the model reached statistical significance, including the two three-way interactions. In this table, the reference level for the intercept is ACCENT=no, CONSTITUENT=N1, and BRANCHING=Left; the coefficients for categorical variables express the change in duration of the associated factor level in relatio to the reference level. The probability of the t values was obtained using the Satterthwaite approximation of the degrees of freedom using the `lmerTest` package[5].

---

[4] `http://cran.r-project.org/web/packages/influence.ME/index.html`

[5] `http://cran.r-project.org/web/packages/lmerTest/index.html`

Table 4: Fixed effects in final regression model predicting Duration ($N = 1418$; reference level: Accent=no, Constituent=N1, Branching=Left).

	B	Std. Error	d.f.	t	$P(>\|t\|)$	
(Intercept)	0.364	0.020	1202.7	18.306	<0.001	***
Constituent=N2	0.042	0.023	1385.0	1.842	0.066	
Constituent=N3	0.028	0.024	1385.4	1.206	0.228	
Branching=Right	-0.012	0.020	1385.1	-0.576	0.565	
log Frequency	-0.004	0.001	1385.0	-2.811	0.005	**
log BigramFreqN1N2	0.000	0.003	1385.1	0.128	0.898	
log BigramFreqN2N3	-0.002	0.002	1385.0	-1.267	0.205	
PhonoLength	0.038	0.001	1385.3	47.252	<0.001	***
Accent=yes	0.034	0.004	1385.0	7.959	<0.001	***
Pitch range	0.001	0.000	1394.1	2.571	0.010	*
Constituent=N2 : Branching=Right	0.012	0.028	1385.0	0.417	0.677	
Constituent=N3 : Branching=Right	0.029	0.028	1385.1	1.037	0.300	
Constituent=N2 : log BigramFreqN1N2	-0.006	0.004	1385.0	-1.668	0.096	
Constituent=N3 : log BigramFreqN1N2	0.009	0.004	1385.2	2.265	0.024	*
Constituent=N2 : log BigramFreqN2N3	0.003	0.003	1385.0	1.119	0.263	
Constituent=N3 : log BigramFreqN2N3	-0.010	0.003	1385.0	-3.535	<0.001	***
Branching=Right : log BigramFreqN1N2	-0.007	0.004	1385.1	-1.894	0.058	
Branching=Right : log BigramFreqN2N3	0.006	0.003	1385.0	2.226	0.026	*
Constituent=N2 : Branching=Right : log BigramFreqN1N2	0.006	0.005	1385.0	1.079	0.281	
Constituent=N3 : Branching=Right : log BigramFreqN1N2	-0.020	0.005	1385.1	-3.601	<0.001	***
Constituent=N2 : Branching=Right : log BigramFreqN2N3	-0.015	0.004	1385.0	-3.661	<0.001	***
Constituent=N3 : Branching=Right : log BigramFreqN2N3	0.005	0.004	1385.0	1.212	0.226	

For the predictors that do not enter higher-order interactions, we find the expected directions of effect. There is a significant negative coefficient for log FREQUENCY: frequent words tend to have a shorter acoustic duration. Likewise, the significant positive coefficient for PHONOLENGTH reflects the expected pattern of phonological length on word duration: with increasing phonological length of the constituent, the duration of the constituent also increases significantly. We also find significant coefficients for ACCENT and PITCH RANGE, which reflect the lengthening effect of pitch accents. On average, a constituent that is expected to bear a pitch accent is 0.034 s longer than those which are expected to be unaccented, which is similar to the lengthenings reported e.g. in Sluijter and Heuven (1996) and Turk and White (1999). The significant effect of pitch range may also be interpreted as a reflection of accentuation: by the very nature of pitch accents as targeted excursions of the intonation contour, a constituent with a large pitch range is more likely to be accented, and hence, more likely to be longer due to accentual lengthening, than an unaccented constituent.

The partial effects of FREQUENCY, PHONOLENGTH, ACCENT, and PITCHRANGE are displayed in figure 1. Below the plots for the numeric predictors, box-whisker plot show the distribution of values of each predictor in the data. The thick line indicates the second quartile (i.e. the median), and the left and right edge of the box are the location of the first and third quartile. Thus, each box contains the 50 percent of the predictor values around the median, i.e. those values that may typically be encountered for that predictor. There is a single statistical outlier with regard to constituent length, as indicated by the dot on the right end of the scale: there is one constituent that is exceptionally long (*manufacturer* in *silicon chip manufacturer*). However, the data points relating to this outlier do not show unusually high Cook's distances – apparently, the observations for this constituent to not exert unduly high leverage on the model estimates. As all plots use the same vertical scaling, the slope of the partial effect may be interpreted as an indication of effect strengths. As expected, the effect of phonological length is by far greater than that of the other main effects.

Central for the three predictions in (5) are the two significant three-way interactions CONSTITUENT × BRANCHING × BIGRAMFREQN1N2 and CONSTITUENT × BRANCHING × BIGRAMFREQN2N3. The nature of these interactions are illustrated in figure 2 for the log bigram frequency N1N2 and in figure 3 for the log bigram frequency N2N3. These partial plots show the estimated effect of the involved variables while holding the effect of all other variables constant. In each figure, the left panel shows the effect of the respective bigram frequency for left-branching compounds, and the right panel shows the effect for right-branching compounds (i.e. with the internal struc-

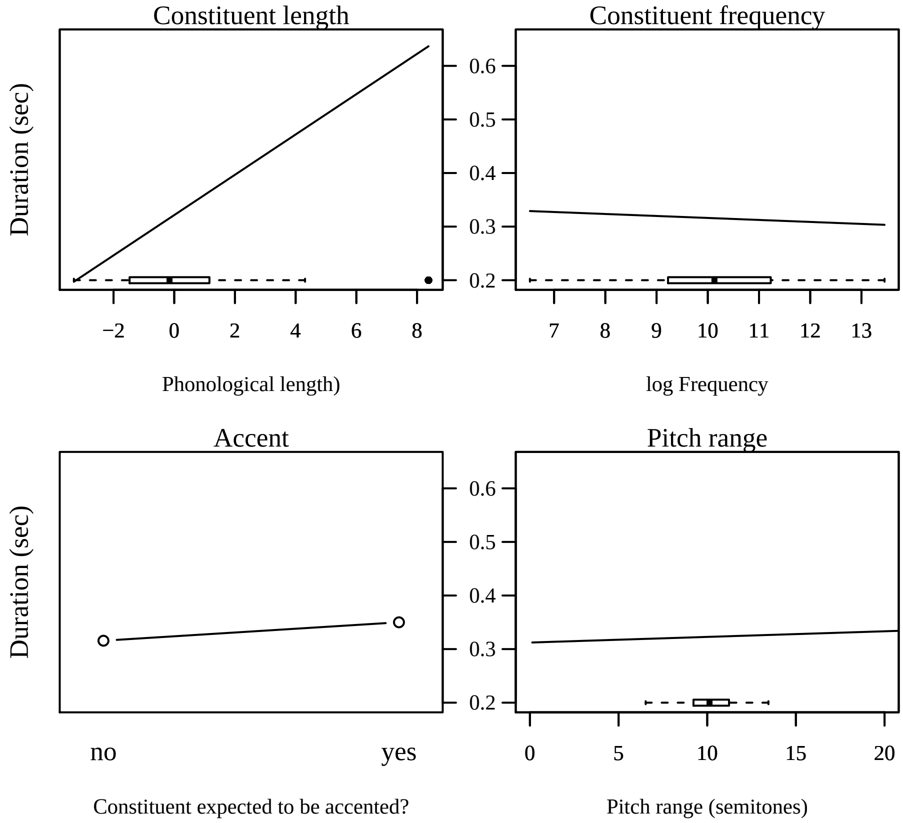


Figure 1: Partial effects of PhonoLength (upper-left panel), Accent (upper-right panel), and log Frequency (bottom panel).

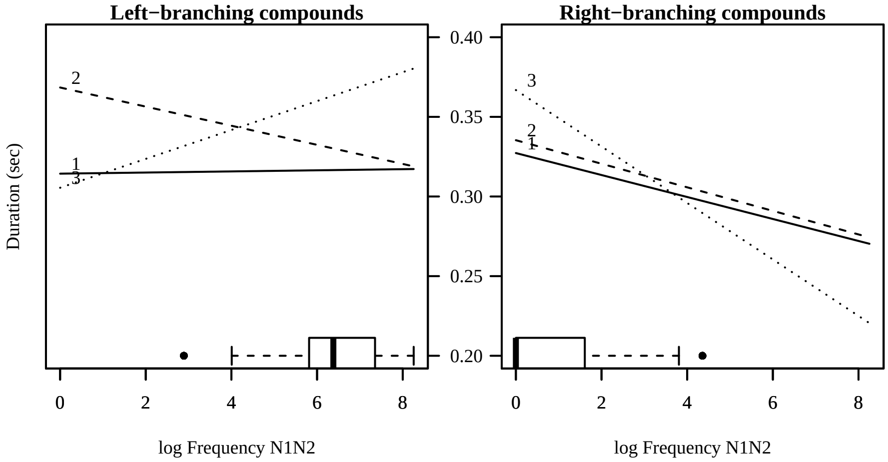


Figure 2: Partial effects of the three-way interaction involving BigramFreqN1N2 and constituent number. Left panel: left-branching compounds, right panel: right-branching compounds.

ture N[NN]). The three lines in each panel represent the estimated changes of the three constituent durations with changes of the bigram frequency for the respective branching type. As in the graph for Length in figure 1, box-whiskers plots at the bottom of each panel show the distribution of the bigram frequencies, split by left-branching and right-branching compounds. We will start with a closer look at the figure for BigramFreqN1N2.

The two panels in figure 2 illustrate that the bigram frequency of N1 and N2 affects the constituent durations in the compounds in our data set, but this effect also depends on the branching direction of the compound. In the panel for left-branching compounds (i.e. compounds with the internal structure [NN]N), BigramFreqN1N2 reflects the frequency of the embedded constituent. Here, the regression line for the first constituent N1 (solid line) is parallel to the horizontal axis. This means that the duration of N1 in left-branching compounds is invariant with regard to the bigram frequency of N1 and N2. This is not true for N2 (dashed line), for which we find a negative correlation with this frequency: with increasing bigram frequency of N1 and N2, the duration of N2 decreases. In other words, the frequency of the embedded compound manifests itself mostly in a reduced duration of the second constituent, while the first constituent is largely unaffected by it. At the same time, this bigram frequency also has a lengthening effect on the duration of the free constituent N3 (dotted line). Higher frequencies of the embedded complex constituent systematically co-occur with longer durations of N3. Starting from bigram frequencies of about 4.0, the duration difference between N2 and N3 becomes progressively more pronounced in the direction predicted by the Embedded Reduction Hypothesis. Given that the box-whisker plot indicates that frequencies below that point are rather uncommon, the panel shows that, other things being equal, in most [NN]N compounds the free constituent N3 is either similar or longer than the duration of the adjacent embedded constituent N2. For bigram frequencies of the embedded compound at the median (i.e. for half of all compounds), the free constituent N3 is estimated to be longer by 0.035 s than the adjacent embedded constituent N2 (and even longer than N1); for the 75-percent quartile, the difference already amounts to 0.048 s.

For right-branching compounds (i.e. with the internal structure N[NN]) shown in the right panel of figure 2, where BigramFreqN1N2 reflects the cross-boundary frequency, the effect on the constituent duration turns out to be rather different. Here, the duration of constituents N1 and N2 are both negatively correlated with the bigram frequency of N1 and N2; their duration decreases with increasing bigram frequencies. Such a negative correlation between bigram frequency and constituent duration can also be observed for the third constituent N3, but here, the slope of the partial regression line is much steeper. In other words, in right-branching compounds, the duration of all three constituents becomes shorter with increasing values of BigramFreqN1N2, and this effect is particularly pronounced for the third constituent. When comparing the two branching types, it is noteworthy that the effect of the bigram frequency on the duration of the third constituent is reversed: in left-branching compounds, higher values of BigramFreqN1N2 lead to particularly long durations of N3, but in right-branching compounds, the duration of N3 becomes particularly short. Note that due to the way the triconstituents were constructed (cf. section 3.2), the distribution of the bigram frequencies of N1 and N2 is very skewed for right-branching compounds: 75 percent of the compounds have a median bigram frequency N1N2 of less than 5. This means that the estimation of the effect of BigramFreqN1N2 in right-branching compounds may not be as robust as the estimation in left-branching compounds.

Turning now to the three-way interaction involving the bigram frequency of N2 and N3, some similarities, but also some differences to the previous interaction become visible. In the left-branching compounds in the left panel of figure 3, the partial regression line for N1 (solid line) does not change notably with increasing values of BigramFreqN2N3. Apparently, the first noun of triconstituent compounds retains its duration, and as the dashed line shows, the duration of N2 is also unaffected by bigram frequency of N2 and N3 (indeed, the coefficients for both constituents are not significant in table 4). The only constituent that is affected by this bigram in left-branching

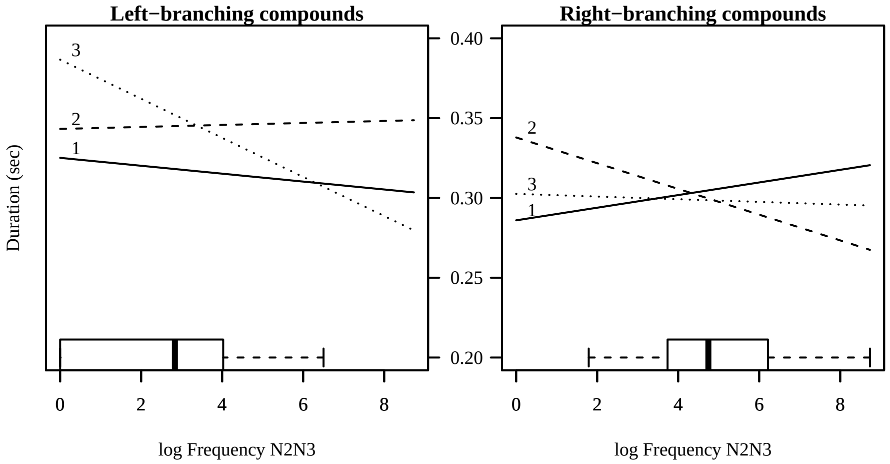


Figure 3: Partial effects of the three-way interaction involving BIGRAM-FREQN2N3 and constituent number. Left panel: left-branching compounds, right panel: right-branching compounds.

compounds is N3: the third constituent becomes significantly shorter with increasing bigram frequencies. Apparently, for typical values of BIGRAM-FREQN2N3 within the range of the box-whisker plot, the duration of the free constituent N1 and the adjacent embedded constituent N2 are rather similar, while the other embedded constituent N3 is longer than either, but becomes similar in duration with increasing bigram frequency of N2 and N3.

A different pattern is visible in the right panel for right-branching compounds, in which the bigram frequency of N2 and N3 corresponds to the frequency of the embedded compound. Here, the slope of the partial regression line for the free constituent N1 is significantly positive: with increasing bigram frequencies of N2 and N3, the duration of N1 becomes longer. The reverse effect is found for the adjacent embedded constituent N2, as the negative slope of the regression line reveals. While all constituents have very similar average durations at the median bigram frequency, the duration difference between N1 and N2 amounts to about 0.025 s at the right edge of the box in the box-whisker plot, i.e. at the 75-percent quantile. The duration of the remaining embedded constituent N3 remains largely unchanged across the whole value range.

Notably, the effect of the frequency of the embedded compound on the duration of the constituents in right-branching compounds is very similar to that observed in the left panel of figure 2, which shows the effect of the frequency of the embedded compound in left-branching compounds. In both cases, the duration of both the free constituent and the adjacent embedded constituent are sensitive to the bigram frequency of the embedded constituents, with the free constituent increasing and the embedded constituent decreasing in duration with increasing frequency. The non-adjacent constituent is mostly unaffected by the bigram frequency in either branching direction. However, while the duration of N1 in left-branching compounds is notably shorter than the duration of the free constituent, the duration of the three constituents in right-branching compounds is relatively similar for a large proportion of observations.

Summarizing the two three-way interaction with regard to the predictions following from the Embedded Reduction Hypothesis in (5), the following durational patterns emerge: The embedded constituents N1 and N2 in left-branching compounds have a shorter duration than the free constituent N3 for typical values of BigramFreqN1N2 (i.e. the frequency of the embedded compound) and BigramFreqN2N3 (i.e. the cross-boundary bigram frequency). The duration difference between the adjacent embedded constituent N2 and the free constituent N3 increases with increasing values of BigramFreqN1N2. This durational pattern is predicted by the Embedded Reduction Hypothesis. The effect is somewhat diminished for very high values of BigramFreqN2N3, but as noted above, the robustness of this finding needs to be tested against data sets with less skewed distribution of the cross-boundary bigram frequency.

In right-branching compounds, all constituent have very similar durations for lower to medium values of BigramFreqN2N3 (i.e. the frequency of the embedded compound). With increasing values of BigramFreqN2N3, we find a similar pattern as in left-branching compounds, namely that the free constituent (N1) becomes longer and the adjacent embedded constituent (N2) becomes shorter. Thus, the durational pattern of these two constituents is again as predicted by the Embedded Reduction Hypothesis, but only for higher values of BigramFreqN2N3.

Unpredicted by the hypothesis, the non-adjacent embedded constituent N3 in right-branching compounds is relatively long, in particular for compounds with low values of the cross-boundary frequency BigramFreqN1N2. In left-branching compounds, the non-adjacent embedded constituent N1 is relatively short, regardless of any of the bigram frequencies. Thus, the last prediction of the Embedded Reduction Hypothesis that the effect of the morphological structure on the durational pattern is the same regardless of the branching direction, is only partly supported by the present results.

6 Discussion

In this paper, we tested the predictions based on the Embedded Reduction Hypothesis as formulated above in (2). Specifically, we predicted in (5) that the embedded constituents should be relatively short and the free constituents should be relatively long. Furthermore, we predicted that this generalization should hold both for left-branching and right-branching compounds.

The acoustic analysis of our data set provides support for these predictions, even if with some important modification due to the concomitant influence of the two bigram frequencies, most notably of the bigram frequency of the complex embedded compound. The predicted durational pattern emerges more clearly if this frequency is very high, and is less visible or may disappear for rare embedded compounds. Yet, the observed changes in duration cannot simply be reduced to an effect of lexical frequency: the non-adjacent constituent is largely unaffected by frequency changes, and it is the adjacent embedded constituent in particular that becomes shorter if the embedded compound is more frequent. It is noteworthy that the duration of the free constituent (N3 in left-branching compounds, N1 in right-branching compounds) is also significantly affected by this frequency: if the embedded compound is very frequent, the duration of the free constituent is estimated to be particularly long. In other words, we find a clear effect of the lexical frequency of the complex constituent on the acoustic realization of the free constituent in triconstituent compounds.

These findings imply a very interesting relation between the strength of the morphological boundary within the complex constituent and the durational properties of the cross-boundary constituents. The bigram frequency of the embedded compound (BigramFreqN1N2 in left-branching compounds, BigramFreqN2N3 in right-branching compounds) may be interpreted as a measure of the boundary strength between the embedded constituents: bigrams with a relatively high bigram frequency such as *city hall* (3154 tokens in COCA) will feature a weaker boundary between the two constituents than bigrams with a relatively low bigram frequency such as *weather station* (124 tokens), because speakers will more clearly recognize the two constituents in the latter than in the former. Apparently, this differential of internal boundary strengths is reflected most strongly at the outer morphological boundary, i.e. the boundary between the adjacent embedded constituent and the free constituent: the weaker the internal boundary of the complex constituent is, the longer the free constituent and the shorter the adjacent embedded constituent become. At a theoretical level these effects mean that the durational properties of the cross-boundary constituents are indicative of the hierarchical structure and the strength of the internal boundaries of triconstituent compounds: the higher the bigram frequency of the complex constituent (and thus, the weaker its internal morphological boundary), the more salient the boundary between the complex constituent and the remaining constituent becomes (as marked by the durational properties of the cross-boundary constituents).

The interpretation of the role of cross-boundary bigram frequency in our analysis is less clear. In general, increasing cross-boundary frequencies (i.e. BigramFreqN1N2 in right-branching compounds, and BigramFreqN2N3 in left-branching compounds) have a shortening effect on constituent durations. This effect is rather strong on the third constituent N3 (regardless of branching direction), and much weaker (in right-branching compounds) or even non-existing (in left-branching compounds) on the other two constituents N1 and N2. As noted above, this may be a by-product of the skewed nature of this bigram frequency, and the effect of cross-boundary frequency certainly merits further study with NNN compounds that are better suited for such an investigation than the present set.

Returning to the Embedded Reduction Hypothesis, we can state that its predictions are going in the right direction: acoustic duration correlates with morphological embedding. In particular, the free constituent tends to be relatively long, regardless of branching direction. While lexical factors codetermine the durational properties of complex words and their constituents, the durational properties of its constituents are nevertheless indicative of the internal morphological structure of the word we listen to. Thus morphological structure can be read off the speech signal.

This result has important implications for linguistic theory. The interplay between the frequency of the embedded constituent on the one hand, and the durations of the free constituent and its adjacent embedded constituent on the other hand, is not easily accountable by any theory that assigns phonetic reduction to the post-lexical level. In Lexical Phonology, for example, Bracket Erasure occurs at the end of the stratum in which the NNN compound is formed. Hence, any information on the structure of the embedded constituent, including embedded frequency, should be unavailable at the time at which changes to the phonetic duration of the constituents are applied. Yet, the present analysis strongly suggests that the post-lexical stage must still have access to the internal structure of the morphologically complex word, including lexical information such as the bigram frequency of the embedded constituents.

As an alternative account, it may appear promising to look at the effect of prosodic phrasing on durational differences: other things being equal, there is solid evidence that the speech material around a prosodic boundary is acoustically lengthened (for instance, see Oller 1973, Wightman et al. 1992, Turk and Shattuck-Hufnagel 2000, Byrd et al. 2006, or Turk and Shattuck-Hufnagel 2007, and for a recent overview, White 2014). Using the same data set as the present analysis, Kösling et al. (2013) suggest that there may be a prosodic boundary between N2 and N3 in the left-branching compounds (but not in right-branching compounds). Is the longer acoustic duration of N3 in left-branching compounds simply a case of lengthening at the boundary of a prosodic phrase that is unrelated to embedding? The durational patterns shown in the left panel of figure 2 does not agree well with such an explanation: the lengthening effect of prosodic boundaries is usually much stronger on the pre-boundary word than on the following (Oller 1973, Wightman et al. 1992, Byrd et al. 2006), but in the present data, we find a notable lengthening of N3 (i.e. the post-boundary constituent). More to the point, N2 experiences acoustical shortening rather than lengthening at the boundary, which is also in conflict with the previous accounts of the temporal effect of prosodic boundaries (see, in particular, White 2014 who emphasizes that prosodic structure is expressed through lengthenings, and not shortenings).

However, acoustical lengthening at prosodic boundaries may help explain a different facet of the data. As described above, the target compounds were in object position of the carrier sentences, followed by an adjunct. This sentence type was chosen in order to avoid sentence-final boundary tones on the final constituent of the compound, but nevertheless, speakers may still mark the phrasal boundary between the triconstituent compound and the following adverbial by boundary tones. The effect of such a boundary tone on the duration of the constituents of the compound is, however, not altogether clear. In his review of the relevant literature on lengthening effects of tonal events, White (2014) identifies as the locus of phrase-final lengthening the accented word of the phrase. In the present case, then, a lengthening effect of the phrase boundary would coincide with the corresponding effects of accentuation. If that is so, the statistical model used above should be able to account for boundary effects on length. Yet, it is also possible that in compounds like the ones in the present case, the lengthening effect of a phrasal boundary not only causes the accented constituents to be longer, but may also lengthen the last constituent of the NNN compound that immediately precedes the boundary. Under this account, the regression lines for N3 in the interaction plots would be shifted upwards in relation to the other regression lines due to phrase-final lengthening. This may explain the asymmetry between left- and right-branching compounds with regard to the non-adjacent embedded constituent: while non-adjacent N1 in left-branching compounds was consistently found to be the shortest constituent, non-adjacent N3 in right-branching compounds was not notably shorter than the other constituents, which contradicts one prediction of the Embedded Reduction Hypothesis. If phrase-final lengthening, which affects N3 in right-branching compounds, but not N1 in left-branching compounds, was factored out, the duration of these two non-adjacent constituents may be more similar than in the present analysis. Thus, future research should investigate whether this prediction is borne out by empirical data using triconstituent compounds that are not located at the boundary between sentence constituents.

The analysis shows that the durational difference which is attributable to embedding can amount to 50 milliseconds or more. Such a durational difference seems large enough to be noticeable by listeners (see Klatt and Cooper 1975; Klatt 1976 for discussions of the minimum durational differences that listeners perceive in individual segments). Previous research suggests that it is not unlikely that these differences facilitate processing of the complex words. For example, Shatzman and McQueen (2006) report that listeners use duration as a cue to the boundary between words, while Kemps et al. (2005) have shown that listeners are sensitive to prosodic differences (including durational cues) between an unaffixed stem (i.e., for example *book* as *in a book*) and a stem with a suffix (i.e. *book-* as in *these books*). At present, however, it is not clear whether listeners make use of durational differences when processing more than one boundary within a single morphologically complex word.

Appendix 1: Triconstituent compounds

L1, L2, and L3 refer to the length in number of phonemes.

Branching	N1	N2	N3	L1	L2	L3	Bigram frequencies	
							N1 N2	N2 N3
left	city	hall	restoration	4	3	9	3154	3
left	coffee	table	designer	4	4	7	1807	0
left	cotton	candy	maker	4	5	5	295	20
left	cream	cheese	recipe	4	3	6	1410	3
left	day	care	center	2	3	6	2148	666
left	diamond	ring	exhibition	7	3	8	352	0
left	family	planning	clinic	6	6	6	1462	18
left	field	hockey	player	4	4	5	318	198
left	gene	therapy	technology	3	6	9	474	0
left	hay	fever	treatment	2	5	8	152	0
left	kidney	stone	removal	5	4	6	67	0
left	lung	cancer	surgery	3	6	6	1525	126
left	maple	syrup	production	4	5	8	948	8
left	money	market	fund	4	5	4	495	93
left	science	fiction	book	5	5	3	1838	28
left	security	guard	service	9	3	5	1551	55
left	sign	language	class	3	7	4	565	40
left	silicon	chip	manufacturer	7	3	13	47	16
left	silver	jubilee	gift	6	6	4	17	1
left	weather	station	data	5	5	4	124	10
right	adult	jogging	suit	5	5	3	0	61
right	baby	lemon	tea	4	5	2	1	22
right	business	credit	card	6	6	3	12	5190
right	celebrity	golf	tournament	9	4	8	44	500
right	company	internet	page	7	8	3	0	5
right	conference	time	sheet	9	3	3	7	14
right	family	christmas	dinner	6	8	5	77	228
right	passenger	test	flight	8	4	4	0	117
right	piano	sheet	music	5	3	6	0	350
right	pilot	leather	jacket	5	5	5	0	752
right	pizza	home	delivery	5	3	8	0	93
right	prisoner	community	service	7	9	5	0	1693
right	restaurant	tourist	guide	8	6	3	0	21
right	student	season	ticket	7	4	5	0	202
right	student	string	orchestra	7	5	7	3	40
right	team	locker	room	3	5	3	4	2506
right	tennis	grass	court	5	4	3	0	17
right	tennis	group	practice	5	4	7	5	79
right	visitor	name	tag	7	3	3	0	290
right	woman	fruit	cocktail	5	4	6	0	79

Appendix 2: List of experimental sentences

The following list shows the list of target sentences and filler items used in Kösling (2013) to elicit the triconstituent compounds that were re-used in the present analysis. The sentences are given in the order of presentation.

He watched an old western movie last night.
He rented a nice beach apartment last summer.

She started hay fever treatment last year.
He organized an exciting adventure vacation last week.
He wrote a restaurant tourist guide last month.
He booked a non-stop transatlantic flight last week.
He sold a cotton candy maker last month.
She bought a small wooden horse for her daughter last week.
She voted for a city hall restoration last month.
She attended a Spanish and French class last semester.
He had a lung cancer surgery last year.
She played a Japanese card game with her family last night.
He ordered an adult jogging suit last week.
She made a beautiful farewell gift for her friend last night.
She founded a student string orchestra last month.
She looked at a black and white photo of her grandmother last night.
She read about a coffee table designer last week.
She saw a little green bird last Friday.
He signed up for a business credit card last month.
She was a good and ambitious student last year.
He bought a science fiction book last Friday.
She ordered a non-alcoholic cocktail last night.
He started tennis group practice last month.
She married a big hairy guy last weekend.
She attended a sign language class last week.
He interviewed a track and field athlete last Tuesday.
She worked at a day care centre last year.
She visited her Mexican-American family last June.
She participated in a passenger test flight last year.
She played hide and seek with her children last weekend.
He participated in a celebrity golf tournament last year.
He lost his hand-made scarf last Monday.
He spoke to a silicon chip manufacturer last month.
She learned a second language last year.
He missed the family Christmas dinner last night.
He cancelled an important business appointment last month.
He tasted some baby lemon tea last week.
He met a crazy pop artist last Monday.
He received some weather station data last night.
She listened to a new country band last evening.
He had a kidney stone removal last Friday.
She had good and bad times last year.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics.* Cambridge University Press, Cambridge.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effect modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4):390–412.

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX lexical database.* Linguistic Data Consortium, Philadelphia, PA.

Bauer, L., Lieber, R., and Plag, I. (2013). *The Oxford reference guide to English morphology.* Oxford University Press, Oxford.

Boersma, P. and Weenink, D. (2013). Praat: doing phonetics by computer (version 5.3.41). Computer program. Retrieved 1 March 2013, from `http://www.praat.org/`.

Bullon, S., editor (2002). *Longman Dictionary of American English. Your complete guide to American English.* Longman, White Plains, NY.

Bybee, J. (2001). *Phonology and language use.* Cambridge University Press, Cambridge.

Byrd, D., Krivokapić, J., and Lee, S. (2006). How far, how long: on the temporal scope of prosodic boundary effects. *Journal of the Acoustical Society of America*, 120(3):1589–1599.

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

Davies, M. (2008). The Corpus of Contemporary American English (COCA): 450 million words, 1990-present. `http://www.americancorpus.org`.

Dressler, W. U. (1985). On the predictiveness of Natural Morphology. *Journal of Linguistics*, 21:321–337.

Ernestus, M. and Warner, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics*, 39:253–260.

Farnetani, E., Torsello, C. T., and Cosi, P. (1988). English compound versus non-compound noun phrases in discourse: an acoustic and perceptual study. *Language and Speech*, 31(2):157–180.

Giegerich, H. (2009). The compound stress myth. *Word Structure*, 2:1–17.

Hanique, I. and Ernestus, M. (2012). The role of morphology in acoustic reduction. *Lingue e linguaggio*, 2012(2):147–164.

Hanique, I., Ernestus, M., and Schuppler, B. (2013). Informal speech processes can be categorical in nature, even if they affect many different words. *The Journal of the Acoustical Society of America*, 133(3):1644–1655.

Hay, J. (2007). The phonetics of 'un'. In Munat, J., editor, *Lexical creativity, texts and contexts*, pages 39–57. John Benjamins, Amsterdam / Philadelphia.

Hay, J. and Plag, I. (2004). What constrains possible suffix combinations? On the interaction of grammatical and processing restrictions in derivational morphology. *Natural Language & Linguistic Theory*, 22(3):565–596.

Hay, J. B. (2003). *Causes and consequences of word structure.* Routledge, London.

Hornby, A. S., editor (1983). *Oxford Student's Dictionary of American English.* Oxford University Press, Oxford.

Hornby, A. S. (1995). *Oxford Advanced Learner's Dictionary.* Oxford University Press, Oxford.

Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. D. (2001). Probabilistic Relations between words: evidence from reduction in lexical production. In Bybee, J. and Hopper, P., editors, *Frequency and the emergence of linguistic structure*, pages 229–254. John Benjamins, Amsterdam.

Kaisse, E. M. and Shaw, P. A. (1985). On the theory of Lexical Phonology. *Phonology*, 2(01):1–30.

Kemps, R. J. J. K., Wurm, L. H., Ernestus, M., Schreuder, R., and Baayen, R. H. (2005). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes*, 20(1-2):43–73.

Kiparsky, P. (1982). From cyclic phonology to lexical phonology. In van der Hulst, H. and Smith, N., editors, *The structure of phonological representations*, pages 131–176. Foris, Dordrecht.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59(5):1208–1221.

Klatt, D. H. and Cooper, W. E. (1975). Perception of segment duration in sentence contexts. In Cohen, A. and Nooteboom, S. G., editors, *Structure and process in speech perception*, pages 69–89. Springer, Berlin.

Kösling, K. (2013). *Prominence assignment in English triconstituent compounds*. PhD thesis, Universität Siegen, Siegen.

Kösling, K., Kunter, G., Baayen, H. R., and Plag, I. (2013). Prominence in triconstituent compounds: Pitch contours and linguistic theory. *Language and Speech*, 56(4):529–554.

Kösling, K. and Plag, I. (2009). Does branching direction determine prominence assignment? An empirical investigation of triconstituent compounds in English. *Corpus Linguistics and Linguistic Theory*, 5(2):201–239.

Kunter, G. (2011). *Compound Stress in English - The Phonetics and Phonology of Prosodic Prominence*, volume 539 of *Linguistische Arbeiten*. de Gruyter Mouton, Berlin.

Kunter, G. (2015). Coquery: A corpus query tool. `http://www.coquery.org`.

Kunter, G. and Plag, I. (2007). What is compound stress? In *Proceedings of the 16th International Congress of Phonetic Sciences*, pages 1005–1008, Saarbrücken.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*, 51:2018–2024.

Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22:1–38.

Liberman, M. Y. and Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2):249–336.

Mohanan, K. P. (1986). *The theory of lexical phonology*. Reidel, Dordrecht.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *The Journal of the Acoustical Society of America*, 54(5):1235–1247.

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. and Baayen, H. (2009). Suffix ordering and morphological processing. *Language*, 85(1):109–152.

Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, 118(4):2561–2569.

Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2006). Effects of word frequency on the acoustic durations of affixes. In *Proceedings of the 9th International Conference on Spoken Language (INTERSPEECH 2006)*, pages 953–956, Pittsburg, PA.

Pluymaekers, M., Ernestus, M., Baayen, R. H., and Booij, G. (2010). Morphological effects in fine phonetic detail: The case of Dutch -igheid. In Fougeron, C., editor, *Laboratory Phonology 10*, Phonology and phonetics, pages 511–531. Mouton de Gruyter, Berlin, New York.

Shatzman, K. B. and McQueen, J. M. (2006). Segment duration as a cue to word boundaries in spoken-word recognition. *Perception & Psychophysics*, 68(1):1–16.

Sluijter, A. M. and Heuven, V. J. v. (1996). Acoustic correlates of linguistic stress and accent in Dutch and American English. In *Proceedings of the 4th International Conference on Spoken Language Processing, October 3-6*, volume 3, pages 630–633, Philadelphia, PA.

Smith, R., Baker, R., and Hawkins, S. (2012). Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics*, 40(5):689 – 705.

Sproat, R. (1993). Looking into words. In Hargus, S. and Kaisse, E. M., editors, *Studies in lexical phonology*, pages 173–195. San Diego Academic Press, San Diego.

Sproat, R. and Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, 21:291–311.

Summers, Della, editor (2000). *Longman Advanced American Dictionary.* Longman, Harlow, Essex.

Trask, L. (1996). *A dictionary of phonetics and phonology.* Taylor and Francis Routledge, London.

Turk, A. E. and Shattuck-Hufnagel, S. (2000). Word-boundary-related duration patterns in English. *Journal of Phonetics*, 28(4):397–440.

Turk, A. E. and Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, 35(4):445–472.

Turk, A. E. and White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics*, 27(2):171–206.

White, L. (2014). Communicative function and prosodic form in speech timing. *Speech Communication*, 63–64:38 – 54.

Wightman, C. W., Shattuck Hufnagel, S., Ostendorf, M., and Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, 91(3):1707–1717.

Zirkel, L. (2010). Prefix combinations in English: structural and processing factors. *Morphology*, 20(1):239–266.