Do information status and sentence position affect the prominence patterns of English compounds?[1]

[1] Acknowledgements to be added.

ABSTRACT

English nominal compounds are usually pronounced in such a way that either the first or the second constituent has the highest perceptual prominence. While the factors and mechanisms that determine the prominence patterns across different compound types are by now rather well-understood, only little is known about within-type variation, that is, instances in which some tokens of the same compound type are produced with highest prominence on the first constituent, and others with highest prominence on the second one. Yet, this within-type variation is a fairly frequent phenomenon of English compounds.

The present study investigates two context-dependent variables that are potential determinants of within-type variation: sentence position (initial or final) and information status (new or given). Using the pitch information obtained from speakers in a production experiment as a cue of perceptual prominence, it is shown that neither information status nor sentence position affect the prominence pattern of compounds in a categorical way, but appear to affect individual compound types differently than others. This finding is in line with grammatical models which highlight the importance of the relations between individual types in the mental lexicon.

1 INTRODUCTION

Two distinct patterns of perceptual prominence can occur in English noun-noun compounds. In the most frequent pattern, the first (or left) constituent is perceived as more prominent than the second (or right) constituent. This pattern occurs in roughly 70 to 80 percent of noun-noun compounds (see Liberman & Sproat 1992, Plag et al. 2008), and has traditionally been considered the prototypical prominence pattern for English compounds (hence the term 'compound stress'). Examples are given in (1); the most prominent syllable is marked by an acute accent:

(1) ópera glasses cóffee maker wátch man

The alternative pattern, in which the right constituent is perceptually more prominent than the left, has often been considered as exceptional, or restricted to small, more or less well-defined subsets. For instance, compounds with an internal locative, temporal, or compositional semantic relation are claimed to be categorically right-prominent in Fudge (1984) or Liberman & Sproat (1992). (2) gives examples of compounds with right prominence that feature these semantic relations.

(2) Boston márathon morning páper silk tíe

Other accounts assume a structural difference between left-prominent and right-prominent compounds. For example, Payne & Huddleston (2002) or Giegerich (2004) argue that compounds in which the left constituent serves as an argument to the right constituent (e.g. *cóffee maker*) will be generally left-prominent. Conversely, right prominence is restricted to compounds in which the left constituents modify the head (e.g. *Boston márathon*). This claim has been refined in Giegerich (2009), who confines right prominence to compounds with a clear modifier-head relation, and in particular to

those in which the relation between modifier and head is ascriptive (e.g. *corn óil*) rather than associative (e.g. *éngine oil*).

Underlying these accounts is the assumption that the prominence patterns occurring in English compounds can be described in categorical terms once the relevant categories are identified. Most authors acknowledge that some compounds may not show the expected prominence pattern, but such exceptions are usually considered to be rare. It is occasionally noted that exceptions appear to be paradigmatic in that they share the same left or right constituent (see Schmerling 1971, Liberman & Sproat 1992, Giegerich 2004). For instance, compounds with *avenue* as the right constituent are generally right-prominent, but those with *street* as the right constituent are generally left-prominent, resulting in pairings such as *Mádison Street – Madison Ávenue* or *Fífth Street – Fifth Ávenue* with contrasting prominence patterns even though the structural and semantic relations are not different. Yet, the scope of such cases of prominence assignment by analogy has usually been considered as rather limited, accounting only for a constrained number of exceptional prominence patterns.

In recent years, there has been renewed interest in researching the mechanisms underlying the distribution of prominence patterns in English noun-noun compounds. Using large data sets obtained from diverse sources such as the CELEX lexical database (Baayen et al. 1995) analysed in Plag et al. (2007), the Boston University Radio Speech Corpus (Ostendorf et al. 1996) analysed in Plag et al. (2008) and Kunter (2011), or experimentally obtained data (Plag 2006, Bell 2008), it has been shown that the distribution of prominence patterns defies a description that relies on categorically determined, syntactic or semantic subsets. Instead, there is an increasing amount of

evidence that the structure of the mental lexicon and the existing forms stored in it play a central role. Lappe & Plag (2007), Arndt-Lappe (2011), and Arndt-Lappe & Bell (manuscript) argue for an exemplar-based account of prominence assignment in English compounds: speakers choose the prominence pattern for a given compound on the basis of the prominence patterns of other, similar compounds in the lexicon of the speaker. In particular, the computational models in Lappe & Plag (2007) and Arndt-Lappe (2011) suggest that compounds belonging to the same constituent family have a strong tendency to show the same prominence pattern. Thus, these models expand vastly on the earlier notion that analogical effects can account only for a small number of exceptional cases. Additionally, Plag & Kunter (2010) show that this effect of constituent family increases in strength with increasing constituent family size, and Bell & Plag (2012) present an experiment in which the informativeness of the two constituents, that is, the degree of predictability that one noun will co-occur with another noun as constituents in a compound, can be used to predict the distribution of left and right prominence: the more informative a constituent is in relation to the other, the more likely it is to attract prominence in the compound. In sum, these studies provide support for a model of prominence assignment that moves away from categorical descriptions toward a mechanism that is mostly driven by analogical effects in the mental lexicon, like in the very recent formalization in Arndt-Lappe & Bell (manuscript) that incorporates constituent family effects as well as semantic similarities.

Yet, there is one issue that further complicates the picture. It has been occasionally noted (e.g. in Bloomfield 1933, Bauer 1983, Ladd 1984, Bell & Plag 2012) that the same compound does not always receive the same prominence pattern when produced

by different speakers, and even a single speaker may produce the same compound with different prominence patterns. The small-scale study in Kunter (2011: chapter 8) suggests that variability in compound prominence patterns is a fairly wide-spread phenomenon. 13 out of the 32 compound types in the Boston corpus that were produced at least six times by the same speaker showed variable prominence patterns, as did 23 out of the 41 compound types that were produced by at least six different speakers. In other words, in about 40 percent of the investigated compound types, the prominence pattern varied within the same speaker, and in about 45 percent of the investigated compound types, the prominence pattern varied between different speakers. Even if the proportion of compound types with varying prominence reported in Bell (2013) is notably lower (about 20 percent), her production data also shows that prominence in compounds is subject to a large degree of variation. These findings raise the question of what motivates the variation. Crucially in this respect, Kunter (2011) found that a given compound type can demonstrate a notable amount of across-speaker variation in some sentences, but no or hardly any variation in other contexts. This observation clearly points toward compound-external factors that may facilitate variation in prominence patterns, but the scope of the study was too limited to isolate these potential factors.

The present paper reports on a production experiment that tests the influence of two potential compound-external factors: sentence position and information status. It is generally assumed that linguistic expressions that introduce new information to the discourse tend to be accented, while expressions that repeat old information already known to the interlocutors are more likely to be unaccented (Chafe 1974, Prince 1981, Hirschberg 2004, Steedman 2014). In his cross-linguistic comparison, Cruttenden

(2006) shows English to be a languages in which de-accentuation of old information is nearly obligatory, and Watson et al. (2008, 2010) find that predictable information is produced in an acoustically less prominent way than unpredictable information. Thus, differences in information status may be a possible explanation for the reported prominence variation. However, as noted already by Liberman & Sproat (1992: 134), the effect of information status, focus, and contrast as determinants of the prominence patterns in English compounds is still not well understood. For Ladd (1984), left prominence in compounds results from de-accentuation of the right constituent that occurs in cases where the right constituent expresses semantically predictable information. An extension of this account to differences in information status might account in part for the observed prominence variation: a compound that is referring to given information in a specific context may undergo similar de-accentuation as proposed by Ladd (1984) for compounds with a predictable right constituent, and hence, show a different prominence pattern in comparison to contexts in which the same compound expresses new information.

Regarding sentence position as a potential factor influencing within-type variation, the pilot study in Kunter (2011) found the tendency that compounds occurring in sentence-initial position were more constant in their prominence pattern than those occurring in other positions, but with the limited amount of available data, this tendency did not reach statistical significance. However, as old information tends to occur later in a sentence than new information (see, for instance, Ward et al. 2002), an effect of sentence position may additionally be confounded by an effect of information status, with unclear effects on the prominence patterns of compounds. The present production

experiment disentangles these two factors, and investigates in how far they affect the prominence pattern of English noun-noun compounds. As the next section shows, pitch is the primary acoustic correlate of these prominence patterns, so the analysis described below will compare the pitch contours occurring in environments controlled for information status and sentential position.

2 Prominence and pitch

In English, prominence relations in a linguistic construction larger than a single word are usually expressed by the distribution of intonational accents across the construction. In the autosegmental-metrical framework of intonation (for general overviews, see Ladd 1996, Gussenhoven 2004; see Pierrehumbert 1980, Beckman & Pierrehumbert 1986 for detailed descriptions of the English intonational system), a pitch accent is considered an tonal target that is aligned with a stressed syllable of a word. English tonal targets can be either high or low (usually transcribed as H* and L* in transcription systems like ToBI, Beckman et al. 2005), or combinations of a high and a low target (L*+H, L+H*, with the asterisk indicating the primary target). In general, a word that is accented is perceptually more prominent than an unaccented word. If there is more than one accent within a phrase, it is the last accent (the "nuclear accent") that is usually perceived with highest prominence (Beckman & Edwards 1994). The pitch accent type that is generally associated with discourse-new information is H* (Pierrehumbert & Hirschberg 1990), but L+H* seems also to be possible (Katz & Selkirk 2011, see also the detailed discussion in Steedman 2014).

With regard to English compounds, several acoustic studies (Farnetani et al. 1988, Kunter & Plag 2007, Kunter 2011 for noun-noun compounds, Kösling et al. 2014 for tri-

constituent noun compounds, Morrill 2012 for adjective-noun compounds) have provided empirical evidence that the two prominence patterns exemplified in (1) and (2) are realized by two different accentuation patterns: the left-prominent compounds in (1) have a single pitch accent on the first constituent, while in the right-prominent compounds in (2), both constituents are accented.

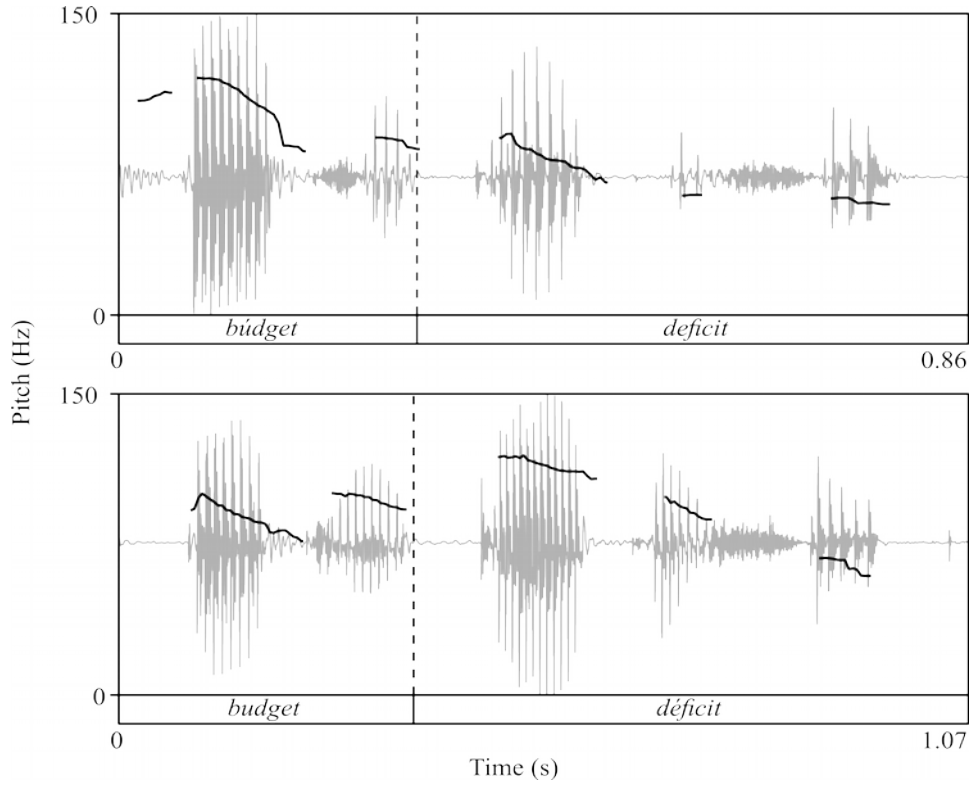

Figure 1

Pitch track for budget deficit *with left prominence (top) and right prominence (bottom)*

Figure 1 illustrates the pitch contours of left- and right-prominent compounds. Superimposed over the respective waveforms, the panels show pitch tracks for two tokens of *budget deficit* spoken by the same male speaker from the Boston corpus (Ostendorf et al. 1996, files M4BS09P4 and M4BS16P2). The dashed vertical lines indicate the boundaries between the left and right constituents. The token in the upper

panel is left-prominent: there is a pitch accent on the stressed syllable of *budget,* with a clear peak in the pitch contour, and a steady decrease after this peak towards the end of the compound, suggesting that the right constituent is unaccented. In contrast to this, the right-prominent token in the lower panel has a markedly different pitch contour. While the relatively high pitch contour in *budget* implies that this constituent is again accented, there is a second pitch accent on the right constituent *deficit*, as the pitch contour has another peak on the stressed syllable of *deficit*, and falls only after that. Apparently, these two tokens are an instance of within-speaker variability: the same compound type is produced by the same speaker once with left prominence and once with right prominence.

Can this variation be explained by sentential position, or by information status? The contexts of the two tokens are given in (3):

(3) (a) The Baileys' daughter is one of four hundred multi-handicapped patients who will be shut out of the so-called turning twenty-two program because of the **búdget deficit**.

(b) Left unchecked, the cost of nursing home care will soon consume much of state revenues, leaving little available for other human service programs. An ominous portend is found in this year's **budget déficit**.

While both tokens occur in sentence-final position, there seems to be a difference in the information status. In (3a), *budget deficit* is the head of a definite noun phrase, thus apparently expressing discourse-old information. This is a context in in which de-accentuation is assumed to occur. In contrast, *budget deficit* is discourse-new in (3b), so no de-accentuation is expected. Since there are two pitch accents present in the

discourse-new pitch contour at the bottom of Figure 1, but only one accent in the discourse-old contour at the top, the two tokens may imply that de-accentuation indeed occurs within English compounds that express old information. Thus, if a change in information status affects the prominence pattern in compounds as a general rule, a similar difference of pitch contours like the one illustrated in Figure 1 should also be visible if larger groups of compounds expressing old and new information are compared.

3 METHOD

The task of the present production experiment was to read aloud sentences containing noun-noun compounds in different sentential positions and with different information status. The utterances were recorded, and the pitch contours across the different conditions were analysed using a statistical model that generalized the pitch information across speakers. If information status and/or sentential position have a systematic influence on the prominence pattern in compounds, there should be a systematic difference between the pitch contours in some conditions. The following subsections explain in more detail the design and methodology of the experiment.

3.1 Compound types

In the study investigating variable prominence patterns in compounds presented in Kunter (2011), some compounds displayed a considerable amount of variation. For instance, only about half of the tokens of *state troopers* or *access barriers* were produced with left prominence in the Boston corpus. However, other types showed an invariant prominence pattern across all tokens (for example, *home phones* was always right-prominent, or *retirement age* was always left-prominent). In order to ensure that

the compounds used as experimental items potentially allowed for prominence variability, a list was compiled containing those compounds that have been found to be variable with regard to their prominence pattern in previous studies (Bauer 1983, Bell 2008, and Kunter 2011).

In some of these compounds, the lexical category of the left constituent was ambiguous. For example, the left constituents of *amber signal* and *back step* may be interpreted as nouns, but have also a possible adjectival reading (*amber*) or adjectival and adverbial readings (*back*). However, if the lexical category of the left constituent is ambiguous, the morpho-syntactic status of the whole construction can vary between speakers: *amber signal* may be interpreted as a noun-noun compound by some speakers, but as a simplex noun modified by an adjective by others. As the phrasal head is generally held to be produced with higher prominence than the modifier, this ambiguity could be a source of variation that is unrelated to information status and sentential position. Accordingly, constructions with ambiguous lexical categories were excluded.

Additionally, compounds in which the left element functioned as a syntactic argument to the head such as *engine overhaul* or *food processor* were excluded from the list, as Kunter (2011) found that this class of compounds may be more immune to prominence variation than other compounds. Also, constructions that seemed to be more frequent in British English, but rare or non-existing in American Englishes were not taken into consideration. Examples of the latter type of compounds included *caster sugar*, *council housing*, and *patch lead*. Finally, highly infrequent compounds such as *jean jacket* with only three tokens in COCA (Davies 2008-) were also removed from the list in order to reduce the rate of hesitation pauses, false starts and similar events during

the production of the sentences.

From the remaining nearly 50 types, a final list of 20 target compounds was chosen as experimental items. Preference was given to those types which could be easily embedded into naturally-sounding environments. The types are listed in (4):

(4)

access barriers	budget deficits	church service
club members	computer genius	dinner potato
farm cats	fish steaks	floor materials
hall lights	killer shark	league team
life sentences	meteor crater	pencil figure
pincer attack	police helmets	state officials
student card	task force	

Some constituents in the target compounds have primary stress not on the first syllable (e.g. *budget* or *helmets*) but on the second (e.g. *computer* or *attack*). As pitch accents are associated with primary-stressed syllables, the shape of the pitch contour in these compounds will be somewhat different. As described below, the statistical models contained additional control variables to incorporate these differences.

3.2 Environments

In order to allow a principled investigation of the effect of sentence position and information status, each of the 20 target compound types was embedded in four different contextual environments in which these variables were varied in a systematic fashion. Each environment consisted of two sentences. The first sentence provided a general context, while the second sentence acted as the carrier sentence for the compound itself. In all carrier sentences, the compound was embedded into a noun

phrase that was headed by the compound. The noun phrases were controlled for number and definiteness: the noun phrases headed by the compound were always indefinite, and a given compound occurred either always as a singular or always as a plural noun. If the compound was singular, the noun phrase was determined by an indefinite article. For any target compound, the noun phrase was held constant across all four environments.

The position of the noun phrase could be either initial or final. Noun phrases with compounds in initial position functioned as the subject of the carrier sentences. Noun phrases with compounds in final position were usually the head of an adverbial prepositional phrase occurring always as the last element of the sentence.

The different environments for information status were constructed using the definition of discourse-old and discourse-new information from Ward et al. (2002), who distinguish

> between information that is old or new with respect to the discourse, and information that is old or new with respect to the addressee. [...] Discourse-old status applies not only to elements that have themselves been explicitly evoked in the prior discourse, but also to those that stand in some salient and relevant relationship to elements that have been evoked (Ward et al. 2002: 1368)[.]

Under this definition, discourse-old status was ensured by introducing concepts in the context sentence that stood in a part-whole relation to the target compound, or which were otherwise semantically linked to it, sometimes by mentioning explicitly one or two of the compound constituents in the context sentence. Conversely, in discourse-new environments, any semantic link between the concepts in the carrier sentence and the target compound was avoided. In order to exclude repetition as an additional source of

variability, the target compound never occurred in full form in any context sentences, not even in discourse-given contexts.

Combining the two sentence positions INITIAL and FINAL with the two types of information status OLD and NEW yielded four different conditions in which a compound occurred. The examples in (5) and (6) feature the compound *state officials* (in bold typeface) in sentence-initial and sentence-final position, respectively. The compounds under (a) represent new information, and the compounds under (b) represent old information.

(5) (a) INITIAL/NEW

There was a plane crash near the airport. **State officials** say that everybody survived.

(b) INITIAL/OLD

The reporters were waiting for a statement from the government. **State officials** were present in the newsroom.

(6) (a) FINAL/NEW

There was a huge crowd in front of the library. Many reporters were waiting for the arrival of **state officials**.

(b) FINAL/OLD

Representatives from several government departments had sent out letters to the school. Teachers and students were invited to meet **state officials**.

The environments under (a) are designated as NEW because the context sentence does not evoke in any obvious sense the concept of *state officials*, which is hence considered

to express discourse-new information. This is not the case in the environments under (b), designated as OLD. Here, the concept *government,* which stands in a metonymic relation to *state officials,* is explicitly mentioned. Under the definition provided by Ward et al. (2002), this is sufficient to consider these occurrences of the compound as representing discourse-old information. Similar environments were constructed for all 20 target compounds, yielding 80 different environments in total (see Appendix).

3.3 Participants and procedure

The 80 test sentences were presented to 23 participants that were recruited from the student pool at the University of Alberta, Edmonton. All of them were native speakers of Western Canadian English; the majority was born in Alberta. 20 participants were female, three participants were male. None of the participants reported any hearing impediments. During participation, they were naive with respect to the aim of the investigation, but the experiment's purpose was disclosed to them after participation. Pparticipants were seated in front of a computer screen in a sound-attenuated booth. They were equipped with a head-worn microphone that was located at approximately three centimetres distance at the left of the participant's mouth. The recordings were digitized at 44.1 kHz using an Alesis digital recorder.

The participants were instructed on the computer screen to read aloud several short news items in a clear and expressive voice that they considered to be suitable for radio broadcasting. Each contextual environment was presented in two stages. During the preparation stage, participants were invited to quietly read the text that was displayed in order to become familiar with it. After pressing a key on the keyboard, the text was erased from the screen, and a five-second countdown timer lead to the beginning of the

recording stage. After the participants had read out loud the text, they proceeded to the preparation stage of the next item by another key press.

Each experimental run started with two training items (which did not contain a target compound) in order to familiarize the participants with the experimental procedure, and to adjust the sensitivity of the microphone to match the participant's speaking volume. After training, the 80 environments containing the target compounds were presented in a pseudo-randomized order, with the restriction that after the presentation of a given compound type, seven or more environments containing different compound types had to be presented before another environment containing the given compound occurred. This constraint was introduced to minimize persistence effects that could potentially arise if tokens of the same type were repeated within an overly short time span.

3.4 Pitch measurements

Within each recorded environment, the beginning and end of the target compound as well as the boundary between the two constituents were manually marked using the phonetic software tool *Praat* (Boersma & Weenink 2013). Conventional segmentation criteria were applied to identify the segment boundaries (see, for instance, Ladefoged 2003 or Turk et al. 2006 for overviews). 13 recordings were excluded from the analysis because the target compound was either mispronounced or contained a notable pause between the two constituents.

Pitch measurements from the remaining 1,827 tokens were obtained similarly as in Kösling et al. (2014) using an unsupervised *Praat* script. For any given environment, the script first extracted a pitch object with pitch floor and ceiling set to 100 Hz and

500 Hz for female speakers, and to 75 Hz and 300 Hz for male speakers. If the compound occurred in final position, a pitch object was extracted for the whole environment; if it occurred in initial position, only the sentence containing the target was considered in the analysis.

Any gap in the resulting pitch contour (caused, for instance, by unvoiced speech segments) was filled by linear interpolation between the closest available pitch measurements surrounding the gap. The rationale behind this step is the assumption that the intonational contour of an utterance is primarily determined by the distribution of accents within it (see section 2). Accordingly, any interruption of the contour introduced by unvoiced segments should be considered a consequence of the phonetic nature of the segments within the compounds under investigation, but these interruptions should not be considered relevant for the resulting prominence patterns. This assumption is compatible with the view expressed for instance in Hermes (2006), who argues that pitch contours are essentially perceived as continuous, and that interruptions during unvoiced segments is not noticed by listeners as affecting the perceived shape of the pitch contour. Interpolating gaps has two advantages. First, it reduces the number of missing values in cases where the pitch tracker was not able to determine a periodic signal. Second, interpolating unvoiced portions reduces compound-specific characteristics of the pitch contour, thus allowing a better comparison between items.

Next, the interpolated pitch object was smoothed in *Praat* with a bandwidth of 10 Hz, giving a more natural shape to the transitions between regular and interpolated portions. The smoother also removed localized misreadings of the pitch tracking algorithm, and reduced the effect of microprosodic changes in the pitch contour, such as

the decrease of F0 in the vicinity of unvoiced obstruents (see, for example, Kingston & Diehl 1994).

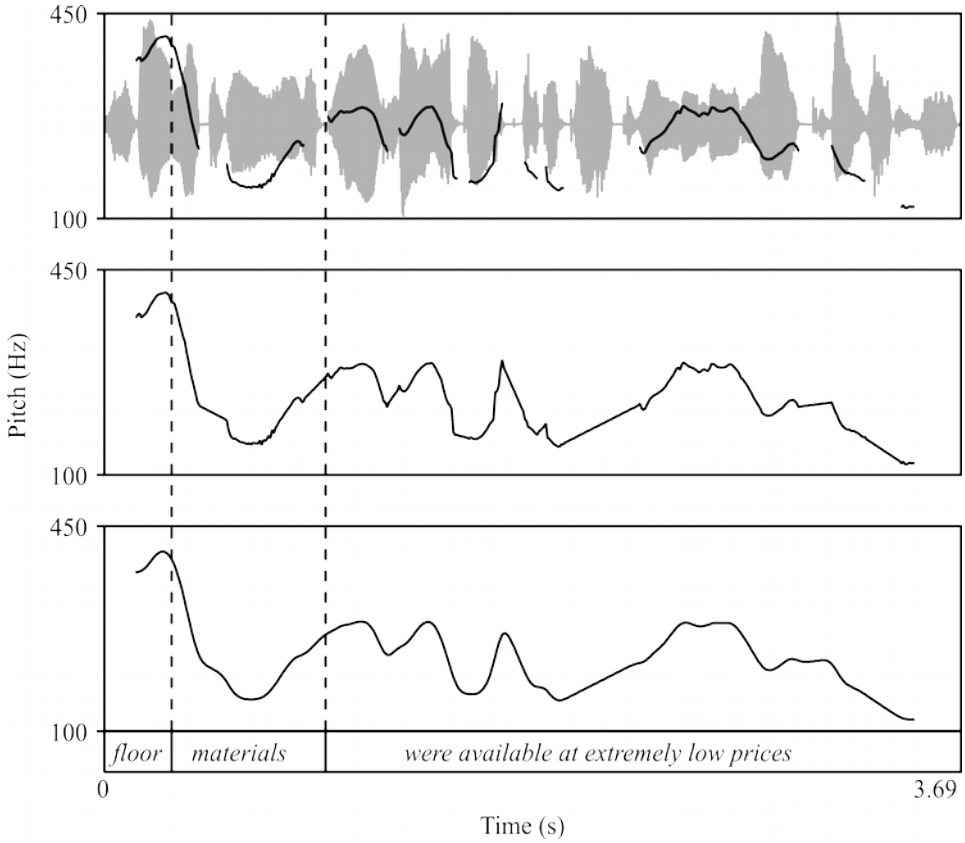

Figure 2

Pitch contour for Floor materials were available at extremely low prices *(top: unprocessed, middle: interpolated, bottom: smoothed)*

Figure 2 illustrates the effect of this procedure on the pitch contour for the target sentence *Floor materials were available at extremely low prices* produced by a female speaker (condition INITIAL/NEW). The upper panel of the figure shows the original, unmodified contour, with the waveform underneath in grey. There are several portions in the pitch contour for which no pitch could be found by the tracker. Within the portion associated with the target, these parts correspond to unvoiced segments: the initial [f] in *floor* as well as the intervocalic [t^h] and the word-final [s] in *materials*, which is realized

as an unvoiced fricative in this recording. The interpolated pitch contour in the middle panel removes the gap in the second element of the target compound. However, interpolation does not provide pitch information for the onset of *floor*, which is the expected outcome, as there is no preceding pitch information that could be used to estimate the pitch contour within the unvoiced fricative. Yet, the interpolation introduces very unnatural sharp edges to the contour at the boundaries of unvoiced segments such as [t^h], or at [gst] in *extremely* in the second half of the sentence. The smoothed contour in the bottom panel removes these edges, as well as any minor changes in the pitch contour contributing to the overall shape.

After processing the contour, the interval associated with the compound was divided into 100 time steps of equal duration. Then, the average pitch for each time step was derived from the smoothed contour. This measurement was stored together with the number of the time step at which it was taken (*t*, ranging from 0 to 99). The time step *t* was used in the analysis below to model the change of pitch over the course of the compound. By obtaining the same number of measurements from all compounds, all durational differences between compound tokens are normalized (for similar uses of time-normalizating procedures, see, for example, Liu & Xu 2007). The theoretical motivation behind this decision is that the pitch contour in compounds is principally determined by the placement of pitch accents: the primary-stressed syllable of each constituent may act as an anchor point for a pitch accent, while any other syllable will contribute only marginally to the pitch contour. Using time-normalized pitch measurements may also reduce compression effects that are introduced by carrier words with different lengths (see, for example, Grabe et al. 2000).

For the majority of compounds in the experiment, the automatic procedure just outlined produced reliable pitch measurements. With 1,827 valid tokens, the total number of measurement points was 182,700. However, there were measurement points at which the pitch tracker failed to detect a periodic signal throughout the whole target compound even after interpolation. There are two primary sources for this failure: unvoiced segments at the left or right edge of the compound, and the use of creaky voice phonation. As can be seen at the left edge of Figure 2, no pitch information could be obtained for word-initial voiceless consonants in sentence-initial position (unless, of course, they are preceded by an article, see section 3.2), and likewise, for word-final voiceless consonants in sentence-final positions. In word-initial voiceless consonants in sentence-final position and in word-final voiceless consonants in sentence-initial position, measurements were obtained from the interpolated pitch contour.

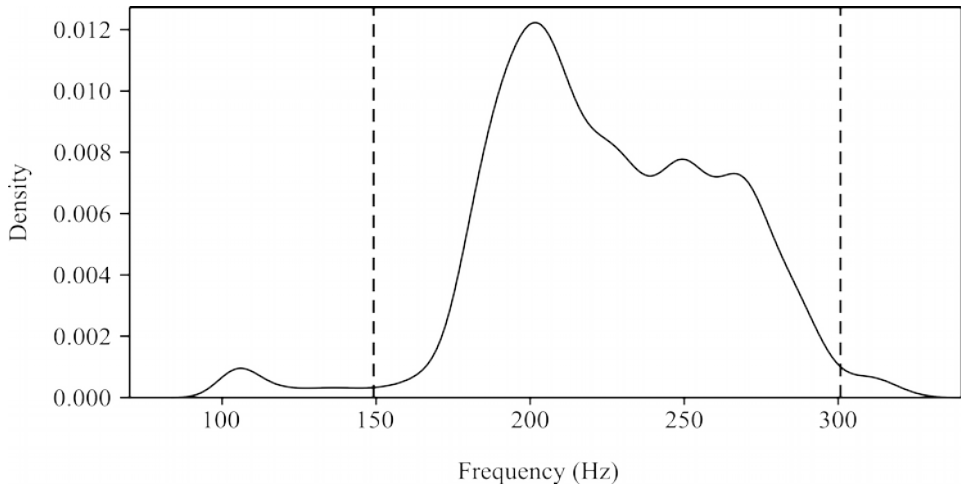

Figure 3

Density function of pitch measurements for one female speaker (2 SD units above and below mean are indicated by dashed lines).

More problematic was obtaining pitch measurements in constituents produced with

creaky voice. In these cases, the pitch tracker either returned no measurement at all, or, if it managed to detect the very low vibration rate of this phonation mode, the obtained pitch was usually well outside of the normal pitch range for the respective speaker. A look at the distributions of pitch measurements separately for each speaker suggests that such measurements should not be analysed together with data obtained from normal phonation. For example, Figure 3 displays an estimation of the density function for one of the female participants. The bulk of pitch measurements are distributed between 150 and 304 Hz, as indicated by the dashed lines at two standard deviation units above and below the mean pitch for this speaker (M = 227.1 Hz, SD = 38.5 Hz). There is a small number of pitch measurements above 304 Hz, but more notably, there is an isolated cluster of very low values located around 105 Hz, which results in a secondary peak in the density estimate at this frequency. Nearly all of these low values were measured from compounds that were produced with creaky voice. As most statistical models are not well-suited to fit data that is so obviously discontinuous, including these very low values in the analysis might stress the models to the point of producing invalid results. Consequently, the distribution of pitch measurements was visually inspected for each speaker. If there was a secondary peak at a low frequency similar to that in Figure 3, all pitch measurements below the lower boundary of the main distribution were excluded (in total, 2,588 measurements from 10 female and one male speaker).

Next, all 80 environments were investigated for time steps at which only very few pitch measurements were available. For example, due to the voiceless coda [ts] of *hall lights*, no speaker produced glottal vibration in FINAL/NEW condition from time step 77 onwards, and only three speakers produced at least some glottal vibration between time

steps 65 to 76 in this type. In other words, all available pitch information for the last third of the contour of this type is based on data from only a small number of speakers. In order to reduce an overly large influence of these speakers on the overall contour for *hall lights*, their data for time steps 65 and onward were also discarded in this example. For other types, such as *league team* in FINAL/NEW condition, a sufficiently large number of pitch values was available for all 100 time steps, so no additional data pruning was necessary.

All in all, 140,905 pitch measurements were available in the final data set, distributed across 1,827 compounds (on average, 77.1 measurements per token). The pitch measurements were converted from Hertz to semitones relative to the lowest pitch measurement in the data set, thus eliminating the non-linear perceptual nature of the Hertz scale (see Nolan 2003 for details). With the lowest pitch measurement at 75.7 Hz, a transformed pitch measurement of 0.0 semitones corresponds to 75.7 Hz, 12.0 semitones to 151.4 Hz, 24.0 semitones to 302.8 Hz, and so on.

3.5 Analysis

The transformed pitch measurements were analysed in a similar way as in Kösling et al. (2014), who used General Additive Models (GAMs, Wood 2006) to compare the pitch contour in four different types of tripartite compounds. GAMs are extensions of linear regression models in that they allow, in the present case, a separate approximation of the pitch contour for each environment in which the compounds occurred. The results in Kösling et al. (2014) have shown that this statistical method can be superior in comparison to studies which analyse the prominence patterns in compounds on the basis of averaged pitch data (as in Štekauer et al. 2007 or Kösling and Plag 2009), or which

use localized pitch readings such as the loci of F0 maxima and minima as approximations of the overall pitch contour (e.g. in Katz & Selkirk 2011).

The dependent variable of the general additive model was the transformed pitch measurements. In order to obtain approximations of the shape of the pitch contour in the four different condition environments across all compound types, separate smooths of time step *t* were fitted for the four conditions INITIAL/NEW, INITIAL/OLD, FINAL/NEW, and FINAL/OLD. By comparing these smooths, the hypothesis can be tested that sentence position and information status affect the prominence pattern in compounds: if the prominence pattern of compounds is affected by information status, that is, if a compound expressing old information is more prone to de-accentuation than a compound conveying new information, there should be systematic differences between the smooths for these two environments. Likewise, if compound prominence patterns are influenced by their position within an utterance, a difference between the smooth for compounds in sentence-initial and sentence-final position is to be expected.

In addition to the smooths central for testing these hypotheses, the GAM also included two further smooths that estimate the adjustments necessary if the primary stress did not fall on the first syllable in either the left or the right constituent. The linear regression part of the GAM contained the factor SEX, which incorporated the different average pitch frequency of male and female speakers, as well as an interaction between the factors POSITION and INFORMATION. The model also included random intercepts for SPEAKER and TYPE, which controlled for variation of the average pitch frequencies between different participants and different compounds. For example, high vowels are generally produced with higher intrinsic pitch than low vowels (Whalen & Levitt 1995),

so it might be the case that the pitch contour for a compound such as *computer genius* is, on average, higher than for *access barriers* just because of the presence of a larger number of high vowels in the former. The random effect for TYPE attempts to reduce such variation.

In order to obtain the final, fully specified GAM, an initial baseline model was fitted that estimated only an intercept for pitch, that is, the grand average of pitch across all speakers, compounds, conditions, and time steps. Following the procedure in Kösling et al. (2014), all predictors described in this section were added in a sequential model comparison. Each addition significantly improved the fit, and lead to a notable reduction of the AIC.[2] Crucially, the addition of smooths for the experimental condition greatly reduces the AIC by 31230.0 to 650487.7, indicating that the fit of the model improves strongly by this addition.

4 RESULTS

Before reporting the results of the final model, this section investigates whether the experimental design was generally successful in eliciting variable prominence patterns from the participants on the basis of a subset of the produced compounds. In order to do this, the 92 tokens of the compound *police helmets* were acoustically inspected by the author, and classified as either left-prominent or right-prominent.

[2] The Akaike Information Criterion (Akaike 1974) is a statistic that can be used to assess the goodness of fit of two models: the model with the *better* fit is the one that has a *lower* AIC, and which hence has a *higher* AIC reduction.

Position	Information status	Speaker ID with right-prominent token															#	percent
		1	5	6	7	9	11	12	15	16	18	19	22	26	27	29		
INITIAL	NEW						×			×	×		×			×	5	21.7%
	OLD		×			×	×						×			×	5	21.7%
FINAL	NEW			×	×		×	×				×			×		6	26.1%
	OLD	×			×			×	×					×			5	21.7%
Total		1	1	1	2	1	3	2	1	1	1	1	2	1	1	2	21	22.8%

Table 1

Occurrences of right-prominent compound tokens for police helmets

Table 1 shows that right prominence is not infrequent for this compound. As indicated by the marks in the speaker ID columns, 15 of the 23 participants produced at least one out of the four tokens with right prominence. Some participants (IDs 7, 12, 22, 29) produced two or even three right-prominent tokens (ID 11). The remaining 8 participants produced only left-prominent tokens. All in all, 21 tokens (22.8 percent) have a second pitch accent on *helmets*; these tokens were perceived as right-prominent, which corresponds very closely to the 25 percent right-prominent tokens reported for this type in Bell (2013). As the percentages in the last column show, right-prominent tokens are distributed fairly evenly across the four conditions. Apparently, while the experiment did elicit tokens with varying prominence patterns, no condition seems to favour or disfavour variation, at least for the compound *police helmets*. But perhaps the statistical model can provide evidence for a general effect of condition on the distribution of prominence patterns in the experimental tokens?

Table 2 summarizes the estimated linear coefficients and the estimated significance of the smooth terms in the final GAM.

Estimated linear coefficients	Estimate	Std. Error	*t*	*p*
(Intercept)	20.493	0.315	65.05	< 0.001
SEX=MALE	−9.753	0.827	−11.80	< 0.001
POSITION=FINAL	−4.337	0.022	−198.00	< 0.001
INFORMATION=NEW	−0.094	0.018	−5.12	< 0.001
POSITION=FINAL, INFORMATION=NEW	0.791	0.030	25.98	< 0.001

Estimated smooth term	est. Df	Ref. Df	*F*	*p*
s(SPEAKER)	20.988	21.000	2843.34	< 0.001
s(TYPE)	16.894	17.000	254.88	< 0.001
s(t, CONDITION=INITIAL/NEW)	8.379	8.881	1984.12	< 0.001
s(t, CONDITION=INITIAL/OLD)	8.257	8.835	1815.94	< 0.001
s(t, CONDITION=FINAL/NEW)	7.864	8.609	610.13	< 0.001
s(*t*, CONDITION=FINAL/OLD)	7.616	8.449	219.75	< 0.001
s(*t*, LEFTPRIMARYSTRESS=SECOND)	9.762	9.978	71.27	< 0.001
s(*t*, RIGHTPRIMARYSTRESS=SECOND)	8.240	9.231	20.70	< 0.001

Table 2

General Additive Model fitting pitch measurements in four different environments

All terms are significant, and the model accounts for 75 percent of the observed variation of pitch values ($R^2_{adj} = 0.75$, AIC = 650487.7). The intercept of the linear part at 20.493 expresses the average pitch (in semitones) for compounds that occurred sentence-initially and that expressed old information, spoken by female speakers. The negative coefficient for SEX implies that on average, male speakers have a pitch that is 9.783 semitones lower than that of female speakers. There is a significant interaction between POSITION and INFORMATION, as well as significant main effects of POSITION and INFORMATION. Apparently, the locations of the pitch contours are affected by this interaction. The nature of this interaction becomes visible if the four CONDITION smooths are also taken into consideration, as shown in figure 4.

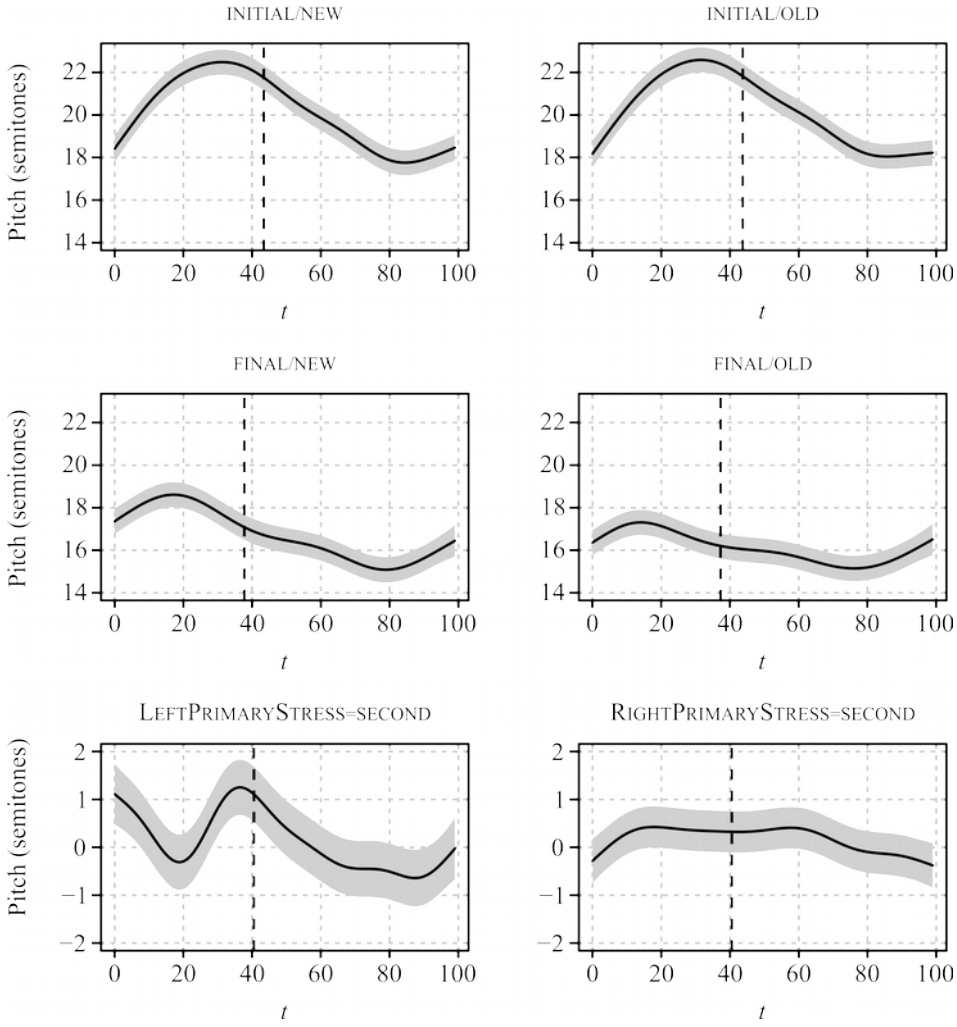

Figure 4

Estimated smooths of pitch by time step t *for environments and locations of primary stress.*

In each panel of figure 4, the curved line corresponds to the estimated change of pitch over *t* for the different predictors. The shaded areas indicate the 95 percent confidence intervals, and the dashed vertical lines indicate the average location of the constituent boundary. All pitch ranges are displayed for female speakers; for male speakers, the smooths are shifted by the liner coefficient for SEX=MALE, that is, by -9.753 semitones.

Row 1 and 2 show estimated smooths for compounds in INITIAL and FINAL position expressing NEW (left) and OLD (right) information. The last row shows corrective smooths for those compounds in which the second syllable of the left constituent (LEFTPRIMARYSTRESS=SECOND) or the right constituent (RIGHTPRIMARYSTRESS=SECOND) receives primary stress; as the two smooths are only included as adjustments for different locations of primary stress, they are not discussed here.

In all four environments, the left constituent is associated with a clear peak in the pitch contour. There is no corresponding peak in the right constituent in any of the four conditions. Under the assumption that the different prominence patterns in compounds are realized by the presence or absence of pitch accents, the model suggests that on average, compounds are predominantly left-prominent in all four conditions: there is always a pitch accent on the left constituent, but usually no pitch accent on the right constituent.

Information status does not appear to have a great influence on the shape of the pitch contour. A comparison of the two smooths for INITIAL position in the first row suggests that there is hardly any notable difference between the two contours. Apart from a small rise at the right end of NEW compounds, the two smooths are nearly identical, and they also cover the same pitch range. This similarity is confirmed by a GAM in which the factor levels INITIAL/NEW and INITIAL/OLD are conflated: the AIC increases only by 64.6 from 650487.7 to 650552.3 using the simplified factor. This very small change of the AIC suggests that the distinction between INITIAL/NEW and INITIAL/OLD is only negligibly helpful.

For FINAL position, the difference between NEW and OLD is somewhat more pronounced, but still minor. Here, the peak in the left constituent is higher by about 1 semitone if the compound expresses new information. Collapsing FINAL/NEW and FINAL/OLD increases the AIC by 487.3 to 650975.0, which implies that the distinction between new and old information is more obvious in sentence-final than in sentence-initial position. A model that estimates only smooths for initial and final position ignoring the information status increases the AIC by 551.6 to 651039.3.

In comparison to the very minor effect of information status, the overall effect of position is much more significant. Here, a model that ignores the distinction between INITIAL and FINAL leads to an AIC increase by 6124.6 to 656612.3, which is much larger than the increase introduced by the model that ignored the information status. In other words, the pitch contour is rather different between compounds in sentence-initial and sentence-final position, but not very different between compounds expressing new and old information.

When comparing INITIAL and FINAL (rows 1 and 2 in figure 4), two differences immediately become visible. The first difference is that the pitch contours are situated at much higher frequencies in INITIAL than in FINAL position, as indicated by the significant linear predictor POSITION. This is in line with the well-documented tendency of speakers to start an utterance with a relatively high pitch that steadily decreases toward the end of the utterance (see, for instance, Strik & Boves 1995 or Gussenhoven 2004: chapter 6 for more detailed accounts on pitch downtrend). The second difference lies in the different pitch ranges for INITIAL and FINAL compounds. In both INITIAL conditions, the pitch contour covers a range from approximately 22 semitones in the

high peak in the left constituent to 18 semitones in the low valley in the right constituent, spanning an interval of 4 semitones. The span is roughly 3 semitones (ranging from 18 to 15 semitones) for FINAL/NEW and 2 semitones (ranging from 17 to 15 semitones) for FINAL/OLD. In other words, the pitch range is smaller at the end of sentences, less so if the compound expresses new information, more so if it expresses old information. As the shape of the pitch contour suggests, the different pitch range in final position is mostly restricted to the first constituent; from $t = 60$ onwards, there is virtually no difference between the two shapes.

5 DISCUSSION AND CONCLUSION

The aim of the current research was to determine whether the prominence pattern in compounds varies systematically depending on the position of the compound within a sentence and on the information status of it. The pitch contours estimated by the GAM provide no evidence for such a systematic variation. In sentence-initial position, the shapes of pitch contours for new and old information are virtually identical. Apparently, pitch contours in this position are not shaped by information status, but rather by other factors that override any effect that discourse-old information might have on the intonation of the compounds: in general, compounds occurring in subject position have the same accentuation pattern irrespectively of whether they express new or old information. Judging from the shape of the pitch contour in the first row of figure 4, the accentuation pattern in sentence-initial compounds typically and consistently appears as what Gussenhoven (2004: 297) calls a ‘fall-rise’: after a high tonal target on the accented syllable (here: the primary-stressed syllable of the left constituent of the compound), the pitch contour falls steeply during the unaccented remaining syllables of

the compound, until it rises again for the trailing boundary tone at the right edge of the compound. This pitch contour (which corresponds to H* L-H% in ToBI notation) is also found in the pitch track for the sentence-initial compound *floor materials* in figure 2 above. In line with Brazil (1997: 68), Gussenhoven associates the fall-rise with given information, and expects a fall without a subsequent rise (H* L-L% in ToBI) for new information (Gussenhoven 2004: 297). The present experiment provides only very weak support for this proposed distinction. The smooth for INITIAL/NEW in figure 4 has a slightly more notable rise at the right edge of the compound than the smooth for INITIAL/OLD, but this difference reaches only marginal significance.

The pitch contour of compounds in sentence-final position is generally more compressed than that occurring in sentence-initial position. This positional difference between contours is to be expected: the peak frequency of pitch accents usually decreases over the course of the utterance; the high tonal target of a pitch accent early in a sentence has a higher frequency than a pitch accent late in the sentence (see Gussenhoven 2004: chapter 6). In addition to this largely physiologically determined downtrend, speakers may also use a compressed pitch range during the latter part of a sentence to signal to the listener that they are approximating the end of a turn (‘final lowering’, see Hirschberg 2002: 37). Unlike sentence-initial compounds, sentence-final compounds also show a difference that is conditioned by the information status: the pitch contour of OLD compounds is more compressed than that of NEW compounds. Apparently, speakers are more careful to retain the shape of the pitch contour in compounds in sentence-final position if the information is new, whereas the overall pitch movement is only small in this position if the information is already discourse-old.

This agrees with Liu & Xu (2007), who found that the pitch range is more compressed in utterances that occurred without focus, and for which an implicit informational givenness may be assumed.

Yet, crucial for the present study, the observed differences between INITIAL and FINAL position, and between FINAL/NEW and FINAL/OLD all indicate only scaling differences between the pitch contours, but there is no indication that the four conditions show a different distribution of tonal targets: The overall shape of the pitch contours is very similar in all conditions, which means that the speakers did neither consistently place a second pitch accent on the right constituent nor consistently remove the pitch accent on the first constituent in any of the four environments. This result of the statistical model corresponds to the distribution of left and right prominence in the tokens of *police helmets* given in table 1, where it was noted that all four conditions showed a comparable number of right-prominent tokens.

So, do these findings imply that variance in the prominence patterns of English nominal compounds is not determined at all by position or information status? The statistical model presented above pooled all compound types in one analysis, which assumed a uniform behaviour for all of them. Yet, this assumption might not be justified: it has been frequently reported in the literature (e.g. in Bauer 1983, Kunter 2011, Bell 2013, Arndt-Lappe & Bell, manuscript) that the degree of prominence variation in compounds can vary between types, and this may also be the case for the items used in the present experiment. Indeed, a subsequent analysis in which separate GAMs were fitted separately for each compound type reveals some notable by-type deviations from the pitch patterns that were displayed in figure 4. Each by-type GAM

included the same predictors as before, except the random effect for ITEM and the adjustment smooths for the position of primary stress within the constituents, as these predictors are invariant in a by-type analysis.

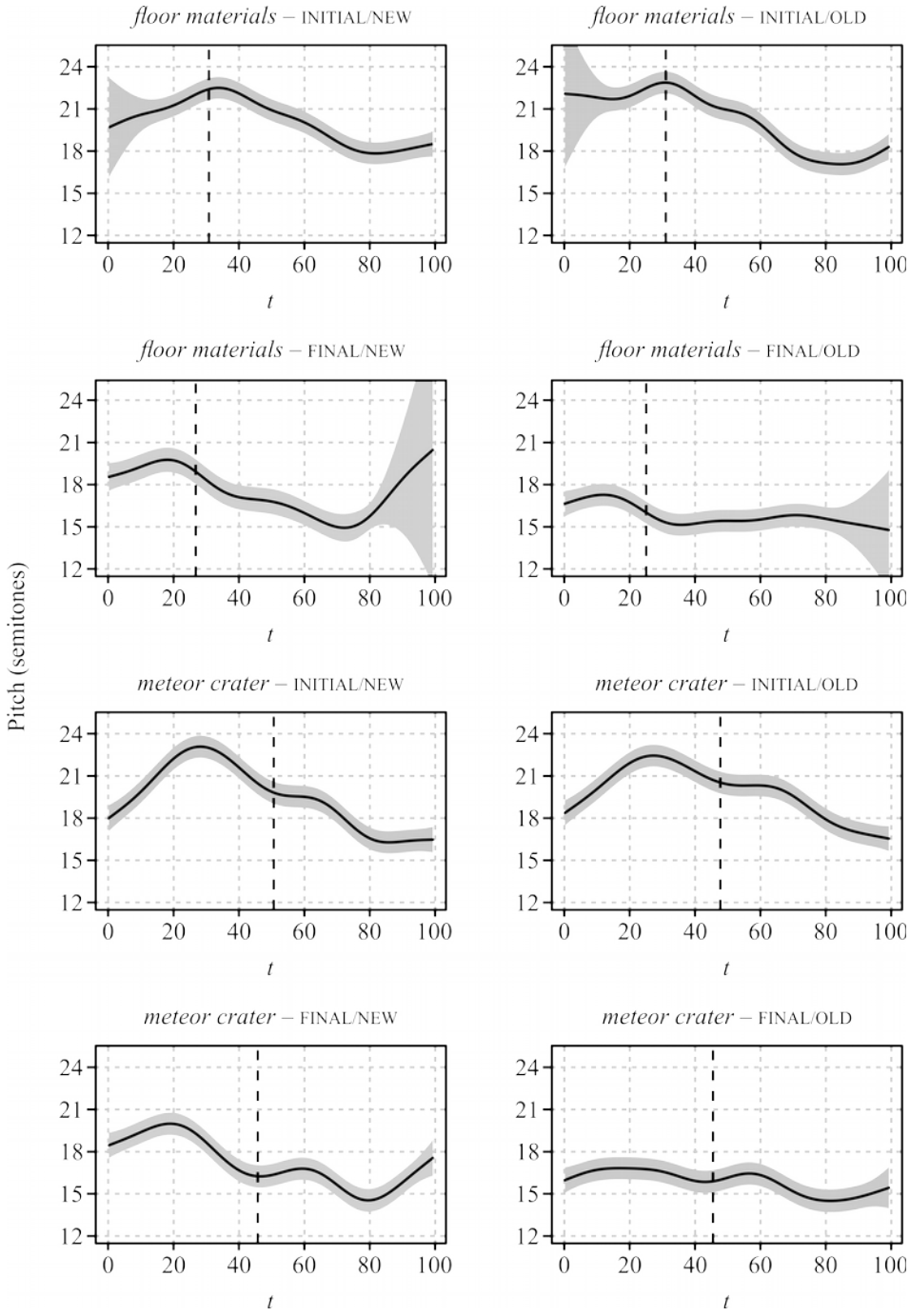

Figure 5

Estimated smooths of pitch by time step t *for compounds* floor materials *(top) and* meteor crater *(bottom).*

Figure 5 shows the estimated smooths of pitch by CONDITION from two of these by-

item GAMs. The two first rows show the smooths for the compound type *floor materials*,[3] and the last two rows show the smooths for *meteor crater.* As generalized above for all compounds, both smooths in initial position (rows 1 and 3) are nearly overlapping for both compounds. Yet, for final position, there are notable differences between the smooths for *floor materials* (row 2) and *meteor crater* (row 4). The former shows roughly the generalized pattern in that both smooths have a clear peak associated with the left constituent, but the pitch range for FINAL/OLD is more compressed than that for FINAL/NEW. For the latter, there is only a minimal pitch movement in the FINAL/OLD environment, and the left constituent does not show any notable pitch excursion. Apparently, speakers do not regularly produce an accent on any constituent of *meteor crater* in this environment – here, the compound type is completely de-accented. Contrary to this, *floor materials* retains a pitch accent on the left constituent in all environments. The by-item analysis found similar instances of complete de-accentuation in the FINAL/OLD environment, but not in any of the other environments, in *dinner potato*, *league team*, *life sentences*, and *student card.* In contrast, *access barriers*, *church services*, *floor materials, pincer attack*, and *police helmets,* are further examples in which the accent on the left constituent is retained.

The effect of de-accentuation on the perception of prominence patterns in FINAL/OLD environments is not fully clear. In those cases in which the left constituent is

[3] The figure shows a strong inflation of confidence bands in *floor materials* towards the left edge of sentence-initial compounds, and towards the right edge of sentence-final compounds. This is due to the distribution of unvoiced segments in the compound, as discussed above in section 3.4: as the estimation is based on a smaller number of measurements in these locations, the estimation uncertainty increases.

accented, but in which no accent can be found on the right constituent (that is, in INITIAL position, and FINAL/NEW position for all, and FINAL/OLD position for many types), the accentuation pattern should clearly trigger the perception of left prominence, as this pattern has been consistently associated with left-prominent compounds in the phonetic literature (Farnetani et al. 1988, Plag 2006, Kunter & Plag 2007, Kunter 2011, Morrill 2012). Right-prominent compounds have been found to have comparable pitch peaks or pitch averages in both constituents, but these studies did not differentiate between those cases in which the pitch in both constituents was comparable because both were accented, and those in which the similar pitch was caused by a de-accentuation of the whole compound. An acoustic inspection by the author of the tokens of *meteor crate*r, the compound that the by-item GAM found to be consistently de-accented in FINAL/OLD position, suggests that the lack of an accent on the left constituent raises the relative prominence of the right constituent even if it is also unaccented. Yet, a principled perception experiment will be necessary to determine how listeners perceive the prominence pattern in compounds that are devoid of any accent.

The finding that some FINAL/OLD compounds were consistently de-accented while others were usually produced with an accent on the first constituent might be a residual of the experimental design, as the context sentences or the carrier sentences of de-accenting environments might, for example, place the focus within the sentence on an element preceding the target compound. Focussed elements within a sentence usually receive a nuclear accent (see Hirschberg 2004: 525), which should cause the de-accentuation of all following sentence elements, including the target compound. However, a look at the environments (provided in the Appendix) does not indicate a

systematic difference in focus between de-accentuation and accent-retention environments, nor are there clear differences in the syntactic structure or the semantic relations. In both types of environments, the target compound occurs either as the head of a prepositional phrase or as the object of a subordinate clause, and they can occur either in singular preceded by an article or in plural forms without an article. The context sentence in both sets either mentions none, one, or both constituents explicitly, but never the whole compound. Finally, all environments appear to provide a context in which the target compound stands 'in some salient and relevant relationship to elements that have been evoked', and are thus discourse-old in terms of Ward et al. (2002: 1358)'s definition. In short, there is no apparent contextual reason why the nuclear accent should be consistently placed outside of the target compound in some, but not in all FINAL/OLD environments.

An alternative to a contextual explanation of these differences may lie in the compound types themselves. In his corpus study, Kunter (2011: chapter 8) has found evidence for type-specific constraints on the degree of variability: while some compound types showed a large degree of variation in their prominence patterns, shifting between left and right prominence both between and within speakers, other types appeared to be resistant to variation, and received the same pattern in all instances. The present result resembles these findings in so far as, here as in the earlier investigation, the prosodic pattern within compounds showed notable differences between types that could not readily be explained by contextual variables. This type-based interpretation agrees well with the growing empirical evidence outlined in the introduction that the distribution of prominence patterns in compounds is strongly

determined by similarity relations between compounds stored in the mental lexicon of speakers. A generic model of prominence pattern assignment in English compounds that uses these relations is presented in Arndt-Lappe & Bell (manuscript), who show that the information stored about exemplars in the lexicon can be used to model type-specific degrees of prominence variability in English compounds to a large degree. Here, type-specific behaviour is not unexpected at all – instead, it is an inherent feature of their model of grammar in which grammatical patterns are not governed by abstract linguistic rules, or by the effect of a definable set of linguistic features on such rules, but in which such patterns emerge from the relations between stored exemplars in the lexicon. Such a model may also be useful in accounting for the type-specific effect of pragmatic factors on the realization of the pitch contour found in the present production experiment.

In conclusion, the present study has found no evidence that prominence variation in English compounds is systematically triggered either by the position of the compound within a sentence, by its information status, or an interaction between the two factors. The pitch contour of compounds in initial position covers a greater pitch range than those in final position, but given that this is probably a general effect of sentence intonation, this difference is not expected to cause different perceived prominence patterns. In sentence-initial position, there does not seem to be any perceivable difference of the pitch contour between discourse-old and discourse-new compounds. This is also true for many compounds in sentence-final position. Yet, some, but not all compounds are systematically de-accented if they express old information and occur in sentence-final position, and this de-accentuation may lead to a shift in the perceived prominence pattern, but this perceptual shift requires more experimental attention

before it is fully understood.

APPENDIX

List of compounds and environments used in the experiment

Target compound	Position	Information	Environment
access barriers	INITIAL	NEW	The company's new web page was not well received. **Access barriers** made it difficult to read for people with visual impairments.
		OLD	Wheelchair drivers complained about high steps and doors that were difficult to manoeuver. **Access barriers** were the most severe disadvantages of the building.
	FINAL	NEW	Several regulations were introduced at the beginning of the year. For instance, building management had to list **access barriers**.
		OLD	The web site was not considered barrier-free. The colour scheme and the small font imposed **access barriers**.
budget deficits	INITIAL	NEW	The public opinion about the government was not positive. **Budget deficits** made the situation even worse.
		OLD	Last year the company spent more than they earned. **Budget deficits** almost always have occurred since 1999.
	FINAL	NEW	The new governor made his first appearance yesterday. He warned that the current situation might lead to **budget deficits**.
		OLD	Public spendings have exceeded the income for several years now. The regional government have come to terms with **budget deficits**.
church service	INITIAL	NEW	The village always celebrates the day it was founded. A **church service** starts the festivities.
		OLD	Several parents protested against all children attending the traditional mass at the end of the school term. A **church service** was considered discriminatory against pupils of different faiths.
	FINAL	NEW	A group of girls had opened the wrong door in the dark corridor. They had accidentally stumbled into a **church service**.
		OLD	The villagers went to church every Sunday. A weekend just didn't feel right without a **church service**.
club members	INITIAL	NEW	The picnic at the arena was a great success. **Club members** celebrated the victory together with their spouses.

		OLD	The Groucho Club was flooded with new people after a T.V. report. **Club members** didn't appreciate the new popularity.
	FINAL	NEW	The prices were not the same for everybody. There was a discount for **club members**.
		OLD	The fan club recently invited all members to a meeting. They hoped to improve communication between **club members**.
computer genius	INITIAL	NEW	The research team was facing a big problem. A **computer genius** had to be asked to rescue some of their data after a hard-disk failure.
		OLD	As a notorious hacker, Steve needed less than five minutes to crack the password. A **computer genius** usually finds his way into even the most secure systems.
	FINAL	NEW	Sarah's parents always believed she would choose an artistic profession. Instead, she turned out to be a **computer genius**.
		OLD	Even though the system was badly broken, the company's expert fixed the computer in no time. That showed the management the advantages of employing a **computer genius**.
dinner potato	INITIAL	NEW	The local biotech company made sweeping changes to their corporate identity. A **dinner potato** was presented as the new logo.
		OLD	Food scientists emphasized the worth of potatos as a central ingredient of dinner and other meals. A **dinner potato** was found to reduce the risk of liver disease.
	FINAL	NEW	Mandy had a brilliant idea for the fancy-dress ball. She planned to dress up as a **dinner potato**.
		OLD	At dinner, little Tom made a mess with his potato dish. Even his mother was hit by chunks from a **dinner potato**.
farm cats	INITIAL	NEW	The grumpy old man was notorious for his shooting frenzies. **Farm cats** were his favourite targets.
		OLD	At the age of seventeen, Tiger was the oldest cat in the whole village. **Farm cats** only occasionally reach this age.
	FINAL	NEW	The house had been abandoned a year ago. The last remaining inhabitants were **farm cats**.
		OLD	Mary realized too late that the old tomcat from the farm was infested by fleas. Her mother had warned her not to play with **farm cats**.
fish steaks	INITIAL	NEW	John surprised everyone with his choice. **Fish steaks** filled his plate.

		OLD	It was obvious that Mary would not order the tuna. **Fish steaks** always upset her stomach.
	FINAL	NEW	John annoyed the host at yesterday's party. He provoked her by asking for **fish steaks**.
		OLD	Mary had to give her piece of grilled tuna to the cat. She couldn't stand the smell of **fish steaks**.
floor materials	INITIAL	NEW	Mike found a great addition to the stock of his bargain shop. **Floor materials** were available at extremely low prices.
		OLD	Richard gave only a cursory glance at the carpet. **Floor materials** had never really interested him.
	FINAL	NEW	The salvage team was surprised when they examined their bounty. The whole container was filled with **floor materials**.
		OLD	None of the carpets and glazed tiles on sale were appealing to Mary. She was well aware of her old-fashioned taste in **floor materials**.
hall lights	INITIAL	NEW	The vandals that had broken into the office building had chosen a very particular target. **Hall lights** had been destroyed throughout the place.
		OLD	The old lighting in the hall of the museum had recently been replaced by a new system. **Hall lights** were now used to direct the attention of visitors.
	FINAL	NEW	Sam believed that his engineering class would like the new project he had assigned to them. They were supposed to develop the next generation of **hall lights**.
		OLD	Vera always made sure to turn off the light in the hall. She had read that a lot of energy was wasted by **hall lights**.
killer shark	INITIAL	NEW	The board of tourists had to send a warning to visitors. A **killer shark** had been spotted in the bay.
		OLD	The doctor did not believe that a shark had killed the diver. A **killer shark** leaves bite marks on its victim.
	FINAL	NEW	The new sea side resort had to go through very hard times initially. The local press had spread rumours of a **killer shark**.
		OLD	Several divers had died in encounters with sharks. Consequently, the tourist board posted warnings about a **killer shark**.
league team	INITIAL	NEW	The opening of the new shopping mall featured an unexpected highlight. A **league team** signed autographs for all customers.
		OLD	The performance of the team in the last match of the season was very disappointing. A **league team**

			may be expected to play better.
	FINAL	NEW	All children were very excited when they came back from school last Monday. Their school had been visited by a **league team**.
		OLD	The performance shown in the last match by the Eagles was the worst ever in the league. Everyone had expected much more spirit from a **league team**.
life sentences	INITIAL	NEW	The conclusion of John's analysis of the encyclopedia of great jazz drummers was surprising. **Life sentences** had ended the career of most of them.
		OLD	After the ruling, it was clear to everyone in court that Richard would not spend the rest of his days in prison. **Life sentences** were usually suspended after a certain amount of time.
	FINAL	NEW	Reviewers lauded Chris on the tone of his poems. His most sensitive poem was about **life sentences**.
		OLD	It was rather unlikely that any of the drug barons would ever again get out of prison. They had to come to terms with their **life sentences**.
meteor crater	INITIAL	NEW	The famous director was looking for a location for his next film. A **meteor crater** was found that provided the right ambience.
		OLD	The impact of the meteor had left a huge hole in the surface of the plateau. A **meteor crater** is a great attraction for tourists.
	FINAL	NEW	To the critics, the picture looked like a mess of blurred lines. The artist explained that it showed a **meteor crater**.
		OLD	Geologists rushed to the area where the meteor had left a huge hole in the ground. There were fixed routines for the investigation of a **meteor crater**.
pencil figure	INITIAL	NEW	Historians were able to reconstruct the ancient Greek ideal of beauty for women. A **pencil figure** had not been appreciated by most men.
		OLD	To the fashion designer, all the applicants at the casting looked far too skinny. A **pencil figure** was not really fashionable any more, he thought.
	FINAL	NEW	When the Browns met their future daughter-in-law, they were quite surprised. They had never expected their son to fall for a girl with a **pencil figure**.
		OLD	Lisa was determined to lose a drastic amount of weight before summer. Her doctor warned her about the health risks of a **pencil figure**.
pincer attack	INITIAL	NEW	They didn't know what to do. A **pincer attack** ended the stalemate between the armies.

		OLD	The tactician suggested crushing the enemy between two flanking armies. A **pincer attack** has very often resolved a deadlock in military conflicts.
	FINAL	NEW	The land looked peaceful and calm. Suddenly, the peace was shattered by a **pincer attack**.
		OLD	The general sent out two armies in order to crush the enemy between them. Some advisors had spoken against launching a **pincer attack**.
police helmets	INITIAL	NEW	The internet auction site banned several items from sale. **Police helmets** were on the list.
		OLD	The headgear worn by the police was not penetrated by the stones thrown at them. **Police helmets** are specifically designed to deflect such missiles.
	FINAL	NEW	The summit was accompanied by a number of counter events. Perhaps the most original protest was a group of senior citizens who all wore **police helmets**.
		OLD	The new headwear regulations issued to the police were met with lukewarm reactions. By now, however, officers have come to terms with **police helmets**.
state officials	INITIAL	NEW	There was a plane crash near the airport. **State officials** say that everybody survived.
		OLD	The reporters were waiting for a statement from the government. **State officials** were present in the newsroom.
	FINAL	NEW	There was a huge crowd in front of the library. Many reporters were waiting for the arrival of **state officals**.
		OLD	Representatives from several government departments had sent out letters to the school. Teachers and students were invited to meet **state officials**.
student card	INITIAL	NEW	Yesterday, police discovered a dead body by the riverside. A **student card** was used to identify the corpse.
		OLD	The university help desk always reminded applicants to bring a photograph for registration. A **student card** could not be issued without a picture of the student.
	FINAL	NEW	The burglar was very disappointed about the contents of the safe. All it contained was a **student card**.
		OLD	The university office informed new students that a photograph was necessary for registration. A picture was a prerequisite for issuing a **student card**.

task force	INITIAL	NEW	No information about the earthquake had been available. A **task force** arrived in the region in the afternoon.
		OLD	The board of directors discussed the size of the team necessary to counter the financial crisis. A **task force** should consist of seven people, the directors agreed.
	FINAL	NEW	The board of directors decided to pursue a new course of action this time. Therefore, all operational decisions were assigned to a **task force**.
		OLD	The president's suggestion was to form a group of experts to respond to the financial crisis. The board of directors, however, did not see the point of a **task force**.

REFERENCES

Akaike, Hirotugo. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6), 716–23.

Arndt-Lappe, Sabine. 2011. Towards an exemplar-based model of stress in English noun-noun compounds. *Journal of Linguistics* 47(3), 549–85.

Arndt-Lappe, Sabine & Melanie J. Bell. (manuscript). *An analogical theory of morpho-phonology. Generalisation and variability in English compound stress*. Submitted for publication. Available from: http://www.phil-fak.uni-duesseldorf.de/anglistik3/arndt-lappe/publications_arndt_lappe/

Baayen, R. Harald, Richard Piepenbrock & Leon Gulikers. 1995. *The CELEX lexical database*. Philadelphia, PA: Linguistic Data Consortium.

Bauer, Laurie. 1983. Stress in compounds: a rejoinder. *English Studies* 64, 47–53.

Beckman, Mary E. & Jan Edwards. 1994. Articulatory evidence for differentiating stress categories. In Patricia A. Keating (ed.), *Phonological structure and phonetic form: papers in laboratory phonology III*, 7–33. Cambridge: Cambridge University Press.

Beckman, Mary E. & Janet B. Pierrehumbert. 1986. Intonational structure in English and Japanese. *Phonology Yearbook* 3, 255–309.

Beckman, Mary E., Julia Hirschberg & Stefanie Shattuck-Hufnagel. 2005. *The original ToBI system and the evolution of the ToBI framework.* In Sun-Ah Jun (ed.), *Prosodic typology: the phonology of intonation and phrasing*, 9-54. Oxford: Oxford University Press.

Bell, Melanie. 2008. *Noun noun constructions and the assignment of stress*. Presented at the 1st Conference of the International Society for the Linguistics of English,

Freiburg, 8-11 October.

Bell, Melanie & Ingo Plag. 2012. Informativeness is a determinant of compound stress in English. *Journal of Linguistics* 48(3), 485–520.

Bell, Melanie. 2013. *The English noun noun construct: its prosody and structure*. Ph. D. dissertation, University of Cambridge.

Bloomfield, Leonard. 1933. *Language*. New York: Holt.

Boersma, Paul & David Weenink. 2013. *Praat: doing phonetics by computer (Version 5.3.41)*. Computer program. Retrieved 1 March 2013, from http://www.praat.org/.

Brazil, David. 1997. *The communicative value of intonation in English*. Cambridge: Cambridge University Press.

Chafe, Wallace L. 1974. Language and consciousness. *Language* 50(1), 111–33.

Cruttenden, Alan. 2006. The de-accenting of given information. A cognitive universal? In Guiliano Bernini & Marcia L. Schmartz (eds.), *Pragmatic organization of discourse in the languages of Europe*, 311–56. Berlin: Mouton de Gruyter.

Davies, Mark. 2008-. *The Corpus of Contemporary American English (COCA): 450 million words, 1990-present*. Available online at: http://www.americancorpus.org.

Farnetani, Edda, Carol Taylor Torsello & Piero Cosi. 1988. English compound versus non-compound noun phrases in discourse: an acoustic and perceptual study. *Language and Speech* 31(2), 157–80.

Fudge, Erik. 1984. *English word-stress*. London: George Allen & Unwin.

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

Giegerich, Heinz J. 2009. The English compound stress myth. *Word Structure* 2(1), 1–

17.

Grabe, Esther, Brechtje Post, Francis Nolan & Kimberley Farrar. 2000. Pitch accent realization in four varieties of British English. *Journal of Phonetics* 28(2), 161–85.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Hermes, Dik J. 2006. Stylization of Pitch Contours. In Sudhoff et al. (eds.), 29–61.

Hirschberg, Julia. 2004. Pragmatics and intonation. In Laurence R. Horn (ed.), *Handbook of pragmatics*, 515–37. Malden, MA: Blackwell.

Hirschberg, Julia. 2002. Communication and prosody: functional aspects of prosody. *Speech Communication* 36(1), 31–43.

Huddleston, Rodney D. & Geoffrey K. Pullum (eds.). 2002. *The Cambridge Grammar of the English Language*, Cambridge: Cambridge University Press.

Kingston, John & Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70(3), 419–54.

Kösling, Kristina, Gero Kunter, R. Harald Baayen & Ingo Plag. 2014. Modelling pitch contours in English tripartite compounds. *Language and Speech* 56, 529–54.

Kösling, Kristina & Ingo Plag. 2009. Does branching direction determine prominence assignment? An empirical investigation of triconstituent compounds in English. *Corpus Linguistics and Linguistic Theory* 5(2), 201–39.

Katz, Jonah & Elisabeth Selkirk. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language* 87(4), 771–816.

Kochanski, Greg, Esther Grabe, John Coleman & Burton Rosner. 2005. Loudness predicts prominence: fundamental frequency lends little. *Journal of the Acoustical Society of America* 118(2), 1038–54.

Kunter, Gero. 2011. *Compound stress in English. The phonetics and phonology of prosodic prominence*. Berlin: de Gruyter.

Kunter, Gero & Ingo Plag. 2007. *What is compound stress?*. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, 1005–8.

Ladd, D. Robert. 1984. English compound stress. In Dafydd Gibbon & Helmut Richter (eds.), *Intonation, accent and rhythm*, 253–66. Berlin: de Gruyter.

Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Ladefoged, Peter. 2003. *Phonetic data analysis*. Malden, MA: Blackwell.

Lappe, Sabine & Ingo Plag. 2007. *The variability of compound stress in English: towards an exemplar-based alternative of the compound stress rule*. *Proceedings of the ESSLLI workshop on exemplar-based models of language acquisition and use*, Dublin, 8–23.

Liberman, Mark Y. & Richard Sproat. 1992. The stress and structure of modified noun phrases in English. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical matters*, 131–81. Stanford, CA: Center for the Study of Language and Information.

Liu, Fang & Yi Xu. 2007. *Question intonation as affected by word stress and focus in English*. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, 1189–92.

Morrill, Tuuli. 2012. Acoustic Correlates of Stress in English Adjective–Noun Compounds. *Language and Speech* 55(2), 167–201.

Nolan, Francis. 2003. *Intonational equivalence: an experimental evaluation of pitch scales*. *Proceedings of the 15th International Congress of Phonetic Sciences*,

Barcelona, 771–4.

Ostendorf, Mari, Patti Price & Stefanie Shattuck-Hufnagel. 1996. *Boston University Radio Speech Corpus*. Philadelphia, PA: Linguistic Data Consortium.

Payne, John & Rodney D. Huddleston. 2002. Nouns and noun phrases. In Rodney D. Huddleston & Geoffrey K. Pullum (eds.), 323–524.

Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation*. Ph. D. dissertation, Massachusetts Institute of Technology.

Pierrehumbert, Janet B. & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry L. Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 271–311. Cambridge, MA: MIT Press.

Plag, Ingo. 2006. The variability of compound stress in English: structural, semantic, and analogical factors. *English Language and Linguistics* 10(1), 143–72.

Plag, Ingo & Gero Kunter. 2010. Constituent Family Size and Compound Stress Assignment in English. In Susan Olsen (ed.), *New Impulses in Word-Formation*, 349–82. Hamburg: Buske.

Plag, Ingo, Gero Kunter, Maria Braun & Sabine Lappe. 2008. The role of semantics, argument structure, and lexicalization in compound stress assignment in English. *Language* 84(4), 760–94.

Plag, Ingo, Gero Kunter & Sabine Lappe. 2007. Testing hypotheses about compound stress assignment in English: a corpus-based investigation. *Corpus Linguistics and Linguistic Theory* 3(2), 199–232.

Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In Peter Cole (ed.), *Radical pragmatics*, 223–55. New York: Academic Press.

Schmerling, Susan F. 1971. A stress mess. *Studies in the Linguistic Sciences* 1, 52–66.

Štekauer, Pavol, Július Zimmermann & Renáta Gregová. 2007. Stress in compounds: an experimental research. *Acta Linguistica Hungarica* 54(2), 193–215.

Steedman, Mark. 2014. The surface-compositional semantics of English intonation. *Language* 90(1), 2–57.

Strik, Helmer & Louis Boves. 1995. Downtrend in F_0 and P_{sb}. *Journal of Phonetics* 23(1-2), 203–20.

Sudhoff, Stefan, Denisa Lenertová, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter & Johannes Schließer (eds.). 2006. *Methods in empirical prosody research*. Berlin: Walter de Gruyter.

Turk, Alice E., Satsuki Nakai & Mariko Sugahara. 2006. Acoustic segment durations in prosodic research: a practical guide. In Sudhoff et al. (eds.), 1–28.

Ward, Gregory, Betty Birner & Rodney D. Huddleston. 2002. Information packaging. In Rodney D. Huddleston & Geoffrey K. Pullum (eds.), 1363–449.

Watson, Duane G., Jennifer E. Arnold & Michael K. Tanenhaus. 2008. Tic Tac TOE: effects of predictability and importance on acoustic prominence in language production. *Cognition* 106(3), 1548–57.

Watson, Duane G., Jennifer E. Arnold & Michael K. Tanenhaus. 2010. Corrigendum to Tic Tac TOE: effects of predictability and importance on acoustic prominence in language production. *Cognition* 114(3), 462–3.

Whalen, Douglas H. & Andrea G. Levitt. 1995. The universality of intrinsic F_0 of vowels. *Journal of Phonetics* 23(3), 349–66.

Wood, Simon N. 2006. *Generalized Additive Models. An introduction with R*. Boca

Raton, FL: Chapman & Hall/CRC.