Marcel Schlechtweg

How stress reflects meaning

The interplay of prosodic prominence and semantic (non-) compositionality in non-lexicalized English adjective-noun combinations

Abstract: The paper discusses the relation between stress and meaning in *non-lexicalized* English adjective-noun (AN) combinations. Native speakers of American English were recorded in a production study while reading sentences containing AN constructions such as *black tram*. These items could be interpreted in either a compositional (e.g., a tram that is black) or a non-compositional way (e.g., a tram that runs only during the night). The objective of the experiment was twofold. First, it aimed at examining whether non-lexicalized constructions with a non-compositional meaning were stressed differently than their compositional counterparts. Second, it was investigated whether stress assignment in non-compositional items further depended on whether the non-compositional meaning was explicitly marked by the immediate context. Possible acoustic correlates of stress, i.e., fundamental frequency, duration, and intensity were measured and analyzed. Overall, while the items with implied non-compositional semantics showed a clear tendency towards initial stress, the combinations with compositional meanings did not. Moreover, the constructions whose non-compositional semantics were explicitly marked by the immediate context tended not to carry initial stress either. I argue that initial stress seems to mark non-compositional semantics only if the non-compositional meaning is not explicitly marked by a different means already. The results are interpreted against the background of the interaction of semantic and phonetic aspects in language production.

1 Introduction

It is well known that some English AN constructions have different meanings. For instance, while *green house* is semantically compositional because its entire meaning is the sum of its constituent meanings, *greenhouse* is non-compositional because parts of its meaning are hidden, i.e., not overtly given (see, e.g., Zwitserlood 1994: 366). Although the overall meanings of *green house* and *greenhouse* differ, the same constituents are combined. Nevertheless, the

https://doi.org/10.1515/9783110501933-119

two forms are not identical because they can be distinguished on, for example, prosodic grounds, i.e., the main prominence is placed on the adjective in *greenhouse* but not in *green house* (see Gussenhoven 2004: 19).

While the prosody of lexicalized English AN constructions has been the subject of both comprehension and production research (e.g., McCauley, Hestvik, and Vogel 2012; Morrill 2012; Vogel and Raimy 2002), the prosody of *non-lexicalized* constructions has been investigated in comprehension (e.g., Schlechtweg 2018) but not in production experiments. The current analysis is a first pilot study to fill this gap and aims at contributing to the understanding of how prosodic and semantic aspects interact. The first question to be answered is whether the main prominence appears on the adjective if a construction is semantically non-compositional. If this is the case, the second issue to be addressed is whether initial stress is also used if the non-compositional semantics are explicitly marked by the immediate context.

The article is structured as follows. Section 2 presents the theoretical foundations of the paper. Semantic non-compositionality and means that explicitly mark it are discussed before we turn to the notion of stress, functions of stress, and reasons for stress variation in complex items in English. Section 3 describes a production study, which is still exploratory in nature but provides first evidence for the interplay of the semantic characteristics and the phonetic form of non-lexicalized AN combinations. Section 4 concludes the present paper.

2 Theoretical foundations

2.1 Semantic non-compositionality and means to mark it in English

As mentioned above, some English AN combinations have different meanings. The examples in (1) illustrate the phenomenon:

(1) a. *a green house*
'a house that is green'
b. *a greenhouse*
'a house made of glass that is used for growing plants'

The two spoken versions are distinguished; i.e., in (1a), both *green* and *house* carry an accent but in (1b) only *green* does so (Gussenhoven 2004: 277). Note that orthographic differences are ignored here as the paper focuses on prosodic

contrasts. The question arises, however, what language users do if they deal with non-lexicalized constructions. Consider (2):

(2) a. *a black tram*
'a tram that is black'
b. *a black tram*
'a tram that runs only during the night'

How could one mark the non-compositional semantics of *black tram* (see 2b), which differ from the compositional meaning (see 2a)? On the one hand, the immediate context can explicitly mark a non-compositional meaning. One way of marking non-compositionality of meaning is by explicitly referring to the first constituent as a naming unit (see Härtl 2016). An example of an explicit marker is the phrase *called so because* (see 3a). On the other hand, prosodic modifications can be used in spoken language. For instance, as shown in (3b), while *black tram* with the compositional meaning would probably carry an accent on both constituents, the accent on the noun might be deleted if the non-compositional reading is intended, leaving only the accent on the adjective (see Gussenhoven 2004: 277).

(3) a. *A black tram is called so because it is a tram that runs only during the night.*
b. *A BLACK tram is a tram that runs only during the night.*

So far, however, no study has investigated whether this is actually the case. That means we can only speculate that the prosodic structure is changed as in (3b) to mark the non-compositional meaning, but we do not know for sure. The present paper aims at filling this gap. Moreover, another interesting question remains if one considers (3a): What is the prosodic structure of *black tram* here? Put differently, do speakers also adjust the prosodic structure of non-compositional constructions if their meaning is explicitly marked as non-compositional by the immediate context? Answering this question represents the second objective of the current article. If non-compositional semantics trigger initial stress, this might happen independently of the immediate context. Alternatively, however, initial stress might be used only if the non-compositional meaning is not marked explicitly. The latter scenario would mean that language users avoid, so to speak, redundancy while producing non-lexicalized constructions and rely on a single means to explicitly mark non-compositionality. This would be compatible with results discussed in Härtl (2016), who found that non-compositional German AN compounds were less likely to occur with *sogenannt* ('so-called') or quotation marks than non-compositional AN phrases. German compounds are marked by their nature

because they lack the inflectional adjectival suffix of phrases and have initial rather than non-initial, phrasal stress. Therefore, in contrast to phrases, they do not seem to depend on additional means that mark their non-compositional meaning.

2.2 Stress in complex constructions in English

2.2.1 Abstract versus concrete approaches to prosody

As Ladd and Cutler (1983) illustrate, prosody research can be roughly classified into two approaches. On the one hand, abstract approaches theoretically describe prosody and its connection to other domains of grammar. On the other hand, concrete approaches examine the physical characteristics of prosody by investigating its different acoustic correlates such as fundamental frequency (F0), duration, and intensity. The present article combines the two perspectives. That is, the influence of semantics on the prominence pattern of complex constructions is discussed, and it is assumed that prominence can be expressed through the physical variables F0, duration, and intensity (see also, e.g., Kunter 2011; Lehiste 1970; Plag, et al. 2008).

2.2.2 Stress versus accent

It has often been claimed that complex constructions in English are stressed either on the first or on the second element. Chomsky and Halle (1968: 94), for instance, argue that primary stress in phrases is assigned "to the rightmost sonority peak in the string under consideration" (= Nuclear Stress Rule). In contrast, primary stress in compounds is located on "the leftmost sonority peak" (= Compound Rule) (see also, e.g., Liberman and Prince 1977: 257). Gussenhoven (2004: 19) takes up the distinction between the two prominence patterns but defines them differently. He states that both elements of a complex construction are stressed; however, while only the first element of a compound is accented, both elements of a phrase are accented. Gussenhoven's differentiation connects to the view that unstressed syllables are never accented but stressed syllables are accented or not (Bolinger 1958, 1986; Vanderslice and Ladefoged 1972). Stress represents here a feature of the lexical level, and pitch-accents are "added" at the phrasal level. In the present article, the term "stress" is used to mean main prominence, and it is measured in terms of its acoustic correlates F0, duration, and intensity. The study remains agnostic as to whether

the relevant level of grammatical computation is lexical or phrasal. Throughout the present paper, the term "initial stress" refers to what is usually known as "compound stress / prominence" and the term "non-initial stress" refers to what is usually known as "phrasal / nuclear stress / prominence".

2.2.3 Functions of stress and reasons for stress variation

Assuming that English AN constructions have either initial or non-initial stress, we must ask what determines whether a specific combination is stressed in one or the other way. One basic function of stress is to structure the information of an utterance according to the speaker's and listener's needs at a specific moment in their communication (see, e.g., Bell and Plag 2013; Ladd 1984). That means, for instance, while information that is in focus or has not been introduced before during a conversation is typically prominent, non-focused or given information is usually not. Stress can also be used to contrast several alternatives. The example in (4) shows that stress serves to, first, contrast different colors and, second, introduce new information, namely the color red, to the current communication.

(4) *Speaker A: I know that you wore the green shirt yesterday night.*
Speaker B: No, I wore the RED shirt.

Furthermore, initial stress in phrasal or compound constructs is often considered to be a reflex of lexicalization (see, e.g., Plag et al. 2008). For example, while the lexicalized *greenhouse* has initial stress, the non-lexicalized *green house* bears non-initial stress. Apart from the aforementioned factors, several other reasons for stress variation exist. These include within- and across-speaker-related factors (see, e.g., Kunter 2011: Chapter 8), dialectal influences (see, e.g., Trudgill and Hannah 2008: 57), sentence type (see, e.g., Morrill 2012), sentence position (see, e.g., Farnetani, Torsello, and Cosi 1988), the surrounding material (see, e.g., Gussenhoven 2004), and analogy (see, e.g., Plag 2006; Plag, Kunter, and Lappe 2007). In the experiment to be reported later, all the factors mentioned so far are controlled for (see Section 3) in order to examine whether another factor, namely semantic (non-)compositionality, has an influence on the stress pattern of non-lexicalized AN constructions in English.

Considering all English AN constructions, i.e., both lexicalized and non-lexicalized ones, we observe that, first, non-initial stress is more frequent than initial stress and, second, that compositionality is more common than non-compositionality as AN combinations typically fulfill a simple descriptive func-

tion (see, e.g., Liberman and Sproat 1992: 134). Hence, on the one hand, there seems to be a connection between non-initial stress and semantic compositionality. The idea finds further support in Giegerich (2004), who argues that noun-noun (NN) attribute-head constructions, whose semantics overlap with those of prototypical AN items to a large extent, usually carry non-initial stress. On the other hand, NN constructions in particular show that initial stress seems to be linked to semantic non-compositionality. Since the semantic relation between the two nouns is not overtly expressed, parts of the semantics are hidden and, thus, the meaning of NN combinations is non-compositional (see, e.g., Zwitserlood 1994). Apart from being non-compositional, NN constructions typically bear initial stress: Based on the investigation of different corpora, several authors claim that approximately 67 percent (Plag and Kunter 2010: 357), around 75 percent (Liberman and Sproat 1992: 134), almost 90 percent (Plag and Kunter 2010: 357; Plag et al. 2007: 207–208) or even approximately 94 percent (Berg 2012: 11; Plag and Kunter 2010: 357) of English NN constructions have initial stress. The aforementioned observations point to the connection both between initial stress and non-compositionality and between non-initial stress and compositionality, which has also been investigated in further experimental studies. Using a lexical-decision task, McCauley et al. (2012) showed that non-compositional AN items in English were responded to more accurately if they were presented with initial stress in comparison to non-initial stress. Compositional constructions, however, showed higher accuracy rates with non-initial stress. Vogel and Raimy (2002) and Hall and Moore (1997) observed that adults were more likely to select a picture representing a non-compositional interpretation when they heard English AN combinations with initial stress. In contrast, hearing non-initial stress, participants favored pictures expressing compositional meanings. Focusing on production rather than comprehension, Farnetani et al. (1988) and Morrill (2012) found that compounds, i.e., non-compositional constructions, were typically pronounced with initial stress, but phrases, i.e., compositional items, with non-initial stress. Overall, the findings of these studies underline the link between stress and semantic compositionality. However, previous research generally suffers from the fact that the non-compositional items were not only non-compositional but also lexicalized. Therefore, it is difficult, if not impossible, to state that semantics, rather than lexicalization, is really responsible for the effects. Investigating non-lexicalized constructions represents an appropriate alternative that enables us to concentrate on semantics while controlling for lexicalization. In Schlechtweg (2018), for instance, only non-lexicalized items were tested in a lexical-decision experiment and it was shown that non-compositional AN combinations in English were perceived

more efficiently with initial than with non-initial stress. The major concern of the current article is to see whether the connection between semantics and stress in non-lexicalized constructions can be confirmed in a production study.

3 (Non-)compositionality and stress in non-lexicalized English AN combinations: Insights from a production study

3.1 Objectives and hypotheses

The study asks, first, whether non-lexicalized and semantically non-compositional AN constructions in English are pronounced with initial stress, as opposed to their compositional counterparts. If this is the case, the experiment further aims at investigating whether non-compositional combinations also have initial stress if their meaning is explicitly marked as non-compositional by the immediate context. Specifically, *called so because* is used in the present study for this purpose. It is examined whether initial stress occurs independently of the immediate context or whether the explicit marker *called so because* inhibits the realization of initial stress (see also Section 2.1). In order to address these issues, the three conditions given in Table 1 are investigated (for further examples, see Table 6).

Tab. 1: The three conditions of the study

Condition	Example
Implied compositional semantics	*Thomas took a black tram again, which has a color he likes.*
Implied non-compositional semantics	*Thomas took a black tram again, which is a tram that runs only during the night.*
Explicitly marked non-compositional semantics	*Thomas took a black tram again, which is called so because it is a tram that runs only during the night.*

The study was designed to test the following hypotheses:

1. The items with implied non-compositional semantics, but not the items with implied compositional semantics, were expected to show initial stress

because this prosodic pattern is regarded as a means to mark non-compositionality. The hypotheses further below are formulated under the assumption that Hypothesis (1) is met. If Hypothesis (1) is not met, it must be concluded that semantic (non-)compositionality does not seem to have an influence on the stress pattern.

2. With regard to the comparison of the items with implied non-compositional semantics and the same items whose non-compositional meaning is explicitly marked, the following outcomes are possible:
   a. There is no effect, and items in both conditions have initial stress. This would mean that non-compositional semantics always trigger initial stress, independently of how explicitly non-compositionality in meaning is marked in the immediate context.
   b. The items with implied non-compositional semantics have initial stress but the items whose non-compositional semantics are explicitly marked do not. This would mean that non-compositionality triggers initial stress only if no other explicit marker of non-compositionality is present. Since *called so because* already explicitly marks the non-compositional semantics, it would be redundant to, additionally, modify the stress pattern.
3. With respect to the comparison of the items whose non-compositional semantics are explicitly marked and the items with implied compositional semantics, the following outcomes are possible:
   a. As opposed to the items with implied compositional semantics, the items whose semantics are explicitly marked were expected to have initial stress if Hypothesis (2a), in addition to Hypothesis (1), was correct.
   b. If Hypothesis (2b) was correct, the items whose non-compositional semantics are explicitly marked would, like the items with implied compositional semantics, carry non-initial stress.

3.2 Method

3.2.1 Participants

Six native speakers of American English, four females and two males, were tested in the study. Their mean age was 26 years (age range: 21–36, standard deviation: 5.9), and they were university students.

3.2.2 Materials

Six non-lexicalized complex AN constructions were created. Each AN combination was embedded in three different sentences and conditions (see Table 1 above and Table 6). The compositional version of an item was always tested prior to the non-compositional variants of the same item because it seemed likely that the compositional interpretation was less accessible once the new non-compositional one had been introduced. The version with implied non-compositional semantics preceded the one whose non-compositional meaning was explicitly marked in 50 percent of the items; in the remaining 50 percent, the order was reversed.

Several potentially confounding variables were controlled for in the experiment (see also Section 2.2.3). In order to reduce the influence of analogy, the non-compositional meanings were based on non-existent relations between the adjective and the noun. For instance, *black tram* with non-compositional semantics refers to a tram that runs only during the night. Although lexicalized AN constructions with the adjective *black* exist in English, there is no combination in which *black* represents the concept NIGHT. The nouns, in turn, did not appear in any lexicalized AN construction anyway. Furthermore, lexicalization effects were ruled out by investigating non-lexicalized items only. Moreover, since the same combinations were used in the three conditions, the informativity of the constituents and the phonetic environment were identical across conditions. The AN items were embedded in the same sentence type and position in each condition. That means, for instance, that *black tram* always occurred in the main clause *Thomas took a black tram again*, which, in turn, was followed by a relative clause starting with *which*. In order to minimize the influence of individual differences between language users and dialects, all subjects spoke each AN item in each condition and all were American speakers. Finally, as will be explained below Table 2, information structure was controlled for as well.

Apart from the 18 test sentences, 42 filler sentences, which increased the distance between the test items in one condition and the same items in another condition, were included in the experiment. Subjects always read at least 20 other sentences between the sentence with an item in one condition and the sentence with the same item in another condition.

3.2.3 Procedure

Subjects sat in a silent room approximately 40 centimeters (16 inches) from a large diaphragm condenser USB microphone and 70 centimeters (28 inches) from a computer screen. Participants saw one sentence from Table 1 or Table 6 at a time on the screen, read it silently first, and said "Okay" once they had read and understood it. After their reaction, a yes-no comprehension question referring to the sentence was shown on the screen and participants were asked to give the correct answer. Table 2 gives the respective questions for the sentences already presented in Table 1. The overall accuracy rate was 97 percent; only sentences associated with correct answers were later analyzed. After the response, the sentence appeared on the screen again, subjects read it aloud, and were recorded with Praat (Boersma and Weenink 2016).

Tab. 2: The yes-no questions

Sentence	Yes-no question
Thomas took a black tram again, which has a color he likes.	Is a black tram a tram that is bright? (Answer: No)
Thomas took a black tram again, which is a tram that runs only during the night.	Is a black tram a tram that goes to the graveyard? (Answer: No)
Thomas took a black tram again, which is called so because it is a tram that runs only during the night.	Is a black tram called so because it is a tram that goes to the graveyard? (Answer: No)

It is well known that information structure can have an influence on prosody (see Section 2.2.3). Therefore, the three conditions under investigation must be comparable with regard to information structure. Apart from the presence / absence of *called so because it is*, the sentences and questions used in the two non-compositional conditions were identical and, hence, information structure was controlled for. It was, however, equally important to ensure that information structure did not vary between the compositional and non-compositional conditions. The yes-no questions played a decisive role in this respect. As Table 2 shows, the focus always laid on the same element, for example, on the noun *tram*. Moreover, the amount of given and new information was identical across the conditions. For instance, *black* was used once and *tram* twice in the question of each condition. Hence, *black* was, so to speak, less given in the context of each condition. Generally speaking, this might increase the

likelihood of initial stress because new, but not given, information is normally made prominent (Ladd 1984); however, the crucial point is that the information structure is balanced across all conditions.

3.3 Data analysis

A vital question in a production study is how one determines whether a construction has initial or non-initial stress. The present analysis is based on the measurement of three potential acoustic correlates of stress, namely F0, duration, and intensity. Generally speaking, higher F0, longer duration, and higher intensity have been traditionally associated with stressed syllables (see Lehiste 1970). Since the present paper cannot give a detailed and general discussion of these parameters (for a review, see, e.g., Kunter 2011: 57–69; Terken and Hermes 2000), we focus on three studies whose methodology is similar to that of the experiment reported in the present section. Plag (2006) analyzed the F0 in compounds. The author defined initial and non-initial stress in the following way: He first calculated the F0 difference between the first and second element of different compound types and then compared these differences. If compound type A showed a greater difference than compound type B, compound type A was considered to have initial stress and compound type B to carry non-initial stress. Farnetani et al. (1988) and Morrill (2012) not only looked at F0 but also at duration and intensity. Further, they used not only differences but also ratios. That means, for instance, that the authors examined minimal pairs and regarded greater F0 differences, higher duration ratios, and / or greater intensity differences between the first and second element of one construction in comparison to another construction as an indication of initial stress in the first and non-initial stress in the second construction. The present experiment adopts and slightly expands this approach, following previous work that has shown that the methodology can be successfully applied to the study of prominence in complex constructions. That is, it is assumed here that a statistically significant difference between the ratios (adjective values divided by noun values) and differences (adjective values minus noun values) of two conditions reflects the phonological categorical distinction between initial and non-initial stress. For instance, if construction A shows a greater F0 ratio and difference than construction B, this can be an indication that construction A carries a pitch accent on the first constituent only, i.e., it has initial stress, and construction B bears a pitch accent on both the first and the second constituent, i.e., it has non-initial stress (see Gussenhoven 2004: 277).

Before one can calculate ratios and differences, however, one has to measure the acoustics of the adjectives and nouns. The vowels of these constituents together with, if available in an item, liquids and glides were separated from the rest of the recordings and used as the intervals for the following measurements (for a detailed overview of segmentation criteria, see Turk, Nakai, and Sugahara 2006). The duration (= D) of each interval was obtained from the oscillogram and the maximum intensity (= I) was measured with the "Get maximum intensity" function in Praat. Moreover, the maximum F0 (= F0) was retrieved with the "Get maximum pitch" function and Praat's autocorrelation method. A pitch range of 75 to 300 Hertz (Hz) (males) and 100 to 500 Hz (females) was chosen and individually adjusted if necessary. Since extreme outliers were excluded from the analyses and since the constituents of the AN items were not associated with boundary tones, maximum F0s were used because they reflect the pitch contour more appropriately than mean F0s or F0s at the mid-points of vowels (Kunter 2011: 74–75). The target items were not placed in clause-, statement-, or question-final position; instead, each item was put before the word *again*, "which was expected to carry all boundary-related tonal elements" (Plag, Kunter, and Schramm 2011: 364).

For all adjectives and nouns, F0, D, and I were obtained. Afterwards, the ratio and difference of the adjective and noun value was calculated resulting in the six dependent variables fundamental frequency ratio ($F0_r$), fundamental frequency difference ($F0_d$), duration ratio (D_r), duration difference (D_d), intensity ratio (I_r), and intensity difference (I_d). Repeated-measures ANOVAs by subject (F_1) and by item (F_2) were conducted using $F0_r$, $F0_d$, D_r, D_d, I_r, and I_d as dependent variables. Homogeneity of variances, an assumption of the ANOVA, was given in F_1 and F_2 in the analyses of D_r, D_d, I_r, and I_d; both the Bartlett test, which assumes a normal distribution of the data, and the Levene test, which does not assume a normal distribution of the data, showed that the variances were equal ($p > .05$). In the analysis of $F0_r$, these tests revealed equal variances at least in F_1. In the analysis of $F0_d$, at least the Levene test indicated equal variances in both F_1 and F_2. The independent and fixed variable, SEMANTIC COMPOSITIONALITY, was a within-subject / item factor and had the three levels implied compositional semantics, implied non-compositional semantics, and explicitly marked non-compositional semantics. SUBJECT and ITEM were included as random variables.

3.4 Results

Note that, in this result section, the three conditions are abbreviated in the following way: C = implied compositional semantics, N = implied non-compositional semantics, S = explicitly marked non-compositional semantics.

3.4.1 Fundamental frequency

The analysis of $F0_r$ revealed a significant main effect of SEMANTIC COMPOSITIONALITY ($F_1(2,10) = 6.83$, $p < .05$; $F_2(2,10) = 5.66$, $p < .05$). Post-hoc comparisons showed that N and C significantly differed (Difference of means $[DM]_1 = 0.104$, $t_1 = 2.79$, $p_1 < .05$; $DM_2 = 0.137$, $t_2 = 2.56$, $p_2 < .05$). Significance was also reached in the comparison of S and N ($DM_1 = -0.130$, $t_1 = -3.50$, $p_1 < .01$; $DM_2 = -0.170$, $t_2 = -3.17$, $p_2 = .01$), but not between S and C.

The analysis of $F0_d$ showed a significant main effect of SEMANTIC COMPOSITIONALITY ($F_1(2,10) = 7.45$, $p = .01$; $F_2(2,10) = 6.50$, $p < .05$). Not only N and C ($DM_1 = 20.63$, $t_1 = 3.40$, $p_1 < .01$; $DM_2 = 25.33$, $t_2 = 3.11$, $p_2 < .05$) but also S and N significantly differed ($DM_1 = -19.88$, $t_1 = -3.28$, $p_1 < .01$; $DM_2 = -25.48$, $t_2 = -3.13$, $p_2 < .05$). The difference between S and C did not reach significance. Overall, Hypotheses 1, 2b, and 3b were confirmed. Descriptive statistics are summarized in Table 3 ($F0_r$ and $F0_d$) and displayed in Figure 1 (only $F0_d$).

Tab. 3: Descriptive statistics of $F0_r$ / $F0_d$, subject analysis (F_1) (item analysis [F_2] in brackets)[1]

	C	N	S
N of $F0_r$	6 (6)	6 (6)	6 (6)
N of $F0_d$	6 (6)	6 (6)	6 (6)
M of $F0_r$	1.003 (1.001)	1.106 (1.138)	0.976 (0.968)
M of $F0_d$	-4.61 (-5.05)	16.03 (20.28)	-3.85 (-5.20)
SD of $F0_r$	0.091 (0.147)	0.145 (0.136)	0.044 (0.039)
SD of $F0_d$	6.68 (21.73)	21.32 (21.81)	5.86 (6.59)

1 N (in first column) = Number of observations, M = Mean, SD = Standard deviation.

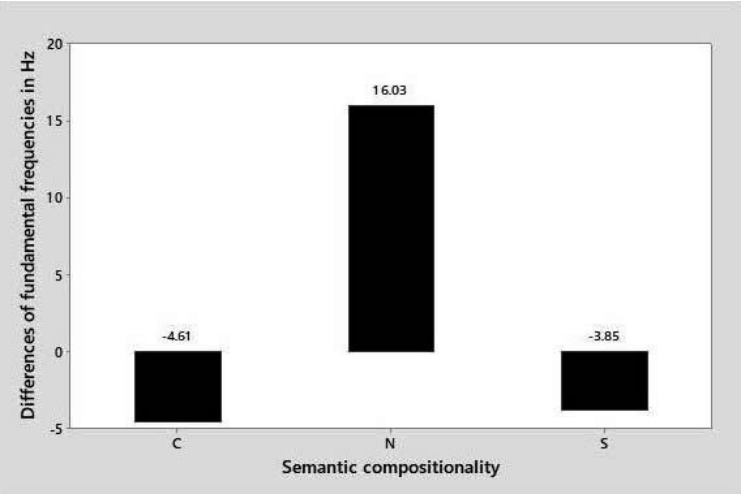

Fig. 1: Mean $F0_d$ in subject analysis (F_1)

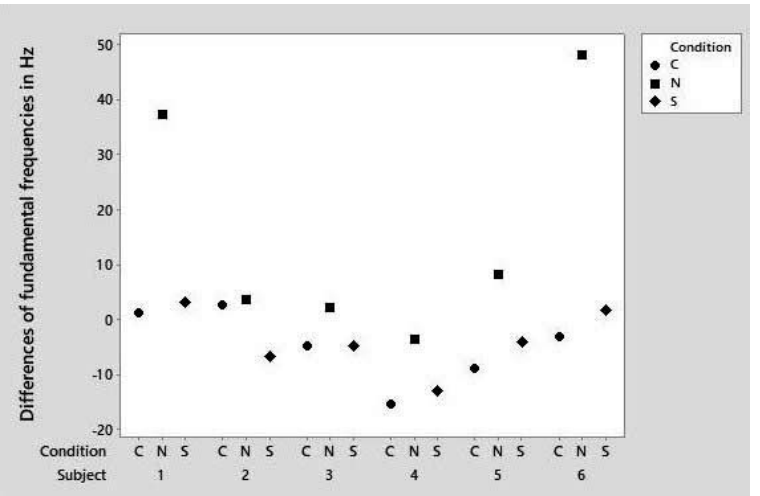

Fig. 2: Distribution of values of $F0_d$ in subject analysis (F_1)

Figures 2 and 3 illustrate that the individual values that contributed to the overall means given in Table 3 were higher in condition N than in conditions C and S, which, in turn, were often closer together, for all subjects and items. Since the

analysis of $F0_r$ revealed a similar pattern, the distributions are not presented here.

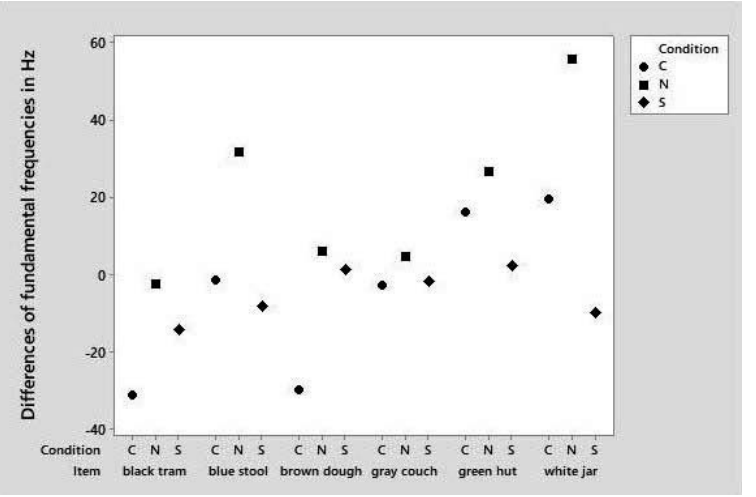

Fig. 3: Distribution of values of $F0_d$ in item analysis (F_2)

3.4.2 Duration

The analysis of D_r showed a significant main effect of SEMANTIC COMPOSITIONALITY ($F_1(2,10) = 11.57$, $p < .01$; $F_2(2,10) = 8.20$, $p < .01$). Post-hoc comparisons revealed a significant difference between N and C ($DM_1 = 0.257$, $t_1 = 4.77$, $p_1 = .001$; $DM_2 = 0.214$, $t_2 = 4.05$, $p_2 < .01$). The difference between S and N was significant in the subject analysis and marginally significant in the item analysis ($DM_1 = -0.157$, $t_1 = -2.91$, $p_1 < .05$; $DM_2 = -0.109$, $t_2 = -2.05$, $p_2 = .067$). S and C did not significantly differ.

The analysis of D_d showed a significant main effect of SEMANTIC COMPOSITIONALITY ($F_1(2,10) = 10.29$, $p < .01$; $F_2(2,10) = 9.87$, $p < .01$). N and C significantly differed ($DM_1 = 35.60$, $t_1 = 4.44$, $p_1 = .001$; $DM_2 = 31.45$, $t_2 = 4.38$, $p_2 = .001$). A significant difference was also detected between S and N ($DM_1 = -24.41$, $t_1 = -3.04$, $p_1 < .05$; $DM_2 = -20.27$, $t_2 = -2.82$, $p_2 < .05$) but not between S and C. Overall, again, Hypotheses 1, 2b, and 3b were confirmed. Descriptive statistics are summarized in Table 4 (D_r and D_d) and displayed in Figure 4 (D_d only).

Tab. 4: Descriptive statistics of D_r / D_d, subject analysis (F_1) (item analysis [F_2] in brackets)

	C	N	S
N of D_r	6 (6)	6 (6)	6 (6)
N of D_d	6 (6)	6 (6)	6 (6)
M of D_r	1.173 (1.181)	1.430 (1.395)	1.273 (1.287)
M of D_d	19.53 (20.49)	55.13 (51.95)	30.72 (31.68)
SD of D_r	0.103 (0.326)	0.111 (0.424)	0.171 (0.384)
SD of D_d	18.25 (61.40)	9.23 (60.40)	26.40 (65.50)

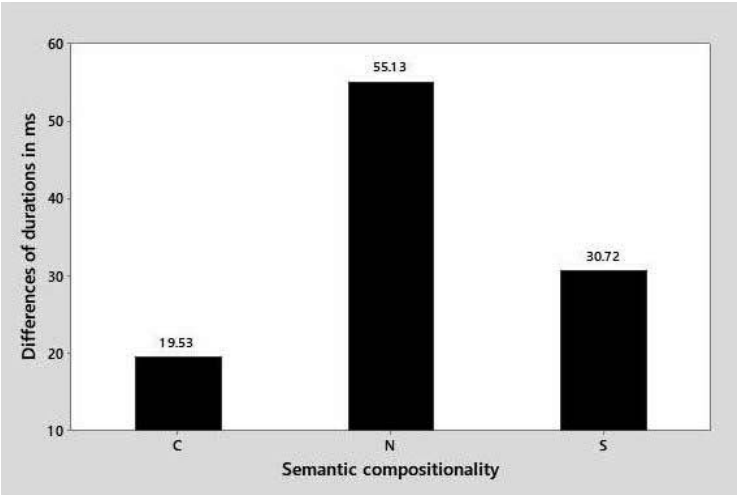

Fig. 4: Mean D_d in subject analysis (F_1)

Some of the standard deviations in Table 4 are high. Figures 5 and 6 present the distributions of the values included in the calculation of the means of D_d. The graphs illustrate that D_d of N is higher than that of C and S for five of the six subjects and items. The analysis of D_r revealed a similar pattern. Figure 5 also shows that N was the most stable condition across subjects. Figure 6 illustrates that the large standard deviations are based on the different phonetic nature of the items and that the standard deviations of all conditions are similar.

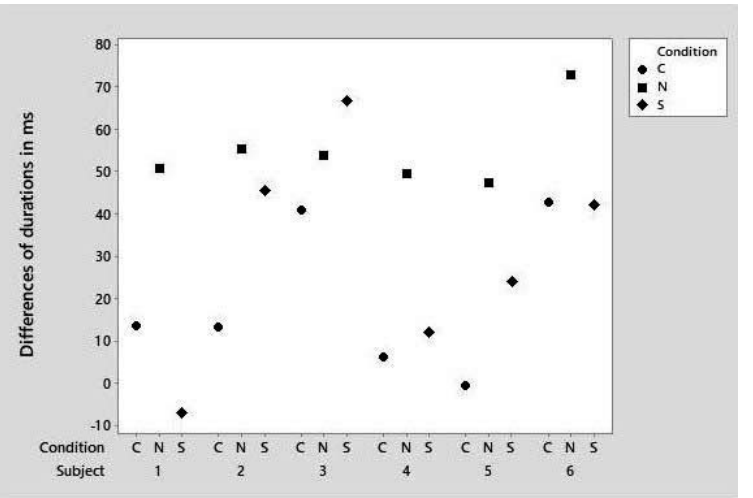

Fig. 5: Distribution of values of D_d in subject analysis (F_1)

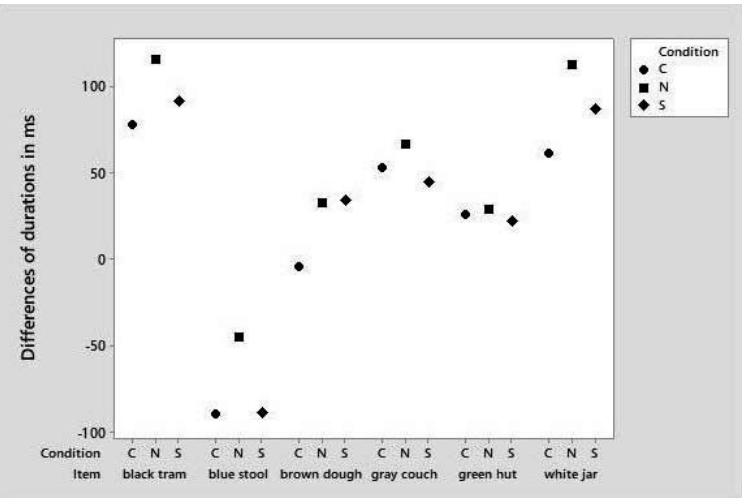

Fig. 6: Distribution of values of D_d in item analysis (F_2)

3.4.3 Intensity

The analyses of I_r and I_d did not show a main effect of SEMANTIC COMPOSITIONALITY. Descriptive statistics are presented in Table 5.

Tab. 5: Descriptive statistics of I_r / I_d, subject analysis (F_1) (item analysis [F_2] in brackets)

	C	N	S
N of I_r	6 (6)	6 (6)	6 (6)
N of I_d	6 (6)	6 (6)	6 (6)
M of I_r	1.012 (1.012)	1.020 (1.021)	1.012 (1.010)
M of I_d	0.76 (0.74)	1.26 (1.30)	0.81 (0.60)
SD of I_r	0.012 (0.010)	0.016 (0.020)	0.015 (0.012)
SD of I_d	0.69 (0.66)	0.98 (1.39)	1.03 (0.78)

3.4.4 Summary of results

In sum, only the AN constructions with implied non-compositional semantics showed a clear tendency towards initial stress. In contrast, the combinations with explicitly marked non-compositional semantics, as well as the items with implied compositional semantics, tended to carry non-initial stress. Crucially, both measures taken (ratios and differences) for two of the three acoustic parameters measured (F0 and duration) showed robust effects in the expected direction. The only parameter that did not show an effect of semantic compositionality was intensity; this, however, is in line with evidence from the literature that suggests that intensity may in fact not always be a reliable cue to stress (for a review, see Cutler 2005).

3.5 Theoretical discussion

Speakers have different means at their disposal to mark that the meaning of a complex construction goes beyond the sum of the meanings of the individual constituents. On the prosodic side, initial stress represents a typical marker of semantic non-compositionality in English. The findings of the present pilot study support the idea that language users place more prosodic prominence on the initial constituent if the semantics of a complex construction deviate from the compositional interpretation. Specifically, it is shown that native speakers of English lengthen the initial syllables and pronounce them at a higher pitch level. The latter finding is compatible with Gussenhoven's (2004: 277) proposal that non-compositional items such as compounds carry an accent only on the

first constituent but compositional constructions such as phrases bear an accent on both constituents. On the non-prosodic level, *called so because* is one way to explicitly mark a non-compositional meaning. The analyses of the current experiment indicate that speakers, if exposed to a non-lexicalized complex construction whose non-compositional semantics are explicitly marked, tend not to use initial stress and, instead, seem to favor the standard prosodic structure of compositional semantics. Therefore, this first pilot study suggests that native speakers of English might rely on either a prosodic or a non-prosodic means, but not on both, to mark meaning deviations when they produce non-lexicalized constructions in their language.

Overall, the present findings are, on the one hand, similar to the results of previous studies such as Farnetani et al. (1988) and Morrill (2012). That means specifically that F0 and duration turned out to be reliable correlates in the distinction between initial and non-initial stress. For example, Morrill's (2012) analysis revealed a greater duration ratio for compounds compared to phrases, indicating that the former carry initial stress but the latter non-initial stress. Keeping in mind that the compounds of her study were non-compositional and the phrases compositional, one can see that the findings are similar to those of the present experiment. On the other hand, however, three differences have to be emphasized as well. First, the other authors investigated lexicalized items and, thus, their effects might also be based on lexicalization rather than semantic non-compositionality. Second, intensity played a much greater role in the other two studies than in the present experiment. A potential explanation is the low number of participants and items examined in my study. Third, some of the other authors' results are connected to the fact that they looked at different sentence positions such as subject position, question-, statement-, or clause-final position. For instance, Morrill (2012) found a higher F0 in the second constituent in comparison to the first one in compounds in question-final position. This effect has its roots in the rising intonation of this environment. In contrast, the present study focused on a single position in which the acoustic properties of items were not influenced by boundary phenomena found in final positions.

A result of the analysis reported here is that only items with implied non-compositional semantics but not items whose non-compositional meaning was explicitly marked by the immediate context showed a clear tendency to carry initial stress. This finding partly connects to the study described in Härtl (2016), who investigated German AN compounds and phrases of comparable frequencies and found that phrases occurred more often with *sogenannt* ('so-called') than compounds. Crucially, the German AN compounds of Härtl's experiment resemble the items with implied non-compositional semantics of the present

study in two respects: First, both are non-compositional and, second, both carry initial stress. The German AN phrases examined by Härtl and the non-compositional constructions that occurred with *called so because* in the current experiment share two characteristics as well: First, and again, both are non-compositional and, second, both have non-initial stress. Taken together, both Härtl's and my own study show that the non-compositional semantics of AN constructions are less likely to be marked by means of initial stress if an explicit and non-prosodic marker such as *sogenannt* or *called so because* is present. Nevertheless, these issues have to be investigated further because, strictly speaking, *sogenannt* and *called so because* differ in crucial respects. The latter, but not the former, has to be followed by an expression that explains the function of the modifier. Further, while *called so because* in the present study focuses on the contribution of the adjectival semantics to the entire meaning only, *sogenannt* refers to the whole AN construction. Whether these syntactic and semantic differences affect stress has not been tackled in the current work.

Finally, it has to be emphasized again that the number of participants and items was, in comparison to other studies such as Morrill (2012), quite small. As a consequence, the results are still rather exploratory in nature and have to be confirmed in subsequent and more comprehensive studies.

4 Conclusion

The current paper investigated non-lexicalized AN combinations in English and addressed two questions: Are non-compositional constructions stressed differently in comparison to compositional items and, if so, does the explicit marking of non-compositionality in the immediate context have an influence on stress distribution? Although the dataset is rather small, this analysis suggests that the answer to both of these questions might be "Yes": Non-compositionality seems to trigger initial stress in non-lexicalized items, but only if no other device to mark non-compositionality, such as *called so because*, is used.

5 References

Bell, Melanie J. & Ingo Plag. 2013. Informativity and analogy in English compound stress. *Word Structure* 6(2). 129–155.

Berg, Thomas. 2012. The cohesiveness of English and German compounds. *The Mental Lexicon* 7(1). 1–31.

Boersma, Paul & David Weenink. 2016. *Praat: doing phonetics by computer [Computer program]: Version 6.0.21*. http://www.praat.org/ (accessed 20 October 2016).
Bolinger, Dwight. 1958. A theory of pitch accent in English. *Word* 14. 109–149.
Bolinger, Dwight. 1986. *Intonation and its parts: The melody of language*. Stanford, CA: Stanford University Press.
Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York, NY: Harper & Row.
Cutler, Anne. 2005. Word stress. In David B. Pisoni & Robert E. Remez (eds.), *The handbook of speech perception*, 264–289. Malden, MA: Blackwell.
Farnetani, Edda, Carol Taylor Torsello & Piero Cosi. 1988. English compounds versus non-compound noun phrases in discourse: An acoustic and perceptual study. *Language and Speech* 31(2). 157–180.
Giegerich, Heinz J. 2004. Compound or phrase? English noun-plus-noun constructions and the stress criterion. *English Language and Linguistics* 8(1). 1–24.
Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
Hall, D. Geoffrey & Catherine E. Moore. 1997. Red bluebirds and black greenflies: Preschoolers' understanding of the semantics of adjectives and count nouns. *Journal of Experimental Child Psychology* 67(2). 236–267.
Härtl, Holden. 2016. Normality at the boundary between word-formation and syntax. *Linguistische Berichte*. Sonderheft 22. 71–98.
Kunter, Gero. 2011. *Compound stress in English: The phonetics and phonology of prosodic prominence* (Linguistische Arbeiten 539). Berlin: De Gruyter.
Ladd, D. Robert. 1984. English compound stress. In Dafydd Gibbon & Helmut Richter (eds.), *Intonation, accent and rhythm: Studies in discourse phonology* (Research in Text Theory 8), 253–266. Berlin: Walter de Gruyter.
Ladd, D. Robert & Anne Cutler. 1983. Introduction: Models and measurements in the study of prosody. In D. Robert Ladd & Anne Cutler (eds.), *Models and measurements in the study of prosody* (Springer Series in Language and Communication 14), 1–10. Berlin: Springer.
Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.
Liberman, Mark & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8(2). 249–336.
Liberman, Mark & Richard Sproat. 1992. The stress and structure of modified noun phrases in English. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical matters*, 131–181. Leland Stanford Junior University: Center for the Study of Language and Information.
McCauley, Stewart M., Arild Hestvik & Irene Vogel. 2012. Perception and bias in the processing of compound versus phrasal stress: Evidence from event-related brain potentials. *Language and Speech* 56(1). 23–44.
Morrill, Tuuli. 2012. Acoustic correlates of stress in English adjective-noun compounds. *Language and Speech* 55(2). 167–201.
Plag, Ingo. 2006. The variability of compound stress in English: Structural, semantic, and analogical factors. *English Language and Linguistics* 10(1). 143–172.
Plag, Ingo & Gero Kunter. 2010. Constituent family size and compound stress assignment in English. In Susan Olsen (ed.), *New impulses in word-formation* (Linguistische Berichte, Sonderheft 17), 349–382. Hamburg: Buske.

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.

Plag, Ingo, Gero Kunter, Sabine Lappe & Maria Braun. 2008. The role of semantics, argument structure, and lexicalization in compound stress assignment in English. *Language* 84(4). 760–794.

Plag, Ingo, Gero Kunter & Mareile Schramm. 2011. Acoustic correlates of primary and secondary stress in North American English. *Journal of Phonetics* 39(3). 362–374.

Schlechtweg, Marcel. 2018. *Memorization and the compound-phrase distinction: An investigation of complex constructions in German, French and English* (Studia Grammatica 82). Berlin: De Gruyter.

Terken, Jacques & Dirk J. Hermes. 2000. The perception of prosodic prominence. In Merle Horne (ed.), *Prosody: Theory and experiment: Studies presented to Gösta Bruce*, 89–127. Dordrecht: Kluwer Academic Publishers.

Trudgill, Peter & Jean Hannah. [5]2008. *International English: A guide to the varieties of Standard English*. London: Routledge.

Turk, Alice, Satsuki Nakai & Mariko Sugahara. 2006. Acoustic segment durations in prosodic research: A practical guide. In Stefan Sudhoff, Denisa Lenertová, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter & Johannes Schließer (eds.), *Methods in empirical prosody research* (Language, Context, and Cognition 3), 1–28. Berlin: Walter de Gruyter.

Vanderslice, Ralph & Peter Ladefoged. 1972. Binary suprasegmental features and transformational word-accentuation rules. *Language* 48. 819–838.

Vogel, Irene & Eric Raimy. 2002. The acquisition of compound vs. phrasal stress: The role of prosodic constituents. *Journal of Child Language* 29(2). 225–250.

Zwitserlood, Pienie. 1994. The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes* 9(3). 341–368.

Appendix: Remaining test sentences in the three conditions

Tab. 6: Remaining test sentences in the three conditions[2]

Condition	Test sentence
Compositionality (C)	Nicole used a white jar again, which has a nice color.
Non-compositionality (N, S)	Nicole used a white jar again, which (is called so because it) is a jar used to store sugar.
Compositionality (C)	Lucy sat on a blue stool again, which has a color she likes.
Non-compositionality (N, S)	Lucy sat on a blue stool again, which (is called so because it) is a stool used for therapies in the water.
Compositionality (C)	Sarah slept in a green hut again, which has a color she likes.
Non-compositionality (N, S)	Sarah slept in a green hut again, which (is called so because it) is a hut we find in a garden.
Compositionality (C)	Steven made a brown dough again, which has a nice color.
Non-compositionality (N, S)	Steven made a brown dough again, which (is called so because it) is a dough made of chocolate.
Compositionality (C)	Sally relaxed on a gray couch again, which has a color she likes.
Non-compositionality (N, S)	Sally relaxed on a gray couch again, which (is called so because it) is a couch made of cement.

2 The two non-compositional conditions are given in the same lines. What appears between brackets belongs to the condition with *called so because* (= S) only.