Interfixation in German compounds: What factors govern acceptability judgements?

Gary Libben, Monika Boniecki, Marlies Martha, Karin Mittermann, Katharina Korecky-Kröll & Wolfgang U. Dressler

Compounding in German is an extremely productive word formation process. While the majority of German biconstituent compounds are formed through the concatenation of unsuffixed roots, many left constituents in German require the presence of an interfix. This brings structural complexity into what might otherwise be considered a relatively simple morphological structure. For the noun-noun German compounds that we examine, at least five interfixation patterns are discernable. Our goal was to investigate which factors may be at play in native speakers' determinations of whether a given compound contains the correct interfixation pattern. We employed a well-formedness judgement task in which participants were required to decide, as quickly as possible, whether the compound presented on the computer screen was a well-formed German compound. Our findings reveal that, among the five interfixation patterns, uninterfixed structures, and those with the *-n-* interfix are the most likely to be judged acceptable. Compound frequency as well as the positional family size of the initial constituent and the number of compounds sharing the exact interfixation pattern for that constituent also affected judgements.

1. Introduction

One of the key characteristics of compounding as a word formation process is the extent to which it is relatively free from morphological well-formedness constraints. In English, for example, virtually any open-class monomorphemic word can combine with any other to create a new biconstituent compound word. Of course, this ability to create new compounds is not fully exploited in the language. Not all potential compound words (e.g., *planetbrick, floorbottom, keypeach*) can be given useful interpretations at specific times in specific cultures. Yet, all such compounds are morphologically acceptable and there is at least the possibility that they will be attested in some culture using the English language at some future time. Nor is this freedom in compound word formation particularly associated with English. Indeed, cross-linguistically, it is very likely that this characteristic morphological freedom associated with compounding is at the core of its prevalence and productivity across the world's languages (cf. Libben & Jarema 2006).

Rivista di Linguistica 21.1 (2009), pp. 149-180 *(ricevuto nel marzo 2009)*

It is important to note that although, cross-linguistically, compounding is relatively unconstrained morphologically, it is more constrained in some languages. In German, which is the focus of this report, compounding shows a rather complex set of well-formedness constraints associated with its patterns of interfixation (cf. Ortner & Müller-Bollhagen 1991, Fuhrhop 1998).

Interfixes (also called linking elements in English and Fugenmorpheme in German) refer to the phonetic material sometimes found in compound words at the constituent boundary. In Greek and Polish, for example, all compounds must have the segment *-o-* between the two constituents. Dutch has a more complex system (Booij 2002) in which interfixation is not present in all compounds and has two interfix forms. German shows even greater interfixation complexity, with a larger number of interfix forms (Fleischer 1976, Ortner & Müller-Bollhagen 1991, Fuhrhop 1998).

The majority of German compounds are formed without interfixes. The compound *Autobahn* ('highway'), for example, is formed through the straightforward combination of the noun constituents *Auto* + *Bahn* ('automobile'+'track'). Similarly, the constituents *Hand* + *Schuh* ('hand'+'shoe') are combined to yield the uninterfixed compound *Handschuh* ('glove').

German allows four main types of interfixes. They are most often analyzed as attaching primarily to the initial constituent of biconstituent compounds. Nevertheless, to highlight the role that they play between the two constituent, we will use the double-hyphen notation *-x-* for the distinct interfix forms. These are: *-e-*, *-(e)n-*, *-er-*, and *-s-*. Some examples of the interfix patterns found in German are shown in Table 1.

Table 1. General interfixation patterns in German

CONSTITUENT 1	INTERFIX	CONSTITUENT 2	COMPOUND WORD	GLOSS
Puppe	-(e)n-	Theater	Puppentheater	Puppet theatre
Kind	-er-	Garten	Kindergarten	Kindergarten
Weg	-e-	Netz	Wegenetz	Path network
Land	-(e)s-	Kunde	Landeskunde	Geography
Hand	-----	Schuh	Handschuh	Glove
Sprache	truncate-	Familie	Sprachfamilie	Language family

As can be seen in Table 1, there is considerable variation in how and whether a German compound is interfixed. Yet, native speakers of German show very little difficulty in choosing the correct interfixation pattern. Indeed, Stark et al. (submitted) have found that the ability to employ the correct interfixation pattern in production is surprisingly well preserved among native speakers of German with aphasia. Yet very little is known about what information is employed by native speakers in their determination of correct interfixation patterns and whether the ability to distinguish between correct and incorrect interfixation forms is equivalent across types of interfixation. These issues are at the center of our investigation.

We might begin with what seems to have already been established: The choice of the correct interfix is governed by the properties of the initial compound constituent. But what factors determine which interfixation patterns go with which initial constituents? As has been discussed by Dressler et al. (2001), very good predictions about the choice of interfixation form can often be made on the basis of the initial constituent's grammatical and phonological properties. For example, a feminine noun ending in schwa will almost always take the *-n-* interfix. For other patterns, however, the interfix pattern can be simply associated with a particular lexical stem. Such lexically based patterns can even run counter to both morphological and semantic factors. It is quite natural, for example, for native speakers of German to suppose that a compound such as *Zukunftsangst* ('future'+ *-s-* + ' fear' = 'fear of the future') takes the *-s-* interfix because this interfix is homophonous with the genitive suffix *-s* because the compound appears to have the genitive meaning 'fear <u>of</u> the future'. But, in fact, this could not be the case. The genitive suffix *-s* only attaches to masculine and neuter nouns. The noun *Zukunft,* however*,* is feminine. It appears, then, that to a large extent, interfixation choice in German is lexically specified, including in this case the word-final shape of the first member.

We must add a further caveat, however, to this claim of lexical specificity. It cannot be the case that interfixation choice is always specified by the characteristics of the initial constituent because of the simple fact that some German compounds can have alternative interfixation forms. For example, the compound *Sprachlabor* ('language laboratory') in German is formed from the constituents *Sprache* ('language') and *Labor* ('laboratory'). To form the compound, the final *e* in *Sprache* is deleted, so that the full compound is composed of the truncated form *Sprach* and the full form *Labor*. German also attests an alternative form *Sprachenlabor*, which contains the interfix *-n-*.

Similarly, German shows two compound forms for 'beef' (which are regionally defined). The first is formed from the constituent *Rind* ('beef') and *Fleisch* ('meat') to create the compound *Rindfleisch* ('beef'). The alternative form *Rinderfleisch* contains the interfix *-er-*.

Such variation is relatively rare in the language and seems to create challenges for native speakers of German. This was reported by Dressler et al. (2001) who used a partial nonsense word paradigm to investigate the extent to which native speakers of German know which interfixation patterns are associated with particular initial compound constituents. Participants were presented with a 'compound puzzle' which consisted of a real initial constituent and a final nonsense word constituent. Their task was to form a compound word from these constituents. In general, participants supplied the appropriate interfixes when required, with an overall accuracy rate of 82%. They were considerably less consistent and less accurate, however, when the initial constituent could have more than one interfixation pattern (mean accuracy = 46). Libben et al. (2002) incorporated this interfixation pattern variability into their analysis of German compound production and recognition data. They found that interfixation in German has a processing cost associated with it, such that compounds without interfixes were both easier to recognize and easier to produce. Across types of interfixation, as well, there were differences in processing ease. Finally, Libben et al. (2002) found that when initial compound constituents were associated with more than one interfixation pattern, production performance was decreased. The effects of this interfixation variability were not categorical, but rather were scalar – it made a difference how many different compounds existed in the language with a particular initial constituent and whether the interfixation pattern that was presented to participants in a particular stimulus constituted the more common or the less common pattern. Thus, for example, a compound such as *Feuersbrunst* (Feuer + -s- + Brunst; 'fire' + 'blaze') was more difficult to produce because although many German compounds begin with the constituent *Feuer* ('fire'), most of them are uninterfixed. In other words, the stimulus form *Feuersbrunst,* shows the minority interfixation pattern.

These findings form the background to our investigation. In the studies cited above, interfixation pattern acceptability was examined only indirectly and inferred from recognition and production latencies. Our goal was to examine interfixation pattern acceptability explicitly and more systematically. In so doing, we sought to determine the extent to which native speakers of German are able to distinguish between acceptable and unacceptable interfixation patterns, the extent to which this ability might differ across types of interfixes that are defined in terms of differences in phonological form and productivity, and, finally, the extent to which putatively nongrammatical factors such as frequency and family size play a role in acceptability judgements.

Our approach to the investigation of these factors involved the use on an online metalinguistic judgement task that targeted interfixation choices. Participants were presented with German compounds with both appropriate and inappropriate interfixation patterns. Their task was to judge the grammatical acceptability of the forms as quickly as possible, thus allowing us to analyze both the accuracy and latency of their responses. The compound stimuli set that was presented to participants was composed of five groups of acceptable interfixation patterns and five groups of compounds which contain inappropriate interfixation forms. These are groups are summarized in Table 2.

Table 2. The five interfixation interfixation patterns employed in the experiment

Group	Interfix	Acceptable form	Unacceptable form	Change that creates unacceptability
1.	no interfix	Zahnspange 'tooth'+'clasp'	Sternkoch 'star'+'cook'	Remove -e-
2.	-n-	Puppentheater 'puppet'+'theatre'	Hundenzucht 'dog'+'breeding'	Add -n- to -e-
3.	-e-	Wegenetz 'path'+'network'	Liedeabend 'song'+'evening'	Add -e-
4.	-er-	Bildergalerie 'picture'+'gallery'	Pferderkoppel 'horse'+'pasture'	Add -er-
5.	truncation	Farbpalette 'colour'+'pallette'	Blumebett 'flower'+'bed'	no interfix -n-

As can be seen in Table 2, our stimulus groups do not contain the *-s-* interfixed forms that are found in German. The reason for this is that we sought, in our study, to eliminate the potentially problematic factor of whether the interfix is interpreted as though it were an inflectional suffix. Clahsen et al. (1996) have argued that the interfixes can be best analyzed as inflectional suffixes. However, as Dressler et al. (2001:206) have pointed out, interfixes have representations and functions that are distinct from those of inflectional suffixes. Nevertheless interfixed first compound elements that are homophonous with plurals of the corresponding autonomous words can be assigned plural meaning (Dressler & Merlini Barbaresi 1991, Fuhrhop 1998:192f, Dressler et al. 2001:210, cf. Schreuder et al. 1998 for Dutch). Moreover certain neologisms are clearly coined with a plural meaning in mind, e.g. the Austrian *Sprache-n-dienst* ('language'+'service' = 'service for (foreign) languages'). However, the great majority of interfixed first compound elements which are homophonous with plural forms, have definitely no plural meaning, e.g. *Garage-n-besitzer* ('garage + 'owner').

In order to eliminate the potential interference of the intervening variable of plural meaning of the first compound element, we systematically limited our selection of test items to compounds whose first element normally has a plural meaning in its concept structure, independent of the form of the first element. This restriction enhanced the semantic homogeneity of the compound stimulus set. It also required that the *-s-* interfix, which is homophonous with the genitive suffix for neuter and masculine nouns not be used.

Against the background of our efforts in homogenizing the characteristics of our stimulus set in terms of the plural meaning of the first constituent, the phonological characteristics of compounds, their lexical category (all noun + noun), and their relative frequency, there remains considerable variation within each of the compound types. The uninterfixed compounds, for example, differ in terms of how their plurals are formed outside of compound structures. Some are formed with the suffix *-e*, others are formed with the suffix *-er*. Because plurality was a central component of our stimulus characteristics, we created two groups of uninterfixed compounds, so that this potential difference could be analyzed more systematically. Accordingly, as will be discussed in the Method section below, uninterfixed compounds comprise two groups, 1a and 1b.

Related differences associated with plurality (or at least plural homophony) are also to be found within the categories of *-e-* interfixed and *-er-* interfixed compounds. Here the issue is umlaut. For some nouns, *-e-* is also associated with umlaut (e.g., *Gast* → *Gäste*; 'guest/guests'). Others do not involve umlaut. The same variation is found among *-er-* interfixed nouns (e.g., *Land, Länder*; 'land/lands'). Our goal was to examine whether any of these category internal differences play a role in judgement choices or latencies.

A final source of variation with the compound categories concerns matters of lexical statistics and constitutes a core aspect of our investigation. As we have noted at various points above, native speakers of German seem to show very accurate and robust knowledge of which interfixes go with which initial compound constituents. Because this knowledge could have only been acquired from experience, it seemed likely that variation is associated with estimates of that experience. Thus, despite the fact that participants in this study are asked to judge compounds solely in terms of their well-formedness, it is quite possible that those judgements are influenced by their experience with the compound word (its whole-word frequency) and their experience with the initial constituent (its whole-word frequency). A related and intriguing possibility is that judgements are influenced by experiential factors related to a compound's structure. Such factors would include the first constituent's positional family size (i.e., the number of biconstituent different compounds in which it is an initial constituent), the extent to which there is variation in the interfixation patterns that the initial constituent can assume, and the extent to which the pattern actually presented in the stimulus compound constitutes the minority or majority case. A consideration of these lexical statistical variables is crucial to our ability to examine the psychological dynamics that may be at work in the well-formedness judgement task. If, for example, real word judgements are made primarily on the basis of the stimulus word as a whole, then whole-word frequency should play an important role. In such a case, we might assume that participants answer "yes" to the questions "Is this compound grammatically well-formed" simply when they know that they have seen or heard it before. Presumably, under this interpretation of the task, the more often participants would have seen or heard the stimulus before, the more consistent and quick their responses would be.

On the other hand, if participants focus their judgements not on the compound as a whole, but rather on the relationship between a particular initial constituent and a particular interfix, then we should expect that the dominant lexical statistical variable determining response patterns would be the positional family size of the initial constituent and its interfixation consistency. The reason for this is that the participant would then be interpreting the task as: "How sure am I that this interfix should go with this initial compound constituent?"

Finally, to relate these issues of task interpretation back to our category-based interfixation groups, it is also possible that although speakers of German relate individual constituents to individual interfixes, this results in generalizations about which interfixes attach to which types of constituents, and to generalizations about which interfixes are more productive. Our data analysis was designed to target the influence of both these category-based and item-based variables on decision choices and latencies.

2. Method

2.1. Participants

Twenty-two students from the University of Vienna (9 female, 13 male) participated as volunteers in the experiment. All participants were monolingual native speakers of German with a mean of 24 years of age (ranging from 19 – 31). They were students in disciplines other than linguistics or languages with an average of nine semesters of university education (range: 3-20).

2.2. Materials

The stimulus set was composed of 144 German compound nouns. These were divided into five major categories of interfixation. These groups are described and ordered below in terms of our expectations of their relative acceptability.

2.2.1. Group 1 real words

This group captures the simplest means by which German compounds are constructed, namely, by the simple concatenation of monomorphemic words. As Krott et al. (2007) note, this is also the most common form of compounding in German. Also, when there are competing patterns of interfixation for an initial constituent, it is most often the case that an interfixed form competes with this uninterfixed pattern (rather than another interfix type). Both Dressler et al. (2001) and Libben et al. (2002) have found that these uninterfixed compounds are the easiest to process. As is the case for all compounds in the stimulus set, there is plural meaning associated with the initial constituent. Because, as has been noted above, these uninterfixed compounds differ in terms of the suffix that would be used to form their plurals, we constructed two sub-lists (of 12 stimuli each). Group 1a contains compounds for which the plural is formed with *-e* suffixation. Group 1b contains compounds for which the plural is formed with *-er* suffixation.

2.2.2. Group 1 ill-formed stimuli

These stimuli are formed from compounds whose initial constituents would have the *-e-* or *-er-* interfix. The removal of these interfixes results in the ill-formed stimuli for Group 1a and 1b respectively. Within these groups there was another subdivision that we considered to be of potential interest. This is whether the ill-formed stimulus contains an umlauted initial constituent. The removal of the *-e-* or *-er-* interfix from initial constituents such as *Gänse* ('geese') or *Räde*r ('wheels') leaves non-words as a result (i.e., *Gäns, Räd)*. In contrast, the non-interfixed forms, non-umlauted ill-formed stimuli in these groups are existing words of German.

2.2.3. Group 2 real words

Our second compound group represents the most consistent and productive forms of interfixation in German – ones which involve the attachment of the *-n-* interfix to initial constituents ending in schwa. An example of a member of this category is the compound *Puppentheater* ('Puppet theatre'). This is formed as *Puppe* + *-n-* + *Theater*. There are very few lexical exceptions to the pattern of having the *-n-* interfix follow initial constituents that end in schwa. One such exception is *Kohle+bergwerk* 'coal + mine'; but even in such cases, in Austria, *-n-* interfixation is gaining ground, and in fact *Kohle+n+bergwerk* has become the more frequent variant.

2.2.4. Group 2 ill-formed stimuli

The appropriate cases of *-n-* interfixation were contrasted with the inappropriate presence of the same interfix *-n-* in stimuli for which a "no" well-formedness judgement was expected. These inappropriate interfixation stimuli were created by taking existing compounds that have the interfix *-e-* (e.g., *Hund + e + Zucht* ('dog + breeding') and then inserting an *-n-* after the *-e-* interfix to form an incorrect compound form *Hundenzucht*. Our expectation in a manipulation such as this one was that the *-n-* interfix, being the most productive and consistent in German, would show the greatest accuracy rates and the lowest latencies (i.e., the greatest number of "yes" well-formedness decisions and the fastest response latencies. The inappropriate cases of *-n-* interfixation should, all other things being equal, also show the greatest number of "yes" responses (because of the frequency of this type of interfixation in the language). Because these inappropriate compound forms were created by altering the form of existing compounds, we also had the opportunity to examine what role the frequency of the original compound plays in decision choices and latencies, as well as the extent to which the frequency and consistency of the interfixation pattern for the initial constituents plays a role.

2.2.5. Group 3 real words

Stimuli in Group 3 all had the *-er-* interfix. Because this interfix is often associated with umlaut, we ensured that half the stimuli within the group had an umlaut form (e.g., *Häusermarkt*; 'house' + 'market') and half did not. All real-word interfixed forms are homophonous with the unproductive plural forms of the respective first elements (11 neuter, 1 masculine). Our expectation was that, due to this pattern's relatively low productivity in German, relative to the uninterfixed pattern and the *-n-* interfixation pattern, Group 3 stimuli would be less likely to be accepted than either the uninterfixed or the *-n-* interfixed compounds.

2.2.6. Group 3 ill-formed stimuli

The ill-formed stimuli were created by changing *-e-* interfixed compounds to *-er-* interfixed ones. Here too, half of the stimuli were interfixed and half were uninterfixed. We expected that the ill-formed *-er-* interfixed would be less likely to be accepted then either the uninterfixed or *-n-* interfixed forms. In other words, they should show higher decision accuracy rates.

2.2.7. Group 4 real words

These stimuli contained the unproductive *-e-* interfix. As in Group 3, initial constituents were homophonous with their free-standing plural forms. Although we hold that these forms are distinct from their homophonous plurals, it is possible that familiarity with these strings as well as the fact that all compounds have plural first constituent meaning might play a role for these stimuli.

2.2.8. Group 4 ill-formed stimuli

As in Group 3, these stimuli were created by changing initial constituents so that they were homophonous with their plural forms.

2.2.9. Group 5 real words

Compounds in this group are characterized by truncated initial constituents. An example is the compound *Sprachfamilie* ('language' + 'family'). It is formed from the nouns *Sprache* and *Familie*. Note, however, that the initial noun is truncated in the compound *Sprach* (through the deletion of schwa). This pattern of compounding was identified by Dressler et al. (2001) as particularly difficult to process and it is the category that we would consider to be the most marked. Accordingly, we expected the real-word Group 5 compounds to show the lowest rate of acceptance.

2.2.10. Group 5 ill-formed stimuli

As part of Group 5, these compounds were also truncated. This was not done, however, by reducing the root, but rather by truncating *-n-* interfixed compounds so that only the schwa-final root remained (e.g., *Tomatensalat* → *Tomatesalat;* 'tomato' + 'salad'). This method for creating ill-formed stimuli led us to the expectation that this ill-formed category would particularly easy to reject because is involves the non-application of the very common pattern of attaching the *-n-* interfix to schwa-final initial constituents, as noted in our discussion of Group 2 compounds above.

2.3. Procedure

The task presented to participants was an online morphological well-formedness judgement task. In this task, participants were presented with a German compound in the center of a computer screen and were asked to decide whether or not the stimulus string was a well-formed word of German. In total, participants were presented with 144 critical test items, of which 72 were well-formed attested German compounds and 72 were created by violating the normal interfixation patterns associated with the first constituent of the biconstituent compound, as described in the materials section above.

Each experimental session consisted of the presentation of task instructions, four practice trials, a further opportunity to ask for task clarification, the 144 experimental trials (presented in a single block of trials), and a final debriefing session. The instructions provided to participants operationalized the online morphological well-formedness judgement task. Participants were instructed as follows: "You will see single words appearing on the computer screen. Your task is to decide whether the word is correctly formed or not. The meaning of the word is not at issue, but rather whether or not it is grammatically correct. If the word is grammatically correct, press the green key. If it is not correct, press the red key. Please decide as quickly and as accurately as possible."

The choice of an explicitly grammatical decision for this task was motivated by two factors: First, because many of the core stimuli were German compounds of very low frequency, a straightforward lexical decision task would have been problematic both in its interpretation by participants (posing the problem of whether the compound had in fact been seen or heard, or whether it could be produced by a native speaker of German) and by the experimenters (creating a high probability of having well-formed as well as ill-formed compounds in the category of "no" lexical decisions). The second reason for the use of this explicitly grammatical judgement was to maximize the probability that participants' judgements would be as related as possible to the core focus of our study: patterns of interfixation. In fact, it is only through the appropriate or inappropriate choice of interfixation patterns that compounds with two open-class constituents can be deemed to be grammatically ill-formed. Our goal then was to employ a task that would be maximally sensitive to this aspect of compounding in the context of an experimental paradigm that is, behaviourally, virtually identical to simple lexical decision. Finally, our third goal in the choice of this task was to examine whether effects of interfix pattern frequency, productivity and family size would be present even under conditions in which participants were instructed to make judgements on the basis of grammaticality only.

Each trial in the experiment proceeded in the following manner: One hundred milliseconds after the onset of the trial, an asterisk in 28 point font appeared in the center of the screen for 150 milliseconds. The fixation point disappeared and the screen remained blank for 100 milliseconds after which the stimulus word appeared in the center of the screen in 24 point font. The stimulus remained on the screen until either the “yes” or “no” key was pressed. The experiment was conducted on a Macintosh laptop running Psyscope 1.02 (Cohen et al. 1993). The “yes” and “no” keys were marked on the laptop’s keyboard. The leftmost and rightmost keys of the bottom row of the keyboard were used, so that the “yes” key was pressed using the index finger of the left hand and the “no” key was pressed using the index finger of the right hand.

3. Results and Discussion

At the core of our analysis was the online grammaticality judgement of well-formed and ill-formed interfixation patterns. Of the initial 22 participants, two (one male and one female) showed extraordinarily long response latencies (in the order of 20,000 milliseconds). The data from these two participants were not analyzed because they could no longer be considered to be ‘online’.

A second treatment of the responses associated with our desire to ensure that response times were online was to set a time-out for responses at 3000 milliseconds. Thus, response latencies above 3000 milliseconds were not analyzed. This affected 191 of the 2,850 responses or 6.6%. One stimulus, *Erdbeereis*, which was in fact trimorphemic rather than bimorphemic was deleted.

The real compound stimuli were analyzed separately from those that were ill-formed, and latency analyses were only applied to correct responses (i.e., "yes" responses for the real word stimuli and "no" responses for the ill-formed stimuli). It should be noted that we use the terms "correct" and "incorrect" to refer to our *a priori* expectations concerning the judgement patterns. Because this was a judgement task, and many of the stimuli were of very low frequency, participants might show judgement patterns that varied considerably from our expectations. Generally, however, participants' response patterns accorded with our pre-classification. The percentage of "yes" responses for the compound stimuli that we classified as real words was 92%. The percentage of "no" responses for the compound stimuli that we classified as ill-formed was 86%. This accuracy difference was significant (p<.01). This analysis, as well as almost all analyses that we report below, was conducted using linear mixed effects modeling, in which both participants and items were included as random factors in the analysis (Baayen 2008).

3.1. The five compound groups

The accuracy rates for the five compound groups are presented in Table 3. As can be seen in this table, the real *-n-* interfixed compounds showed an acceptance rate of almost 100%. Ill-formed *-n-* interfixed compounds showed an acceptance rate of 20%. This pattern is consistent with our expectations based on the reasoning that *-n-* interfixation should be the most consistently accepted of all the interfix forms. This would manifest itself as high accuracy for the real forms, for which "yes" is the correct response, and low accuracy for the ill-formed stimuli, for which "yes" is the incorrect response.

Table 3. Response accuracy for each of the compound groups

	Response Accuracy			
Compound Group	Ill-formed Stimuli		Real Stimuli	
	Mean	Median	Mean	Median
Group1a-no interfix	71	85	99	100
Group 1b- no interfix	88	95	91	95
Group 2-n interfix	80	85	99	100
Group 3-er interfix	90	90	97	100
Group 4-e interfix	93	95	76	82
Group 5-truncation	92	95	88	95
Average	86	92	92	97

The data for the uninterfixed forms was less clear. Group 1a, which is composed of uninterfixed compounds whose initial constituents form their plurals with *-e*, shows accuracy patterns that are comparable to those in Group 2. This is essentially what we expected. Uninterfixed compounds are the most frequent type in German and, of the interfixed ones, *-n-* interfixation is the most common form. So, they should pattern together as the most likely to be accepted. This is not seen for Group 1b, however, which showed only 90% acceptance rates for real words and 12% acceptance rates for ill-formed stimuli. Closer inspection of this group reveals, however, that the higher values were driven by just a few stimuli. For the real words, the two outlying stimuli were two of the 12 stimuli, *Lidcreme* and *Dingwelt*. When these were eliminated from the analysis, the accuracy rates for Group 1b rose to 96%.

The compounds *Lidcreme* ('eyelid'+ 'creme') and *Dingwelt* ('thing'+' world' = 'world of things') are relatively rare for uninterfixed compounds. But, the inclusion of such relatively rare uninterfixed compounds was necessary to maintain the overall frequency balance for the entire compound stimulus set, which, on the whole, contains low-frequency German forms. Moreover, there are individual characteristics of *Lidcreme* and *Dingwelt* that could play a role in the participants' judgements. Both *Lid* and *Creme* are loanwords in German. Neither *Ding* nor *Lid* enter into compounding as initial constituents very commonly. Participants, therefore, would have had little experience with interfixation patterns associated with the initial constituents of these stimuli.

3.2. Frequency and positional family size

The comments above regarding the individual characteristics of *Lidcreme* and *Dingwelt* bring us to a more general discussion of individual item variation within the dataset. As can be seen by inspecting the accuracy rates for individual stimuli, which is provided in Appendix, many of the group mean differences are driven by extreme values within the group. Accordingly, in Table 3, median values are provided along with the mean accuracy values for each compound group. But there is another question that we must ask: Why are the values within a compound category, so diverse? Our tentative answer to this is that the pattern of experience that participants have with the stimuli and their components may play a very large role in how stimuli such as ours are perceived and judged. In addition, an exploration such as this allows us an opportunity to explore the relation of these factors to this particular type of judgement task and the processing of the compound stimuli in a more fine-grained manner.

This exploration of the role of frequency and family size effects most naturally begins with a consideration of whole-word frequency. As we have noted in the introduction to this study, it would seem reasonable to suppose that the likelihood of acceptance of a compound stimulus would be increased when it is of relatively high frequency. The more frequent a compound is, the more is might seem to be obviously well-formed. We might also expect that "yes" decision latencies for high frequency compounds would be further lowered by the processing component of the task. In order to judge a stimulus, a participant also needs to process it. Thus, "yes" latencies should be expected to include a component that relates to processing ease as well as a component that relates to decision ease. For "yes" responses to real words, both of these components would be lead to lower latencies for more frequent words.

The expectations for the role of whole-word frequency in the processing of ill-formed stimuli are less straightforward. In this case, the frequency effect would derive not from the stimulus itself, but from the real word that was changed in order to render the stimulus ill-formed. The reason that a frequency effect might be expected here is that a participant could use whole-word frequency to perform the task more effectively. When confronted with an ill-formed stimulus string, a participant might adopt a strategy of determining whether there exists another (real word) compound with those constituents. If there is a real word compound, then it is most likely that the stimulus string is not well-formed (because, as we have noted above, although alternative interfixation patterns for the same constituents do exist in German, they are very rare). The effect of frequency in this case would be involved in ease with which a participant could activate the original real-word compound. Thus, we might expect correct "no" responses to be faster for ill-formed stimuli that are formed from high frequency compounds.[1]

One might imagine that if this task requires the decomposition of stimuli into their morphological components, the frequency of those components might also play a role. We investigated this possibility as well as the possibility that the most relevant aspect of the participant's morphological experience in this task centers on the role that an initial constituent plays in compounds across the language and, in particular, as Libben et al. (2002) found, whether the participant's experience with the interfixation pattern for this initial constituent is consistent.

The results of our analyses are provided in Tables, 4, 5, 6 and 7. Table 4 presents the accuracy analyses for real words. Table 5 presents the latency analyses for real words. Tables 6 and 7 do the same for the ill-formed stimuli. Linear mixed effects modeling was used in which both participants and items are treated as random factors. For the morphological acceptability judgements in Table 4 and Table 6, a generalized linear mixed model with the logit link function and binomial variance were used.

Table 4. Logistic regression results for the accuracy scores for real-word stimuli.. Independent variables are: compound group, whole-word frequency, constituent frequency, positional family size of the first constituent, and the subfamily size of words that have an interfixation pattern that is identical to that of the stimulus

	Estimate	Std.Error	*z*-value	Pr(>|z|)	
Intercept (Gp1a-0)	10.0008	1.7612	5.6790	4.25E-08	***
Gp1b-0	-1.2599	0.8569	-1.4700	0.1415	
Gp2-n	-0.2098	0.9639	-0.2180	0.8277	
Gp3-er	-0.3111	0.9525	-0.3270	0.7440	
Gp4-e	-2.3591	0.8335	-2.8300	0.0046	**
Gp5-trunc	-3.4558	0.8287	-4.1700	3.05E-05	***
CompoundFrequencyClass	-0.3157	0.0816	-3.8680	0.0001	***
C1 Freq (CELEX)	0.0001	0.0001	0.6950	0.4872	
C2 Freq (CELEX)	-0.0004	0.0001	-2.7000	0.0070	**
C1 Pos family size	-0.0318	0.0127	-2.5000	0.0124	*
Same C1+IF Family size	0.2706	0.0870	3.1090	0.0019	**

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05
Note: The standard deviation for the by-participant and by-item random intercepts were 0.20 and 0.37 respectively.

In Table 4, Compound Group 1a is on the intercept. The estimates indicate that all other compound groups were less accurate (as indicated by the negative values). However, only Groups 4 and 5 were significantly less accurate. Whole word compound frequency facilitated accuracy (see Footnote 1 for an explanation of the Compound Frequency Class values). The positional family size of the compound initial constituent had a negative effect on accuracy. This suggests that when many compounds begin with a particular constituent, the opportunity for variation and therefore uncertainty is greater. This conclusion seems consistent with the observation that accuracy was increased by elevations in the number of compounds that have an interfixation pattern that is identical to that of the stimulus word.

An additional interesting finding shown in Table 4 concerns the role of constituent frequency. There was no effect of the frequency of the initial constituent as an independent word. This suggests what is important is not experience with the initial constituent as an independent word, but rather, experience with it as an initial compound constituent. There was, however, a small but reliable effect of second constituent frequency for these stimuli. The effect was inhibitory, suggesting that, for example, *puppet theatre* and *theatre* might compete to the detriment of performance on this task.

The analysis of response time for real words is shown in Table 5. The analysis was conducted with linear mixed effects modeling in which both participants and items were included as random factors and log-transformed response time served as the dependent variable. As can be seen in Table 5, the response time data yielded fewer significant effects as compared to the accuracy analysis. Here, only the truncated Group 5 compounds were slower than the other groups. As in the accuracy analysis, whole-word frequency had a facilitating effect, as did the number of compounds with the exact same initial constituent and interfix pattern as the stimulus. It should be noted

Table 5. Estimated coefficients, standard errors and t-values for response times to correct responses to real-word stimuli

	Estimate	Std. Error	*t* value	
Intercept (Gp1a-0)	6.36E+00	1.83E-01	34.67	
Gp1b-0	2.74E-02	6.60E-02	0.42	
Gp2-n	-2.81E-02	6.02E-02	-0.47	
Gp3-er	5.33E-02	6.93E-02	0.77	
Gp4-e	1.63E-01	7.32E-02	2.23	*
Gp5-trunc	9.59E-02	6.15E-02	1.56	
CompoundFrequencyClass	3.75E-02	9.19E-03	4.08	*
C1 Freq (CELEX)	-1.05E-05	1.78E-05	-0.59	
C2 Freq (CELEX)	2.15E-05	2.45E-05	0.88	
C1 Pos family size	3.02E-03	1.72E-03	1.75	
Same C1+IF Family size	-6.41E-03	2.80E-03	-2.29	*

Note: The standard deviation for the by-participant and by-item random intercepts were 0.15 and 0.13 respectively. The standard deviation for the residual error was 0.27.

that this analysis, as opposed to the accuracy analysis was on correct "yes" responses only.

To summarize the real word findings, it seems that all the independent variables that we expected to play a role did in fact do so in affecting judgement accuracy and latency. The compound group does make a difference, as does the frequency of the entire compound. In addition, the greater the number of compounds with the same pattern as the stimulus, the more likely participants are to judge the stimulus as well-formed and the faster their judgement.

These findings suggest that, in this task, a rather diverse set of factors are at play. Although the task is explained to participants as a metalinguistic judgement task rather than a recognition task, we see both accuracy and latency effects that are related to whole-word frequency. The family size effects that we found suggest, in addition, that there is broad activation across the lexicon while participants believe themselves to be focusing on the judgement of a single stimulus. In our analysis of responses to ill-formed stimuli, we were particularly interested in whether category-based and frequency-based effects also obtained.

In our analysis of ill-formed stimuli, we also included the factor of whether the ill-formed stimuli contained an umlaut. The reason for this is that, in the formation of ill-formed stimuli, umlauted forms were employed in half the stimuli for all groups except Group 5. The Group 5 stimuli were truncated by deleting an interfix to leave a schwa final root. An additional way in which the analysis of ill-formed stimuli differed from the analysis of real words was in the removal of the factor "number of words with the exact constituent and interfixation pattern". Ill-formed stimuli were designed with the intention of keeping the value for that factor at zero.

The results of our analysis of response accuracy are shown in Table 6. Here we see that whole-word frequency and positional family size did not have a significant effect on accuracy rates. We did, however, find a difference between the umlauted and non-umlauted stimuli, such that the presence of umlaut made stimuli less acceptable. The Group 5 stimuli, which were truncated through the deletion of the *-n-* interfix, were also less acceptable. Thus, we see additional evidence for the privileged position of the *-n-* interfix in German. Real words with this interfix are more easily judged as acceptable. Participants find its removal to be the greatest of the interfix violations that we have examined in this study.

Table 6. Logistic regression results for the accuracy scores for ill-formed stimuli. Independent variables are: compound group, whole-word frequency, constituent frequency, positional family size of the first constituent

	Estimate	Std. Error	z value	Pr(>|z|)	
Intercept (Gp1a-0, Stimulus Has Umlaut (No)	1.3669	1.8246	0.7490	0.4538	
Stimulus Has Umlaut (yes)	1.4379	0.4478	3.2110	0.0013	**
Gp1b-0	0.7199	0.6394	1.1260	0.2602	
Gp2-n	0.7895	0.6721	1.1750	0.2401	
Gp3-er	0.6444	0.6733	0.9570	0.3385	
Gp4-e	1.1412	0.6829	1.6710	0.0947	.
Gp5-trunc	2.0072	0.7113	2.8220	0.0048	**
CompoundFrequencyClass	-0.0026	0.0855	-0.0310	0.9754	
C1_CELEX_FREQ_Mann	0.0001	0.0003	0.3360	0.7365	
C2_CELEX_FREQ_Mann	0.0003	0.0004	0.6820	0.4953	

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
Note: The standard deviation for the by-participant and by-item random intercepts were 1.01 and 1.10 respectively.

Table 7. Estimated coefficients, standard errors and t-values for response times for correct responses to ill-formed stimuli

	Estimate	Std. Error	t value
Intercept (Gp1a-0, Stimulus Has Umlaut (No)	6.6550	0.1867	35.6400
Stimulus has umlaut (yes)	-0.0718	0.0442	-1.6300
Gp1b-0	-0.0292	0.0646	-0.4500
Gp2-n	0.0339	0.0727	0.4700
Gp3-er	0.0990	0.0685	1.4500
Gp4-e	-0.0534	0.0678	-0.7900
Gp5-trunc	-0.1339	0.0731	-1.8300
CompoundFrequencyClass	0.0377	0.0085	4.4600
C1 Freq (CELEX)	0.0000	0.0000	-1.0100
C2 Freq (CELEX)	0.0000	0.0000	0.7400

Note: The standard deviation for the by-participant and by-item random intercepts were 0.18 and 0.13 respectively. The SD for the residual error was 0.25.

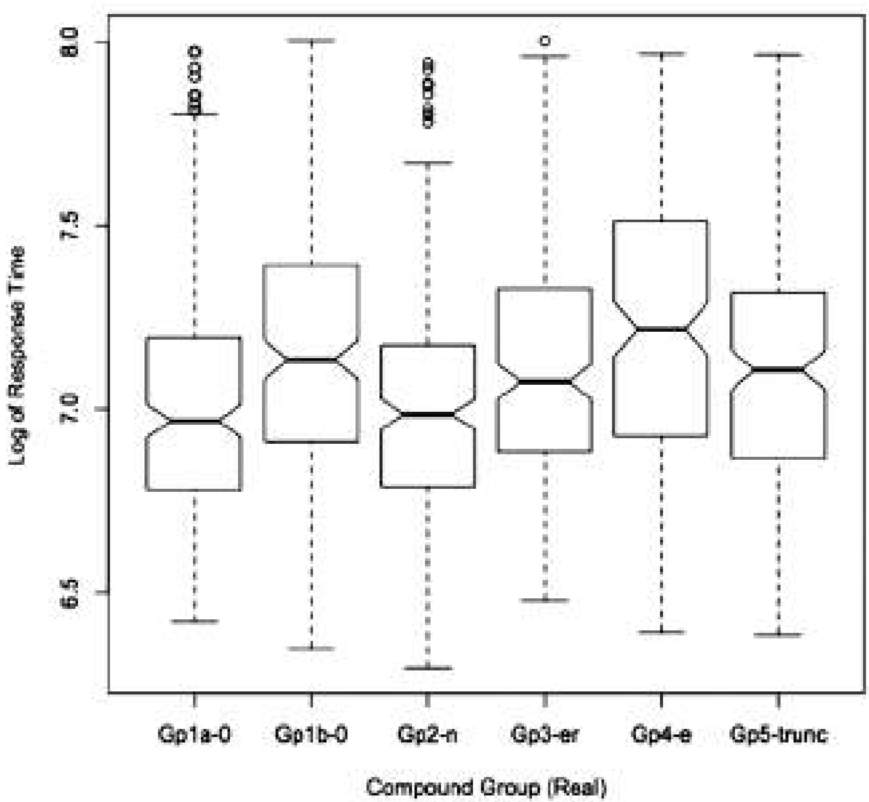

Figure 1. Log-transformed response times by compound category for the real compound stimuli. Group labels indicate the compound group and the interfix (e.g., Gp1a-0 indicates group 1a – no interfix). The bottom edge of the box represents the first quartile. The top edge of the box represents the third quartile. The vertical line within the box represents the median. The upper whisker is the value that is 1.5 times the third quartile. The dots above the upper whisker represent outliers.

Our final analysis focused on the response latency for correct "no" responses to ill-formed stimuli. As we have noted above, there is reason to suppose that whole-word frequency might play a role here, if a component of participants' judgements involves accessing the real word compound from which the ill-formed stimulus was created. The results are shown in Table 7. There was a significant effect of whole-word frequency, so the more frequent the source compound (i.e., the lower its frequency class), the faster participants were to reject the ill-formed stimulus that was derived from it. This suggests that access to

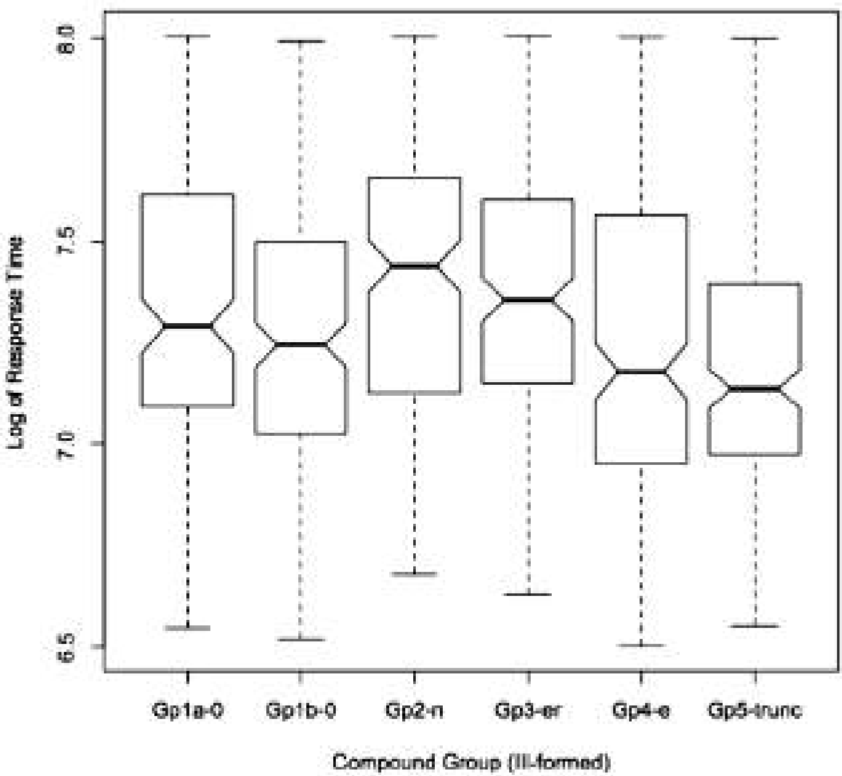

Figure 2. Log-transformed response times by compound category for the ill-formed compound stimuli. The boxplot is created following conventions described for Figure 1.

the source compound is a key component of the judgement task. There were no significant differences associated with compound group. However, as can be seen in Figure 2, the general pattern of response times is along expected lines, with Group 2 compounds being the slowest to reject, and Group 5 compounds being the fastest to reject.

4. Conclusions and implications

Now, taken together, what do these data tell us about the roles of interfixation patterns in the processing of German compounds? The general pattern of findings is presented in Table 8. As this table suggests, real German compounds with different interfixation patterns

Table 8. The overall pattern of results. A tick mark (√) indicates the presence of a significant effect. The absence of a significant effect is indicated by 'X'

	ACCURACY		LATENCY	
REAL WORDS	Comp group	√	Comp group	√
	Word freq	√	Word freq	√
	Pattern support	√	Pattern support	√
	C1-famsize	√	C1-famsize	√
	C1-freq	X	C1-freq	X
	C2-freq	√	C2-freq	X
ILL-FORMED	Comp group	√	Comp group	X
	Word freq	X	Word freq	√
	Pattern support	N/A	Pattern support	N/A
	C1-famsize	X	C1-famsize	X
	C1-freq	X	C1-freq	X
	C2-freq	X	C2-freq	X

are not judged by native speakers of German in the same manner. In both accuracy and latency measures, there is a facilitating effect of whole word frequency. On the latency side, it is possible that this effect reflects the fact that any lexical judgement task is also a lexical processing task. Lexical frequency is the most robust of all lexical processing effects (Balota & Chumbley 1985). It is therefore unsurprising that it would play a role in this task. On the accuracy side, the effect of frequency might be related to the role that it would play in a conventional lexical decision task. If participants perceive a stimulus to be a real word, they are very unlikely to also judge it as ill-formed. Thus, again, while we instructed participants to perform a type of grammaticality judgement, in fact, the task must be understood to have a substantial lexical recognition component.[2]

First constituent frequency did not play an important role in this task. This suggests that although the well-formedness of the stimulus compound is a matter of whether the first constituent and the interfixation pattern fit, it seems that it is only the form as a constituent that plays a role, not the form as a free-standing word. In one of our analyses, second constituent frequency did play a role. However, it should be noted that this role was inhibitory. This too suggests that a compound constituent, even though it is often homophonous with a free-standing morpheme, is distinct from that free standing morpheme. The claim here is that the mental lexicon contains distinct representations for compound constituents and free-standing roots. Moreover, the compound representations would be positionally defined (Libben 2008) so that we could represent the mental lexicon as having representations *Puppe* ('puppet') as a free morpheme, *Puppe-*

as an initial constituent (as in *Puppentheate*r), and *puppe* as a final constituent (as in *Handpuppe,* 'hand puppet'). It should be noted that these are not redundant representations. The processing and representational properties of all these cases are related, but distinct.

Our approach to the relationship between interfixed constituents and their homophonous forms which are inflected free-standing words is quite similar. In our view, just as *Puppe-* as an initial constituent is related to the free form *Puppe* but is distinct from it by virtue of its bound function as an initial compound constituent, the bound initial constituent *Puppen-* is distinct from the free-standing plural *Puppen*, with which it is homophonous and homographic.

Given the above, our finding of a facilitating effect of first constituent family size on acceptance rates for the real word stimuli is to be expected. The more experience that a participant has with a particular form as an initial constituent, the easier it is to perform the task correctly.

This effect of first constituent family size was not found for the ill-formed stimuli. Although our data do not allow us to test this hypothesis, we suspect that, for the ill-formed stimuli, there are both positive and negative effects of first constituent positional family size that might cancel each other out. On the positive side, more experience is better. On the negative side, the larger the positional family, the greater the possibility of variation, and the less certain one might be that the ill-formed stimulus presented is in fact not acceptable.

As we have noted above, findings such as these lead us to a view in which responses to the presentation of a particular stimulus involve not only that stimulus, but a much larger set of lexical items that share properties with the stimulus. For example, we have seen that it is not only the size of the initial constituent family that plays a role. The number of compounds within that family that have an interfix pattern that is identical to that of the stimulus also plays a facilitating role. In Table 8, we have labeled this as *Pattern support*. By definition, this factor can only be assessed for real words. There are no real-word compounds with interfix patterns that are identical to those of the ill-formed stimulus

An extremely interesting question is how such 'lexicon-wide' effects can best be understood. One way to do this is to imagine that during the task, there is an actual spread of activation to the structurally related forms, resulting in either inhibition or facilitation, depending on their relationship to the stimulus word. Another way to conceive of the effect is to imagine that, over time, related representations influence each other's representations. Under this view, the fact that many compounds begin with *Puppe*, and all of them have the form *Puppen*, would influence the functional representation for *Puppentheater*. Although our data do not speak in a direct and unequivocal manner to these alternatives, we suggest the latter view, one that has a relatively rich representation for compound constituents and their patterning. This allows for an explanatory framework within which both our constituent family size effects, whole-word frequency effects, and compound category effects can be accommodated.

Under such a view, the influence of compound groups is seen as another effect of the broader pattern within the lexicon on the representation of individual members. However, in the case of compound categories, reference is not only made to other compounds with the identical first constituent, but also to other members sharing the interfixation pattern and at least one other characteristic. Under this approach, it follows easily that *-n-* interfixation (Group 2) is advantaged. The *-n-* interfixation pattern is very productive for nouns ending in schwa, and virtually ubiquitous for feminine nouns ending in schwa. Although uninterfixed forms are perhaps the default unmarked form in German, they do not have an identifiable structural description of this sort. This may account for the heterogeneity across Groups 1a and 1b.

We would like to note, however, that our explanation for why Group 2 real-word stimuli, for example, are advantaged may be different in kind from an explanation of why Group 5 real-word stimuli, for example, might be disadvantaged. Although we have labeled this as a single factor, our view is that its role may have a number of sources. For real words, the factor 'compound group' may reflect the markedness relations and relative productivity among interfixation patterns. For ill-formed stimuli, it may also reflect such influences as well as the effects of the interfixation pattern of the source compounds from which the ill-formed stimuli were created.

Thus, we may find, as we often do in psycholinguistic inquiry, that descriptive distinctions in language patterns (in this case the patterns of interfixation among German compounds) help us to begin an investigation of the psychological factors that play a role in how the different structures are perceived and used. Our investigation has revealed that differences in interfixation patterns are associated with differences in rates of acceptance for both well-formed and ill-formed stimuli. The general pattern of interfixation (as represented by the compound groups) counts. But it is not the only thing that counts. Participants' judgements are influenced by the relative frequency of the compound to be judged; they are influenced by the frequency with which a particular initial constituent participates in compounding.

Finally, they are influenced by the frequency with which the specific pattern of interfixation that they see for a particular initial constituent is attested in other compounds in the language.

Addresses of the Authors:

Gary Libben, Department of Linguistics, University of Calgary, 2500 University Drive, N.W., Administrative Building, Executive Suite, A100, Calgary, Alberta, Canada T2N 1N4 <glibben@ucalgary.ca>

Monika Boniecki, Marlies Martha, Karin Mittermann, Department of Linguistics, University of Vienna, Berggasse 11 A, 1090 Wien, Austria

Katharina Korecky-Kröll, Wolfgang U. Dressler, Department of Linguistics, Austrian Academy of Sciences, Kegelgasse 27 A, 1030 Wien, Austria

Notes

1 We should note again at this point that compound stimuli that we used were of relatively low frequency. Thus, differentiated frequency measures could not be obtained from CELEX (Baayen et al. 1995). Instead, we used the frequency classes employed in the Wortschatz database (Quasthoff, Richter & Biemann 2006). These classes are numbered so that they are higher as frequency decreases. For example, the stimulus *Puppentheatre* has a Frequency class of 16. This indicates that the most frequent word of German (the definite article *der*) has a frequency that is 2^{16} that of *Puppentheatre*.

2 There are possible cases in which well-formedness judgement and real-word judgement would not correspond. In English, for example, many speakers would judge the word *irregardless* to be ill-formed, but would likely also judge it to be an existing English word. Our stimulus set did not contain stimuli of this sort.

Bibliographical References

Baayen R. Harald 2007. *Analyzing Linguistic Data. A Practical Introduction to Statistics*. Cambridge: Cambridge University Press.

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

Balota David A. & James I. Chumbley 1985. The locus of word-frequency effect in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language* 24. 89-106.

Booij Geert E. 2002. *The morphology of Dutch*. Oxford: Oxford University Press.

CHAMBERS John M., William S. CLEVELAND, Beat KLEINER & Paul A. TUKEY 1983. *Graphical Methods for Data Analysis.* Pacific Grove: Wadsworth & Brooks/Cole.

CLAHSEN Harald, Gary MARCUS, Susanne BARTKE & Richard WIESE 1996. Compounding and inflection in German child language. *Yearbook of Morphology* 1995. 115-142.

COHEN Jonathan, Brian MACWHINNEY, Matthew FLATT & Jefferson PROVOST 1993. PsyScope: an interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments and Computers* 25.2. 257-271.

DRESSLER Wolfgang U. 2003. Degrees of grammatical productivity in inflectional morphology. *Italian Journal of Linguistics* 15. 31-62.

DRESSLER Wolfgang U. 2007. Productivity in Word Formation. In JAREMA Gonia & Gary LIBBEN (eds.). *Core perspectives on the mental lexicon*. Oxford: Elsevier. 155-180.

DRESSLER Wolfgang U. & Mária LADÁNYI 2000. Productivity in Word Formation: a morphological approach. *Acta Linguistica Hungarica* 47. 103-144.

DRESSLER Wolfgang U., Gary LIBBEN, Gonia JAREMA, Jacqueline STARK & Christiane PONS 2005. Produttività nel processamento di composti: esempi tedeschi con e senza interfissi. In GROSSMANN Maria & Anna M. THORNTON (eds.). *La formazione delle parole. Atti del 27. congresso della Società di Linguistica Italiana*. Roma: Bulzoni. 153-162.

DRESSLER Wolfgang U., Gary LIBBEN, Jacqueline STARK, Christiane PONS & Gonia JAREMA 2001. The processing of interfixed German compounds. *Yearbook of Morphology* 1999: 185-220.

DRESSLER Wolfgang U. & Lavinia MERLINI BARBARESI 1991. Interradical interfixes: contact and contrast. In IVIR Vladimir & Damir KALOGJERA (eds.). *Languages in Contact and Contrast*. Berlin: Mouton de Gruyter. 133-145.

FLEISCHER Wolfgang 1976. *Wortbildung der deutschen Gegenwartssprache*. Leipzig: Bibliographisches Institut.

FUHRHOP Nanna 1998. *Grenzfälle morphologischer Einheiten.* Tübingen: Stauffenburg.

KROTT Andrea, Robert SCHREUDER, R. Harald BAAYEN & Wolfgang U. DRESSLER 2007. Analogical effects on linking elements in German compound words. *Language and Cognitive Processes* 22. 25-57.

LAAHA Sabine, Dorit RAVID, Katharina KORECKY-KRÖLL, Gregor LAAHA & Wolfgang U. DRESSLER 2006. Early noun plurals in German: regularity, productivity or default? *Journal of Child Language* 33. 271-302.

LIBBEN Gary 2008. How do we parse compound words? In SRINIVASAN Narayanan, A. K. GUPTA & Janak PANDEY (eds.). *Advances in Cognitive Science.* New Delhi: Sage Publications. 154-172.

LIBBEN Gary & Gonia JAREMA 2006. *The Representation and Processing of Compound Words.* Oxford: Oxford University Press.

LIBBEN Gary, Gonia JAREMA, Wolfgang U. DRESSLER, Jacqueline STARK & Christiane PONS 2002. Triangulating the effects of interfixation in the processing of German compounds. *Folia Linguistica* 26. 23-43.

ORTNER Lorelies & Elgin MÜLLER-BOLLHAGEN 1991. *Substantivkomposita*. Berlin: de Gruyter.

QUASTHOFF Uwe, Matthias RICHTER & Christian BIEMANN 2006. Corpus Portal for Search in Monolingual Corpora. In *Proceedings of the fifth international conference on Language Resources and Evaluation. LREC 2006, Genoa*. 1799-1802. http://corpora.informatik.uni-leipzig.de/

SCHREUDER Robert, Anneke NEIJT, Femke VAN DER WEIDE & R. Harald BAAYEN 1998. Regular plurals in Dutch compounds: linking graphemes or morphemes? *Language and Cognitive Processes* 13. 551-573.

STARK Jacqueline, Christiane PONS, Karin NAULT, Gary LIBBEN, Gonia JAREMA, Wolfgang U. DRESSLER (submitted). *The composition of German compounds: evidence from aphasia.*

Appendix

Real Words

Group	Real Words	Gloss	% correct
1a-0	BAUMGRENZE	tree+border	95
1a-0	FISCHTEICH	fish+pond	100
1a-0	KEKSDOSE	cookie+box	95
1a-0	KLANGWOLKE	sound+cloud	100
1a-0	LIFTFIRMA	elevator+company	100
1a-0	PASSAMT	passport+office	100
1a-0	SCHUHKARTON	shoe+box	100
1a-0	STOFFGESCHÄFT	cloth+business	100
1a-0	TERMINCHAOS	deadline+chaos	100
1a-0	TONSTUDIO	sound+studio	100
1a-0	WURSTLAGER	sausage+warehouse	100
1a-0	ZAHNSPANGE	tooth+clasp (=braces)	100
1b-0	BANDAPPARAT	ligament+apparatus	95
1b-0	BUCHGEWERBE	book+trade	95
1b-0	DINGWELT	thing+world	60
1b-0	FELDGRUPPE	field+group	100
1b-0	HORNTIER	horn+animal	95
1b-0	LIDCREME	eyelid+crème	75
1b-0	LOCHGITTER	hole+fence	95
1b-0	SCHWERTPAAR	sword+pair	84
1b-0	SKIWECHSEL	ski+change	95
1b-0	TUCHHANDEL	cloth+trade	100
1b-0	WORTLISTE	word+list	100
1b-0	WURMMITTEL	worm+substance	100
2-n	GURKENGLAS	pickle+jar	100
2-n	HASENJAGD	rabbit+hunt	100
2-n	LAIENBÜHNE	amateur+stage	100
2-n	LOCKENSTAB	curl+rod	95
2-n	OLIVENHAIN	olive+grove	95

Group	Real Words	Gloss	% correct
2-n	PERLENKETTE	pearl+chain	100
2-n	PUPPENTHEATER	puppet+theatre	100
2-n	ROLLENSPIEL	role+play	100
2-n	ROSENSTRAUCH	rose+bouquet	100
2-n	SPROSSENWAND	sprout+wall	95
2-n	TRAUBENMOST	grape+must	100
2-n	WESPENNEST	wasp+nest	100
3-er	BILDERGALERIE	picture+gallerie	95
3-er	BLÄTTERBERG	leaf+mountain	100
3-er	BRETTERZAUN	plank+fence	100
3-er	DÄCHERMEER	roof+sea	90
3-er	GÜTERLOGISTIK	goods+logistics	90
3-er	HÄUSERMARKT	house+market	100
3-er	KINDERFEST	child+party	100
3-er	LÄNDERKAMPF	country+war	95
3-er	LICHTERGLANZ	light+glow	100
3-er	MÄNNERCLUB	man+club	95
3-er	RINDERFARM	beef+farm	100
3-er	WEIBERHELD	Lady+hero	100
4-e	BRIEFEFLUT	letter+flood	63
4-e	FRÜCHTETEE	fruit+tea	100
4-e	GÄSTEZIMMER	guest+room	100
4-e	KÖCHEVEREIN	cook+club	26
4-e	LÄUSEKAMM	louse+comb	80
4-e	ORNAMENTEBUCH	oranament+book	44
4-e	PUNKTESCHNITT	point+cut	100
4-e	SPIELEMESSE	game+exhibition	84
4-e	STÄDTEBAU	city+construction	95
4-e	STÜHLEKAUF	chair+purchase	55
4-e	TALENTEMANGEL	talent+deficit	85
4-e	WEGENETZ	path+network	75
5-trunc	ACHSSTAND	axle+position	47
5-trunc	ADRESSHEFT	address+book	80

Group	Real Words	Gloss	% correct
5-trunc	BIRNBAUM	pear+tree	58
5-trunc	FARBPALETTE	colour+palette	100
5-trunc	KIRSCHDESSERT	cherry+dessert	95
5-trunc	KONTROLLSTATION	control+station	100
5-trunc	SACHREGISTER	subject-index	100
5-trunc	SCHULINSPEKTOR	school+inspector	100
5-trunc	SPRACHFAMILIE	language+family	95
5-trunc	WUNDARZT	wound+doctor	90
5-trunc	ZELLFUSION	cell+fusion	95

Ill-formed Stimuli

Group	Ill -formed StImulI	GloSS	% correct
1a-0	GÄNGLABYRINTH	corridor+labyrinth	100
1a-0	GÄNSZUG	goose+migration	95
1a-0	KNECHTSCHAR	servant+group	17
1a-0	MÄRKTKONZEPT	market+concept	95
1a-0	RÄTREPUBLIK	council+republic	95
1a-0	RECHTAUKTION	right+auction	75
1a-0	SÄFTTHEORIE	juice+theory	100
1a-0	STÄNDSTAAT	corporate+state	100
1a-0	STARSCHWARM	starring+swarm	28
1a-0	STERNKOCH	star+cook	50
1a-0	WIRTLOBBY	host+lobby	47
1a-0	ZERTIFIKATBRANCHE	cetrificate+branch	32
1b-0	ÄMTDSCHUNGEL	office+jungle	95
1b-0	EILIKÖR	egg+liquour	80
1b-0	FÄCHKANON	area+canon	100
1b-0	GESICHTZAHL	face+number	80
1b-0	GESPENSTSCHLOSS	ghost+castle	95
1b-0	GÖTTBOTE	God+messenger	90
1b-0	GRÄBFELD	grave+field	100
1b-0	KLEIDSCHRANK	dress+closet	85

Group	Ill-formed Stimuli	Gloss	% correct
1b-0	RÄDTAUSCH	wheel+exchange	95
1b-0	SCHILDWALD	sign+forest	45
1b-0	STRÄUCHGARTEN	shrub+garden	95
1b-0	VÖLKMIX	people+mix	95
2-n	AGGREGATENWERK	aggregate+works	68
2-n	ELEMENTENFABRIK	element+factory	63
2-n	GERÄTENRAUM	device+space	85
2-n	GERÜCHTENBÖRSE	rumour+exchange	90
2-n	GESCHÄFTENSCHWUND	business+reduction	90
2-n	GESCHENKENSTRESS	gift+stress	95
2-n	GETRÄNKENSTEUER	drink+tax	85
2-n	HUNDENZUCHT	dog+breeding	100
2-n	KERNENSCHACHTEL	seed+box	63
2-n	OBJEKTENKERAMIK	object+ceramic	85
2-n	ROSSENBALLETT	steed+ballet	58
2-n	WERTENWANDEL	value+change	74
3-er	APPARATERMEDIZIN	device+medicine	100
3-er	ÄRZTERKAMMER	doctor+chamber	90
3-er	HÄNDERDRUCK	hand+squeeze	95
3-er	HENGSTERSCHAU	stallion+show	95
3-er	HÜTERKOLLEKTION	hat+collection	74
3-er	MÄCHTERSYSTEM	power+system	90
3-er	MÄUSERKÖNIG	mouse+king	90
3-er	PFERDERKOPPEL	horse+paddock	84
3-er	PLAKATERKLAU	poster+theft	100
3-er	SCHWEINERTROG	pig+trough	100
3-er	SPRÜCHERCLOWN	expression+clown	85
3-er	TEILERDIEB	part+thief	80
4-e	BÄDETOUR	bath+tour	95
4-e	BIESTEWITZ	beast+joke	89
4-e	GEISTEBAHN	ghost+train	95
4-e	GESCHLECHTEROLLE	sex+role	74
4-e	GLÄSEKLANG	glass+sound	100

Group	Ill-formed Stimuli	Gloss	% correct
4-e	GLIEDEPUPPE	member+puppet	95
4-e	HÜHNESTALL	chicken+stall	95
4-e	KÖRNEMAIS	kernel+corn	100
4-e	KRÄUTEBUTTER	herb+butter	90
4-e	LIEDEABEND	song+evening	95
4-e	NESTESTAPEL	nest+pile	100
4-e	TÄLELOIPE	valley+run	90
5-trunc	ALGEGIFT	algae+poison	100
5-trunc	AUGEBINDE	eye+bandage	95
5-trunc	BLUMEBEET	flower+bed	95
5-trunc	FRANSEHUT	fringe+hat	95
5-trunc	FUNKEFLUG	spark+flight	100
5-trunc	GERSTESAFT	barley+juice	90
5-trunc	KAROTTEBREI	carrot+puree	95
5-trunc	LATTEROST	slat+grid	100
5-trunc	STUNDEHOTEL	hour+hotel	95
5-trunc	TOMATESALAT	tomato+salad	100
5-trunc	WAREKUNDE	ware+knowledge	95
5-trunc	WEIDEALLEE	willow+avenue	40