Paradigms in the mental lexicon: Evidence from German

Ruben van de Vijver [1,*], **Dinah Baer-Henney** [2]

[1] *Heinrich-Heine-Universität, Institute für Sprache und Information, Düsseldorf, Germany*
[2] *Heinrich-Heine-Universität, Institute für Sprache und Information, Düsseldorf, Germany*

Correspondence*:
Ruben van de Vijver
Ruben.Vijver@hhu.de

ABSTRACT

Previous research showed that the mental lexicon is organized morphologically, but the evidence was limited to words that differ only in subphonemic detail.

We investigated whether word forms that are related through morphology but have a different stem vowel affect each other's processing. We focused on two issues in two auditory lexical decision experiments. The first is whether the number of morphologically related word forms with the same stem vowel matters. The second is whether the source of similarity matters.

Word recognition experiments have shown that word forms that are phonologically embedded and related through inflection speed up each other's recognition, suggesting the word forms are represented within one unit in the mental lexicon. Research has further shown that words that are related through derivation, but that are phonologically different, are affected in a different way than words that are related through inflection. We conducted two experiments to further investigate this.

We used three subtypes of one inflectional class of German nouns, which allowed us to study different word forms with a phonological difference, while keeping the morphological relations among the word forms constant. All of these nouns have a plural form that ends in a -ə. They differ in the distribution of front and back vowels in the singular, plural and diminutive. This allows us to investigate the question whether word forms with different phonemes are processed differently with regard to (a) the number of word forms that share a vowel, and (b) the source of the similarity among the word forms; is the processing among word forms related through inflection different from the processing of word forms that are related through derivation?

We found that nonces that are based on word forms with a fronted vowel are mistaken for words when they resemble words in the word family, but not when they are unrelated to words in the word family. This shows that morphological effects in word auditory recognition studies are also found when the word forms differ in a full phoneme. We argue that this can be captured with a network representation, instantiated as a frame.

Keywords: Mental lexicon, word family, German inflectional classes, lexical decision, frame representation

1 INTRODUCTION

The repository of words in memory—the mental lexicon—is organized in intricate ways. Different degrees of similarities, and different dimensions of similarity affect the recognition of words to different degrees, and these differences allow us to draw conclusions about the structure of word forms in the mental lexicon (McQueen and Cutler, 1998; McQueen et al., 1995; McQueen, 2007). In this paper we explore the relationships among inflected words (singulars and plurals), and derived words (diminutives) in German. The umlaut-system of German, in which back vowels are fronted in particular morphological contexts, allows us to investigate morphological relations among word forms that cannot be reduced to phonetic similarity.

Words that sound similar facilitate each other's recognition. Words that share phonological material are considered in parallel for lexical access in all models of spoken word recognition (Weber and Scharenborg, 2012). The Dutch words *kapitaal* 'capital' and *kapitein* 'captain', which share sounds but not meaning, are both considered after hearing the first two syllables (Zwitserlood, 1989) and the word *bone* is activated after hearing *trombone* (Isel and Bacri, 1999).

A similar facilitation has been found for words that share meaning. In a lexical decision experiment Marslen-Wilson and Zwitserlood (1989) found that the prime *honing* 'honey' speeded up the recognition of the semantically related word *bij* 'bee'.

Recent models of word recognition treat words that are similar because of their morphological relatedness in the same way as words that are only phonologically, but not morphologically similar (Weber and Scharenborg, 2012). [1] There is, however, evidence from the literature that morphology should be more strongly incorporated in such models. The facilitatory effects among phonologically and semantically similar words on word recognition come together in morphologically related words. The words *Boot* 'boat' and *Boote* 'boats', for example, affect each other's recognition more strongly than neighbors that are only phonologically related. For example, the recognition of *car* is facilitated by its plural *cars*, but less by the unrelated *card* (Stanners et al., 1979), even though *cars* and *card* both differ in one phoneme from *car*.

Diving deeper into the relationships among morphologically related words in German, Schriefers et al. (1992) found in two experiments that word forms that are members of the same word family often influence each other's response latencies. In a first experiment they investigated relationships among inflected words, and in a second experiment they investigated the relationships among inflected and derived words.

In the first experiment they investigated four types of inflection in adjectives; adjectives with the nominative suffix -e (*klein*-[ə] 'small NOM, F/N/M'; the dative suffix -em of the strong declension of masculine and neuter adjectives (*klein*-[əm] 'small DAT, M/N'); adjectives ending in the suffix -es, which indicates nominative and accusative in the strong declension (klein-[əs] 'small NOM/ACC, N'); adjectives without suffix. They found asymmetries between the suffixes. Uninflected adjectives facilitated the recognition of all inflected adjectives. Adjectives inflected with [ə] facilitated the recognition of uninflected adjectives as well as inflected adjectives. Adjectives inflected with [əm] only facilitated recognition of itself and adjectives inflected with [əs], and, finally, adjectives inflected with [əs] only facilitated recognition of itself and adjectives inflected with [əm].

[1] A model that might incorporate such information is proposed by Gaskell and Marslen-Wilson (1997), whose distributed model of speech perception incorporates phonological and semantic information. However, it is not entirely clear how this should be implemented, and it is beyond the scope of this paper to develop a model of speech recognition.

In a second experiment Schriefers et al. (1992) looked at the response latencies between inflected and derived words. They used the derivational suffixes *-lich*, to create derived adjectives or adverbs (e.g *kleinlich* 'petty'), and *-heit* (e.g. *Kleinheit* 'smallness'), to create abstract nouns. Additionally, they used uninflected adjectives and inflected adjectives ending in *-es*. It turned out that uninflected adjectives prime all other items; adjectives ending in *-es* prime adjectives but not derived items; derived *-heit* items prime uninflected adjectives and themselves, but not other items; derived *-lich* adjectives prime themselves, but not other items. It appears that there is an asymmetry among the derived items: Derived *-heit* items prime uninflected adjectives, but derived *-lich* items do not. Schriefers et al. (1992) speculate that this difference among derived forms is a result of the stem vowel change that accompanies most *-lich* items, called umlaut. For example, the adjective *rot* 'red', has a fronted vowel when it is derived with *-lich*: *rötlich* 'reddish'. This finding suggests that phonologically similar word forms affect each other in a priming study, but when the word forms are not phonologically similar, if they differ in a vowel as the vowel in *rot* and the first vowel in *rötlich* do, they do not facilitate each other's recognition.

The findings of Schriefers et al. (1992) for German were corroborated and extended by Ernestus and Baayen (2007b) for Dutch. They observe that words in a paradigm—words that are related through inflection—are effectively neighbors of each other. Inflected words differ from uninflected words in one or more affixes. This fact alone would make them neighbors, and in addition an inflected word is embedded in an uninflected word, which affects its duration Kemps et al. (2005). It has been shown that words that are embedded in longer words, such as *ham* in *hamster*, are shorter than when they are standing alone. This difference in length is noticeable to listeners (Davis et al., 2002; Ernestus and Baayen, 2007b; Kemps et al., 2005; Salverda et al., 2003). Kemps et al. (2005) showed that participants take longer to decide whether an item is a word when its duration is off: if the string of a singular form [**bek**] 'brook' is given the (shorter) duration of the same string embedded in the plural form [**bek**ə] 'brooks', it takes longer to recognize as a word than when it is presented with its normal duration. If the string of the singular embedded in the plural [**bek**ə] is given the duration of the singular [**bek**] it is also recognized more slowly than when it has its expected duration.

Since speakers are aware of small phonetic differences, the question arises whether such small differences play a role in word recognition in paradigmatically related words. Ernestus and Baayen (2007b) investigated this question for paradigmatically related words in Dutch.

Dutch has final devoicing; the stem-final obstruent in the plural [hɑndə] 'hands' is voiced, whereas its correspondent in the singular [hɑnd̥] 'hand' appears to be voiceless, indicated in IPA by the ring underneath the d̥. However, there are traces of voicing in the singular that are small and subphonemic, but nevertheless noticeable (Warner et al., 2004). The vowel in [hɑnt], in which the final obstruent is devoiced, is slightly longer than it is in comparable words that have no voicing alternation in its paradigm, such as [krɑnt] (Ernestus and Baayen, 2003, 2006, 2007a,b; Warner et al., 2003, 2006, 2004). In words such as *krant* 'newspaper sg.' a completely voiceless stem-final obstruent has a completely voiceless correspondent in the plural [krɑntə] 'newspapers.

In a lexical decision experiment Ernestus and Baayen (2007b) compared judgements about the lexical status of two groups of nonces that were based on existing Dutch words. The nonces in one group, exemplified by *[krɑnd̥], had no support from members in its paradigm. There are no allomorphs that contain the string *[krɑnd]. Nonces in the other group, exemplified by *[hɑnd], do have support from other words in the word family. The plural allomorph of the singular [hɑnt] *hand* 'hand' is [hɑndə] *handen* 'hands'. The singular [hɑnt] shows traces of the voicing of the final obstruent in other forms in its word family (Ernestus and Baayen, 2003; Warner et al., 2004; Ernestus and Baayen, 2007b, 2006). It turned out that nonces that have no support in the paradigm are rejected faster as words than nonces that do have support from other members in the word family.

Ernestus and Baayen's interpretation of the effects is based on the amount of support the nonce word receives in the word family. Since the effect is cumulative, the representation of a word family can be interpreted as a list. If a nonce is embedded in many members of the word family, the effect is stronger than when the nonce is embedded in few or even in no members. However, if a word family is represented as a list, the findings of asymmetrical priming as reported by Schriefers et al. (1992) are difficult to interpret. In a list interpretation, the amount of support for a word form in a word family is crucial, not the source of support for a word form.

The evidence presented from German (Schriefers et al., 1992) and Dutch (Ernestus and Baayen, 2007b) suggests that word forms in the mental lexicon are organized along morphological lines. The word forms of the same word family affect each other's response latencies. However, in all data we have considered so far the word forms that affected each other were very similar; they only differed in small phonetic detail. The word forms were either embedded in each other and therefore had only small subphonemic durational differences in both languages.

This leaves open the question as to the generality of the morphological effect both studies reported. Would word forms that differ in one phoneme, rather than just in subphonemic detail, also affect each others recognition? Another question concerns the results from Schriefers et al. (1992), who found that priming is not equally strong among the members of the word family, which suggest that a word family is not simply a list. This raises the question as to what is the structure of a word family.

Schriefers et al. (1992) analyze their results within a network model, in which the lexicon is made up of nodes for words, morphemes, syllables and phonemes. Stems are morpheme nodes to which each word is connected. Since morphological variants share a stem node they are connected through a shared stem, but not directly through a shared lexical entry. For example, the stem *klein* is present in all inflected forms of the adjective, as well as in the derived forms *kleinlich* and *Kleinheit*. The stem *rot* is not present in the derived form *rötlich*. This model, then, explains their results.

Yet, Schriefers et al.'s network model needs to be modified. Their assumption that stems are stored separately in the mental lexicon is called into question by two sets of findings. First, there is accumulating evidence that complex words are stored and processed as wholes. Schreuder and Baayen (1997) found that reaction times to simplex words are modulated by the frequency of whole complex words, and not by the summed frequency of their individual morphemes. This is true even in agglutinative languages (Lehtonen et al., 2007; Moscoso del Prado Martín et al., 2004; Vannest et al., 2002). This shows that network models are correct in assuming that the mental lexicon is a network of connected nodes; words that share phonological form and meaning through shared morphology are activated simultaneously. But it also shows that complex words are stored as wholes.

Another argument against the centrality of stems in the network model comes from instances of paradigm leveling; members of a paradigm are often adjusted to each other–leveled–in order to make them more similar. An example of such leveling is found in Dutch. In Dutch [n] is normally not pronounced after a [ə]. The infinitive of *lopen* 'to walk' is pronounced [lopə]. Only under very formal circumstances it is pronounced [lopən] (Booij, 1995). The first person singular present tense is *ik loop*, pronounced [ɪk lop], and often analyzed as the stem form. However, in case an infinitive ends in a sequence [ənə], as in *oefenen* [ufənə] 'to practice', the first person singular, present tense is *ik oefen* [ɪk ufən], and not *[ɪk ufə] (Koefoed, 1979). Even though this process is correctly described as blocking of [n]-deletion at the end of a verbal stem (Booij, 1995), this description does not provide an understanding of the blocking. In nouns there is no such blocking. This can be seen by comparing suffixation of the agentive -aar in *ler-aar* [lerar] 'teacher', form the verb *leren* [lerə] 'to teach', with *molen-aar* [molənaar] 'miller' from the noun *molen* [molə] 'mill'. The agentive suffixes appear after the stem form and the form [molənaar] shows that the final [n] is part of the stem of [molən]. In the singular, however, this [n] is deleted; n-deletion is not blocked in nominal stems. This raises the question why [n] deletion only affects nominal stems, but not verbal stems? To answer this question, we propose that the blocking of [n]-deletion in verbs is a case of paradigm leveling; as far as we know this has not been proposed before. The verbal paradigm of [ufənə] has the plural forms *wij, jullie, zij* [ufənə] and the [n] after the first [ə] is therefore preserved in the first person singular [ɪk ufən]. The paradigm of nouns such as [molə] do not contain forms with a final [n]. In short, this argument reinforces the case against a central role of stems in the representation of paradigms.

In addition to providing an argument against the centrality of stems, paradigm leveling also highlights the fact that paradigms have structure and should not be represented as a list. In Dutch paradigm leveling, as we have seen above, plural verbal forms asymmetrically affect the singular forms. Such asymmetrical relations have also been observed for morphological features that make up a paradigm (Blevins, 2016; Haspelmath and Sims, 2010; Seyfarth et al., 2014). In German nouns, for example, it has been observed that in some inflectional classes is a dependency between genitive forms and plural forms, but the reverse is not true. If the genitive of a noun ends in [ən], for example, the plural does as well: the genitive form of *Mensch* [mɛnʃ] 'human being is *des Mensch-en* [dɛs mɛnʃən] 'human being-GEN', and its plural is *die Menschen* [diː ɱɛnʃən] 'human being-PL'. A plural ending in [ən] does not necessarily imply a genitive in [ən]: the plural form *die Staaten* [diː ʃtatən] 'the state-PL' has as genitive *des Staates* [dɛs ʃtatəs] 'the state-GEN' (Eisenberg, 2004; Thieroff and Vogel, 2009). Morphological properties sometimes depend on phonological properties (see also Neef, 1998). For example, if a plural ends in a [ə] its singular ends in a closed syllable. This is true for words such as *Bart* 'beard' *Bärte* 'beard-PL', *Boot* 'boat' *Boote* 'boot-PL' and *Fest Feste* 'party, celebrartion-PL'. The reverse, again, is not always true. Singulars such as *Mensch* or *Staat* have a plural that ends in *en*: *Menschen* and *Staaten*.

In short, the paradigm-as-list model of Ernestus and Baayen (2007b) is insufficient because paradigms are not lists, and the network model of Schriefers et al. (1992) is insufficient because paradigmatic effects go beyond shared stems. A representation of a paradigm needs to capture the dependencies among its word forms. This, then, raises the question as to how paradigms can be represented.

Frame representations allow us to capture the dependencies effects mentioned above (Barsalou, 1992; Gamerschlag et al., 2013; Löbner, 2014; Petersen and Osswald, 2014). In a frame the properties of a central node are represented as attribute-value structures. Attributes are functions that return a value. We will now analyze inflectional classes of German nouns as sets of (recursive) attribute-values pairs.

We propose to represent the inflectional classes of German nouns (Eisenberg, 2004; Köpcke, 1988; Thieroff and Vogel, 2009) as frames. The central node of each class is the category noun, and its attributes and their values are morphological and phonological properties that define an inflectional class. Providing a full overview of all inflectional classes is beyond the scope of this paper. Instead we provide frame representations of the class of nouns that has a plural that ends in a schwa–these nouns will also be at the heart of our experiments. The frame representations of these nouns are illustrated in figures 1, 2, 3 and 4. Each frame represents one subclass of nouns. The central node–the referential node–is indicated by a double circle, that attributes in small caps and their values in italics.

The paradigm of the nouns illustrated in figure 1 are masculine, end in a closed syllable, have a genitive that ends in [əs], and a plural that ends in [ə]. It is exemplified by the word forms [tak] *Tag* 'day' for the nominative, [tagəs] for the genitive and [tagə] for the plural. The paradigm of the nouns illustrated in figure 2 are masculine, end in a closed syllable, have a genitive that ends in [əs], and a plural that ends in [ə] and has a front vowel. It is exemplified by the word forms [bɑɐt] *Bart* 'beard' for the nominative, [baɐtəs] for the genitive and [bɛɐtə], with a front vowel, for the plural. The paradigm of the nouns illustrated in figure 3 are feminine, end in a closed syllable, have a plural that ends in [ə] and has a front vowel. It is exemplified by the word forms [hɑnt] *Hand* 'hand' for the nominative, and [hɛndə], with a front vowel, for the plural. The paradigm of the nouns illustrated in figure 4 are neuter, end in a closed syllable, have a genitive that ends in [əs], and a plural that ends in [ə]. It is exemplified by the word forms [bot] *Boot* 'beard' for the nominative, [botəs] for the genitive and [botə] for the plural.

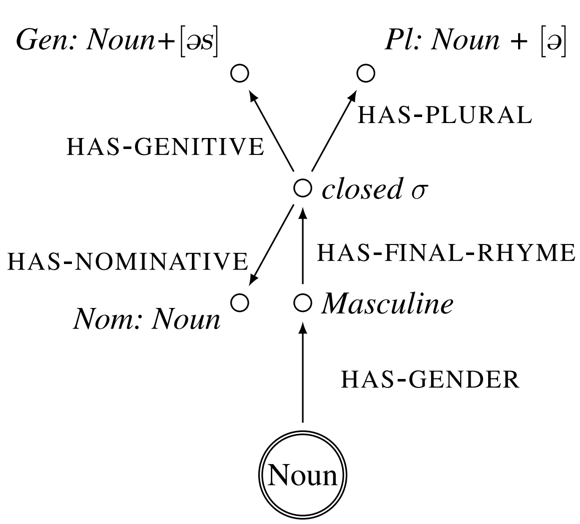

Figure 1. Frame representation of the inflectional class of nouns such as [tak] 'day'. Noun + [əs] indicates that the value of this attributes is [tagəs].

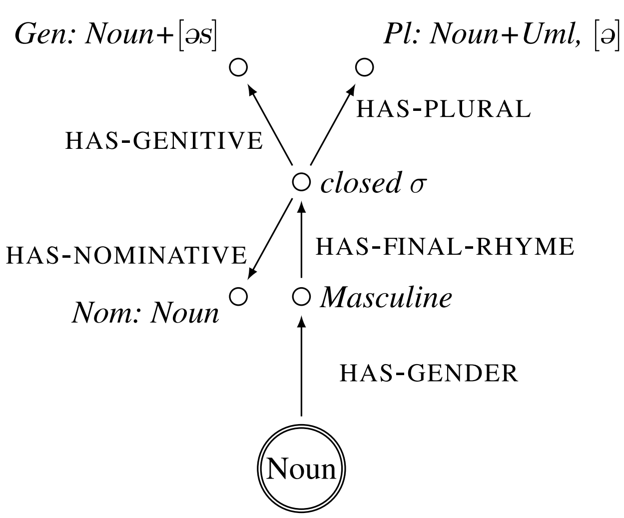

Figure 2. Frame representation of the inflectional class of nouns such as [bɑɐt] 'beard'. Uml(aut) indicates a fronted vowel, such that the pluaral is [bɛɐtə].

Now that the inflectional classes are represented as frames we can add diminutives. Typological work on diminutives shows that that they are lexically different from their base. In an overview of the typology of meaning of diminutives Jurafsky (1996) finds that, in addition to denoting smallness, diminutives can also denote affection, pejorative meanings or even contempt. This also holds for German diminutives. The word form spelled *Bärtchen* may refer to a small beard, either to indicate its smallness or to express a measure of contempt. The word form *Frauchen*, in contrast, can only refer to a woman who owns a pet–usually a dog–irrespective of the size of the woman. The word form *Brötchen*, as a further example, can only refer to a roll, no matter what its size, and never to a small loaf of bread. As these meanings are partly lexicalized they must be stored in the mental lexicon.

The change in meaning associated with derived forms, as with diminutives, is analyzed as a shift of the referent from one node to another (Andreou, 2018; Kawaletz and Plag, 2015). This is illustrated in figure 5. The referent of the noun has shifted to the node that contains the value of the attribute HAS-SIZE. In the figures 1, 2, 3 and 4 a branch with attribute-values for size size was omitted to avoid cluttering the representation. The frames in these figures do include such a branch the crucial difference with the representation of a diminutive as in figure 5 is the referential node, indicated with a double circle. The

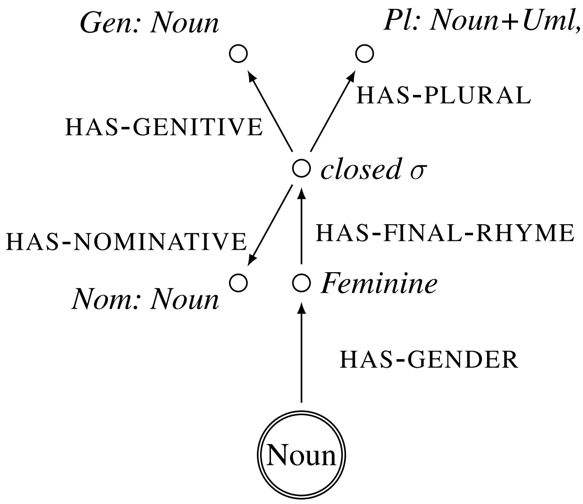

Figure 3. Frame representation of the inflectional class of nouns such as [hɑnt] ’hand’. Uml(aut) indicates a fronted vowel, such that the pluaral is [hɛndə].

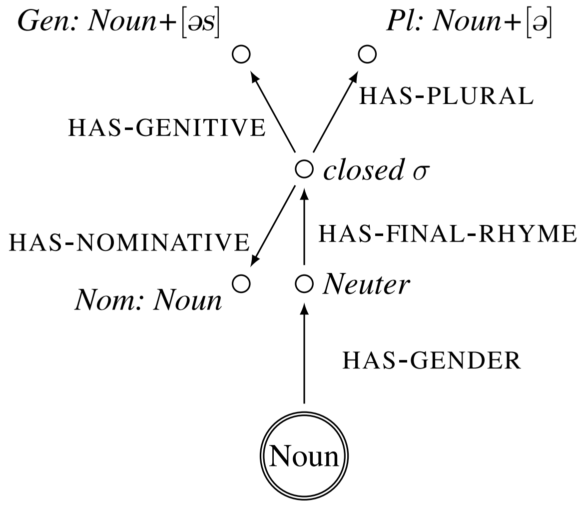

Figure 4. Frame representation of the inflectional class of nouns such as [bot] ’boat’. There are only one or two neuter words in this class with umlaut in the plural (Köpcke, 1988; Thieroff and Vogel, 2009)

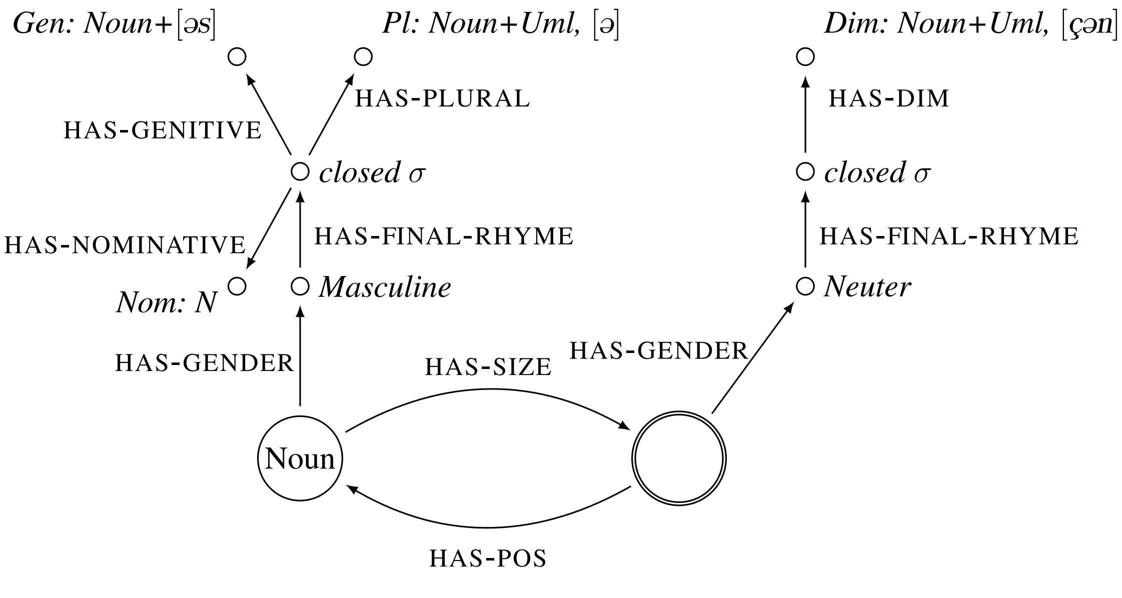

Figure 5. Frame representation the diminutive of the noun [bɑɐt] ’beard’ and the inflected formed of the paradigm of the plain, non-diminutive word. The central node of the frame of the diminutive is the Size-of-N.

referent can be selected dynamically by the speaker or hearer as needed (Kawaletz and Plag, 2015; Andreou, 2018).

To further investigate the role of morphology in word recognition and to test the predictions of our proposed frame representations, we will study the responses latencies in a particular type of German noun in an auditory lexical decision experiment. The nouns of this type are characterized by taking a [ə] in the plural, and their representations as frames are given in figures 1, 2, 3, 4 and 5 above. They can be divided into three subgroups (Köpcke, 1988) (Examples are given in table 1.) In one subclass the nouns have a back vowel in the singular and front vowels in the plural and the diminutive; for example, *Bart* [baɐt] ’beard’, *Bärte* [bɛɐtə] ’beards’ and *Bärtchen* [bɛɐtçən] ’little beard’. We will refer to this group of nouns as *Type 1* nouns (see figure 2 and 3). The nouns in the second subclass of this inflectional class have a back vowel in the singular and the plural, and a front vowel in the diminutive *Boot* [bot] ’boat’, *Boote* [botə] ’boats’ and *Bötchen*, [bøtçən] ’little boat’. We will refer to this group of nouns as *Type 2* nouns (see figure 1 and 4). The nouns in the third subgroup have a front vowel in all three word forms: *Fest* [fɛst] ’party, celebration’, *Feste* [fɛstə] ’parties, celebrations’ and *Festchen* [fɛstçən] ’little party, little celebration’. We will refer to this group of nouns as *Type 3* nouns (see figure 1 and 4).

Table 1. The noun types of the inflectional class in our study. V = Vowel, f = front, b = back

	Type 1	V	Type 2	V	Type 3	V
	Bart		*Boot*		*Fest*	
	[baɐt]	b	[bot]	b	[fɛst]	f
Inflection	[bɛɐtə]	f	[botə]	b	[fɛstə]	f
Derivation	[bɛɐtçən]	f	[bøtçən]	f	[fɛstçən]	f

This class of nouns allows us to address two questions that have arisen from the research summarized above. The first question is: Are morphological effects in word recognition limited to word forms that are embedded in each other, or do they extend to all word forms that are morphologically related; even to word forms that differ in a vowel? The nouns in in our word families are not always embedded in each other; they sometimes have a different vowel (for type 1 and type 2). For example the word form [bɑɐt] is not embedded in the word form [bɛɐtə]. The second question is: What is the structure of the representation of a word family? Since the nouns are not embedded in each other we are able to discern different effects for different sources of similarities, should there be evidence for an asymmetric structure; (if the source of similarity of a nonce is an inflected form, is it processed differently than when the source of similarity is a derived form?) If the word forms in a paradigm are represented together with derived word forms in one frame, as in figure 5, we also expect that inflected forms are more strongly associated with each other than derived word forms with inflected word forms. The derived word form has a different referential node than the inflected word forms.

These nouns form an excellent empirical basis for our investigation. We can create nonces for each type by changing the backness of a word form. For example, in one experiment we changed the word [bɑɐt] to the nonce [bɛɐt], in the other experiment we changed the word [bɛɐtə] to the nonce [bɑɐtə]. These nonces can show us whether the amount of word forms that are similar to the nonce affects it processing, and whether the source of the similar form (an inflected form are a derived form) affects it processing.

This brings us to our expectations. The first set of expectation concerns the role of morphology in word recognition. The evidence provided by Schriefers et al. (1992) and Ernestus and Baayen (2007b) shows that morphologically related word forms affect each other’s response latencies, but their evidence is limited to word forms that differ only in subphonemic duration. To see whether the effect is morphological in nature we will use words that are morphologically related and differ by a phoneme, rather than in subphonemic duration only. We expect that the recognition of nonces of type 1 (see table 1 above) is affected by their relation to existing word forms that are morphologically related, despite their phonological difference with an existing word. The more easily a nonce is mistaken for a word, the more mistakes participants will make in their accuracy and the more their response latencies will be affected.

The second set of expectations relates to the structure of the representations of inflected and derived words in the mental lexicon. If these are stored in the mental lexicon as the specific frame proposed in figure 5, in which diminutives have a different referential node than plain nouns, we expect that diminutives exert less influence on singular and plural nouns than singular and plural nouns on each other; singular and plural nouns share a referential node. This difference in referential nodes will affect both the accuracy and the response latencies.

We ran two auditory lexical decision tasks and measured the accuracy and response latency to words and nonces. Our method is slightly different from the one used in Ernestus and Baayen (2007b). We did not tell our participants to accept a nonce if it occurs in a word, but rather we asked them to judge whether an item is a word or not.

2 EXPERIMENT 1

In the first lexical decision test we investigated whether the accuracy and speed with which a nonce with a front vowel, as in the case of type 1 [bɛɐt], or type 2 [bøt], is judged, and if the accuracy and speed correlate with the amount of such stems in the word family. The nonce [bɛɐt] has support from 2 words in the word family and [bøt] is supported by only one word form.

2.1 Participants

Fifty-six native monolingual German adults took part in the experiment (these participants did not take part in experiment 2.) All of them were students at the University of Düsseldorf and they were given course credit for their participation. Their mean age was 20 years and 5 months. Forty-six women and 10 men participated; 50 of them were right-handed. One participant holds a university degree in a non-linguistic subject and all other participants reported to have a secondary school diploma that qualifies as entrance for a university as their highest educational degree. All participants had normal hearing and normal or corrected vision, and none of them reported any neurological problems.

2.2 Material

The material consisted of 90 German words (they are listed in the Appendix **??**). All material was recorded in a carrier sentence *Ich habe X gesagt.* 'I said X.' to ensure that the words have comparable prosodies. The words were excised from the sentences with Praat (Boersma and Weenink, 2018).

We used thirty Type 1 words: Monosyllabic words with a back stem vowel in the singular (e.g. *Bart* bɑɐt 'beard') and a front vowel in the plural (*Bärte* bɛɐtə) and the diminutive (*Bärtchen* bɛɐtçən). We created thirty nonces by giving the singular a front vowel (e.g. bɛɐt). The nonce has the same vowel as two allomorphs in the paradigm of *Bart*: the plural allomorph and the diminutive allomorph. Apart from the value of the [back] feature nothing in the word was changed in order to preserve its syllable structure.

We further used thirty Type 2 words: Monosyllabic words with a back vowel in the singular (e.g. *Boot* bot 'boat') and the plural (e.g. *Boote* botə) and a front vowel in the diminutive (e.g. *Bötchen* bøtçən). We created thirty nonces by giving the singular a front vowel.(e.g. bøt). This nonce has the same vowel as the diminutive.

The last group of thirty words were Type 3 words. They were also monosyllabic and had either front vowels in the singular, plural and diminutive stem or a back vowel in the singular. Nonces in this group of items were created by inverting the value of the [back] feature of the singular: if the singular had a front vowel, such as *Fest* [fɛst] 'party', the nonce was given a back vowel: [fɔst]; if the singular had a back vowel, such as *Pott* [pɔt] 'mug', the nonce was given a front vowel: [pɛt].

In addition we selected 180 existing monosyllabic words as fillers and 180 nonces based on these fillers. The total amount of items was therefore 540. They are all listed in 5. As filers we used monosyllabic nouns with front vowels from the same inflectional class as the words.

To be able to estimate the effect of frequency on our results, but we found no significant differences in frequency among the types of words in our experiments. We provide the details, therefore, in an appendix 5. We also estimated the neighborhood density of the words in our experiment. Here, too, we found no significant differences among the word types and provide the details in an appendix 5.

We created two lists, A and B, to prevent a sequence of a word and a related filler in the experiment. Half of the words were in list A and the other half was in list B. The nonces based on the words in list A were put in list B and the nonces based on the words in list B were put in list A.

2.3 Procedure

The experiment was programmed with PsyScope (Cohen et al., 1993) and was carried out in a quiet room at the University of Düsseldorf. The stimulus material was presented over headphones.

The experiment started with 16 practice trials half of which consisted of words and the other half of pseudo-words that obeyed the phonotactics of German. In the experiment there were 90 words and 90 nonces that we derived from the existing words. In addition we used 180 fillers; again 90 words and 90 pseudo-words.

After this the experimental items were presented in random order for each participant. Each trial started with a silence of 500 ms. followed by a tone of 500 ms. Then, after a silence of 450 ms., an item was presented and the participants had to decide as quickly as possible whether this was a word or not. The participants were instructed to press a key on the keyboard with a green sticker if they thought it was a word and a key with a red sticker if they thought it was not. For half the participants the green button was on the left side of the keyboard and for the other half it was on the right side of the key board. After the participants had made their choice the next trial started after a 2500 ms silence. The experiment lasted about 25 minutes.

2.4 Results

We first consider the accuracy of the participants to words in order to establish that they understood the task; that they correctly accepted words and did not incorrectly reject them. The raw result is summarized in table 2. The counts in 2 show that the words of all types were correctly accepted in more than 93% of the cases.

Table 2. Proportions of correct answers of *words* in Experiment 1

	Type 1 (baɐt)	Type 2 (bot)	type 3 (fɛst)
Correct	98%	93%	93%

A logistic mixed effects model with accuracy as dependent variable and Type as fixed effect, and random slopes for items and participants shows that the difference in table 2 is significant, as is illustrated in table 3.

We expected that nonces of type 1 are more likely to be mistaken for a word, because they resemble two existing word forms in the paradigm. We expected that nonces of type 2 are, in comparison to type 1

Table 3. Logistic regression analysis of the accuracy of the judgements of the participants to words in in experiment 1

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept = type 1)	5.19	0.46	11.35	0.00
type2	-1.36	0.56	-2.45	0.01
type3	-1.22	0.56	-2.17	0.03

nonces, less likely to be mistaken for an existing word. As type 3 nonces resemble no existing word forms in the paradigm, they should be easiest to recognize as nonces.

The results of the nonces in table 4 show that nonces of type 1 were incorrectly accepted in 14% of the cases, proportionally more than type 2 and type 3 nonces.

Table 4. Proportions of incorrect answers to *nonces* in Experiment 1

	Type 1 (bɛɐt)	Type 2 (bøt)	type 3 (fɔst)
Incorrect	14%	7%	9%

A logistic mixed effects model with accuracy as dependent variable and with Type as fixed effect, and random slopes for items and participants shows that the difference in table 4 is significant, as is illustrated in table 5.

Table 5. Logistic regression analysis of the accuracy of nonces in experiment 1

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept = type 1)	2.67	0.31	8.51	0.00
Type 2	1.02	0.40	2.58	0.01
Type 3	0.79	0.39	2.01	0.04

Nonces of type 1 were more often mistaken for real words than nonces of type 2 or 3. This analysis, then, confirms that nonces of type 1 are more difficult to reject than nonces of type 2 or 3 as expected. In an analysis, which is not shown here, in which type 2 was designated to be the intercept showed that the accuracy of type 2 and 3 nonces is not statistically different.

We will now present the results of a mixed effects model of the log-transformed reaction times of the correctly judged words in experiment 1. The results of a linear mixed effects model with the logarithm of the *Reaction time* as dependent variable and *Type* (type 1, type 2, type 3), as fixed factor, and random slopes for Items and Participants is presented in Table 6.

Table 6. Linear regression analysis of the log-transformed response latencies of the reaction times to correctly accepted words in experiment 1

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept = type 1)	6.34	0.04	131.17	178.22	0.00
Type 2	0.09	0.04	85.68	2.51	0.01
Type 3	0.10	0.04	85.78	2.70	0.01

The results of the analysis, presented in table 6, show that words of type 1 are reacted to fastest and that type 2 and type 3 words are reacted to slightly, but significantly slower. In combination with results of the accuracy to words, presented in table 3, it suggests that type 1 words are recognized most accurately and fastest.

We will end the presentation of the results of experiment 1 with a mixed effects model of the reaction times to the incorrectly identified nonces in experiment 1. The participants thought erroneously that these were words and in that case the paradigm may have been activated to influence the reaction times. The number of items over which this analysis was run, was very small, though, as the participants made relatively few mistakes.

The results of a linear mixed effects model with the logarithm of the *Reaction time* as dependent variable and *Type* (type 1, type 2, type 3), as fixed factor, and random slopes for Items and Participants is presented in Table 7.

Table 7. Linear regression analysis of the log-transformed response latencies of the reaction times to incorrectly accepted nonces in experiment 1

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept = Type 1)	6.67	0.05	75.00	141.58	0.00
Type 2	0.10	0.05	56.70	1.87	0.07
Type 3	0.09	0.05	50.47	1.77	0.08

Even though the reaction times to the incorrectly accepted nonces are not statistically different, there appears to be a tendency to react a bit more slowly to type 2 and type 3 nonces.

We expected that nonces of type 1 were more likely to be mistaken for words, because there is enough support for their assumption within word family of type 1. This expectation turned out to be correct. It was most difficult to correctly reject nonces of type 1 ([bɛɐt]). The difference between making a correct and an incorrect decision is smallest for type 1 nonces and larger for type 2 ([bøt]) and 3 nonces ([fɔst]), where there is either support from a derived word form in the word family (type 2) or no support for the nonce (type 3), and therefore more uncertainty on the part of the participants. The data from the reaction time analysis of nonce items are more inconclusive. The participants were so good at rejecting nonce words, that we had few data on which to base our analysis. The tendency of the data, though, is that nonces of type 1 are reacted to more slowly than type 2 and 3 nonces (see table 7).

In short, experiment 1 showed that there is evidence for a role of morphological information in word recognition that goes beyond small subphonemic differences among the parts of words forms in a word family (Ernestus and Baayen, 2007b; Schriefers et al., 1992). This evidence is given by a reduced accuracy for nonces that are supported by many forms in the word family. This support provides the participants with mistaken certainty that they are, in fact, dealing with a word. The analysis of the words provides additional support for this interpretation. Type 1 words are processed fastest (see table 3) and most accurate (see table 6) of the types in our experiment. The singular of type 1 activates the associated inflected and derived word forms and thus makes it more likely for a participant to mistakingly think that a nonce form of type 1 is an existing word.

A different interpretation cannot be ruled out without further evidence. As experiment 1 showed no difference between nonces of type 2 and type 3, it may also be that the source of support caused our results, rather than the amount of support. In this interpretation type 1 nonces are reacted to differently because they are similar to an inflected form in the word family, whereas the nonces of type 2 are related to a derived word and type 3 nonce are not related to any word in the word family.

A second experiment, in which the amount of support for nonces is kept constant will be able to distinguish these two interpretations.

3 EXPERIMENT 2

The second experiment was a lexical decision experiment as well. Its aim was to investigate whether the source of similarity among word forms in a word family is relevant. Are nonces processed differently if they resemble an inflected word form than when they resemble a derived word form? If they are, we expect differences in accuracy and response latencies among the nonces of different types, correlating with the source of support for a nonce.

3.1 Participants

Fifty-one native monolingual German adults took part the experiment (these participants did not take part in experiment 1.) All of them were students at the University of Düsseldorf and they were given course credit for their participation. Their mean age was 22 years and 5 months. Forty-seven women and 4 men participated. Forty-five participants were right-handed, 6 were left-handed. One participant holds a university degree in a non-linguistic subject and all other participants reported to have a secondary school diploma that qualifies as entrance for a university as their highest educational degree. All participants had normal hearing and normal or corrected vision, and none of them reported any neurological problems.

3.2 Material

We used the bisyllabic plural forms of the German nouns used in experiment 1 and to create nonces we changed the stem vowel of the plural form.

For the words of type 1—bɑɐt, bɛɐtə, bɛɐtçən—we created a nonce by changing the front vowel of the plural word form to back: bɑɐtə. This nonce is only similar to the the singular word form. Words and nonces of type 1 nonces are listed in table 15 in the Appendix (section III).

For the words of type 2—bot, botə, bøtçən—we created a nonce form by changing the back vowel of the plural to front: bøtə. This nonce is only similar to the diminutive word form. Words and nonces of type 2 are listed in table 16 in the Appendix (section III).

For the words of type 3—fɛst, fɛstə, fɛstçən—we created a nonce form by changing the front vowel of the plural to back: fɔstə or by changing the back vowel of the plural to front: [pɛtə]. Neither of these nonces are similar to a word form in the word family of the existing words upon which they are based. Words and nonces of type 3 are listed in table 17 in the Appendix (section III).

In addition we selected as fillers 180 existing bisyllabic plural words from the same inflectional class as the words, and 180 nonces based on these fillers. The total amount of items was therefore 540. They are all listed in the Appendix (section III).

3.3 Procedure

The procedure for experiment 2 was identical to experiment 1.

3.4 Results

We first consider the accuracy of the participants. This establishes that the participants understood the task. The data in table 8 show that words of type 1 were recognized best as words, whereas the percentages correct answers to type 2 and 3 words are very similar. These relatively low percentages show that it was relatively difficult for the participants to recognize the words as existing words. The reason might be that the words in experiment 2 are plurals, which were presented to the participants without context. The participants may have expected singulars by default–since singulars are on average more frequent–and, not finding a fitting singular in their mental lexicon, incorrectly rejected it as a word.

Table 8. Proportions of correct answers of *words* in Experiment 2

	Type 1 (bɛɐtə)	Type 2 (botə)	type 3 (fɛstə)
Correct	1180 (77%)	1078 (71%)	1122 (73%)

A logistic mixed effects model with accuracy as dependent variable and Type as fixed effect, and random slopes for items and participants shows that the difference in table 8 between words of type 1 and 2 is significant. Type 3 words caused more mistakes, but the difference is not significant, as is illustrated in table 9.

Table 9. Logistic regression analysis of the accuracy of the judgements of the participants to words in in experiment 2

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept = type 1)	1.92	0.34	5.67	0.00
Type 2	-0.70	0.23	-3.00	0.00
Type 3	-0.42	0.23	-1.80	0.07

Table 10 is an overview of the incorrect acceptance of the nonces in experiment 2. Most mistakes were made in type 1 and type 2 nonces, while the number of mistakes to type 3 nonces is smaller than to type 1 and 2 nonces.

Table 10. Proportions of incorrect answers of *nonces* in Experiment 2

	Type 1 (bɑɐtə)	Type 2 (bøtə)	type 3 (fɔstə)
Incorrect	481 (32%)	478 (31%)	327 (21%)

The data in table 10 were analyzed in a logistic mixed effects model with accuracy as dependent variable and with Type as fixed effect, and random slopes for items and participants. The analysis confirms that nonces of type 1 and 2 are judged equally accurately, whereas nonces of type 3 are judged with greater accuracy, as is illustrated in table 11.

We expected that the source of support mattered and that nonces that are supported by an inflected form are treated differently from nonces that have support from a diminutive. It turns out, though, that nonces of type 1 and type 2 are both mistaken for words to the same extent, but differently from type 3.[2]

[2] Releveling of our factors showed that this Type 1 and 2 and indeed the same and that they are different from type 3.

Table 11. Logistic regression analysis of the accuracy of the judgements of the participants to nonces in in experiment 2

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept = type 1)	1.02	0.31	3.32	0.00
Type 2	0.02	0.19	0.12	0.90
Type 3	1.04	0.20	5.16	0.00

Let us turn to the analysis of the reaction times. The results of a linear mixed effects model with the logarithm of the *Reaction time* as dependent variable and *Type* (type 1, type 2, type 3), is presented in Table 12. *Item* and *Participants* were given random slopes.

Table 12. Linear regression analysis of the log-transformed response latencies of the judgements to correctly accepted words in experiment 2

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept = type 1)	6.36	0.05	129.05	120.31	0.00
Type 2	0.13	0.06	86.82	2.26	0.03
Type 3	0.10	0.06	86.61	1.75	0.08

Words of type 2 are reacted to slower than words of type 1, and words of type 3 are reacted to a bit slower, but not significantly, than words of type 1. An analysis in which the fixed factor was releveled so as to make type 2 the intercept (the analysis is not shown here), showed that the difference between type 2 and type 3 words is not significant. The reaction time data, too, show that type 1 and type 2 are different from type 3 words.

We also analyzed the accuracy data of incorrectly accepted nonces, that we have presented in table 10. The participants thought erroneously that these were words and in that case the paradigm may have been activated to influence the reaction times.

The results of a linear mixed effects model with the logarithm of the *Reaction time* as dependent variable and *Type* (type 1, type 2, type 3), as fixed factor, and random slopes for Items and Participants is presented in table 13. The analysis shows that the reaction times to items of type 2 and 3 are slightly, but significantly, faster than reaction times to items of type 1.

Table 13. Linear regression analysis of the log-transformed response latencies of the judgements to incorrectly accepted nonces in experiment 2

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept = type 1)	6.72	0.04	75.80	178.83	0.00
Type 2	-0.08	0.03	78.04	-2.94	0.00
Type 3	-0.08	0.03	100.87	-2.69	0.01

Nonce words of type 1 are supported by an inflected form, while nonce words of type 2 are supported by a derived form, and nonce words of type 3 have no support at all in their word family. The reaction time analysis indicate that having support from an inflected form in the word family makes the reaction times slower than having support from a derived form or no support at all.

In combination with the analysis of accuracy, the data indicate that participants are the accuracy of their judgements is not affected by the source of support for a nonce (table 11), but the source of support does affect the time they need to take their erroneous decision. the influence of word forms in a word family is not equal from all forms to all other forms, as a list interpretation of the representation of paradigms in the mental lexicon would lead us to believe.

4 DISCUSSION

On the basis of the findings of (Schriefers et al., 1992; Ernestus and Baayen, 2007b), we set out to investigate two questions. This first was whether word forms that are morphologically related influence each other's recognition, even if they differ in a complete phoneme. The second was whether inflectionally related words exert more influence on each other than derivationally related words on inflected words.

In a first lexical decision experiment we assessed whether nonces that differ in one phoneme and have support from two word forms in the word family are treated differently from nonces that differ from words in one phoneme and have support from one word form, or whether they are treated differently from nonces without any support. We used nouns of three subtypes of the same inflectional class. In the first subtype the plural form has a front vowel (*Bart* 'beard, sg. ' and *Bärte* 'beard, pl.'); the second subtype has a back vowel in the plural (for example *Boot* 'boat, sg.' and *Boote* 'boat, pl.'); the third subtype has a front vowel in the singular and the plural (*Fest* 'party, celebration sg.' and *Feste* 'party, celebration, pl.'.) All three subtypes have diminutives with front vowels: *Bärtchen* 'little beard', *Bötchen* 'littel boat', and *Festchen* 'little party, celebration'. The word forms in these word families sometimes differ by one phoneme, for example vowel in the the singular of *Bart* is back and its counterpart in the plural is front *Bärt*. We used the diminutives to investigate whether inflected forms (singulars and plurals) affect each other more strongly than inflected forms affect derived forms (singulars and plurals as opposed to diminutives.)

We expected that, if morphology plays a role in word recognition, the nonces with support from word forms in the word family would be more likely to be mistaken for a word. As a consequence, such a nonce would be more likely to be erroneously accepted as a word (type 1 nonces in experiment 1). Moreover, we expected that the source of support would affect the reaction times and the accuracy to judgements of nonces, since we hypothesize that not all words forms in a word family affect each other to the same extent.

These expectations were borne out. Participants were more likely to mistake a nonce for a word if the phonological make up of a nonce was supported by two word forms in the word family (see table 4 and 7). However, as the participants made relatively few mistakes, the reaction time data do not allow us a firm conclusion, even though the tendency in the data hints at a faster decision in case a nonce is supported by two forms in the word family. We extend the results from (Schriefers et al., 1992; Ernestus and Baayen, 2007b) by showing that even morphologically related word forms that differ in one phoneme affect each other's response latencies, provided they are morphologically related.

In a second lexical decision experiment we assessed whether a derived item exerts less influence on an inflected item, than inflected items on each other. We expected that a nonce that resembles an inflected form would be more likely to be mistaken for a word than when a nonce resembled a derived form (type 1 and 2 nonces in experiment 2). Moreover, we expected that the difference in response latencies of incorrect reactions to a nonce that resembles an inflected form are different than the response latencies of incorrect answers to a nonce that resembles a derived form (type 1 and 2 nonces in experiment 2).

The expectations were partially borne out. Nonces that are similar to an inflected word are mistaken for a word as often as nonces that are similar to a derived word. This shows that derived words do indeed influence inflected words and that inflected words influence each other, but not that the strength of the influence is determined by the source of the influence. However, the response latencies show that a nonce that has support from an inflected form (nonces of type 1) take longer to be erroneously accepted as a word than a nonce that has support from a derived form (type 2) or a nonce that has no support (type 3).

In combination the results show that morphologically related word forms that differ in a vowel phoneme affects each other, and that the influence of word forms in a paradigm is not equal: Inflected word forms exert a stronger influence on each other than a derived word form on an inflected word form. In short, the results of experiment 1 and 2 together suggest that the frame representation proposed in figure 5 is on the right track.

These results are reflected in the frame representations (see figures 1, 2, 3, 4 and 5): inflected forms share a central node and influence each other more strongly. The influence of derived words on inflected words is smaller because they do not share a central node with inflected word forms.

Ernestus and Baayen (2007b) showed that both inflected words and derived words influence each other, but their items were almost identical and differed only in subphonemic detail. This, it may turn out, is a crucial difference with our study. in order for derived forms to exert a greater influence on inflected forms it may be necessary for them to not only resemble the inflected words semantically, but also phonologically and phonetically. This would also extend to the results of Schriefers et al. (1992).

Our results support network models in which word forms are organized according to morphological affiliation, and phonological and semantic similarity. We have made the morphological organization more specific to include the difference between inflection and derivation as a difference between the referential node within a concept. In processing this difference is reflected by the fact that the influence of inflected words on each other is stronger than the influence of derived forms on inflected forms. Moreover, we have provided an argument to further incorporate word families in models of word recognition.

Moreover, by proposing a frame representation we have connected the psycholinguistically motivated network models (Schriefers et al., 1992; Schreuder and Baayen, 1995) with attribute-value models (Bonami and Crysmann, 2016), in general and frame models in particular (Gamerschlag et al., 2013; Löbner, 2014; Andreou, 2018).

5 CONCLUSION

Our experiments provided further evidence that the mental lexicon is organized along morphological lines. Much evidence in the literature shows that derived word forms themselves for networks of related derived word forms (Lehtonen et al., 2007; Moscoso del Prado Martín et al., 2004; Schreuder and Baayen, 1997; Schriefers et al., 1992; Vannest et al., 2002). Our results extends these findings to inflectionally related word forms and further entrench the theory that inflectionally related words are also represented as a network. This provides evidence for a network of paradigmatic relations, that we represented as a frame in figures 1, 2, 3, 4 and 5. Inflectionally related forms share a referential node, while in derived words the referential node is a different one.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

RvdV: Conception, Design, Analysis, Writing
DBH: Design, Analysis, Writing

FUNDING

The funding of the Strategischer ForschungsFonds (*Frames in Phonology: Paradigms as representations*) of the Heinrich-Heine-Universität, and the financial support of the Deutsche Forschungsgemeinschaft, SFB 991, D05 is gratefully acknowledged.

ETHICS COMPLIANCE STATEMENT

According to information obtained from Dr. H. Weyerts-Schweda (Deutsche Forschungsgemeinschaft), no approval is required for behavioral experiments (reaction time) using standard psycholinguistic stimulus materials (auditorily presented words) without any aversive or emotionally arousing materials.

ACKNOWLEDGMENTS

Thank you Pascal Coenen, Kurt Erbach, Jens Fleischauer, Todor Koev, Sebastian Löbner, Albert Ortmann, Wiebke Petersen, Kim Strütjen, Peter Sutton and Fabian Tomaschek.

REFERENCES

Andreou, M. (2018). A frame-based approach to evaluative morphology. Heinrich-Heine-Universität Düsseldorf

Baayen, R., McQueen, J. M., Dijkstra, A., and Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisting Dutch plurals. In *Morphological Structure in Language Processing*, eds. R. Baayen and R. Schreuder (Berlin: Mouton de Gruyter), Trends in Linguistics. Studies and monographs 115

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). The CELEX lexical database (release 2)[cd-rom]. *Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor]*

Baroni, M., Bernardini, S., Ferraresi, A., and Zanchetta, E. (2009). The wacky wide web: a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation* 43, 209–226

Barsalou, L. (1992). Frames, concepts, and conceptual fields. In *Frames, fields and contrasts: New essays in semantic and lexical organization*, eds. A. Lehrer and K. E.F. (Lawrence Erlbaum Associates, Inc). 21–74

Blevins, J. P. (2016). *Word and paradigm morphology* (Oxford University Press)

Boersma, P. and Weenink, D. (2018). *Praat: Doing phonetics by computer (version 6.0.39) [computer program]*. Tech. rep., Universiteit van Amsterdam

Bonami, O. and Crysmann, B. (2016). The role of morphology in constraint-based lexicalist grammars. In *Cambridge Handbook of Morphology*, eds. A. Hippisley and G. T. Stump (Cambridge University Press). 609–656

Booij, G. (1995). *The Phonology of Dutch* (Oxford: Clarendon Press)

Cohen, J. D., MacWhinney, B., Flatt, M., and Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers* 25, 257–271

Davis, M. H., Marslen-Wilson, W. D., and Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 28, 218

Eisenberg, P. (2004). Grundriss der deutschen grammatik/peter eisenberg.-band1: Das wort.-3.

Ernestus, M. and Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in dutch. *Language* 79, 5–38

Ernestus, M. and Baayen, R. H. (2006). The functionality of incomplete neutralization in Dutch: The case of past-tense formation. *Laboratory Phonology* 8, 27–49

Ernestus, M. and Baayen, R. H. (2007a). Intraparadigmatic effects on the perception of voice. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4* 286, 153–174

Ernestus, M. and Baayen, R. H. (2007b). Paradigmatic effects in auditory word recognition: The case of alternating voice in Dutch. *Language and Cognitive Processes* 22, 1–24

Faaß, G. and Eckart, K. (2013). Sdewac – a corpus of parsable sentences from the web. In *Language Processing and Knowledge in the Web*, eds. I. Gurevych, C. Biemann, and T. Zesch (Berlin, Heidelberg: Springer Berlin Heidelberg), 61–68

Gahl, S. and Strand, J. F. (2016). Many neighborhoods: Phonological and perceptual neighborhood density in lexical production and perception. *Journal of Memory and Language* 89, 162–178

Gamerschlag, T., Gerland, D., Osswald, R., and Petersen, W. (2013). *Frames and concept types: applications in language and philosophy*, vol. 94 (Springer Science & Business Media)

Gaskell, M. G. and Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and cognitive Processes* 12, 613–656

Hall, K. C., Allen, B., Fry, M., Mackie, S., and McAuliffe, M. (2015). *Phonological CorpusTools*. Tech. rep., (https://github.com/PhonologicalCorpusTools/CorpusTools/releases)

Haspelmath, M. and Sims, A. D. (2010). *Understanding Morphology*. Understanding Language Series (Abingdon, Oxon: Routledge), 2^{nd} edn.

Isel, F. and Bacri, N. (1999). Spoken-word recognition: The access to embedded words. *Brain and Language* 68, 61–67

Jurafsky, D. (1996). Universal tendencies in the semantics of the diminutive. *Language* , 533–578

Kawaletz, L. and Plag, I. (2015). Predicting the semantics of english nominalizations: A frame-based analysis of-ment suffixation. In *Semantics of complex words* (Springer). 289–319

Kemps, R. J., Wurm, H., Lee, Ernestus, M., Schreuder, R., and Baayen, R. (2005). Prosodic cues for morphological complexity in dutch and english. *Language and Cognitive Processes* 20, 43–73

Koefoed, G. A. (1979). Paradigmatische invloeden op fonetische processen. *Glot Special* , 51–72

Köpcke, K.-M. (1988). Schemas in German plural formation. *Lingua* 74, 303–335

Lehtonen, M., Cunillera, T., Rodríguez-Fornells, A., Hultén, A., Tuomainen, J., and Laine, M. (2007). Recognition of morphologically complex words in Finnish: Evidence from event-related potentials. *Brain research* 1148, 123–137

Löbner, S. (2014). Evidence for frames from human language. In *Frames and concept types* (Springer). 23–67

Luce, P. A. (1985). *Structural distinctions between high and low frequency words in auditory word recognition*. Ph.D. thesis, Indiana University

Luce, P. A. and Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and hearing* 19, 1

Marslen-Wilson, W. and Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human perception and performance* 15, 576

McQueen, J. M. (2007). Eight questions about spoken-word recognition. *The Oxford handbook of psycholinguistics* , 37–53

McQueen, J. M. and Cutler, A. (1998). *Morphology in word recognition* (Wiley Online Library)

McQueen, J. M., Cutler, A., Briscoe, T., and Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes* 10, 309–331

Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., and Baayen, R. H. (2004). Morphological family size in a morphologically rich language: the case of Finnish compared with Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30, 1271

Neef, M. (1998). The reduced syllable plural in German. In *Models of Inflection*, ed. R. Fabri (Niemeyer Tübingen). 244–265

Petersen, W. and Osswald, T. (2014). Concept composition in frames: Focusing on genitive constructions. In *Frames and concept types* (Springer). 243–266

Pisoni, D. B., Nusbaum, H. C., Luce, P. A., and Slowiaczek, L. M. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech communication* 4, 75–95

Salverda, A. P., Dahan, D., and McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition* 90, 51–89

Schreuder, R. and Baayen, H. R. (1995). Modeling morphological processing. In *Morphological aspects of language processing*, ed. L. B. Feldman (Lawrence Erlbaum Associates). 131–156

Schreuder, R. and Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language* 37, 118–139

Schriefers, H., Friederici, A., and Graetz, P. (1992). Inflectional and derivational morphology in the mental lexicon: Symmetries and asymmetries in repetition priming. *The Quarterly Journal of Experimental Psychology* 44, 373–390

Seyfarth, S., Ackerman, F., and Malouf, R. (2014). Implicative organization facilitates morphological learning. In *Annual Meeting of the Berkeley Linguistics Society*. vol. 40, 480–494

Shaoul, C. and Tomaschek, F. (2013). A phonological database based on CELEX and N-gram frequencies from the SWEDAC corpus. Personal communication

Stanners, R. F., Neiser, J. J., Hernon, W. P., and Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior* 18, 399–412

Thieroff, R. and Vogel, P. M. (2009). *Flexion* (Winter)

Vannest, J., Bertram, R., Järvikivi, J., and Niemi, J. (2002). Counterintuitive cross-linguistic differences: More morphological computation in English than in Finnish. *Journal of Psycholinguistic Research* 31, 83–106

Vitevitch, M. S. and Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* 40, 374–408

Warner, N., Good, E., Jongman, A., and Sereno, J. (2003). Orthography and underlying form in incomplete neutralization. *Acoustical Society of America Journal* 114, 2396–2396

Warner, N., Good, E., Jongman, A., and Sereno, J. (2006). Orthographic vs. morphological incomplete neutralization effects. *Journal of Phonetics* 34, 285–293

Warner, N., Jongman, A., Sereno, J., and Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32, 251–276
Weber, A. and Scharenborg, O. (2012). Models of spoken-word recognition. *Wiley Interdisciplinary Reviews: Cognitive Science* 3, 387–401
Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32, 25–64

Appendices

APPENDIX I. FREQUENCY

As frequency differences among words affects their processing (Baayen et al., 2003), we established the frequencies of the words in our experiments. We estimated their frequency by using the frequencies of these words in the SdeWaC corpus (Faaß and Eckart, 2013). This corpus was created by parsing all sentences from all websites that end in '.de' (Baroni et al., 2009). Shaoul and Tomaschek (2013) then used this corpus to establish the frequency of the words in CELEX (Shaoul and Tomaschek, 2013) that occur in the SdeWaC corpus. Our estimates are based on the occurrences of words in CELEX, but with the more recent frequency counts of the SdeWaC corpus.

We did not find frequency information of all words, in fact for 21% of our data we did not find frequency information (we did not find frequency information on 32% of Type 1 nouns, 19% of Type 2 nouns and 13% of Type 3 nouns).

We these caveats in mind, we calculated a regression model with Number (singular or plural) and Type (Type 1, Type 2 and Type 3) as predictors of the frequency per million. As can be seen in table 14 no main effect nor any interactions reached significance.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.17	3.14	0.05	0.96
Type 2	7.46	4.34	1.72	0.09
Type 3	7.22	4.34	1.66	0.10
Plural	-0.14	4.93	-0.03	0.98
Type 2 * Plural	-7.24	6.55	-1.11	0.27
Type 3 * Plural	-6.12	6.48	-0.94	0.35

Table 14. Statistical comparison of the frequencies of the different type of words in our experiments.

In short, the frequencies if the three types of nouns in our experiments is comparable and any effect that we may find is attributable to factors other than (or, perhaps more accurately, in addition to a similar) frequency effect.

APPENDIX II. NEIGHBORHOOD DENSITY

An inhibitory effect is found among words that are phonologically or phonetically similar, and which do not stand in a morphological relationship to each other. The similarity among words can be measured in several ways (Gahl and Strand, 2016), but often it is done in terms of phonemes. Words that differ one phoneme are called neighbors (Luce, 1985; Gahl and Strand, 2016). For example, the words *sling* and *fling* are neighbors. The response latencies to words with many neighbors is slowed down in comparison to words with a few neighbors (Luce, 1985; Luce and Pisoni, 1998; Pisoni et al., 1985). To ensure that the effects we found can indeed be ascribed to morphology and not on an effect of neighborhood density, we calculated the neighborhood density of our items. We created a data set of German word forms of nouns, verbs and adjectives by extracting 355.625 nouns from the CELEX corpus (Baayen et al., 1995; Shaoul and Tomaschek, 2013). We then created a list that contained all words that we used in our experiments; all singulars, plurals and diminutives. We then used the data set to calculate, for each word in our experiment, how many neighbors each had by using (Hall et al., 2015). For each word in our experiment we counted as neighbor each word in the data set that differed by one phoneme from the experimental word (Vitevitch and Luce, 1999). [3] For example, we found that *Krug* 'mug' has 4 neighbors: *Krugs* 'mug GEN', *trug* 'bear PST', *klug* 'smart' and *Krieg* 'war'. We then used the density in a regression analysis. The density of plurals and singulars is higher than the density of diminutives, but other than that the density are comparable. It is therefore unlikely that differences in neighborhood density among our words have contributed much to our results.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept = Diminutive Type 1)	2.97	1.44	2.06	0.04
Plural	13.67	2.04	6.70	0.00
Singular	11.93	2.04	5.85	0.00
type 2	-1.90	2.04	-0.93	0.35
type 3	-2.47	2.04	-1.21	0.23
Plural × type 2	3.33	2.88	1.16	0.25
Singular × type 2	4.03	2.88	1.40	0.16
Plural × type 3	4.93	2.88	1.71	0.09
Singular × type 3	2.87	2.88	0.99	0.32

APPENDIX III. MATERIAL

[3] There are other methods of establishing neighborhoods (Gahl and Strand, 2016), and we tried them but our results remain the same.

Table 15. Target type 1 items

type 1 targets					
Item	translation	exp.1 word (sg)	exp. 2 word (pl)	exp. 1 nonce	exp. 2 nonce
Gans	*goose*	gans	gɛnzə	gɛns	ganzə
Hand	*hand*	hant	hɛndə	hɛnt	handə
Luft	*air*	lʊft	lʏftə	lʏft	lʊftə
Saft	*juice*	zaft	zɛftə	zɛft	zaftə
Wand	*wall*	vant	vɛndə	vɛnt	vandə
Wolf	*wolf*	vɔlf	vœlfə	vœlf	vɔlfə
Bart	*beard*	baɐt	bɛɐtə	bɛɐt	baɐtə
Korb	*basket*	kɔɐp	kœɐbə	kœɐp	kɔɐbə
Marsch	*march*	maɐʃ	mɛɐʃə	mɛɐʃ	maɐʃə
Wurf	*throw*	vʊɐf	vʏɐfə	vʏɐf	vʊɐfə
Ball	*ball*	bal	bɛlə	bɛl	balə
Bauch	*stomach*	baʊx	bɔʏçə	bɔʏç	baʊxə
Baum	*tree*	baʊm	bɔʏmə	bɔʏm	baʊmə
Damm	*dam*	dam	dɛmə	dɛm	damə
Fuss	*foot*	fuːs	fyːsə	fyːs	fuːsə
Fluch	*curse*	fluːx	flyːçə	flyːç	fluːxə
Kauz	*fogey*	kaʊts	kɔʏtsə	kɔʏts	kaʊtsə
Krug	*jug*	kʁuːk	kʁyːgə	kʁyːk	kʁyːgə
Laus	*louse*	laʊs	lɔʏzə	lɔʏs	laʊzə
Maus	*mouse*	maʊs	mɔʏzə	mɔʏs	maʊzə
Rang	*rank*	ʁaŋ	ʁɛŋə	ʁɛŋ	ʁaŋə
Zahn	*tooth*	tsaːn	tsɛːnə	tsɛːn	tsaːnə
Saal	*hall*	zaːl	zɛːlə	zɛːl	zaːlə
Sack	*sack*	zak	zɛkə	zɛk	zakə
Saum	*seam*	zaʊm	zɔʏmə	zɔʏm	zaʊmə
Schwan	*swan*	ʃvaːn	ʃvɛːnə	ʃvɛːn	ʃvaːnə
Schwung	*momentum*	ʃvʊŋ	ʃvʏŋə	ʃvʏŋ	ʃvʊŋə
Stall	*shed*	ʃtal	ʃtɛlə	ʃtɛl	ʃtalə
Zoll	*custom*	tsɔl	tsœlə	tsœl	tsɔlə
Zopf	*plait*	tsɔpf	tsœpfə	tsœpf	tsɔpfə

Table 16. Target type 2 items

type 2 targets					
Item	translation	exp.1 word (sg)	exp. 2 word (pl)	exp. 1 nonce	exp. 2 nonce
Docht	*wick*	dɔxt	dɔxtə	dœçt	dœçtə
Dolch	*dagger*	dɔlç	dɔlçə	dœlçə	dœlçə
Fund	*find*	fʊnt	fʊndə	fʏnt	fʏndə
Kult	*cult*	kʊlt	kʊltə	kʏlt	kʏltə
Mast	*mast*	mast	mastə	mɛst	mɛstə
Pult	*desk*	pʊlt	pʊltə	pʏlt	pʏltə
Salz	*salt*	zalts	zaltsə	zɛlts	zɛltsə
Takt	*beat*	takt	taktə	tɛkt	tɛktə
Luchs	*lynx*	lʊks	lʊksə	lʏks	lʏksə
Kurs	*class*	kʊɐs	kʊɐzə	kʏɐs	kʏɐzə
Boot	*boat*	boːt	boːtə	bøːt	bøːtə
Boss	*boss*	bɔs	bɔsə	bœs	bœsə
Bus	*bus*	bʊs	bʊsə	bʏs	bʏsə
Gas	*gas*	gaːs	gaːzə	gɛːs	gɛːzə
Huf	*hoof*	huːf	huːfə	hyːf	hyːfə
Kohl	*kale*	koːl	koːlə	køːl	køːlə
Pol	*pole*	poːl	poːlə	pøːl	pøːlə
Ruf	*call*	ruːf	ruːfə	ryːf	ryːfə
Schaf	*sheep*	ʃaːf	ʃaːfə	ʃɛːf	ʃɛːfə
Tag	*day*	taːk	taːgə	tɛːk	tɛːgə
Tod	*death*	toːt	toːdə	tøːt	tøːdə
Haar	*hair*	haːɐ	haːʁə	hɛːɐ	hɛːʁə
Pfad	*path*	pfaːt	pfaːdə	pfɛːt	pfɛːdə
Paar	*pair*	paːɐ	paːʁə	pɛːɐ	pɛːʁə
Tor	*gate*	toːɐ	toːʁə	tøːɐ	tøːʁə
Brot	*bread*	broːt	broːtə	brøːt	brøːtə
Knall	*bang*	knal	knalə	knɛl	knɛlə
Schluck	*swallow*	ʃlʊk	ʃlʊkə	ʃlʏk	ʃlʏkə
Stoff	*fabric*	ʃtɔf	ʃtɔfə	ʃtœf	ʃtœfə
Flur	*hall*	fluːɐ	fluːʁə	flyːɐ	flyːʁə

Table 17. Target type 3 items

type 3 targets					
Item	translation	exp.1 word (sg)	exp. 2 word (pl)	exp. 1 nonce	exp. 2 nonce
Fest	*celebration*	fɛst	fɛstə	fɔst	fɔstə
Film	*film*	fɪlm	fɪlmə	falm	falmə
Heft	*notebook*	hɛft	hɛftə	hʊft	hʊftə
Hirn	*brain*	hɪɐn	hɪɐnə	hʊɐn	hʊɐnə
Hirsch	*deer*	hɪɐʃ	hɪɐʃə	hɔɐʃ	hɔɐʃə
Keks	*cookie*	keːks	keːksə	kuːks	kuːksə
Kelch	*goblet*	kɛlç	kɛlçə	kalç	kalçə
Lift	*lift*	lɪft	lɪftə	laft	laftə
Pferd	*horse*	pfɛɐt	pfɛɐdə	pfʊɐt	pfʊɐdə
Wirt	*host*	vɪɐt	vɪɐtə	vʊɐt	vʊɐtə
Dieb	*thief*	diːp	diːbə	doːp	doːbə
Fett	*fat*	fɛt	fɛtə	fat	fatə
Fisch	*fish*	fɪʃ	fɪʃə	fɔʃ	fɔʃə
Kitz	*fawn*	kɪts	kɪtsə	kʊts	kʊtsə
Reiz	*stimulus*	ʁaɪts	ʁaɪtsə	ʁaʊts	ʁaʊtsə
Ring	*ring*	ʁɪŋ	ʁɪŋə	ʁʊŋ	ʁʊŋə
Sieb	*sieve*	ziːp	ziːbə	zoːp	zoːbə
Sinn	*sense*	zɪn	zɪnə	zʊn	zʊnə
Tisch	*table*	tɪʃ	tɪʃə	tɔʃ	tɔʃə
Pott	*mug*	pɔt	pœtə	pɛt	pɛtə
Schuss	*shot*	ʃʊs	ʃʏsə	ʃɛs	ʃɛsə
Sohn	*son*	zoːn	zøːnə	ziːn	ziːnə
Wall	*rampart*	vɑl	vɑlə	vʊl	vʊlə
Rock	*skirt*	ʁɔk	ʁœkə	ʁak	ʁakə
Zug	*train*	tsuːk	tsyːgə	tseːk	tseːgə
Stier	*bull*	ʃtiːɐ	ʃtiːʁə	ʃtoːɐ	ʃtoːʁə
Trieb	*instinct*	tʁiːp	tʁiːbə	tʁuːp	tʁuːbə
Zweig	*branch*	tsvaɪk	tsvaɪgə	tsvaʊk	tsvaʊgə
Blick	*gaze*	blɪk	blɪkə	blak	blakə
Brief	*letter*	briːf	briːfə	broːf	broːfə

Table 18. Filler words

sg form	translation	sg form	translation	sg form	translation	sg form	translation
Blech	*iron sheet*	Blitz	*lightning*	Gleis	*track*	Greif	*griffin*
Greis	*old man*	Knick	*bend*	Kreis	*circle*	Kreuz	*cross*
Krieg	*war*	Preis	*price*	Schlitz	*groove*	Schmied	*smith*
Schrein	*shrine*	Schwein	*pig*	Spiel	*game*	Spie§	*skewer*
Spitz	*spitz*	Steg	*runway*	Stein	*stone*	Stiel	*handle*
Stil	*style*	Stück	*piece*	Zweck	*purpose*	Gnom	*gnome*
Brauch	*custom*	Braut	*bride*	Draht	*wire*	Flug	*flight*
Fluss	*river*	Frosch	*frog*	Gruss	*greeting*	Klang	*sound*
Kloß	*dumpling*	Klotz	*brick*	Knopf	*button*	Pflock	*plug*
Pflug	*plough*	Plan	*plan*	Platz	*place*	Schlauch	*pipe*
Schluss	*end*	Schwamm	*sponge*	Span	*blade*	Spass	*joke*
Stab	*bar*	Stadt	*city*	Stamm	*stem*	Stock	*stick*
Stoß	*kick*	Stuhl	*chair*	Traum	*dream*	Zwang	*bondage*
Schwur	*vow*	Gen	*gene*	Keil	*liner*	Pfeil	*arrow*
Busch	*bush*	Hut	*hat*	Pass	*passport*	Beet	*bed*
Beil	*axe*	Bein	*leg*	Bier	*beer*	Biss	*bite*
Deich	*dyke*	Ding	*thing*	Fell	*coat*	Föhn	*blow-dryer*
Gel	*gel*	Hain	*grove*	Hecht	*pike*	Heer	*army*
Heim	*home*	Hieb	*flourish*	Keim	*germinal*	Laib	*loaf*
Meer	*sea*	Netz	*net*	Reck	*high bar*	Reim	*rhyme*
Riss	*crack*	Scheich	*sheik*	Schein	*glint*	Schiff	*ship*
Sieg	*victory*	Speer	*javelin*	Teich	*pond*	Teig	*dough*
Teil	*part*	Tier	*animal*	Weg	*way*	Wein	*wine*
Zeug	*stuff*	Ziel	*goal*	Bass	*bass*	Fang	*catch*
Gaul	*horse*	Guss	*shower*	Hang	*slope*	Haut	*skin*
Hof	*court*	Kahn	*barge*	Kamm	*comb*	Koch	*chef*
Kopf	*button*	Lauf	*run*	Lohn	*wage*	Naht	*fissure*
Nuss	*nut*	Pfahl	*pale*	Rat	*advice*	Raum	*room*
Satz	*sentence*	Schopf	*tuft*	Ton	*tone*	Topf	*pot*
Zaun	*fence*	Chor	*choir*	Jahr	*year*	Recht	*law*
Vers	*verse*	Mönch	*monk*	Berg	*mountain*	Feind	*enemy*
Gift	*poison*	Gips	*cement*	Helm	*helmet*	Herd	*stove*
Kerl	*fellow*	Kern	*nucleus*	Nerz	*mink*	Pelz	*fur*
Pilz	*mushroom*	Rest	*rest*	Schelm	*rascal*	Scherz	*joke*
Schild	*shield*	Schirm	*umbrella*	Term	*term*	Werk	*work*
Wert	*value*	Wicht	*goblin*	Wind	*wind*	Wink	*cue*
Witz	*joke*	Zelt	*tent*	Dachs	*badger*	Halm	*blade*
Hund	*dog*	Farn	*brake*	Molch	*newt*	Mond	*moon*
Barsch	*perch*	Garn	*twine*	Gurt	*belt*	Hort	*hoard*
Lurch	*amphibian*	Mord	*murder*	Dampf	*steam*	Duft	*smell*
Faust	*fist*	Fuchs	*fox*	Gast	*guest*	Hals	*neck*
Kampf	*battle*	Lust	*desire*	Macht	*power*	Magd	*maidservant*
Nacht	*night*	Rumpf	*body*	Schacht	*chamber*	Schaft	*bootleg*
Schatz	*treasure*	Sucht	*addiction*	Sumpf	*swamp*	Tanz	*dance*
Wunsch	*wish*	Darm	*bowel*	Sarg	*coffin*	Turm	*tower*