Nominal composition and the demarcation between morphology and syntax: Grammatical, variational, and cognitive factors

Katja Böer, Sven Kotowski & Holden Härtl
Universität Kassel

Abstract
We investigate nominal compounds in English and German against the background of the debate about the boundary between morphological and syntactic structure building in language. After an examination of grammatical, variational as well as functional differences between compounds and phrases, we focus on processing factors to disentangle cognitive differences between the two domains. Crucially, we report on three experimental studies, which are designed to reveal contrasts in the cognitive treatment of compounds versus phrases. At the moment, however, we can only speculate about what implications can be drawn from the results for the architecture of the language system and its different combinatorial levels.

1 Introduction
Composition in language is a central topic in the ongoing debate about the demarcation between syntax, morphology and the lexical system, respectively. While some authors have claimed a boundary to be non-existent and analyzed phrasal products on a par with morphological ones, others have argued for a modular isolation of morphological structure building and word-formation, in particular, cf. Ackema & Neeleman (2004); Härtl (2011); Lieber (1992) for discussion. In this context, for example, nominal compounds provide a valuable test ground for an investigation of the issue. For instance, certain end-stressed noun-noun compounds in English like *summer dréss, steel bridge,*1 have been theorized to be of phrasal provenance, whereas synthetic compounds like *beer drinker* are classically characterized as lexical-morphological units, cf. Giegerich (2006). Here, however, factors of regional variation play an important role as well. For example, in American English a stronger tendency for stress to be placed on the second element in noun-noun compounds (and level stress in general) has been observed and a similar dialectal influence can be stated for Northern British varieties of English, see Giegerich (2004); Plag (2006). As will become clear below, other factors further blur the theoretical distinction between compound and phrase and, thus, we are left with no clear indication as to an answer to the original question of

---

1 Prominent/stressed syllables are marked by an acute accent throughout this paper.

* The current study reports on an extended version of a paper which was originally accepted as “Stress differences in English noun-noun compounds: Regional and cognitive factors” (author: Holden Härtl) for the Anglistentag 2011. Recent investigation and cooperation have required the above modification of the title. Furthermore, we wish to thank Marco Benincasa, Friederike Kreter, Kim-Vivien Lichtlein, and Peter Schöpperle for discussion as well as technical support.
how to disentangle morphological and syntactic structure building. A promising way out of the impasse is to investigate processing aspects, which may give us an insight into the cognitive foundation of operations of morphology and syntax. We report on three experimental studies in this context, which address the question of how (supposedly) morphological vs. syntactic complexes are treated cognitively and what possible implications, if any, we can draw from this for the localization of word-formation in the language system.

2 Linguistic properties of compounds and phrases

In the following section, we recapitulate some of the central properties of compounds and phrases. Whereas classical, more or less superficial features associated with stress and meaning may tempt us to accept a sharp boundary between word-formation and syntax, theoretically based considerations call this into question. One main objection in this respect is based on a circularity argument, stating that for compounds to be morphological products one needs to presuppose a morphological level to be existent in the first place.

2.1 Morphon-syntactic properties and stress distribution

Since English is poor in inflectional morphology, the distinction of syntactic and morphological constructions is not as straightforward as it is in morphologically richer languages such as German. In German, adjective-noun constructions (AN) can be told apart by simply looking at the adjective of the constructions in question. When dealing with a compound the adjective will usually surface uninflected, i.e., as a root, in a phrase it will carry an inflectional suffix, as illustrated below (1):

(1) a. rote Beete ‘beetroot’
    b. Rotkohl ‘red cabbage’

Whereas example (1a) can doubtlessly be identified as a phrase due to the inflectional suffix -e² and example (1b) can clearly be identified as a compound since the adjective rot is missing out on inflectional affixes, this is not possible in corresponding English constructions. In English, inflection will not help to distinguish between compounds and phrases, as shown by the example in (2) below.

(2) a. green house ‘a house of green color’
    b. gréenhouse ‘a building, usually out of glass, used to grow plants’³

² Also on the basis of orthography. In German compound words are mostly written as one word.
³ We disregard the differences in orthography here, since English is known to vary in its spelling of compound words.
Following conventional analysis, the construction in (2a) is a phrase and the construction in (2b) a compound, since the latter’s internal constituents are not accessible to syntactic operations as the one-coordination, as is shown in the following examples:

(3) a. *Henry owns a green house and Mary owns a red one.
b. Henry owns a greenhouse and Mary owns a red one.

As already illustrated in example (2), the two constructions can be told apart by looking at the stress patterns they exhibit. As phrases in English carry phrasal stress, i.e. the second constituent of the construction is assigned primary stress, compounds carry lexical stress, i.e., the first constituent is the most prominent one. This fact has been captured in the so called ‘nuclear stress rule’ and the ‘compound stress rule’ by Chomsky and Halle (cf. 1968: 17). The latter rule, which was argued to be especially accurate for noun-noun-constructions, does, however, in many cases not hold true (see the examples in (4)) – in fact, in so many cases that they can hardly be treated as exceptions, cf. Plag et al. (2008):

(4) geologist-astrónomer Boston máráthon
    scholar-activist apricot crámble
    Michigan hôspital silk tíe
    summer níght aluminium foíl

Some scholars, however, have proposed to consider right-stressed constructions phrases rather than exceptionally stressed compounds (cf. Marchand, 1969; Payne & Huddleston, 2002). In many cases, however, this seems to be based on mere stipulation: If we regard minimal pairs such as Christmas cake and Christmas pudding, there is not much reason for considering the former a compound and the latter, in contrast, a phrase. Thus, stress assignment does not seem to provide a sufficient criterion to distinguish compounds from phrases (cf. Bauer 1998; Spencer 2003). A further independent criterion is desirable.

Considering synthetic compounds like bookseller, various scholars characterize them as “real” compounds; they should, therefore, never display phrasal stress, cf. Giegerich (2004). Giegerich, however, also reports on counterexamples and empirical testing against a large amount of data by Plag et al. (2007, 2008) could not verify the assumption in its entirety, but only for a subset of data. Regarding primary compounds, Giegerich (2004) analyzes the distribution of compound stress in such a way that only modifier-head compounds, such as silk shirt, represent phrases and thus carry phrasal stress. Those modifier-head compounds which carry fore-stress are lexicalized and do no longer figure as phrases in the narrow sense, cf. hándbag, dárkroom). Note, however, that this assumption allows the again curious conclusion that a compound which carries phrasal stress in one variety of English is consequently characterized as a phrase, whereas in another variety in which the same expression has lexical stress it figures as a lexicalized compound. Cross-variety differences of this type occur in
Northern British varieties of English as well as in American English, cf. Giegerich (2004); Plag (2006). While Plag uses dialectal differences in this domain as evidence against a structural account of phrasal stress compounds, Giegerich observes, for a variety of Scots, a much stronger tendency towards end-stress in complex nouns in general, as is evident in place names like Loanhéad, Gorebrídge (ibid.). According to Giegerich, this general tendency may simply disguise stress effects related to lexicalization because stress on the non-head element of a compound is not necessarily induced in Scots in the way it tends to be in Standard English. With its diachronic implication, this view is compatible with analyses which attribute stress variations in the compound domain to a general grammatical change in Modern English. Olsen (2000), for example, argues that the inherited Germanic compound pattern more and more assimilates to the final stress pattern of phrases, on a par with clause structure having deviated more and more from the verb-final pattern (ibid.: 68). Of interest in this regard is the observation that, according to Kastovsky (2006), phrasal stress in noun-noun compounds is an innovation of the Middle English period (ibid.: 253), in which, as is known, the fixation of the SVO order and the loss of the V2 constraint progressed significantly.

As becomes clear from this overview, stress does not seem to be a dependable factor to distinguish between compounds and phrases. Recent studies by Plag (2010) and others suggest that compound stress is likely to be generated by other factors such as analogy, i.e. in the lexicon, and that stress assignment is no reliable indicator for the grammatical status of a construction.

2.2 Semantic and functional aspects

Concerning semantic properties it is commonly assumed that compounds have a name giving function, whereas phrases represent descriptions:

(5)  a. PHRASE:  a white bóard
    b. COMPOUND:  a whiteboard

It is, however, possible to find numerous counterexamples to this generalization. Grüner Tee [‘green tea’] is a phrase in German, yet it is the name for a special kind of tea, whereas compounds such as terror dad are, in turn, not established names, but ad hoc creations, often used in newspaper headlines. We can thus argue that having knowledge only of what a given construction denotes, we cannot decide whether we are dealing with a compound or phrase without looking at further criteria such as in-

---

4 Remember that in German AN-phrases and compounds can easily be told apart, since the adjective in the phrase carries an inflectional suffix, whereas the first constituent of the AN-compounds is realized as the adjectival stem.

5 Note that the abovementioned name giving function of compounds is particularly clear in the example of terror dad, as it is semantically highly intransparent – thus, the source it is taken from does not aim at evoking a reading such as a dad who is a terrorist or who terrorizes, but the interpretation of dad of a terrorist; cf. Maddux & Huebert (2011).
flection, orthography or stress. But do we have to make this decision at all or could we just treat these structures equivalently as they do not appear to fulfill distinct purposes? This solution would certainly provide an easy way out of the dilemma, but it does not do justice to the structural as well as functional differences between the two constructions. Consider the following examples:

(6) a. a drunk driver and a melancholic one
b. *a truck driver and an Audi one

As can be seen in the sentences above, synthetic compounds like truck driver display lexical integrity, i.e., constituents of a complex unit are not accessible to syntactic operations (cf., among others, Booij 2009; Giegerich 2006). This is true for German⁶ as well as English constructions. This is not true for phrases, in which both constituents are visible to syntactic operations, cf. the examples in (3) above also.

Furthermore, indefinite compound nouns are compatible with a kind reading, which does not hold for phrases, as can be seen in the examples below, see Härtl et al. (2011) for details:

(7) a. ?A bottle of beer is green in Germany.
   b. A beer bottle is green in Germany.

It is also striking that compounds tend to have a specialized meaning, i.e., they cannot be semantically interpreted as freely as phrases. This does not mean that they cannot display ambiguity, but that their meaning is more restricted than that of phrases: For example, a phrase like a sweet tákler can receive intersective and non-intersective interpretations, whereas with the corresponding compound a non-intersective interpretation is preferred. Sentence (8a) can be read in two different ways, an intersective reading, in which Max is a talker as well as sweet, or in a non-intersective one, in which Max is somebody who “talks sweetly” (i.e., in order to achieve a certain goal), where sweet functions adverb-like, modifying the act of talking rather than describing a property of the talker. This ambiguity is not present in (8b), which only allows one interpretation, namely the non-intersective one:

(8) a. PHRASE: Max is a sweet tálker.
    b. COMPOUND: Max is a swéet talker.

As just illustrated, a number of differences between compounds and phrases can be observed, yet it has been argued repeatedly in the literature that these cannot be considered precise criteria to distinguish the two constructions from one another – which could be evidence for the assumption that they are not different after all. Another problem is that most of the arguments brought forward by many scholars are circular: In

---

⁶ Compare the following examples also:
(i) Mia drives a Vauxhall. It now has engine failure.
(ii) Mia is an Audi driver. *It now has engine failure.
order to maintain that compounds are not accessible for syntactic operations, we have at first to accept the distinction between morphology and syntax as a premise, cf. Haspelmath (2011). Alternatively, one could also argue phonological factors to be responsible for the apparent grammatical differences between compounds and phrases. In this manner, for example, Kremers (2011) argues for apparent differences between complex words and phrases to be superficial and explainable merely on the basis of the interaction of a global syntactic module with the phonological level of language.

We suggest that a potential way out of this theoretical impasse is to investigate processing aspects related to the cognitive status of the different constructions. In the following section we report on three experimental studies, whose results can be interpreted to support the distinction between distinct structural levels of grammar, whether one wishes to label them “morphology” and “syntax” or not.

3 Cognitive properties of compounds and phrases

3.1 Experiment 1: Memorization study

Building on the hypothetical linguistic differences between phrases and compounds, as described above, a learning study was designed in order to look into possible differences from a cognitive perspective, cf. Schöpperle & Härtl (2011). As Wunderlich (2008) remarks with regard to frequency effects, semantically non-transparent morphological products – a quality which has to be ascribed to the better part of all compounds – can often be processed faster than transparent syntactic expressions as long as they occur frequently enough (cf. ibid.: 252). This observation may well relate to the naming function or quality of compounds as opposed to the, at large, mere descriptive quality of syntactic expressions (cf. Bücking, 2009). Discrepancies of this kind were therefore hypothesized to also be at work in the first part of a series of experiments on processing differences between compounds and phrases.

In the first phase of this study, the learning phase, subjects were asked to memorize pictures of everyday items and objects, such as a saw, an arm, or a comb. The pictures were each labeled with a German adjective-noun expression, either a novel, i.e. unknown AN-compound or an AN-phrase (see Fig. 1). In a recall phase immediately following the learning phase, subjects were then asked to decide whether the picture-label combinations they were presented visually were correct, i.e. matched the learned combinations, or incorrect, i.e. differed from the learned material (see Fig. 2). Besides the accuracy of the answer – being in accordance with the learned picture-label combinations – the response variable in the recall phase was the subjects’ reaction times needed to decide on the correctness of the presented combinations. Crucially, the entire procedure was repeated over three days for all subjects in order to create a suitable testbed for the hypothesized processing differences between compounds and phrases.
Twenty-five test subjects participated in the experiment, all of whom were between 20 and 30 years of age and German native speakers. Subjects were paid 40€ each for completing the whole experiment. The test items to be learned consisted of 12 “prototypical” pictures of common objects taken from the Snodgrass & Vanderwart corpus, each labeled with a combination of an indefinite article (*ein* or *eine*), a dimensional adjective (e.g. *hoch* ['high'], *weit* ['broad'], *lang* ['long'] etc.), and the appropriate head noun in either an AN-compound (e.g. *ein Weitmesser* ['a broad knife’ COMPOUND]) or an AN-phrase combination (e.g. *eine hohe Axt* ['a high axe’ PHRASE]) (6 items each). All items had been tested and balanced for degrees of lexicalization. The experimental input and instructions were in German and test items were presented visually only – due to typical German compound spelling of compounds and separate spelling of AN-phrases, the distinction between the two conditions remained clear-cut throughout. Furthermore, all adjective-noun combinations were designed to feature three syllables to keep the size of the material to be learned constant (compare e.g. *Tief_be_sen*⁷ and *ho_he Axt* above), which was necessary due to inflecting adjectives in German (see above). Subsequent to a training run designed to familiarize them with the procedure, subjects were asked to memorize the presented picture-label combinations. After the learning phase, subjects had to decide on the correctness of a total of 24 presented combinations in pseudo-randomized order, 12 of which matched the learned material, while the other 12 did not. Reaction times were recorded for this decision-making task.

---

⁷ An underscore here indicates syllable boundaries within a word.
The procedure was repeated for all subjects three times with two day breaks in between sessions, i.e. over three days on days 1, 4, and 8.

There are several main effects to be observed in the results, two of which are rather unsurprising. First, learned items were decided faster than unlearned ones ($p < .001$), while, second, test subject performance got better over the course of the three sessions ($p < .001$). These two findings can basically be taken as the legitimization of the experiment as a whole, as they show that there is a medium-term memorization effect in the first place. Besides, phrases were overall decided faster than compounds ($p < .01$), again an expected effect ascribed to the assumed markedness and semantic intransparency of novel compounds.

Neither item type was memorized better over time ($p < .26$), which opposes the above stated assumption for phrases and compounds to be treated differently from a cognitive viewpoint. However, an interesting finding can be extricated from comparing learned and unlearned items in the overall time window, i.e. if a certain picture-label combination had been presented in a learning phase or not, with the distinction between AN-compounds and AN-phrases, see Fig. 3. While neither item type was memorized better over time, we observed the following memorization effect when we analyzed the overall results from all three days (The statistical significance of the interaction for LEARNED × ITEM TYPE was at $p < .09$.): Not learned compounds took, taken together, longer to decide than phrases. This effect is highly significant ($p < .001$). Crucially, this difference disappears with learned compounds: Learned compounds were processed just as fast as phrases ($p < .67$). We interpret this result to be an indication of a stronger memorization effect for novel compounds: While novel compounds are difficult to process, unknown phrases are not, but compounds pronouncedly gain in processability over time as soon as they are learned, such that they can become accessible just as effortlessly as phrases, see Fig. 3.

This memorization effect is also reflected in the error rates, where a comparison of the error numbers for learned compounds as opposed to not learned ones (interaction of LEARNED × ITEM TYPE: $p < .001$) with those for learned and not learned phrases reveals that compounds profited significantly from learning ($p < .05$), whereas phrases did not ($p < .75$), and that learned compounds were decided as correctly as phrases ($p < .99$), see Fig. 4.
Summarizing the major findings of this study, we can say that it took subjects longer to decide upon not learned compounds than upon not learned phrases. This difference between syntactic and morphological expressions, however, disappeared when compounds and phrases had been learned – subjects then performed equally well for compounds as for phrases. Moreover, it can be stated that while there is a clear processing difference between unlearned and learned phrases in terms of reaction times, the difference in this respect between unlearned and learned compounds is significantly larger. Error rates for unlearned compounds were also significantly higher than for unlearned phrases, as opposed to learned items of either type, on which subject performance did not show any discrepancy.

The data at hand clearly suggests processing differences between novel AN-compounds and phrases, which may be of structural and/or semantic provenance. Hence, the study supports the lexicalist view of module-based approaches to the build-up of the language faculty, which assume a separation of syntax and morphology (cf. also Clahsen & Almazan, 2001; Mondini et al., 2002; Wunderlich, 2008 for further implications of the modularity assumption). The reason that novel AN-compounds are more difficult to process than AN-phrases can probably be found in the former’s linguistic markedness. This interpretation is supported by the compounds’ apparent tendency to encode specialized meanings (see above), as can for example be read off from their compatibility with sogenannte-contexts (cf. Bücking, 2009; Schlücker & Hünig, 2009). Memorization, i.e. learning of such compounds, however, evens out this markedness effect, up to the point at which there are no processing differences between learned expressions of either type. The stronger memorization effect can in principle be accounted for by the poor “starting point” of novel compounds in terms of processing performance – starting out as semantically intransparent expressions, they quickly catch up regarding their processability with phrases, which appear to primarily serve a descriptive function. This quality of compounds indicates that they may well
be the prime suggestions and thus be better candidates for lexicalization (cf., among others, Motsch, 2004).

3.2 Experiment 2: Discourse salience

The study reported on in the previous section suggested a stronger memorization effect for novel compounds like Weitmesser in comparison to phrases like weites Messer. Possibly, this difference can be ascribed to the higher degree of linguistic markedness of novel compounds, which may be a distinctive property of AN-compounds and perhaps morphological products (in German) in general. Against this background, a further hypothesis was formulated which upholds that novel compounds display a higher discourse salience in comparison to their phrasal counterparts. To test this we conducted a questionnaire study in German as a pilot, in which we utilized psychological predicates of the following type, cf. Härtl et al. (2011):

(9) The encyclopedia fascinates the student because […]

From a processing vantage point, (causative) psych-verbs like fascinate, frighten etc. as well as (stative) psych-verbs like appreciate, fear – as instances of verbs of implicit causality – have been argued to trigger a strong bias as to which participant role causal attributes are assigned to, cf., among others, Brown & Fish (1983); Härtl (2008). Typically, a sentence like in (9) creates a certain expectation for a because-clause to be continued with an assertion about the encyclopedia, i.e. the stimulus role, and less so about the student, i.e. the experiencer. In (10) causal attribution is indicated by the choice of the corresponding pronoun:

(10) The encyclopedia fascinates the student because […]

✓ it has a certain property.
✗ he has a certain property.

We employed this property of psych-verbs in our study with the hypothesis that novel compounds in the stimulus role elicit an increase in the assignment of causal attributes in comparison to the analogous phrasal complexes. For the study, the compound/phrase material identical to Experiment 1 was used (24 critical sentences: 12 fascinate- and 12 appreciate-verbs each combined with 6 compounds and 6 phrases, and 24 filler sentences were added), as is illustrated in (11):

(11) a. Die flache Säge begeistert Christoph, weil [sie | er] …
   ‘the flat saw fascinates Christoph because [she | he]’

b. Johanna schätzt das Schmalmesser, weil [sie | es] …
   ‘Johanna appreciates the slim-knife because [she | it]’

---

8 Note that Säge (‘saw’) is grammatically feminine in German.
Participants were instructed to complete the sentences using one of the two given pronouns, thus indicating the causal attribution to either stimulus or experiencer. As for the results, first, the statistical analysis indicates (as anticipated) a highly significant main effect for implicit verb causality such that more often pronouns referring to the stimulus were used for completion than those signifying the experiencer ($p < .001$). More importantly, we detected a significant effect for compounds: The probability for the because-sentence to be assigned to the stimulus was significantly higher when the stimulus was a compound ($p < .09$), i.e. in the configuration in (11b). Thus, we are led to accept the above hypothesis that novel compounds exhibit a higher discourse salience in comparison to their phrasal counterparts. To further examine this matter and consolidate it with online processing data, we have designed a third experiment. In this study, we again use sentences as displayed in (11) and examine the behavioral interaction between the occurrence of the stimulus as either compound and phrase, the semantic markedness of the items involved and (self-paced) reading time.

### 3.3 Experiment 3: Discourse salience and reading time

The third experiment is designed to detect differences in discourse salience as well as differences in processing of marked and unmarked phrases and compounds, respectively. In experiment 2 we found an increased tendency to assign causal attributes to a stimulus when the stimulus position is filled by a novel compound word. The result is, however, only marginally significant ($p < 0.9$). The question which arises is whether this effect is due to the structural difference between compounds and phrases or whether the semantic opacity of the novel compounds causes the effect. To disentangle these different linguistic aspects, we designed a third experiment whose results should clearly indicate whether we are dealing with an effect caused by structural or semantic differences.

In this reading time study, participants are asked to read sentences (self-paced reading) and to answer a content question afterwards. All participants are undergraduate students, their native language is German and their ages range between 20 and 25. Four different critical conditions are tested in this experiment. The subject position of the clause is either filled with a novel compound, a semantically deviant, i.e. marked, phrase, an established compound, i.e. unmarked, or an unmarked phrase (all AN). This will enable us to compare the results for marked and unmarked structures as well as the results for the differing structures.

The psych-verbs chosen for this experiment again bias causal attributes to be assigned to the object of the clause (see above), but the sentence will be completed in the unex-

---

9 Note that in this design attributions are not always unambiguous: For example, a sentence (i.e. its completion) such as The saw enrages Tom because he constantly cuts himself with it, although pronouncing the experiencer role, may well be argued to express a causal attribution to the stimulus. Nevertheless, answers like these were counted as experiencer attributions in the analysis.
pected way, namely with a pronoun referring to the subject, i.e. the experiencer role. An example is given below:

(12) Der junge Student bewundert Maria, weil er ein wenig naiv ist.
    ‘the young student admires Maria because he is a bit naïve’

If structural differences are the main cause for the heightened discourse salience detected in experiment 2, we expect that compounds – novel or established – cause a decrease in reading times of the pronoun in the subordinate clause. Should, however, the semantic opacity be the cause for the heightened discourse salience, we expect novel compounds and deviant phrases to be on a par in eliciting shorter reading times. It can, however, well be the case that a combination of the two factors will prove to be responsible for a heightened discourse salience. In this case it will be interesting to see which structure displays the highest discourse salience and which one the lowest, which will possibly allow for conclusions about the interaction of the different factors.

4 Conclusion

This paper aims at contributing to the discussion on the modularization of language. We argued that there are no precise criteria to theoretically distinguish between morphological and syntactic products without a priori assuming a separation of these modules. In order to circumvent this problem we introduced three psycholinguistic experiments which examine possible differences between compounds and phrases concerning their cognitive status. We found empirical evidence which is in large parts compatible with the assumption that we are dealing with two different modules of grammar as is stated in lexicalist approaches towards modeling the language system. Alternative explanations, which support non-lexicalist accounts of the reported effects, require further investigation.

References


Härtl, Holden; Katja Böer; Sven Kotowski & Marco Benincasa (2011). *Nominal Composition and the Demarcation Between Morphology and Syntax*. Talk at Universität Göttingen English Department.


Mondini, Sara; Gonia Jarema; Claudio Luzzatti; Cristina Burani & Carlo Semenzai (2002). Why is Red Cross Different from Yellow Cross? A Neuropsychological Study of Noun-Adjective Agreement within Italian Compounds. *Brain and Language* 81, 621-634.


