THE MENTAL REPRESENTATION OF INFLECTED WORDS:
AN EXPERIMENTAL STUDY OF ADJECTIVES AND VERBS IN GERMAN

HARALD CLAHSEN  SONJA EISENBEISS

University of Essex  Max Planck Institute for Psycholinguistics

MEIKE HADLER  INGRID SONNENSTUHL

University of Düsseldorf

The authors investigate how morphological relationships between inflected word forms are represented in the mental lexicon, focusing on paradigmatic relations between regularly inflected word forms and relationships between different stem forms of the same lexeme. We present results from a series of psycholinguistic experiments investigating German adjectives (which are inflected for case, number, and gender) and the so-called strong verbs of German, which have different stem forms when inflected for person, number, tense, or mood.

Evidence from three lexical-decision experiments indicates that regular affixes are stripped off from their stems for processing purposes. It will be shown that this holds for both unmarked and marked stem forms. Another set of experiments revealed priming effects between different paradigmatically related affixes and between different stem forms of the same lexeme.

We will show that associative models of inflection do not capture these findings, and we explain our results in terms of combinatorial models of inflection in which regular affixes are represented in inflectional paradigms and stem variants are represented in structured lexical entries. We will also argue that the morphosyntactic features of stems and affixes form abstract underspecified entries. The experimental results indicate that the human language processor makes use of these representations.*

1. MORPHOLOGY AND THE MENTAL LEXICON. Much psycholinguistic research has been devoted to the question of whether there is any correspondence between the linguistic structure of a morphologically complex word and the way it is segmented by the speaker-hearer during on-line production and comprehension. Are morphologically complex words that have stem + affix representations (e.g. derived words such as govern-ment or regularly inflected words such as walk-ed) computed via their constituent morphemes? Are irregularly inflected words that cannot be formed through affixation stored unanalyzed in the mental lexicon?

Experimental studies have produced conflicting answers to these questions. Two broad views can be distinguished. Associative models of morphological processing claim that the morphological structure of words plays no role in the way they are produced or perceived and that words are listed as full forms in memory. The key idea is that all morphological patterns, including those that can be decomposed into stems, roots, and affixes, are derived from a network of associative relations. Connectionist networks of inflection (Rumelhart & McClelland 1986, MacWhinney & Leinbach 1991, Plunkett & Marchman 1993, among others) can be seen as modern implementations of associative models of language. In contrast to this, other researchers have argued that the mental lexicon encodes morphological structure and that this information plays a role in comprehension and production. Specifically, the language processor is said to make use of morphological decomposition for dealing with morphologically complex words, in addition to full-form representations (e.g. Laudanna & Burani 1985, 1995, Frauenfelder & Schreuder 1992, Schreuder & Baayen 1995, Pinker & Prince 1991).

* The research reported in this article is supported by the Deutsche Forschungsgemeinschaft (DFG) (German Research Council), DFG grant SFB 282/C7. We thank the members of our research group—Joana Cholin, Kerstin Mauth, and Axel Huth—for assistance in administering the experiments. We are also grateful to Martin Atkinson for a close reading of the article.
A related question that has received much less attention from psycholinguists is how morphological relationships between inflected word forms are represented in the mental lexicon. Consider, for example, the inflected adjectives from German in 1a, which are marked for case, number, and gender, and the inflected verb forms in 1b, which are marked for person, number, and tense.

(1) a. wild-es wild-em wild-er
   ‘wild-NOM.NEUT.SG.’ ‘wild-DAT.MASC.SG.’ ‘wild-NOM.MASC.SG.’

b. (ich) werf-e (du) wirf-st (sie) warf-en
   ‘(I) throw-1SG.PRES.’ ‘(you) throw-2SG.PRES.’ ‘(they) throw-3PL.PRET.’

Both adjectival and verbal agreement affixation are highly regular: -s, -m, and -r can be attached to any adjective and -e, -st, and -n to any verb. Exceptions are a few high frequency forms, such as the suppletive form bin (‘am’ 1 SG. PRES.) and first and third person singular present tense forms of verbs such as müssten ‘to be able to’, dürften ‘to be allowed to’, sollen ‘to be about to’, which (like their corresponding preterite forms) do not have a person/number agreement suffix: ich muss, darf, soll ‘I/he must, may shall’. Thus, decompositional accounts of morphological processing predict that all the inflected forms in 1 are segmented into stems and affixes, yielding two sets of affixes (one for adjectives, one for verbs) and two sets of stems (one invariant adjective stem, wild, and three different stems of the lexeme werfen). This raises the question of how the morphological relationships between affixes such as those in 1a and the relationships between different stem variants such as those in 1b are represented in the mental lexicon and how they are processed.

We can distinguish between two principal views on these questions. In associative models of inflection, in which all inflected word forms are represented in terms of associative networks, morphological relationships such as those in 1 are not directly encoded in the mental lexicon. Rather, they are treated as epiphenomena of associations between contiguous or otherwise similar linguistic elements (see e.g. Bybee 1991, 1995a, MacWhinney et al. 1989, Taraban et al. 1989, Elman et al. 1996). For example, the final segments -s, -m and -r of German adjectives may form an associative pattern in that they share sets of phonological and semantic connections that are repeated across multiple sets of words. Under this account, the affixal properties of these linguistic elements as well as the paradigmatic relations between them are said to emerge from a network of associative patterns without encoding any morphological structure. Similarly, the different stem variants illustrated in 1b might be separately stored and associatively related, with respect to phonological and semantic properties. If this is correct, there would be no need for morphological representations and operations in the mental lexicon. Different variants of associative models of inflection have been suggested in the literature; see §2 for discussion. Here, our focus will be on the questions of how associative models deal with morphological relationships, specifically with related regular affixes and with stem variants of the same lexeme.

The alternative view holds that morphological structure and morphological relationships are encoded in the mental lexicon and are important for understanding on-line morphological processing (see e.g. Marslen-Wilson et al. 1994, Pinker 1999, Clahsen 1999 for review). For many morphologists, regular affixes such as those in 1 constitute inflectional paradigms. In psycholinguistic terms, one can think of a paradigm as a matrix or access system for mapping grammatical information, i.e. morphosyntactic features, to their exponents or affixes. The question, then, is whether there is any empirical evidence that the mental lexicon is organized into paradigms or whether paradigms are epiphenomena that reflect similarities between word forms in meaning and form. A related question concerns the representation of properties that define para-
digms: Are affixes fully specified for their morphosyntactic features, or are they underspecified for redundant features? Linguists have also proposed mechanisms to account for the relationships between stem variants such as those in 1b, e.g. lexical redundancy rules (Chomsky 1970, Jackendoff 1975) and default inheritance representations (Corbett & Fraser 1993, Wunderlich 1996). The purpose of these mechanisms is to capture relationships between lexical items and at the same time to permit the lexicon to avoid listing redundant parts of the word. We will address the question of whether there is any psycholinguistic evidence that lexical representations for stem variants (as well as for affixes) are ‘economical’, in the sense that they are underspecified for redundant features.

In what follows, we will first discuss how different models of the mental lexicon deal with morphological relationships between regular affixes and relationships between stem variants of the same lexeme. We will then report results from two experiments on German adjective inflection examining the processing of regular affixes, and three experiments examining the so-called strong verbs of German to determine how stem variants are represented in the mental lexicon.

2. ASSOCIATIVE MODELS OF INFLECTION. In associative models of inflection, a word’s morphological structure plays no direct role in the way it is produced or perceived. Inflected words are all represented in the same way, that is, by storing them in associative networks and by creating connections among them. Through repeated exposure to many sets of inflected words, an associative network will form patterns that range over sets of connections. In this way, morphological relationships are claimed to be secondary, as they are derivable from associations between words. A well-known example are the so-called SATELLITE MODELS in which word nodes are posited for each morphological variant of a given lexeme, and bidirectional connections between each word node and a corresponding base form, most typically the stem (Lukatela et al. 1978, 1980, 1987, Fowler et al. 1985, Feldman & Fowler 1987). Applying this approach to the inflected word forms in 1 would yield lexical entries such as those in 2 which are composed of a nucleus, the stem in these cases, and satellites representing all other forms.

(2) a.

b.
This model posits a special status to the base form or nucleus. Some experimental evidence supports this view. Lukatela et al. (1978, 1980) showed, for example, that Serbo-Croatian nouns in oblique case forms (instrumentals and datives) elicit longer response times in visual lexical-decision tasks than nominative forms. In their model, the latter are regarded as nucleus forms and the oblique case forms as satellites, i.e. full-form representations connected to the nucleus. Shorter lexical decision times for base forms have also been obtained for other categories and for other languages (see Günther 1988). These results have been taken to indicate that the lexical identity of a word is tied to the nucleus, and that if access is made via a satellite form, extra time is required. Moreover, the inflectional variants of a lexeme are connected to the nucleus, but not necessarily connected to each other. Thus, if this model is correct, we would expect corresponding experimental effects, for example, frequency effects for inflected word forms, rather than for affixes and stems. Furthermore, this model does not predict any experimental differences between the various satellites of a nucleus.

In more recent associative models of inflection, the distinction between a designated base form and inflectional variants has been given up in favour of the view that all word forms are stored in associative memory irrespective of their morphological constituency. This view has been taken in most connectionist models of inflection (see Elman et al. 1996) as well as in schema-based models of inflection (Bybee 1995a, Köpcke 1993, 1998). Inflectional paradigms are claimed to ‘emerge on the basis of associations between cues’ (MacWhinney et al. 1989:274). Along the same lines, Bybee (1995b: 242ff.) says that paradigms should be represented as ‘clusters of highly connected words’. Under this view (see also Taraban et al. 1989), a paradigm is a set of word forms connected more or less strongly in various ways (phonologically, semantically) to other word forms that need not necessarily be members of the same paradigm, in traditional terms. The strength of the connections is largely determined by frequency and similarity. Hence, the nature of the connections between two phonologically related word forms, e.g. car and card, does not differ in any fundamental way from the connections between morphologically related word forms such as those in 1a and 1b. For illustration, consider the associative representation of inflected word forms of the verb werfen in 3.
As illustrated in 3, this model does not attribute any special status to base forms or nuclei, and all inflectional variants (including the base form) have word-level status. Hence we would expect to find word-form frequency effects for all inflectional variants of a given word. For example, if there are processing differences between inflected forms—*wildes* and *wildem* for example—these should be due to different word-form frequencies. Moreover, we should find similarity effects for all inflected word forms. That is, pairs of word forms that exhibit a high degree of formal (= orthographic and/or phonological) or semantic similarity are more closely connected and should therefore exhibit stronger priming effects than pairs of word forms that are less similar in their formal and semantic properties.

3. **Combinatorial approaches to inflection.** In opposition to associative models of the mental lexicon, several other researchers have argued that morphologically complex words are decomposed or parsed, i.e. segmented into smaller morphological units (stems, roots, affixes, e.g.). This view has been supported by a rich body of psycholinguistic literature demonstrating experimental effects of morphological decomposition for derivational morphology and regular inflection (see Marslen-Wilson et al. 1994, Pinker 1999, Clahsen 1999 for review). Consider, for example, results from morphological priming studies. For (semantically and phonologically transparent) derived words, Marslen-Wilson and colleagues (1994) found that nouns such as *govern* and *government* both prime their corresponding form *govern*, but that the two derived nouns do not prime each other. To account for these findings, they proposed decomposed, that is, morphemically based, representations of derived nouns in terms of stems and affixes, such as in 4.

\[
\begin{array}{c}
govern \\
\downarrow \\
\text{-ment} \\
\uparrow \\
\text{-or}
\end{array}
\]

The arrows in 4 are meant to indicate inhibitory links between the two suffixes to capture the lack of priming between -*or* and -*ment* forms of the same stem. Thus, *govern* + *ment* activates the stem *govern*, but at the same time inhibits other derived forms of the same stem. Marslen-Wilson et al. (1994) attribute this to lexical competition: -*ment* and -*or* produce lexical items with distinct meanings that are incompatible with each other, and hence they inhibit each other in priming tasks.

Similar issues arise for regular inflection. Suppose that regularly inflected words are indeed decomposed into stems and affixes (see Stanners et al. 1979, Kempley & Morton 1982, Fowler et al. 1985, Napps 1989, Marslen-Wilson et al. 1993, Sonnenstuhl et al. 1999). This then raises the question of how morphologically related affixes of the same stem are to be represented in the mental lexicon. In most morphological frameworks the relationships between regular affixes such as those in 1 are described in terms of paradigms. An inflectional paradigm is a multidimensional, potentially recursive matrix defined by the morphosyntactic features of word forms or affixes (see Aronoff

---

1 Note, however, that the status of inflectional paradigms is controversial among linguists. There are theories that are based on paradigms (Stump 2001, Wunderlich 1996), others that countenance paradigms (Anderson 1992), and yet others that deny their existence altogether (Bybee 1985, Lieber 1992, Halle & Marantz 1993, and associative models in general).

2 A related linguistic controversy concerns the status of affixes. Wunderlich (1996) and Jackendoff (1997), for example, believe that regular affixes represent lexical entries, and consequently, paradigms are considered to be affix-driven, directly constituted by the ‘combinatory force of the inflectional affixes’ (Wunderlich
A paradigm contains a set of slots defined in terms of morphosyntactic feature values and shows how each slot is to be filled. The result is that any lexeme that belongs to a particular syntactic category has inflected word forms defined by the paradigm. The formation of paradigms is constrained by general principles, such as blocking and specificity (Kiparsky 1982, 1998), and by paradigm structure constraints, such as completeness and uniqueness (Pinker 1984, Wunderlich 1996). Completeness requires that every cell of a paradigm must be occupied, and uniqueness that every cell is uniquely occupied. Blocking and specificity require that if two rules or affixes are in competition for one paradigm slot, the rule that is more specific in its application is preferred over the more general one.

If this is how regular affixes are organized in the mental lexicon, we would expect to find corresponding experimental effects, for example, priming differences between specific and less specific affixes (or rules). Moreover, the idea that regular affixes are represented in inflectional paradigms means that affixes such as those in 1 (or equivalently corresponding realization rules) are stored independently of the stems on which they occur. Hence, inflected word forms containing regular affixes should be decomposed into stem and affix, irrespective of the kind of stem on which they occur. Finally, if regular affixes are stored in the mental lexicon, we would expect them to produce experimental effects characteristic of lexically stored items, for example frequency effects. We examine these predictions in the experiments reported in §4.

With respect to the representation of stem variants such as those in 1b, linguists have developed the notion of lexical redundancy rules (Chomsky 1970) or equivalent mechanisms, e.g. lexical rules (Jackendoff 1975, 1997) or default inheritance hierarchies (Corbett & Fraser 1993, Wunderlich 1996), to capture the morphological relationships between them. The key idea common to these approaches is that stem variants of the same lexeme are not separately and completely listed, but that some of them have an impoverished (underspecified) entry (e.g. wirf and warf) and that when these items are used, their full interpretation and form is filled in from the base entry (= werf). In other words, stem variants may form subnodes within hierarchically structured lexical entries. If this is how stem forms are stored in the mental lexicon, we would expect corresponding experimental effects, for example, stem frequency effects in lexical decision experiments and priming differences between the various stem variants. These predictions were examined in the experiments reported in §5.

4. EXPERIMENTS ON ADJECTIVE INFLECTION. In this section, we will investigate evidence from lexical decision and priming tasks on German adjective inflection in light of the theoretical controversy between associative and combinatorial models of inflection. We begin with a brief description of adjective inflection in German.

4.1. ADJECTIVE INFLECTION IN GERMAN. German attributive adjectives carry a portmanteau affix that expresses the grammatical features gender, number, and case. With respect to the morphological expression of these features, two declension classes are commonly distinguished, weak and strong declension as shown in 5; see e.g. Wunderlich 1987. Adjectives that are used without a determiner or a demonstrative are combined with an uninflected determiner carrying a strong affix, e.g. (ein) kalter Wein ‘(a) cold
wine’, while adjectives that are combined with a strongly inflected determiner take an affix from the weak paradigm, e.g. der kalte Wein ‘the cold wine’, mit dem kalten Wein ‘with the cold wine’, mit einem kalten Wein ‘with a cold wine’.

(5) a. Strong adjective declension

<table>
<thead>
<tr>
<th>Region</th>
<th>SINGULAR</th>
<th>PLURAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOMINATIVE</td>
<td>-r</td>
<td>-e</td>
</tr>
<tr>
<td>ACCUSATIVE</td>
<td>-n</td>
<td>-e</td>
</tr>
<tr>
<td>DATIVE</td>
<td>-m</td>
<td>-n</td>
</tr>
<tr>
<td>GENITIVE</td>
<td>-n</td>
<td>-n</td>
</tr>
</tbody>
</table>

b. Weak adjective declension

<table>
<thead>
<tr>
<th>Region</th>
<th>SINGULAR</th>
<th>PLURAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOMINATIVE</td>
<td>-e</td>
<td>-e</td>
</tr>
<tr>
<td>ACCUSATIVE</td>
<td>-n</td>
<td>-e</td>
</tr>
<tr>
<td>DATIVE</td>
<td>-n</td>
<td>-n</td>
</tr>
<tr>
<td>GENITIVE</td>
<td>-n</td>
<td>-n</td>
</tr>
</tbody>
</table>

As is clear from 5, the affixes differ with respect to the degree of homonym. Both, -e and -n occur in the strong and weak declension, while -s, -r, and -m occur only in the strong declension. Several linguists (Bierwisch 1967, Zwicky 1986, Blevins 1995, 2000, Wunderlich 1997) have analyzed this system using morphological paradigms. In these accounts, morphosyntactic properties are represented in terms of binary features with marked ( = positive) and unmarked ( = negative) values. The distribution of these feature values is governed by a morphological blocking or specificity condition according to which forms with positive feature values take precedence over unmarked forms. Moreover, in order to reduce lexical redundancy, the distinct forms of a paradigm are proposed as having underspecified entries, that is, minimally specified analyses. A common assumption is that only positive feature values are directly specified, while negative feature values are introduced into paradigms in opposition to a marked positive value on the same feature (Wunderlich 1997, Blevins 2000). Plural forms of nouns, for example, are said to be positively specified for number, i.e. [+PL], whereas singular nouns are not directly specified; instead, a noun receives the unmarked singular interpretation, i.e. [−PL], by its paradigmatic opposition to the positively specified forms.

In Bierwisch’s (1967) and Blevins’s (2000) accounts of German adjective inflection, the forms -e and -s have only negative feature specifications, while -m has two positive features, as shown in 6. Thus, in these analyses the -m affix comes out as the most specific form, and -s is specified for more features than -e, indicating that -s is more specific in terms of its feature content than -e (see also Cahill & Gazdar 1997).

(6) -e -s -m

- [−PL] [−PL]
- [−FEM] [−FEM]
- [−MASC] ~
- [−OBL] [−OBL] [+OBL]
- ~ [+DAT]

Wunderlich’s (1997) account relies on specifications of the structural cases in terms of a hierarchy of theta roles as shown in 7 for the four cases of German. Here [+hr]
stands for ‘there is a higher role’ indicating that this case links to a lower theta role, and [+lr] stands for the reverse; [+nominal] means that this case links to the theta role of (relational) nouns.

(7) NOMINATIVE: [ ]
ACCUSATIVE: [+hr]
DATIVE: [+hr, +lr]
GENITIVE: [+hr, +nominal]

Using this system, a comparison of the adjective affixes in 6 reveals that in contrast to all other forms, the -m affix of the strong declension paradigm is restricted to just one case, dative, whereas -s, for example, occurs in nominatives and accusatives. Moreover, in terms of 7, dative requires two features for its specification, whereas nominative is completely unspecified and accusative has just one feature. Thus Wunderlich’s morpheme-based account, like Bierwisch’s and Blevins’s analyses, yields -m as a more specific adjective ending than -e and -s.

To summarize, despite differences among the linguistic treatments of German adjective inflection mentioned above, there seems to be agreement that the -m affix is paradigmatically more specified than any of the other adjective forms, and that the feature content of -s is more specific than that of -e. Note that this also fits with frequency differences between the various adjective forms. Schriefers et al. (1992:376) observed that ‘-e is the suffix with the highest frequency in the inflectional system of German adjectives’, and this corresponds to its low paradigmatic specificity. As far as the adjective affixes -m and -s are concerned, we compared their type and token frequencies in the CELEX database. On both counts, -s comes out considerably more frequent than -m (-m, types/tokens: 1,264/5,965; -s, types/tokens: 2,286/12,828). Hence, the lower paradigmatic specificity of -s relative to -m corresponds to higher frequencies of use.

The empirical question we will address here is whether the linguistic differences among the various adjective forms have any effect on morphological processing. If this is the case, we would expect to find two experimental effects: specific affixes should produce longer lexical-decision times and should be more difficult to prime than less specific affixes. Consequently, in a word/nonword (lexical) decision task, adjectives inflected with -m should exhibit longer response times than the same adjectives inflected with -s. This is because -m is the more specific form, and the mapping of the form to its corresponding feature bundle is likely to cause a longer lexical search. Moreover, we would expect to find that in a priming task adjectives inflected with -m should be less effectively primable than adjectives inflected with -s, and adjectives with -s less effectively than those with -e. This is because forms affixed with -m require the processing of specific, i.e. unprimed, features that should lead to longer response times. Similarly, -s adjective forms contain features that cannot be primed by -e adjectives, whereas -e adjective forms do not have any features that could not be primed by -s. Therefore, if the morphological feature content of these adjective forms matters for processing, we would expect to find corresponding asymmetries between -s and -e adjective forms in priming.

4.2. EXPERIMENT 1: VISUAL LEXICAL DECISION TASK ON INFLECTED ADJECTIVES. The experimental paradigm employed here was a word/nonword discrimination task with
reaction time (RT) as the dependent variable. Lexical decision times on noninflected simplex words have consistently been shown to be affected by word frequency: subjects take less time to decide that high-frequency items are existing words than they do for low-frequency items (see Balota 1994 for review). This is conceived of as a memory effect: as memory traces get stronger with additional exposures, high-frequency entries can be more readily accessed than low-frequency ones. We have used this task to examine the processing of inflected adjectives and to test for potential differences between specific and less specific affixes in their lexical decision times.

We examined adjectives inflected with -m and -s. Recall that the latter form is approximately twice as frequent as the former. However, despite the overall low frequency of the -m affix compared to -s, some adjectives appear more often with -m than with -s; this is, for example, the case for ruhig ‘quiet’, as shown in 8. Compare this with an -s dominant adjective such as rein ‘pure’, which appears more often with -s than with -m, also shown in 8.

(8)  

\[
\begin{array}{cccc}
\text{stem form} & \text{-m} & \text{-s} & \text{stem form} \\
\text{ruhig} & 838 & 51 & 13 \\
\text{rein} & 783 & 14 & 38 \\
\end{array}
\]

By using -m dominant and -s dominant adjectives with both -m and -s in a lexical-decision task, the potential effects of affix frequency and word-form frequency can be teased apart. If inflected adjectives are stored as wholes, we would expect to find word-form frequency effects. Ruhigem for example, should produce shorter lexical-decision times than ruhiges, whereas reinem should produce shorter response times than reinem. But if adjectives are decomposed into stems and affixes, word-form frequencies should be irrelevant, and we would instead expect to find AFFIX-frequency effects, i.e. adjective forms with -s should produce shorter response times than -m adjective forms (where stem frequency is held constant), because -s is the more frequent affix.

The design of this experiment involves four conditions shown in 9 that allow us to compare response times to -m and -s forms of adjectives and to tease apart word form from affix-frequency effects.

(9)  

<table>
<thead>
<tr>
<th>LOW AFFIX FREQUENCY</th>
<th>LOW WORD-FORM FREQUENCY</th>
<th>HIGH WORD-FORM FREQUENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>-s dominant adjective</td>
<td>-m dominant adjective</td>
<td>-s dominant adjective</td>
</tr>
<tr>
<td>with -m</td>
<td>with -m</td>
<td>with -s</td>
</tr>
</tbody>
</table>

We predict that -s forms of adjectives will produce shorter lexical decision times than -m forms, because -m is the most specific adjective form in the paradigm and should therefore require a longer lexical search than the less specific -s form. Associative models of inflection, in contrast, would predict word-form-frequency effects, but no affix-frequency effects. Hence, adjective forms with high word-form frequencies should produce shorter response times than the same adjectives with low word-form frequencies. Differences in affix frequency, however, should not affect lexical-decision times.

MATERIALS

We selected eighteen -m dominant and eighteen -s dominant adjectives, which were matched for stem frequency. Each of these adjectives was presented with -s and with -m. Because no participant should see the same adjective more than once, two experimental
versions were constructed. The eighteen -m dominant and eighteen -s dominant adjectives were divided into two groups, matched for mean stem and word-form frequency (see Appendix A1). The two inflectional variants of the two types of adjectives were then distributed over two experimental versions in a Latin-square design (Winer 1971). Each version included eighteen -m dominant adjectives (nine of them inflected with -m and nine with -s) and eighteen -s dominant adjectives (nine combined with -m and nine with -s).

In addition to the 36 experimental items, we constructed 140 filler items (52 words and 88 pseudowords), which should obscure the regularities in the test items. The list of word fillers consisted of adjectives inflected with the affixes -s, -m, -e, and -r so that each affix appeared equally often within the total list of 88 word stimuli. The pseudoword fillers were created from real words by replacing two adjacent letters in varying positions. The real words used as a basis for the pseudowords showed the same distribution of inflectional affixes as the 36 test items and 52 word fillers. The entire stimulus set for each experimental version consisted of 176 items. The two lists were pseudorandomized, with the same order of test and filler items in both lists.

METHOD

Thirty native speakers of German, all of them students at the University of Düsseldorf (mean age: 28) were paid for their participation in the experiment. None of them participated in more than one experimental version.

Each trial consisted of the presentation of a fixation point in the middle of a computer screen in front of the participant, followed after 1,000 ms. by the stimulus at the same position. The stimuli were presented on a seventeen-inch computer monitor in white letters (Arial 24 pt.) on a dark background and remained on the screen for 200 ms. The participants reacted by pressing a green button (for a word) or a red button (for a pseudoword) on a dual box. The green button was on the right side for right-handed and on the left side for left-handed participants. The next trial was initiated 1,200 ms. after the response.4

Written instructions with a detailed description of the task and examples of inflected adjectives and pseudo-adjectives were given to the participants before the experiment. The experiment itself started with a short practice phase, after which the participants were given the opportunity to ask any remaining questions about the procedure. There were no further breaks during the experiment. The overall duration of an experimental session was about ten minutes.

Errors—nonword responses to existing words and word responses to nonwords—as well as extreme reaction-time values were removed from the data set before any further analysis of the reaction times.5 Incorrect responses and outliers made up less than 8.3% of the data set. Errors were analyzed separately. For the reaction time analysis we computed means for each subject and each item in each condition. These means were entered into two separate ANOVAs for subjects (F1) and for items (F2), with the factors Dominance (-m dominant vs. -s dominant) and Affix (-s vs. -m).

4 For all experiments reported in this paper, the presentation of the stimuli and the measuring of the reaction times were controlled by the NESU software package (Baumann et al. 1993).

5 In all our experiments, we used the same criterion for outliers. For each condition, extreme reaction times, those that exceeded two standard deviations of the subject’s mean in this condition, were removed prior to any further analysis of the reaction time data.
RESULTS

The mean reaction times for the four experimental conditions are shown in Fig. 1. Fig. 1 shows that adjectives inflected with -s consistently produced shorter lexical-decision times than adjectives co-occurring with -m, irrespective of whether word-form frequencies were -m dominant or -s dominant. This was also confirmed statistically. There was a significant main effect of Affix for subjects and for items (F1(1,29) = 29.55, p < .001, F2(1, 17) = 14.55, p = .001). Pairwise statistical comparisons using matched t-tests confirmed that reaction times for adjectives affixed with -s were significantly shorter than reaction times for adjectives with -m, both for -m dominant adjectives and for -s dominant adjectives. The results indicate that both -s and -m access their own representations in the mental lexicon and that (as predicted) accessing of -m produces longer response times than accessing of -s. We attribute this difference to the fact that -m is the more specific and less frequent form. Furthermore, there were no differences in the response times between -m dominant adjectives on the one hand and -s dominant adjectives on the other hand. Statistically, there was no significant effect of Dominance (F1(1,29) < .01, p = .980, F2(1,17) = .02, p = .880) and no significant Dominance × Affix interaction (F1(1,29) = .42, p = .524, F2(1,17) = .02, p = .897). This means that -m dominant adjectives such as ruhig did not produce significantly shorter reaction times when they were presented with -m than when they were presented with -s.

The error analysis revealed that adjectives co-occurring with -m produced significantly more incorrect responses (7.9%) than adjectives affixed with -s (2.0%); there was a significant main effect of Affix (F1(1,29) = 7.18, p = .012; F2(1,17) = 29.21, p = .001).
However, error rates did not differ significantly with respect to Dominance, that is -m dominant adjectives produced almost the same error rate (5.0%) as -s dominant adjectives (5.3%). This pattern is consistent with the observed reaction times.

We thus found that word-form frequency did not affect lexical-decision times or error rates. This finding provides evidence against whole-word-based representations for inflected word forms such as those posited in associative models of inflection (see §2). Instead, reaction times and error rates were influenced by the form of the inflectional ending. The more specific and less frequent affix -m produced longer reaction times and higher error rates than the less specific and more frequent -s ending.

4.3. Experiment 2: Cross-modal Priming Task on Inflected Adjectives. While the results of the lexical-decision experiment provide evidence against an associative full-listing representation of inflected adjectives, they do not allow us to decide whether the processing advantage of the -s affix (over -m) is due to its higher frequency or whether it is caused by its lower paradigmatic specificity. To further investigate the mental representation of inflected adjectives, we used the cross-modal priming paradigm. As will become clear, results from priming experiments are less directly affected by frequency differences than results from the lexical decision task.

In a priming experiment, subjects are asked to perform a word/nonword discrimination, just as in a lexical-decision task, but the context in which the target stimuli are presented is manipulated. Consider, for example, the items in 10. In the experimental condition, a word is presented that is morphologically related to the target, e.g. happiness, followed by the presentation of the target word (happy) to which subjects have to make a word/nonword lexical decision. The response times to the target under this primed condition are then compared with those to an unprimed control condition, happy preceded by careful, for example, and/or with those to an identical repetition of happy. The effect may be facilitative or inhibitory.

We adopted the cross-modal immediate repetition priming paradigm (Marslen-Wilson et al. 1993, 1994) in which subjects hear a spoken prime immediately followed by a visually presented target form for which they make a word/nonword decision. This technique has three main advantages. First, since the task is cross-modal, any priming effects are likely to be due to the lexical representations themselves, rather than to effects of modality-specific access procedures. Second, since all targets are presented immediately at the offset of the prime, the task is likely to tap on-line processes of morphological priming, while unwanted effects of episodic memory are reduced. Third, results from priming experiments are less directly affected by frequency than results from lexical-decision tasks. This is because in a priming experiment we are not directly comparing two different inflected word forms. Instead, morphological priming effects are measured within target sets, that is by comparing response times to the same targets after the presentation of different primes. Thus, in contrast to a lexical-decision task in which response times to -m and -s adjectives have to be compared directly, we can examine priming effects separately for -m and for -s adjective forms. In this way, potential effects of the morphological differences between -m and -s forms can be studied independently of frequency differences.
The priming conditions are illustrated in 11 for the adjective *kariert* ‘chequered’. Each adjective occurred in three morphological variants (*-e, -m, -s*), both as a prime and as a target. Priming condition I provided the baseline for each of the three target forms, identical repetition. In priming conditions II and III, different paradigmatically related adjective forms were presented, with condition III containing more specific forms than those of condition II. Compare, for example, priming conditions II-\textit{e} and III-\textit{e}, with \textit{-e} forms as targets and \textit{-s} versus \textit{-m} forms as primes and recall (§4.1) that \textit{-m} is a paradigmatically more specific form than \textit{-s}.

\begin{tabular}{|c|c|c|}
\hline
\textbf{CONDITIONS} & \textbf{AUDITORY PRIMES} & \textbf{VISUAL TARGETS} \\
\hline
I-\textit{e} (Identity) & \textit{kariert-} & \textit{kariert-} \\
\hline
II-\textit{e} & \textit{kariert-} & \textit{kariert-} \\
\hline
III-\textit{e} & \textit{kariert-em} & \textit{kariert-em} \\
\hline
I-\textit{s} (Identity) & \textit{kariert-} & \textit{kariert-} \\
\hline
II-\textit{s} & \textit{kariert-em} & \textit{kariert-em} \\
\hline
III-\textit{s} & \textit{kariert-em} & \textit{kariert-em} \\
\hline
I-\textit{m} (Identity) & \textit{kariert-em} & \textit{kariert-em} \\
\hline
II-\textit{m} & \textit{kariert-} & \textit{kariert-} \\
\hline
III-\textit{m} & \textit{kariert-} & \textit{kariert-} \\
\hline
\end{tabular}

We expect the priming patterns to correspond to the feature specifications of the affixes involved. Given the highly specific feature content of \textit{-m}, we predict a Target Type effect, i.e. adjectives inflected with \textit{-m} should be less effectively primable than adjectives inflected with \textit{-s} or \textit{-e}. Adjectives inflected with \textit{-m} are specified for the positive features \([+\text{OBL}]\) and \([+\text{DAT}]\) which are not available from any of the primes. The presentation of \textit{-m} adjective forms as targets therefore requires the processing of unprimed features, and this should lead to longer response times. We also expect to find an interaction between prime type and target type. Priming conditions II and III should produce longer response times for the different targets than priming condition I because in conditions II and III, the feature specifications of the affixes in the primes differ from those of the targets, whereas in condition I they are identical.

\section*{MATERIALS}

Eighty-one monomorphemic adjective stems (see Appendix A2) were chosen, all of which had low stem frequencies of less than 100 in the CELEX database (Baayen et al. 1993). For each of these stem forms, nine prime-target pairs were constructed as shown in 11, resulting in 729 experimental prime-target pairs. To ensure that no participant sees the same target-stem more than once, nine experimental versions were constructed in a Latin-square design. The mean stem frequencies and the mean number of syllables were held constant for each of these nine versions. Thus each version included 81 different prime-target pairs (nine from each of the nine conditions shown in 11). No target appeared more than once in any version.

To keep the proportion of related pairs low and to deter the participants from developing strategies based on expectations about likely relations between primes and targets, 669 filler items were included in the experiment. We constructed 294 prime-target pairs in which primes and targets had different adjective stems, but were otherwise parallel to the experimental items, e.g. \textit{abstrakt-es} → \textit{festlich-e} ‘abstract → festive’. These items were included to counterbalance the experimental items in which both primes and targets had the same adjective stem. We also constructed 375 prime-target pairs in which the target adjective was a pseudoword. The pseudo-adjectives were constructed
by changing two or three letters of an existing adjective. In 50 of these distractor items, the stems used as primes and targets were formally similar, e.g. *anonymen* → *anonymstes*, *populäres* → *spopulärem*. The distractors were constructed in such a way that the different adjective endings (-e, -s, -m, -r, and -n) appeared equally often in each combination and that all inflectional affixes appeared equally often as primes and targets.

In total, the stimulus set of each experimental version consisted of 750 prime-target pairs. To eliminate undesired priming effects across items, all prime-target pairs were pseudorandomized, making sure that no semantic associations of any kind existed between consecutive items and that not more than four items of the same type occurred consecutively. Each of the nine versions exhibited the same order of test and filler items.

**METHOD**

Sixty-three native speakers of German (mean age: 25) participated in the experiment for payment. None of them participated in more than one experimental version. The primes were spoken by a female native speaker of German and compiled into separate audio .wav files. The auditory stimuli were presented over headphones. The sequence of stimulus events within each trial was as follows: a short attention tone (250 ms.) preceded the presentation of a fixation point (800 ms.), followed by the auditory prime word. Immediately at the offset of the (spoken) prime, the visual target was presented and remained on the computer screen for 200 ms. Reaction times were measured from the presentation of the target onwards. Three breaks were provided during the experiment. During each break and at the end of the experiment, the participants were asked to read a list of fifteen words, and to mark those words they had heard during the experiment. For each of these lists, nine words had been presented as auditory primes in the preceding experimental phase. The remaining six words did not occur in the experiment at all. The purpose of this task was to ensure that the participants paid attention to the auditory stimuli. The overall duration of the experiment was approximately one hour per subject. All other procedures were parallel to those of experiment 1.

Erroneous lexical decisions as well as extreme reaction times were excluded from the database. The data excluded made up 4.9% of the total responses. Mean response times for each subject and each condition were entered into two separate ANOVAs for subjects and for items with the factors Prime Type and Target Type.

**RESULTS**

Fig. 2 presents the mean lexical-decision times on the three visual target forms -e, -s, and -m and the three auditory prime forms -e, -s, and -m; see 11 for examples.

As expected, the shortest RTs were produced when both prime and target forms were identical. Fig. 2 shows that this was the case for all the target forms tested and also shows that the target affix -m elicited longer RTs than the two other adjective forms. Both ANOVAs produced significant main effects of Target Type ($F_1(2,124) = 4.95$; $p = .009$, $F_2(2,160) = 5.58$; $p = .005$) and a significant Prime Type × Target Type interaction ($F_1(4,248) = 10.96$; $p < .001$, $F_2(4,320) = 13.18$; $p < .001$). There was no significant main effect of Prime Type ($F_1(2,124) = 1.27$; $p = .283$, $F_2(2,160) = .51$; $p = .306$). To examine these effects further, we compared the overall mean RTs shown in Fig. 2 using matched t-tests.

Pairwise statistical comparisons of the RTs to the identity condition of each of the three target affixes revealed no significant differences (-e: 509 ms, -s: 508 ms, -m: 513 ms); see n. 6. This means that the target-type effect cannot be due to different response times to the identity condition. Significant differences were, however, found when
comparing the target affix -m to the two other target affixes: -m adjective forms produced significantly longer RTs when primed by -e than visual targets with -s primed by -e (549 ms. vs. 527 ms.). The same holds for -m as a target affix primed by -s, as opposed to the target affix -e primed by -s (543 ms. vs. 516 ms.). But visual targets with -e or -s produced the same mean response time when primed by -m ( = 533 ms). These comparisons show that the Target-Type effect is due to the longer response times for adjective forms with -m, confirming the prediction that adjectives with -m should be more difficult to prime than the other adjective forms.

Furthermore, the lack of a main effect of prime type indicates that none of the three adjective forms is overall a better or weaker prime than any of the other adjective forms tested. Pairwise comparisons revealed no significant differences between -e and -m forms used as primes for visual targets with -s (527 ms. vs. 533 ms.), and likewise no differences between -e and -s forms used as primes for targets with -m (549 ms. vs. 543 ms.). A significant difference was observed only for target forms with -e, where primes with -m elicited longer RTs than primes with -s (533 ms. vs. 516 ms.).

The results of this experiment indicate that the priming patterns are determined by the feature specifications of the affixes involved. In the identity condition, the feature specifications of the affixes in the visual targets are fully primed. Reduced priming (relative to the identity condition) occurs in cases in which a target form contains unprimed features that are not available from the prime. Example 12 makes this clearer by showing the unprimed features of the target forms for each of the prime-target pairs tested in our experiment; the feature specifications are taken from the analysis in 6 above. Moreover, 12 shows for each prime-target pair the difference in response times to the corresponding identity condition. These differences were also tested statistically using t-tests; an asterisk indicates a significant difference to the identity condition.
Consider the prime-target pairs \(-e \rightarrow -m\) and \(-s \rightarrow -m\) which produced the most significant reduction in priming compared to the identity condition. The targets in these two conditions have unprimed positively specified case features that do not match the unmarked \([{-OBL}]\) feature of the primes, hence the reduced priming effect. Consider next the prime-target pairs \(-e \rightarrow -s\) and \(-m \rightarrow -s\), which also yielded reduced priming. In the case of \(-e \rightarrow -s\), the targets have unprimed gender and number features, and for \(-m \rightarrow -s\) there is an unprimed gender feature \([{-MASC}]\) in the target, in addition to a case feature mismatch between prime and target. In the prime-target pair \(-m \rightarrow -e\), there is a case feature mismatch (\([+OBL]\) vs. \([-OBL]\)), and hence a significant reduction in priming. For \(-s \rightarrow -e\), however, there are no unprimed features in the target form, and correspondingly, \(-s\) adjective forms do not have any inhibitory effect on priming \(-e\) target forms. Note additionally that prime-target pairs that involved \(-m\) target forms produced larger differences to the corresponding identity condition than prime-target pairs with \(-e\) or \(-s\) target forms, suggesting that positively specified features are harder to prime than negatively specified ones.

4.4. Preliminary Summary. The two experiments reported in this section examine how regularly inflected and paradigmatically related word forms are represented in the mental lexicon. We found that inflected adjectives in German are decomposed into stems and affixes and that observed priming patterns depend on the morphosyntactic feature content of the affixes involved. A morphological analysis of German adjective inflection along the lines of 6 can account for the results of both experiments.

5. Experiments on Strong Verb Inflection. In this section, we will investigate the processing of inflected forms of the so-called strong verbs of German. While weak (regular) verbs employ the same stem in all their inflected forms, strong verbs
exhibit stem changes in the preterite and/or in the participle, e.g. werfen-warf-geworfen ‘to throw-threw-thrown’. This makes them parallel to irregular verbs in English, in that in both languages the marked stem forms are largely unpredictable from their corresponding infinitive forms. In contrast to English, however, the marked stems of strong verbs in German can be combined with fully regular person and number affixes in much the same way as weak stems, yielding forms such as warf-st ‘threw-2sg’ and warf-t ‘threw-2pl’.

The first question we will examine is whether inflected forms of strong verbs are processed via their constituent morphemes or are stored as wholes. We will show that the experimental results provide further evidence for morphological decomposition of regularly inflected word forms. The second question concerns the mental representations of different stem forms of the same lexeme. We present results from lexical-decision and priming experiments to show that stem forms are stored separately from the affixes with which they combine and that the various stem forms of a lexeme constitute a hierarchically structured lexical entry. Before turning to the experiments, we will provide a brief description of (strong) verb inflection in German.

5.1. STRONG VERB INFLECTION IN GERMAN. There are about one hundred sixty simplex verbs in German that belong to the strong class. These verbs have marked stems in present tense, preterite, or participle forms. Most of them (155) fall into three minor classes, illustrated in 13; see Wunderlich & Fabri 1995 for a detailed classification. While A-B-A and A-B-C verbs have differently marked stem forms in preterite and participle forms, A-B-B verbs exhibit the same vowel change for both preterites and participles. Moreover, a large number of strong verbs have subjunctive forms with umlauted preterite stems, as shown in 13 for gab/gäb- ‘gave’, flog/flög- ‘flew’ and sang/säng- ‘sang’. There is also a small number of strong verbs in which second and third singular present-tense forms as well as imperatives have fronted vowels (e.g. werfen vs. er wirft ‘to throw’ vs. ‘he throws’). Strong verbs also differ from weak verbs in that their participle forms have the ending -n, rather than the regular participle suffix -t. This holds for all strong verbs and none of the participle forms of weak verbs. Hence, strong verbs have a class-specific participle ending.

<table>
<thead>
<tr>
<th>Class</th>
<th>Infinitive</th>
<th>Preterite</th>
<th>Participle</th>
<th>Subjunctive</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-B-A</td>
<td>geben</td>
<td>gab-</td>
<td>gegeben</td>
<td>gäb-</td>
</tr>
<tr>
<td>A-B-C</td>
<td>singen</td>
<td>sang-</td>
<td>gesungen</td>
<td>säng-</td>
</tr>
<tr>
<td>A-B-B</td>
<td>fliegen</td>
<td>flog-</td>
<td>geflogen</td>
<td>flög-</td>
</tr>
</tbody>
</table>

Except for suppletive forms such as those of sein ‘to be’ and the forms of modals like dürfen ‘to be allowed to’, the marked stems of strong verbs are regularly inflected for person and number in the same way as the unmarked stems of weak verbs. Wunderlich (1996) has suggested a linguistic analysis of German verb inflection in which the morphological relationships between the various stem variants of strong verbs are represented in terms of nonmonotonic default-inheritance hierarchies (Evans and Gazdar 1996). In this account, stem variants constitute subnodes of hierarchically structured lexical entries in which each subnode is defined in terms of a phonological string and a morphological feature, e.g. [ . . . ] vs. [. . . a . . ] and [ ± PRET] for wirf vs. warf. Consider, for illustration, the lexical entry for the German verb werfen ‘to throw’ from Wunderlich 1996:96.
Each node in a structured lexical entry represents a pair (<phonological string, morphological feature value>), and each subnode inherits all information of its mother, except for the features it replaces or adds; for example, the subnode \([. . .a. . .]_{+\text{PRET}}\) inherits the onset \(w-\), the coda \(-rf\) and the categorical feature \([+ V]\) from the higher node. The base or default stem form is \(werf-\), the form that occurs in most present-tense forms and the infinitive. The other stem variants occur under specific circumstances, e.g. \(wirf-\) for second and third singular present-tense forms and in imperatives, \(warf-\) in preterite forms, \((ge)worfen\) in participles and \(wu¨rf-\) in subjunctives. Note that most stem variants have impoverished entries (to avoid unnecessary redundancies) and that the various stem forms are hierarchically structured. According to this account, stem forms are represented separately from the (regular) affixes with which they combine. We would therefore expect stem-frequency effects in lexical decision and priming differences between the various stem variants corresponding to the structure of the lexical entries. But associative models of inflection in which all inflected variants of a lexeme have full-form representations lead us to expect word-form frequency effects for inflected forms of strong verbs and word-form similarity effects. We will examine these predictions in two lexical-decision experiments.

5.2. EXPERIMENT 3: VISUAL LEXICAL-DECISION TASK ON PRETERITE FORMS. To examine whether inflected forms of strong verbs are stored as wholes or are morphologically decomposed for processing purposes, we examined preterite forms of strong A-B-A, A-B-B, and A-B-C verbs in a visual lexical-decision task. Recall from §4.2 that lexical decision times are sensitive to frequency. Thus, if forms such as \(warf-en\) ‘throw-1/3PL’ are stored as wholes, we would expect lexical-decision times to be affected by the frequencies of the word forms; high-frequency word forms should produce shorter response times than low-frequency word forms. If, however, regularly inflected forms of strong verbs are decomposed into stems and affixes, word-form frequencies should be irrelevant, and the response times should be affected by the frequencies of the stem (where affix frequencies are held constant).

MATERIALS

Fifty-two items from the three subclasses of strong verbs were selected, all of which were presented with the same affix, as first/third plural forms, e.g. \(sprachen\) ‘to speak’ (see Appendix A3). To determine effects of preterite stem frequency, the experimental items were arranged in pairs, resulting in two conditions. The items in the first condition had a relatively low mean preterite stem frequency of 79, while the preterite stems in the second condition had a higher mean frequency of 135. To control for other frequency effects, the items in both conditions were matched for their mean word form and their mean verb frequencies. Verb frequencies were calculated as the sum of the token frequencies of each base verb plus all its lexical variants that have separable prefixes, e.g. \(sprechen\) ‘to speak’, \(an-sprechen\) ‘to speak to’, \(vor-sprechen\) ‘to recite’. Variants with inseparable prefixes, e.g. \(besprechen\) ‘to discuss’, \(versprechen\) ‘to promise’, were not taken into account. Preterite stem frequencies included all preterite forms of each
base verb and its forms with separable prefixes, e.g., *an-sprachst, an-sprach*. Word-
form frequencies were the token frequencies of the inflected word forms presented in
the experiment, for example the frequency of the word form *sprachen*.

In addition to the 52 experimental items, we included 348 filler items: 40 weak verbs,
16 verbs of the mixed class (*bracht-en* ‘brought’), 92 other strong verbs, and 200
pseudoverbs that were constructed by changing two or three letters of an existing
weak or strong verb. To neutralize potential affix effects, the filler items, including the
pseudoverbs, had the same ending as the experimental items, i.e. *-en*. The entire stimulus
list consisted of 400 items and was presented in a pseudorandomized order.

**METHOD**

Thirty-four adult native speakers of German (mean age: 26) were tested. The methods
and procedures were parallel to those of experiment 1. Errors as well as extreme reac-
tion-time values were determined using the same criteria as in the previous experiments.
The total amount of excluded data was 11.7%. The mean response times for the remain-
ing test items were compared using matched *t*-tests for subjects and for items.

![Figure 3](image.png)

**RESULTS**

As Fig. 3 illustrates, there is a significant effect of preterite stem frequency: lexical-
decision times for verbs with high preterite stem frequencies were significantly shorter
(49 ms) than recognition times for verbs with low preterite-stem frequencies (*t*(33) =
5.97, *p* < .001 for subjects and *t*(25) = 4.33, *p* < .001 for items).

The error data showed the same difference. For low-frequency preterite stems the
mean rate of incorrect responses was 8.8%, whereas for high-frequency preterite stems
it was 4.5%, a statistically significant difference (*t*(33) = 4.00, *p* < .001).

The experimental effect found here cannot be attributed to the frequencies of the
verbs used in the experiment or the frequencies of the inflected word forms as a whole,
as both were held constant in the two experimental conditions. Instead, the stem-fre-
quency effect indicates that the verb forms presented in this experiment are decomposed
into stem and affix.

**5.3. EXPERIMENT 4: VISUAL LEXICAL-DECISION TASK ON A-B-B VERBS.** In associative
models of inflected word forms, morphosyntactic features and categories are not directly
encoded; see for example 3 in §2. In combinatorial accounts of inflection, however, morphological features and morphosyntactic categories play an important role. In Wunderlich’s (1996) analysis of strong verbs in German, the various subnodes of a hierarchically structured lexical entry, such as the one in 14, are defined in terms of morphological features and morphosyntactic categories (in addition to phonological strings). An interesting test case to assess the role of morphosyntactic features in the processing of inflected verb forms is provided by A-B-B verbs such as *lügen* ‘to lie’, which exhibit the same stem change in preterite and participle forms (*log-*ge-log-en*).

If morphosyntactic features are encoded in the representations of these forms, preterite and participle forms of A-B-B verbs should be stored separately (regardless of their phonological overlap), and there should be corresponding experimental effects. In a lexical-decision experiment, for example, we would expect to find separate frequency effects for the preterite and participle forms of verbs such as *lügen*. Alternatively, the mental representations of verbs may contain phonological but no morphosyntactic features. In this case, A-B-B verbs such as *lügen* would have two stored stem forms (*lüg-* and *log-*), and one would expect lexical decision times for preterite and/or participle forms to vary according to the frequency of the B-stem.

**MATERIALS**

Thirty A-B-B verbs were selected as experimental items, all of which were presented as first/third plural forms, e.g. *logen* ‘lied’ (see Appendix A4). To the 30 experimental items, we added 370 fillers (170 existing verbs and 200 pseudoverbs). As in experiment 3, all verbs and pseudoverbs had the same inflectional ending as the experimental items ( = -en). The experimental items were arranged pairwise in two conditions according to their preterite stem frequencies. The items in the first condition had a lower mean preterite stem frequency ( = 101) than the items in the second condition ( = 165). In addition, the items in both conditions were matched for their mean B-stem frequencies, that is, the mean frequency of all preterite and participle forms of each experimental item.

We determined verb frequencies for the experimental items, in the same way as for experiment 3. In contrast to experiment 3, however, it was not possible to match the experimental items on verb frequency. As shown in Appendix A4, the mean verb frequency of the items in condition I (those with low preterite-stem frequencies) is higher than that for condition II items (those with high preterite-stem frequencies). This means that if verb frequency affects lexical decision, response times should be shorter for condition I than for condition II. But if preterite-stem frequencies count, condition II should yield shorter lexical-decision times than condition I, since condition II has higher preterite-stem frequencies. Finally, given the restricted number of A-B-B verbs, it was also impossible to keep the word-form frequencies constant for each pair. Recall from experiment 3, however, that word-form frequencies did not have an effect on the reaction times during the processing of regularly inflected preterite forms. Moreover, the mean word-form frequencies for the items in the two conditions of experiment 4 were in a similar range (17 for condition I and 22 for condition II), and such a small difference is unlikely to have effects on lexical-decision times.

**METHOD**

Thirty-four adult native speakers of German (mean age: 26) were tested. The methods, procedures and time settings were the same as for experiment 3. Incorrect responses as well as outliers were excluded from the reaction time analysis. The total amount of excluded data was 11%. The remaining mean response times were compared using matched *t*-tests for subjects and for items.
RESULTS

Fig. 4 shows a significant effect of preterite-stem frequency. Recognition times for A-B-B verbs with high preterite-stem frequencies were significantly shorter (54 ms. for subjects and 60 ms. for items) than recognition times for A-B-B verbs with relatively low preterite stem frequencies ($t(33) = 7.11, p < .001$ for subjects and $t(14) = 3.05, p < .001$ for items).

The error data showed the same difference. The mean error rate was 8.6% for experimental items with low preterite-stem frequencies and 3.9% for items with high preterite-stem frequencies, a statistically significant difference ($t(33) = 3.08, p < .001$) that corresponds to the observed differences in reaction times.

These results cannot be attributed to the B-stem frequencies, as these were held constant in both conditions. Neither can they be explained in terms of verb frequency differences. It is true that the items in the two experimental conditions were not matched for verb frequency. If, however, this affected the response times, the results should show the opposite pattern, shorter RTs for the first condition (because of higher verb frequencies) than for the second one. Instead, the observed response-time difference appears to be linked to the different preterite-stem frequencies, indicating that preterite stems of A-B-B verbs access their own mental representations.

5.4. PRELIMINARY SUMMARY. In experiments 3 and 4, we examined marked stem forms of strong verbs in German that were combined with regular person and number affixes. The results from both experiments provide evidence against whole-word based representations for these kinds of inflected verbs and support the view that regularly inflected word forms are decomposed into stems and affixes, even if they contain marked (irregular) stems. The stem-frequency effects we found in both experiments show that marked stem forms are represented separately from the inflectional affixes with which they may occur. The frequency effect for A-B-B verbs in experiment 4 indicates that preterite stems are stored separately from participle forms despite their phonological overlap. This finding indicates that morphosyntactic features form part of the lexical representations of marked stem forms.

5.5. EXPERIMENT 5: CROSS-MODAL PRIMING TASK ON STRONG VERBS. While response times in lexical-decision tasks seem to be sensitive to properties of lexical entries, results from priming tasks provide more direct psycholinguistic measures for examining
relationships between lexical entries or subentries in the mental lexicon. Several previous studies have provided evidence for morphological priming effects in the mental lexicon (Stanners et al. 1979, Marslen-Wilson et al. 1994, Münte et al. 1999 for English; Sonnenstuhl et al. 1999 for German; Orsolini & Marslen-Wilson 1997 for Italian; Lukács & Pléh 1999 for Hungarian). It is clear from these studies that regularly inflected word forms produce priming effects that manifest themselves in shorter lexical-decision times on prime-target pairs such as walked → walk compared to a baseline condition in which the same target is preceded by an unrelated word (spilled → walk).

We made use of the priming technique to examine relationships between the different stem forms of a structured lexical entry focusing on the representation of strong verbs in German. The experimental items were inflected forms of strong verbs that were composed of a stem and the second person plural suffix -t, yielding prime-target pairs such as warft → werft ‘threw-2 PL → throw-2 PL’ vs. werft → warft. Like the past tense -ed in English, the second person plural -t is fully predictable and regular in German. We would therefore expect the -t ending to be stripped off from word forms such as werf-t and warf-t and the priming patterns to depend on the morphological relationships between the stem forms involved. If strong verbs are represented in hierarchically structured lexical entries (e.g. 14 for werfen), one would expect the stems’ feature specifications to affect the priming patterns, such that marked stems (which are further down on the inheritance tree) should be more difficult to prime than less specific stems. We predict that preterite stems ( = warf-t) should be less effectively primable by present tense forms ( = werf-t) than vice versa, because in terms of the representation in 14 the target form warft (preceded by the prime werft) contains an unprimed feature, [+PRET], whereas the target werft (preceded by warft) does not have any unprimed morphological features.

MATERIALS

Thirty-two strong verbs (see Appendix A5) from different subclasses were used. For each of these verbs, four prime-target pairs were constructed as shown in 15.

<table>
<thead>
<tr>
<th>CONDITIONS</th>
<th>AUDITORY PRIMES</th>
<th>VISUAL TARGETS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (pres. → pres.)</td>
<td>ihr helft ‘you help-2pl. pres.’</td>
<td>helft</td>
</tr>
<tr>
<td>II (pret. → pres.)</td>
<td>ihr halft ‘you help-2pl. pret.’</td>
<td>halft</td>
</tr>
<tr>
<td>III (pres. → pret.)</td>
<td>ihr helft ‘you help-2pl. pres.’</td>
<td>helft</td>
</tr>
<tr>
<td>IV (pret. → pret.)</td>
<td>ihr halft ‘you help-2pl. pret.’</td>
<td>halft</td>
</tr>
</tbody>
</table>

All experimental items contained the second person plural suffix -t. Note that second plural present-tense forms of weak verbs and of many strong verbs are identical to third singular forms, e.g. ihr kommt/sie kommt ‘you-pl. come/she comes’. To nullify any effects of this ambiguity, we presented all primes together with the second person plural form of the personal pronoun. Moreover, for the experimental items, we used only verbs with distinct second plural and third singular forms, for example, helfen ‘to help’ for which in the present tense the second person plural is helft and the third singular, hilt. Because no participant should see the same target more than once, four experimental versions were constructed. The prime-target pairs were distributed over four versions in a Latin-square design, so that each version included 32 different prime-target pairs (eight from each of the four conditions), and each verb appeared only in one prime-target pair. In addition to the experimental prime-target pairs, 176 unrelated word/word filler pairs and 208 word/pseudoword filler pairs were constructed. The list
of word/word fillers included German verbs in tense and agreement forms other than those of the experimental items (156 pairs). Moreover, all word/word pairs were counterbalanced so that present tense and preterite forms and the various person and number agreement affixes appeared equally often during the experiment. Parallel to the 208 word/word pairs in the experiment, the same number of word/pseudoword pairs (which were inflected in the same way as the real-word items) was constructed. In twelve of the pseudoword items, the prime was either fully contained within the target (e.g. *ich formte* ‘I formed’, *pluformtelst*) or partially overlapping with the target (e.g. *ich flüsterte* ‘I whispered’, *flüspurtet*). These pairs were added to the stimulus list to ensure that not all phonologically related pairs had real words as targets. The overall list of 416 prime-target pairs was pseudorandomized. Each of the four versions exhibited the same order of experimental and filler items.

**METHOD**

The methods, procedures, and time settings were the same as in experiment 2. Participants were forty-four native speakers of German (mean age: 26). In addition to errors and extremely long reaction times, one item (*trefft/traft* ‘meet/met’) had to be entirely excluded from any further analysis because of an error in the preparation of the stimulus list. The total amount of excluded data was 7.2%. To test the prediction that preterite stems are less effectively primable by present tense forms than vice versa, two separate ANOVAs were performed (for subjects and items) with the factors Prime Type (identity vs. test) and Target Type (preterite vs. present).

**RESULTS**

Fig. 5 presents the overall mean RTs to the visual targets in the four experimental conditions. Conditions I and IV (in which primes and targets were identical) produced shorter response times than the two morphological priming conditions II and III (in which primes and targets were different). Moreover, condition III where priming went
from present tense forms to preterite targets produced much less priming compared to the identity condition (77 ms.) than condition II (25 ms.) where priming went from preterite to present tense forms. These differences were confirmed statistically. The ANOVAs revealed a Prime Type × Target Type interaction which was significant for subjects (F₁(1,43) = 6.46, p = .015) and marginally significant for items (F₂(1,30) = 3.77, p = .062). In both subject and item analyses, there were also significant main effects of Prime Type, F₁(1,43) = 28.49, p < .001, F₂(1,30) = 44.95, p < .001, and Target Type, F₁(1,43) = 21.88, p < .001, F₂(1,30) = 31.72, p < .001. Matched t-tests revealed that the differences between conditions I and II, as well as those between III and IV, are statistically significant (see n. 6). Moreover, the differences between the priming effects in the two morphological priming conditions II and III (25 ms. vs. 77 ms.) were statistically significant in the subjects analysis (p = 0.015) and marginally significant for items (p = 0.062). This confirms the predicted priming patterns: preterite forms prime present tense forms better than present tense forms prime a preterite target.

In addition to the reaction times, we also analyzed the distribution of errors. A comparison of the mean error rates for conditions I and II to those for conditions III and IV revealed that preterite forms produced a significantly higher error rate than present tense forms (6.3% vs. 0.4%, t(43) = 3.21, p = .003). This difference is consistent with the reaction-time data, in particular with the finding that preterite forms are less easily primable than present tense forms.

To explain the observed priming patterns, we can rule out the possibility that the form of the inflectional affixes had an effect on the response times. Recall that all experimental items in all conditions had the same person and number affix (the second person plural -t). Hence, the form of the affix cannot account for the observed differences between present tense and preterite forms. It is more likely that they are caused by the morphological feature specifications of the stem forms involved, in the following way. Suppose that the regular second plural -t affix is stripped off from all the inflected word forms we presented and that the remaining stem forms are accessed separately. When a marked stem such as half- is presented as the visual target preceded by an instance of the unmarked stem (helf-), the target contains a feature, [+PRET], that is unavailable from the prime, and this unprimed feature produces the increased target response times. But when an unmarked stem (helf-) is the target preceded by a marked stem (half-), the target does not contain any unprimed features, and hence there are significantly shorter target response times than for preterite targets.

6. DISCUSSION. In discussing the controversy between combinatorial and associative approaches to inflection, our focus will be on three core aspects of morphological processing and representation: (i) morphological decomposition of inflected word forms, (ii) the representation of morphologically related word forms, and (iii) the specification of morphological form in entries of the mental lexicon.

6.1. EVIDENCE FOR MORPHOLOGICAL DECOMPOSITION. Research on morphological processing indicates that regularly inflected word forms are decomposed into smaller morphological units, such as stems, roots, and affixes (see Clahsen 1999 for review). Evidence for this comes, for example, from cross-modal priming experiments on past participles and noun plurals in German (Sonnenstuhl et al. 1999). We found full stem priming for regularly inflected -t participles (gelacht → lach ‘laughed → laugh’) and regular -s plurals (Autos → Auto ‘cars → car’), but reduced stem priming for irregular (-n) participles (geschlafen → schlaf ‘slept → sleep’) and irregular (-er) plurals (Kinder...
Kind ‘children → child’). We explained these differences in morphological terms: -s plural and -t participle word forms are decomposed into stem + affix, and can thus prime their base stem directly. Irregular plurals and participles, however, access full-form entries stored in memory and cannot directly activate their corresponding base entries; therefore the priming route is less direct. At the same time, alternative nondecompositional accounts of the observed priming differences could be ruled out, since the prime-target pairs in the regular and the irregular conditions were matched with respect to their frequencies and their phonological/orthographical overlap.

The findings reported here provide further support for morphological decomposition. In three lexical-decision experiments, we found separate affix and stem-frequency effects indicating that affixes access their own representations in the mental lexicon. Experiment 1 showed that adjectives inflected with (the more frequent affix) -s yielded shorter response times than the same adjectives inflected with (the less frequent affix) -m. At the same time, the word-form frequencies of the items had no effect on the participants’ response times. In experiments 3 and 4, we found stem-frequency effects for regularly inflected word forms of strong verbs. Preterite forms of A-B-A, A-B-C, and A-B-B verbs with high preterite-stem frequencies elicited significantly shorter response times than corresponding forms with low preterite-stem frequencies. Again, as in experiment 1, word-form frequencies—the frequency of an inflected preterite form as a whole—did not influence the response times.

Interestingly, lexical-decision tasks examining participle forms of A-B-A and A-B-B verbs produced different results (Clahsen et al. 1997), even though the design of these experiments was parallel to that employed in experiments 3 and 4. Lexical-decision times for high-frequency participle forms of both subclasses of strong verbs were found to be shorter than those of low-frequency participle forms. At the same time, stem and verb frequencies did not affect response times. The participle forms of A-B-B verbs tested by Clahsen and colleagues (1997) were matched with respect to B-stem frequencies and verb frequencies in the same way as in experiment 4. Response times to participles with low word-form frequencies, however, showed a significant 58-ms. delay compared to participles with high word-form frequencies.7 These findings show that the lexical-decision times of participle forms of strong verbs are sensitive to word-form frequency, rather than to stem frequency.

To account for the observed differences between the processing properties of preterite and participle forms, recall that the preterite forms from experiments 3 and 4 were suffixed with a regular person and number suffix, while the strong participle forms used in Clahsen et al.’s (1997) lexical-decision experiments (and those used in Sonnenstuhl et al.’s (1999) priming experiments) were -(e)n participles such as gelog-en ‘lied’ in which -(e)n is a class-specific ending that is unique to members of the strong class. Wunderlich (1996) posited different subnodes within the lexical entries of A-B-B and A-B-A verbs to capture this difference. Consider, for example, the lexical entry of the A-B-B verb lügen ‘to lie’ in 16; the preterite stem is log-, the participle form gelogen, and the subjunctive has a fronted vowel formed on the basis of the preterite stem (lög-).

7 Note that this was not the case for -t participle forms of weak verbs such as gekauft-t ‘bought’. In contrast to the strong participle forms, there were no word-form frequency effects for weak participle forms. This was taken to indicate that regular -t participles are decomposed into (weak) stems and a -t participle affix (see Clahsen 1999 for discussion).
Our experimental findings indicate that preterite stems and participle forms are stored in separate subnodes, as indicated in 16. The word-form frequency effects for strong participle forms show that these forms do not have independent stem representations, in contrast to strong preterites for which stem frequency effects were found. Moreover, our results indicate that the presence of a regular affix is what triggers morphological decomposition. This was the case in the preterite forms tested in Experiments 3 and 4, but not in the strong participle forms examined in Clahsen et al. (1997). The form of the stem, though, does not determine whether an inflected word is decomposed or stored as a whole. Decomposition effects are found for inflected word forms such as *walked* or *gezeigt* ‘shown’ which have both a regular inflectional ending and an unmarked (regular) stem form, as well as for items that contain a regular (agreement) ending with a marked (irregular) stem form (see experiments 3 and 4). Hence, it seems (as one would expect from combinatorial models of inflection) that regular inflectional affixes are stripped off inflected words, irrespective of the form of the stem.

6.2. MORPHOLOGICAL RELATIONSHIPS IN THE MENTAL LEXICON. To address the question of how relationships between different inflectional affixes and between different stem forms of the same lexeme are mentally represented, we investigated inflected adjectives and preterite stem forms of strong verbs in two cross-modal priming tasks. The empirical generalization that comes out of the results of these experiments is that reduced priming occurs in cases in which the target forms contain feature specifications that are not available from the primes.

A paradigmatic analysis of German adjective inflection was shown to account for the priming patterns observed in experiment 2. Recall, for example, that *-m* target forms were more difficult to prime than any other adjective form. We interpreted this as a SPECIFICITY EFFECT: *-m* is specified for the morphosyntactic features [ + OBL] and [ + DAT], and these features are not available from any other adjective form. For this reason, *-m* forms are hard to prime by other adjective forms, and hence the longer response times for *-m* target forms in our experiment. We also found an asymmetry in the priming patterns of *-s* and *-e* forms: *-s* primes a target form *-e* as effectively as an identical prime; there was no statistically reliable difference between *-s → -e* and *-e → -e*. In contrast, *-e* adjective forms prime *-s* adjective targets much less effectively than identical targets. These priming differences are also determined by the feature contents of the two affixes. The prime-target pairs *-e → -s* contain unprimed target features (for gender and number), whereas the target forms in *-s → -e* pairs do not have any features that are not already available from the primes.

Evidence about how relationships between different stem forms of the same lexeme are mentally represented comes from the results of experiment 5. We found that preterite stems (e.g. *warf* ‘threw’) are less effectively primable by present tense stems (= *werf*) than vice versa. This priming difference can be explained in terms of structured lexical entries, such as those in 14 and 16. In these entries, preterite stems are marked forms that are represented further down on the inheritance hierarchy than present tense stems.
Preterite stems (primed by present tense forms) are left with an unprimed \([ + \text{PRET}]\) feature in the target, whereas present-tense forms (primed by preterites) do not have any unprimed target features, and hence the different priming patterns. Thus, despite different representations for stems and affixes, the observed priming patterns are parallel, in that for affixes and stems reduced priming was found when the target forms contained unprimed morphological features.

### 6.3. Underspecification in the Mental Lexicon

The notion of underspecification is used in phonological and morphological representations of lexical entries to reduce lexical redundancies. In phonological representations, for example, redundant feature values that do not contrast two segments are left unspecified (Archangeli 1988, for psycholinguistic evidence see Lahiri & Marslen-Wilson 1991).

Underspecification is also used in morphological representations, in inflectional paradigms, for example, to reduce the number of syncretic forms. Our experimental findings indicate that the mental lexicon contains entries that are underspecified for morphosyntactic features. Inflectional affixes, for example, were assumed to have impoverished entries in which only positive (marked) feature values are directly specified, while negative (unmarked) feature values are introduced into paradigms in opposition to a positive value on the same feature. With these assumptions, the priming patterns observed in experiment 2 could be explained. The alternative possibility that all affixes have fully specified representations including positive and negative feature values, does not account for our findings. Consider, for example, the priming asymmetries between \(-e\) and \(-s\) adjective forms. If both these affixes had fully specified entries, we would have unprimed features in both \(-e\) and \(-s\) target forms, and the observed priming difference between the prime-target pairs \(-e \rightarrow -s\) (= reduced priming) and \(-s \rightarrow -e\) (= full priming) would be left unexplained.

Priming asymmetries were also found for preterite and present-tense forms of strong verbs; preterite forms prime present-tense forms more effectively than present-tense forms prime preterites. This is compatible with the view that (like inflectional affixes) stem forms also have impoverished underspecified entries in which only preterite stems are specified for a tense feature. If all stem forms were fully specified (e.g. present tense forms as \([- \text{Pret}]\) and preterite forms as \([ + \text{Pret}]\)), there would be an unprimed tense feature in the targets of both prime-target pairs, since the \([-\text{Pret}]\) feature of present-tense targets cannot be primed by preterite forms, and the \([+\text{Pret}]\) feature of preterite targets cannot be primed by present-tense forms. The observed priming asymmetries are not compatible with this account. Our findings support the hypothesis that the mental representation of lexical forms uses underspecified representations.

### 6.4. Explaining the Findings within Associative Models of Inflection

It was argued that combinatorial models of inflection account for the results from both lexical-decision and cross-modal priming. We will explore how associative models of the mental lexicon, such as those presented in §2, might explain the experimental results, and will show that some of our findings present difficulties for these approaches.

In associative models of inflection, inflected word forms are stored as wholes. We would therefore have expected word-form frequency effects in the three lexical-decision experiments; shorter response times for high-frequency adjective and preterite forms than for those with low word-form frequencies. One would not predict affix- and stem-frequency effects. The results of experiments 1, 3, and 4 showed the opposite pattern. Affix-frequency effects paired with the lack of word-form-frequency effects are prob-
lematic for any account that assumes whole-word representations for regularly inflected word forms (see Clahsen 1999 for additional evidence from noun and verb inflection).

As for the results from the two priming experiments, the satellite model (see 2, §2) provides a partial account of our findings. This model distinguishes between a base form (nucleus) of a lexeme and its inflected variants (satellites), which are connected to the nucleus. Thus, if lexical access is made via a satellite form, the corresponding nucleus is also activated. A base form, however, does not necessarily activate all its satellites. With these assumptions, the priming asymmetries reported in experiment 5 can be accounted for. Strong preterite forms are satellites in this model and activate their corresponding base forms; hence they prime present-tense forms more effectively than vice versa. Note, however, that the stem- and affix-frequency effects found in the three lexical-decision experiments are incompatible with the full-form representations the satellite model posits for inflected word forms. The results of experiment 2 on adjective priming are also hard to explain for the satellite model. Recall the priming differences between various inflected forms of the same adjective, for example priming asymmetries between -s and -e forms and differences between -m and -sl/-e target forms. These experimental differences between the various satellites of a nucleus are left unexplained by the satellite model.

In other associative models of inflection, all inflectional variants (including the base form) are assumed to have word-level status and to be associatively connected to each other (see 3, §2). Experimental differences between inflected word forms are explained in terms of frequency and similarity differences of the forms involved. Schriefers et al. (1992) argued along these lines, suggesting that inflected adjectives in German have associative full-listing representations in the mental lexicon. The evidence they reported comes from a unimodal (visual) delayed repetition priming task examining adjectives in four morphological variants, for example gute, gutem, gutes, gut ‘good’. Their main finding was an asymmetric priming pattern for -m adjectives: -m forms primed other adjective forms as effectively as an identical prime, but when -m adjectives were the targets and other adjective forms appeared as primes, priming was reduced. Schriefers and his colleagues reasoned that the decomposition hypothesis ‘would not only require full priming of a stem by its morphological variant, but also . . . full priming between its morphological variants’ (1992:375). As this was not confirmed in their experiment, they argued instead that each morphological variant of a stem is fully represented as a word node and that the observed priming asymmetries follow from frequency differences, claiming that -m target forms are difficult to prime because of the relatively low frequency of these forms.

We do not think that these conclusions are borne out. According to decompositional accounts of morphological processing, stems and affixes of regularly inflected words access their own mental representations. Hence, in prime-target pairs in which the stems are identical but the affixes (or rather their feature contents) contain unprimed features in the targets, we would expect to find reduced priming (and not full priming). This is indeed what our experiment 2 showed. But our experiment did not replicate Schriefers et al.’s finding that -m forms used as primes produced the same target response times as an identical prime. Their finding is probably an episodic memory effect, caused by the long delay (of 10 to 14 unrelated items) they introduced between primes and targets. Episodic memory effects are based on a participant’s remembering a prior event and have been shown to be stronger when there is a long delay between prime and target, and when the participant has to react in the same way to both elements of a related stimulus pair (Tenpenny 1995). Both points apply to the Schriefers et al. experiment.
It is therefore conceivable that when presented with a target such as *akutes*, participants remember having seen another instance of the same adjective 10 to 14 items earlier without necessarily remembering whether it was *akutes* or *akutem*, and therefore Schriefers and his colleagues were unable to find a difference between these two conditions. The cross-modal immediate repetition task we used in experiment 2 is less affected by episodic memory effects (since there is no delay between primes and targets and participants react only to the visually presented targets while listening to auditory primes), and in this experiment we found reduced priming in all cases in which the target forms contained unprimed features.

Schriefers et al. attempted to explain priming effects in terms of frequency differences. The finding (both theirs and ours from experiment 2) that -m forms are less easily primable than other adjective forms might be explained in this way. If, however, frequency was the determining factor for the target response times, we would also expect -e targets to be primed more effectively than -s targets, since -e is the most frequent adjective ending in German, as Schriefers et al. pointed out themselves (1992: 376). This prediction is not confirmed, either in our experiment 2 or in Schriefers et al.’s data. Fig. 2 shows that the response times to -e targets (primed by -m) are identical to the reaction times to -s targets (primed by -m); 533 ms. in both cases. The same holds for Schriefers et al.’s data; 688 ms. for -m → -e and 689 ms. for -m → -s. Frequency is unlikely to be the decisive factor in accounting for the observed priming differences.

In addition to frequency, proponents of associative models of inflection argue that priming is determined by the degree of surface form similarity—the orthographic and phonetic overlap between prime and target. Regular past-tense forms in English are, for example, orthographically and phonologically more similar to their base forms than many irregular past-tense forms (compare *walked* → *walk* versus *taught* → *teach*), and it has been argued that this is the reason why regular past-tense forms are more effective primes for their corresponding stems than irregular ones (Rueckl et al. 1997). This surface similarity account, however, can handle only a small subset of the observed priming patterns reported in the present study. Consider, for example, the prime-target pair -s → -e which in our experiment produced a nonsignificant 7-ms. increase in response times compared to the identity condition (see 12). One might argue that effective priming in this case is due to the high degree of orthographic and formal overlap between prime and target, as the target form (e.g. *karierte*) is fully contained in the form of the prime (e.g. *kariertem* → *karierte*), even though primes and targets show the same degree of surface similarity as -s → -e. This shows that a simple surface-based account does not explain the experimental findings.

### 7. Conclusion.

In investigating how morphological relationships between inflected word forms are represented in the mental lexicon, we focused on paradigmatic relations between regularly inflected word forms and relationships between different stem forms of the same lexeme. We examined whether inflected word forms are morphologically decomposed or stored as wholes. We presented results from five psycholinguistic experiments that examined the processing of inflected adjective and verb forms of German, explaining our findings in terms of combinatorial approaches to inflection, with morphological paradigms to represent regular affixes and structured lexical entries to represent different stem forms of the same lexeme. For both stems and affixes, we relied on linguistic analyses that posit underspecified lexical entries, that is, minimally specified
analyses in which only positive (marked) feature values are directly specified. To the extent that our interpretation can be maintained, the experimental findings provide psycholinguistic support for these theoretical notions. Associative models of inflection, in which morphological structure and morphosyntactic features and categories are not encoded, provide only partial explanations of our findings.

APPENDIX: EXPERIMENTAL ITEMS

A1. Adjectives used in experiment 1

Frequency counts are given in parentheses (stem / word-form / preterite stem)

(i) -m dominant adjectives: **mean frequencies (402/16/9)**

<table>
<thead>
<tr>
<th>Adjective</th>
<th>Frequency Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>pur 'pure'</td>
<td>(25/3/1), grell 'glaring' (35/3/1), spiz 'pointed' (46/4/1), sauer 'sour' (53/3/0), steif 'stiff' (82/7/0), schräg 'slipping' (104/5/0), flüssig 'liquid' (111/12/7), nass 'wet' (113/7/4), lebhaft 'lively' (281/27/19), dicht 'dense' (339/11/6), hell 'bright' (345/13/9), kalt 'cold' (503/24/17), rasch 'quick' (522/8/5), scharf 'sharp' (640/15/9), direkt 'direct' (706/14/10), ruhig 'quiet' (838/51/13), gering 'little' (929/31/9), frei 'free' (1560/42/38)</td>
</tr>
</tbody>
</table>

(ii) -s dominant adjectives: **mean frequencies (397/7/17)**

<table>
<thead>
<tr>
<th>Adjective</th>
<th>Frequency Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>stumpf 'blunt' (25/1/4), 'sharp' (640/15/9), 'absurd', 'clear' (1510/16/3), hart 'hard' (633/19/25), leer 'empty' (315/6/15), stabil 'stable' (101/1/4), starr 'stiff', schärf 'sharpen' (8/0/3)</td>
<td></td>
</tr>
</tbody>
</table>

A2. Adjectives used in experiment 2

<table>
<thead>
<tr>
<th>Adjective</th>
<th>Frequency Count</th>
</tr>
</thead>
</table>

A3. Verb forms used in experiment 3

Frequency counts are given in parentheses (stem / word-form / preterite stem)

(i) verbs with low preterite stem frequency: **mean frequencies (411/16/79)**

<table>
<thead>
<tr>
<th>Verbal Forms</th>
<th>Frequency Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>soffen 'drank' (8/0/3), wanden 'wound' (24/3/6), schliffen 'ground' (30/2/6), schwanden 'dwindled' (33/0/4), stanken 'stank' (40/29), fochten 'fenced' (42/3/6)</td>
<td></td>
</tr>
<tr>
<td>priesen 'raised' (52/2/12), schworen 'swore' (69/6/19), verdarben 'spoilt' (76/0/8), gruben 'dug' (87/6/19), logen 'lied' (105/0/18), warben 'recruited' (124/4/14)</td>
<td></td>
</tr>
<tr>
<td>pfiffen 'whistled' (135/13/58), stritten 'quarreled' (144/16/30), scheidten 'separated' (276/10/39), banden 'tied' (385/7/39), schnitten 'cut' (386/20/87), sangen 'sang' (450/34/116), tranken 'drank' (474/39/146), laden 'loaded' (646/22/100), rissen 'tore' (661/53/282), brachen 'broke' (830/40/207), wuchsen 'grew' (935/43/165), glichen 'resembled' (1014/8/40), schaffen 'created' (1479/48/149)</td>
<td></td>
</tr>
<tr>
<td>schrieben 'wrote' (2160/41/473)</td>
<td></td>
</tr>
</tbody>
</table>

(ii) verbs with high preterite stem frequency: **mean frequencies (415/19/135)**

<table>
<thead>
<tr>
<th>Verbal Forms</th>
<th>Frequency Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>kniffen 'pinched' (22/0/8), schlagen 'wrapped' (27/3/19), fluchen 'plaited' (24/2/10), sannen 'pondered' (31/1/15), misslingen 'failed' (30/1/14), mieden 'avoided' (40/6/11), liehen 'lent' (66/2/14), schlichen 'sneaked' (74/11/40), erschraken 'scared' (72/3/51), stachen 'stung' (82/8/39), rißen 'rubbed' (96/8/46), bliesen 'blew' (105/4/41), rochen 'smelt' (135/6/73), stritten 'strde' (207/23/70), litten 'suffered' (254/23/78), empfinden 'recommended' (382/6/68), schwiegen 'remained silent' (360/18/133), empfingen 'received' (450/21/156), sprangen 'jumped' (516/42/229), zwangen 'forced' (668/25/112), trieben 'drifted' (631/40/160)</td>
<td></td>
</tr>
<tr>
<td>stießen 'pushed' (815/64/322), fingen 'caught' (971/39/257), griffen 'grabbed' (1021/54/337), riefen 'called' (1514/45/524), schienen 'shone' (2050/36/634)</td>
<td></td>
</tr>
</tbody>
</table>
A4. Verb forms used in experiment 4

Frequency counts are given in parentheses (stem / word-form / B-stem / preterite stem)

(i) verbs with low preterite stem frequency: **mean frequencies (72/17/253/101)**
- *woben* ‘wove’ (15/0/9/2),
- *schliffen* ‘ground’ (30/2/17/6),
- *schmissen* ‘threw’ (33/2/20/7),
- *fochten* ‘fenced’ (42/3/23/6),
- *priesen* ‘praised’ (52/2/27/12),
- *schworen* ‘wore’ (69/14/38/19),
- *logen* ‘lied’ (105/0/48/18),
- *bissen* ‘bit’ (144/47/52/101),
- *strichen* ‘painted’ (222/7/137/66),
- *schieden* ‘separated’ (276/10/134/39),
- *schnitten* ‘cut’ (386/20/218/87),
- *glichen* ‘resembled’ (1014/6/86/40),
- *boten* ‘offered’ (1643/50/651/291),
- *schrieben* ‘wrote’ (2160/41/1071/473),
- *schlossen* ‘closed’ (2393/90/1235/404)

(ii) verbs with high preterite stem frequency: **mean frequencies (444/22/251/165)**
- *kniffen* ‘pinched’ (22/0/10/8),
- *flochten* ‘plaited’ (24/2/18/10),
- *mieden* ‘avoided’ (40/6/19/11),
- *schmolzen* ‘melted’ (62/4/26/13),
- *wogen* ‘weighed’ (80/2/32/18),
- *ritten* ‘rode’ (128/9/44/26),
- *stritten* ‘quarreled’ (144/16/42/30),
- *rochen* ‘smelt’ (135/6/79/73),
- *schwiegen* ‘remained silent’ (360/18/155/133),
- *litten* ‘suffered’ (254/23/122/78),
- *trieben* ‘drifted’ (631/39/254/160),
- *flossen* ‘flowed’ (245/30/81/60),
- *hoben* ‘lifted’ (1256/31/754/409),
- *stiegen* ‘climbed’ (1580/145/938/558),
- *zogen* ‘pulled’ (1693/5/1192/891)

A5. Verbs used in experiment 5

- *befehlen* ‘to order’,
- *bergen* ‘to recover’,
- *bersten* ‘to crack’,
- *brechen* ‘to break’,
- *empfangen* ‘to receive’,
- *empfehlen* ‘to recommend’,
- *essen* ‘to eat’,
- *fangen* ‘to catch’,
- *fressen* ‘to feed’,
- *geben* ‘to give’,
- *halten* ‘to hold’,
- *helfen* ‘to help’,
- *laden* ‘to load’,
- *lesen* ‘to read’,
- *messen* ‘to measure’,
- *nehmen* ‘to take’,
- *raten* ‘to guess’,
- *schlagen* ‘to hit’,
- *sehen* ‘to see’,
- *schlafen* ‘to sleep’,
- *schlagen* ‘to speak’,
- *schichten* ‘to steal’,
- *stechen* ‘to steal’,
- *sterben* ‘to die’,
- *stossen* ‘to push’,
- *treffen* ‘to meet’,
- *treten* ‘to step’,
- *werben* ‘to recruit’,
- *werfen* ‘to throw’.

REFERENCES

THE MENTAL REPRESENTATION OF INFLECTED WORDS:

Clahsen, Harald; Sonia Eisenbeiss; and Ingrid Sonnenstuhl. 1997. Morphological structure and the processing of inflected words. Theoretical Linguistics 23.201–49.
Elman, Jeffrey L.; Elizabeth A. Bates; Mark H. Johnson; Annette Karmiloff-Smith; Domenico Parisi; and Kim Plunkett. 1996. Rethinking innateness: A connectionist perspective on development. Cambridge, MA: MIT Press.
Lukatela, Georgie; Claudia Carello; and Michael T. Turvey. 1987. Lexical represen-
tation of regular and irregular inflected nouns. Language and Cognitive Processes 2.1–17.


LUKATELA, GEORGIJE; Z. MANDIC; B. GLIGORJEVIC; ALEKSANDAR KOSTIC; MILAN SAVIC; and MICHAEL T. TURVEY. 1978. Lexical decision for inflected nouns. Language and Speech 21.166–73.


Clahsen
Dept. of Linguistics
University of Essex
Colchester CO4 35Q
UK
[harald@essex.ac.uk]

[Received 20 September 2000; revision received 14 February 2001; accepted 1 April 2001]