Edited by

Michael Hammond
Department of Linguistics
University of Wisconsin-Milwaukee
Milwaukee, Wisconsin

Michael Noonan
Department of English
University of Wisconsin-Milwaukee
Milwaukee, Wisconsin

Theoretical Morphology

Approaches in Modern Linguistics

1988

Academic Press, Inc.
Harcourt Brace Jovanovich, Publishers
San Diego New York Berkeley Boston
London Sydney Tokyo Toronto

Robinson, O. (1975). Abstract phonology and the history of umlaut. *Lingua* **37**, 1–29.
Sapir, E. (1915). Notes on Judeo-German phonology. *The Jewish Quarterly Review*, n.s., **6**, 231–266. Reprinted in "Selected Writings of Edward Sapir in Language, Culture, and Personality" (David G. Mandelbaum, ed.). Univ. of California Press, Berkeley.
Weinreich, U. (1965). "College Yiddish"[4] Yivo Institute for Jewish Research, New York.
Weinreich, U. (1968). "Modern English-Yiddish Yiddish-English Dictionary." Yivo Institute for Jewish Research and McGraw-Hill, New York.
Zarecki, A. (1929). "Yidiše Gramatik." Vilner Farlag fun B. Kleckin, Vilne.

Chapter 6

Are Inflected Forms Stored in the Lexicon?

Joseph Paul Stemberger & Brian MacWhinney

Lexical items are a fundamental part of a speaker's knowledge of language. However, it is not always clear what items should be listed in the lexicon. There has been major disagreement in both linguistic theory and psycholinguistics over the status of inflected forms. Should all members of an inflectional paradigm be given separate entries in the lexicon? Or should there be only one entry for each paradigm, with the individual inflected forms being created by applying an inflectional rule to a base form? We examine this issue from a psycholinguistic perspective, bringing to bear data that we hope resolve this issue for language production. In addition to addressing the question of whether inflected forms are stored, we also address the question of how they might be stored.

Proposals about the representation of inflected forms fall into three main "generic" approaches. One approach is full suppletion: all inflected forms are stored in the lexicon as single units, as if they were monomorphemic (Leben and Robinson, 1977; Bybee, this volume). There may additionally be some form of redundancy rule that relates an inflected form to the other members of its paradigm and to other forms with the same affix. Regular and irregular forms are parallel in this approach, as in (1).

(1)
{FREEZE} → /frouz//[+past]
|
/fri:z/ elsewhere
{WALK} → /wa:kt//[+past]
|
/wa:k/ elsewhere

An opposite approach is possible, where no inflected forms are stored in the lexicon, but are rather produced by lexical-item-free rules that take a base form as input and give an inflected form as output. This is possible with regularly inflected forms as in (2), but not with irregular forms, where some form of lexical conditioning is required.

(2) Regulars: v] T v]/[+past]

The last generic position is an intermediate one; it holds that at least some regularly inflected forms are stored in the lexicon, but that they are stored as two items. Minimally, this would involve storing the words as a sequence of two morphemes (Jackendoff, 1975), but could also involve the use of lexically conditioned rules, as in (3).

(3)
$$\begin{array}{l} \text{æ} \\ \quad | \\ [\ldots \text{V } C_i \text{ v}]/[+\text{past}],\{\text{SING, STINK, RUN, } \ldots\} \\ \\ \text{v] T v}]/[+\text{past}],\{\text{WALK, ASK, SNEEZE, } \ldots\} \end{array}$$

Of course, there is no reason why one must take any one of these positions for all inflected forms. It is fairly popular to take one position for irregular forms (generally suppletion) and a different one for regular forms (generally lexical-item-free rules) (Bybee and Slobin, 1982; Kiparsky, 1982). There is very little data that allow us to choose between these options. All of them can account for the basic linguistic data, differing primarily on simplicity and what generalizations are captured. And there is little agreement on what is simple or on what generalizations should be captured. (In order to account for productivity, the full-suppletion approach requires an explicit formulation of analogy or the assumption that rules exist but are used only to relate forms in the lexicon and to create forms not stored in the lexicon (Aronoff, 1976; Butterworth, 1983)).

Psycholinguistics has only been able to partially resolve the issue (see Butterworth, 1983). It has generally been assumed that irregular forms are stored in the lexicon, but it is not clear how. Perceptual studies have often maintained that the base and inflection of regularly inflected forms are perceived independently (Stanners *et al.*, 1979; Kempley and Morton, 1982), but other researchers have suggested that the results are compatible with suppletive storage, as long as members or a paradigm are clustered in some special way (Bradley, 1980; Lukatela *et al.*, 1980). There is some evidence that young children may store inflected forms (and syntactic phrases) as unanalyzed units (MacWhinney, 1978; Peters, 1983), but no evidence that such storage continues into adulthood (or even past the age of 6). Walsh and Parker (1983) report that the /s/ of the plural morpheme is 9 msec longer than a nonmorphemic /s/ in homophonous words, suggesting that inflected forms are at least analyzed into two morphemes. Stemberger (1985b) reviews evidence from

speech errors and argues that regularly inflected forms appear to be controlled as two morphemes in language production, and that inflectional rules are needed to account for errors where irregular forms are regularized (e.g., *choosed* instead of *chose*), at least within one type of language-production model. Thus, there is some evidence for inflectional rules, but none that really addresses the question of whether inflected forms are stored in the lexicon. Since lexical storage and rules are compatible (the third position outlined above), this leaves us uncertain about what is stored in the lexicon.

Psychology provides us with two clear tests of whether forms are stored as units or not. (a) Given storage, units that are of high frequency should be produced faster and more accurately than units that are of low frequency (Atkinson and Shiffrin, 1968; MacKay, 1982). Frequency effects derive from storage and cannot differentiate two items if neither is stored. (Of course, if a frequency effect is present, it only tells us that high-frequency forms are stored. It gives no information about whether low-frequency forms are stored, since storage is assumed to increase the speed and accuracy of processing relative to no storage.) (b) Given storage, items that are similar will tend to reinforce each other and lead to faster, more accurate performance (Rumelhart and McClelland, 1982; McClelland and Elman, 1986). This is a form of analogy and has sometimes been termed EXTENDED ANALOGY (e.g., Stemberger, 1985b). These "gangs" of similar items also tend to interfere with the processing of other lexical items that are similar to the members of the gang. Such "gang effects" must be present if items are stored, and cannot be present if items are not stored.

Frequency effects are present if inflected forms are stored in the lexicon, but absent if no inflected forms are stored in the lexicon. Some inflected forms are of high frequency and are used very often, for example, *ended* (Francis and Kucera, 1982). Other inflected forms are of low frequency and are used comparatively rarely but are nonetheless quite familiar, for example, *mended* (Francis and Kucera, 1982). If both are stored in the lexicon, the higher frequency of *ended* will entail that it is produced more rapidly than *mended* is, and that it will be less susceptible to error than *mended* is. This is especially true of a particular type of error that we call a NO-MARKING ERROR, where the speakers errs by accessing the base form where an inflected form was required, for example, *end* or *mend*. (We focus on no-marking errors in this chapter, since this is the most common type of error on regularly inflected forms in English.) The high frequency of *ended* guarantees more accurate access of the inflected form and fewer errors where the base form is accessed instead. If neither inflected form is stored, the difference in frequency of the inflected form is irrelevant, since this frequency is not represented anywhere in the system. The only frequencies that can matter are the frequency of the inflection and the frequency of the base. The frequency of the inflection is relevant only to access of the inflection, with more accessing failures

(no-marking errors) on low-frequency inflections than on high-frequency inflections; Stemberger (1985a) demonstrates that this is at least partly the case. However, it cannot differentiate *ended* and *mended*, since the same inflection (and the same allomorph of that inflection) is involved in both forms. The frequency of the base should matter to the access of the base, with high-frequency base forms showing fewer accessing errors than low-frequency base forms (Stemberger, 1984; Stemberger and MacWhinney, 1986a). However, there is no reason to expect that success or failure with accessing the base should in any way influence the success or failure with accessing the inflection, since the two morphemes are being accessed independently. Thus, the high frequency of a base such as *end* could not lead to a more accurate access of the following *-ed* than the low frequency base *mend*. (We argue this point in more detail below.) Thus, we expect to find frequency effects if inflected forms are stored in the lexicon in some fashion, but no frequency effects if they are not stored in the lexicon.

Study 1: Naturally Occurring Error Data

Naturally occurring speech errors were taken from the first author's corpus of 7220 errors that occurred spontaneously in natural speech. All errors were made by adult native speakers of English. (For details of collection procedures, see Stemberger 1984, 1985a.) No-marking errors on the past and perfect forms of all verbs (except for the auxiliaries *was*, *were*, *did* and *had*) were identified. In all cases, the linguistic and/or extralinguistic context clearly required the verb to be in the past or perfect form, but it was produced with the simple base form of the verb instead, as in (4).

(4) a. *Boy, that* ***draw*** *him out—drew him out.*
b. *What was it you just* ***sing****? (sang)*
c. *So we* ***test*** *'em on it. (tested)*
d. *That's what I* ***need*** *to do. (needed)*

Frequency values per million words of printed text for each of the past and perfect forms of English were derived from Francis and Kucera (1982). The verbs were divided into regular and irregular verbs. The regular and irregular verb groups were then divided into high-frequency and low-frequency groups. The high-frequency groups contained those inflected forms with a frequency of 35 occurrences per million words or greater; the low-frequency groups contained the inflected forms with a frequency less than 35. A frequency of 35 was chosen because it is close to the midpoint of the frequency distribution of all inflected forms, such that half the tokens of inflected forms are accounted for by forms with a frequency of 35 or greater, and half by forms below 35. The frequency distribution for any given inflection differs from

this, but 35 was used to ensure uniform treatment across inflections. The number of verbs in each group was summed, giving an estimate of chance for each group. The raw number of errors on verbs in each group was divided by the group frequency to yield an approximation of the error rate.

The results for the irregular verbs are given in (5).

(5)

Individual frequency	Group frequency	Number of errors	Rate
low	1735	17	.00980
high	15012	39	.00260

The first column tells if the past tense verb is of high or low frequency. The second column gives the number of verb tokens in these groups in Francis and Kucera (1982). The third column gives the number of observed errors in the corpus, and the fourth gives the approximation of an error rate discussed above. Most tokens of irregular past and perfect forms belong to high-frequency verbs. While there are more errors on high-frequency verbs, this is due solely to the greater level of chance. In fact, 30.4% of the errors occurred on low-frequency forms, which is significantly greater than the chance rate of 10.4% (χ^2 (1) = 23.90, $p < .0005$). This frequency effect entails that at least high-frequency irregular forms are stored in the lexicon.

The results for the regularly inflected forms are given in (6).

(6)

Individual frequency	Group frequency	Number of errors	Rate
low	21305	24	.00113
high	16315	11	.00072

Somewhat fewer than half of the tokens of regular past and perfect forms are associated with high-frequency forms. There is a greater error rate on low-frequency forms than on high-frequency forms, but this failed to reach significance (χ^2 (1) = 2.03, $p < .10$). The observed differences in error rate imply that at least high-frequency regularly inflected forms are stored in the lexicon, but we cannot be certain of this, since the differences were not significant. Note that the error rate on the low-frequency regular forms is nonetheless lower than the error rate on the high-frequency irregular forms (χ^2 (1) = 10.97, $p < .001$), suggesting that regular forms are easier to access than irregular forms. This might water down the effect of frequency differences through a floor effect; the production of regular forms is so easy that low-frequency forms are not at that great a disadvantage, and there were consequently too few errors in the corpus to detect a significant difference.

To more reliably test our hypothesis, we turn to an experimental task that can yield errors more quickly and under more reliable conditions (MacKay, 1976; Bybee and Slobin, 1982; Stemberger and MacWhinney, 1984, 1986a,b): having subjects produce a given member of a paradigm on presentation of a different member of the same paradigm. In Study 2, we contrast high-frequency and low-frequency verbs in this experimental task.

Study 2: Frequency and Regular Forms

A list of 40 English verbs was constructed, consisting of 10 high-frequency verbs, 10 low-frequency verbs, and 20 distractors. The high- and low-frequency verbs all ended in /t/ or /d/, since such verbs have fairly high error rates (Stemberger, 1981; Bybee and Slobin, 1982; Stemberger and MacWhinney, 1986b). All verbs were monosyllabic and were balanced for frequency and length. We chose the 10 highest-frequency regular past tense forms that fit these criteria and 10 very-low-frequency past tense forms.

The verbs were presented one at a time in the frame *was* ____*ing* in the center of the CRT display screen of an IBM personal computer. The subject read each verb silently, then spoke the past tense form of the verb out loud into a microphone. The instructions emphasized the need to react as quickly as possible. The subject's verbal responses were recorded on audiotape and analyzed for all errors. Subjects were 75 undergraduate students at Carnegie-Mellon University, receiving credit in an introductory psychology course for their participation in the experiment. Note that, unlike with the naturally occurring errors above, the exact error rate can be determined by dividing the number of errors by the total number of trials.

The results are shown in (7).

(7)

Verb type	Number of errors	Number of trials	Rate
low frequency	28	700	.037
high frequency	13	700	.017

There are significantly more errors on low-frequency forms than on high-frequency forms, using a Wilcoxon signed ranks test (two-tailed) over subjects ($p < .05$). This implies that at least high-frequency regularly inflected forms are stored in the lexicon in some fasion. The results are compatible with low-frequency forms being stored in the lexicon, but are also compatible with low-frequency forms being created on-line by a lexical-item-free rule.

There is only one other possible interpretation of these data that does not assume that high-frequency regularly inflected forms are stored in the lexicon, and this interpretation is clearly incorrect. In order for the access of the base form to affect the accessing of the inflection, one must make the access of the inflection a secondary action that follows the access of the base. One must also assume that there is a limited amount of attentional resources available for the access of both base and inflection, so that if all the resources are used up on the base, there will be none left for accessing the inflection. Since low-frequency bases use more resources, it is more likely that none will be left for the inflection, and a no-marking error will result. This predicts a greater error rate on inflections than on base forms, and implies that an error on the base form will invariably lead to an error on the inflection. These predictions are false for normal language production (Garrett, 1980; Stemberger, 1984, 1985b) and for jargon aphasia (Butterworth, 1983). The data from both normal and aphasic behavior imply a certain amount of autonomy between the access of the base and the access of the proper inflected form of that base. Since this alternative makes clearly incorrect predictions about language production, we must conclude that at least high-frequency regularly inflected forms are stored in the lexicon.

There are some data that suggest that high-frequency forms are nonetheless stored as a sequence of two morphemes. Stemberger and MacWhinney (1986a) review the natural speech-error data that suggest that inflected forms have two morphemes in them, in particular the occurrence of shifts of affixes to an incorrect position in the sentence, such as *tell-us-ing* for 'telling us.' They report that such shifts are actually more common with high-frequency inflected forms than with low-frequency inflected forms, consistent with the finding that high-frequency lexical items are differentially involved in shifts (Stemberger, 1984). Insofar as shifts are a good argument for analysis into two morphemes (e.g., they are common with inflections but very rare with similar nonmorphemic sequences), it appears that high-frequency items are so analyzed. Since shifts are generally viewed as a postlexical phenomenon (Garrett, 1980; Stemberger, 1984), this entails that inflected forms are treated as two units even postlexically; we return to this implication below.

While these data allow us to draw conclusions about high-frequency regularly inflected forms, they do not allow us to draw any conclusions about low-frequency forms. To address the question of how low-frequency forms are treated in language production, we must turn to another phenomenon that is sensitive to the presence or absence of items from the lexicon: gang effects (Rumelhart and McClelland, 1982; McClelland and Elman, 1986). Gang effects arise in interactive activation models of cognitive processing. Since such models are beyond the scope of this chapter, we just describe the effects without going into detail on how they arise. The interested reader is referred to the articles cited above.

In a gang effect, several words in the lexicon that are similar in form reinforce the patterns of phonemes or letters that they have in common. The strength of the gang (i.e., the degree to which they reinforce the shared pattern) is a function of two things. First, the effect is greater when more phonemes or letters are shared. All other things being equal, a gang with three shared phonemes has more effect than a gang with two shared phonemes. Similarity of the phonemes involved in terms of position in the word and in terms of contiguity also has an effect. Homogeneous gangs that have the shared phonemes or letters in identical positions with the same items contiguous are stronger gangs. Second, the number of members in the gang has an effect. If two gangs are identical in terms of similarity of phoneme or letter patterns, the stronger gang is the one that has more words in it.

Gang effects can influence the pronunciation of nonwords such as *mave*.

The letter sequence *ave* can be pronounced /æ:v/ (as in *have*) or /eiv/ (as in *cave*). When we examine the lexicon, we find that there is only one word where *ave* is pronounced /æ:v/, while there are 14 monosyllabic words where it is pronounced /eiv/. All 14 /eiv/ words reinforce that pronunciation, while only *have* reinforces the alternative pronunciation. The greater size of the /eiv/ gang leads to greater reinforcement of the pronunciation /eiv/ and leads to the result that subjects are far more likely to pronounce *mave* as /meiv/ than as /mæ:v/. In essence, gang effects are a new form of analogy in which a word that is being processed is compared to a larger number of words than in traditional analogy (where it is compared to only a single word), and these words affect processing by each contributing a small amount of reinforcement to bias the system toward a certain outcome. By their very nature, gang effects are only possible if forms are stored in the lexicon, and must be present if forms are stored in the lexicon.

Gang effects may be able to account for some of the known phenomena in inflectional processing. Since gang effects are the result of interactions among lexical items, it is necessary to assume that inflected forms are stored in the lexicon in order to obtain any gang effects. Novel words can be made into, for example, a past tense form by interaction with known past tense forms in the form of gangs. The past tense form of *bick* will be *bicked* because a final /t/ is reinforced by a gang made up of past tense forms like *tricked, picked, kicked*. Note that there is no need to explicitly build in any connection between base and inflected forms such as *trick* and *tricked* to get this effect; it derives solely from the semantic and phonological similarities between different inflected forms. In Rumelhart and McClelland's (1986) computer model of the acquisition of past tense forms, they show that this is possible. They also argue that no-marking errors on regular verbs can arise through interference from gangs of irregular past tense forms. For example, their model produced a high rate of no-marking errors on the verb *kid* because of its resemblance to a gang of irregular verbs like *rid* that do not show an *-ed* in the past tense. The presence of gang effects with inflected forms requires lexical storage of inflected forms, while the absence of gang effects implies that inflected forms are created by lexical-item-free rules.

We first demonstrate experimentally that gangs of past tense forms can and do arise, and that these gangs can affect the processing of known regularly inflected forms. We begin by looking for the effect of irregular forms on regular forms. There are a number of regularly inflected forms with bases that resemble irregular forms. For example, the regular verb *spank* resembles the irregular past tense forms *drank*, *sank*, and *stank*. We would predict that an *-ank* gang would form and reinforce their shared phonological characteristics, including the absence of a final /t/. The word *spank* should thus have a high rate of errors without a final /t/, that is, of no-marking errors. There are other regular verbs which share fewer phonemes with the

irregular forms that they resemble. For example, the regular verb *snore* shares only two phonemes with the irregular past tense forms *wore*, *bore*, *swore*, and *tore*. The regular verb *chew* shares only a single phoneme with the irregular past tense forms *knew*, *grew*, *flew*, *drew*, *blew*, *threw*, and *slew*. Gang effects predict that the gangs will be stronger and have greater effects on processing as a function of the number of phonemes shared by the members of a gang, given that the gangs are of roughly equal size (as they are in the examples given here). We thus predict that the rate of no-marking errors will be greatest for verbs like *spank* that share three phonemes with their gang, intermediate for verbs like *snore* that share two phonemes with their gang, and least for verbs like *chew* that share only one phoneme with their gang. Study 3 examined this question.

Study 3: Regular Verbs that Resemble Gangs of Irregulars

Three lists of verbs were made up, one for each of the types of regular verbs mentioned above. Each list contained 16 monosyllabic regular verbs that resemble irregular past tense forms, 16 regular verbs that do not resemble irregular past tense forms, 16 regular verbs that do not resemble any irregular past tense forms very closely, and 16 unrelated irregular verbs that acted as distractors; the two groups of regular verbs in each list were balanced for length and frequency. List 1 contained regular verbs that shared an average 3.1 phoemes with the members of their gang; half had the vowel /æ:/ and half had the vowel /ʌ/, and all ended in /ŋk/ (as in *spank* and *flunk*). List 2 contained regular verbs that shared an average 2.1 phonemes with the members of their gang; all had the vowel /ou/ and ended in /k/, /r/, or /z/ (as in *choke*, *snore*, and *doze*). List 3 contained regular verbs that shared an average 1.1 phonemes, with the members of their gang; all ended in /u:/ (as in *chew*). The procedure was the same as in Study 2. A different group of subjects were run on each list, all from the same subject pool as in Study 2.

No-marking errors were identified and analyzed. The error rate on each group of verbs is presented in Fig. 6.1. There were significantly more no-marking errors on verbs that resembled irregular past tense forms than on verbs that did not for those forms that shared approximately two or three phonemes with those irregular forms. There were also more no-marking errors on regular verbs that shared only one phoneme with irregular forms than on other regular verbs, but this did not reach significance. There was a significant effect of the number of phonemes shared with the irregular forms ($p < .05$), with more errors on forms that shared more phonemes. In fact, the error rate was almost a linear function of the number of shared phonemes.

To further reinforce the effect of shared phonemes, we can do a post hoc examination of the performance of the subjects who saw List 1 on other

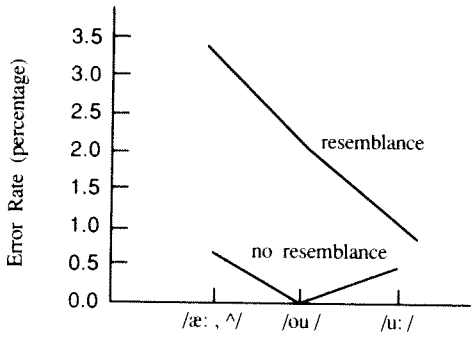

Fig. 6.1: Error rate for verbs in Study 3.

regular verbs that also had the vowels /æ:/ and /ʌ/ but shared few other phonemes with irregular forms. Each subject saw 11 such verbs, for example, *pass* and *bump*, sharing an average 1.2 phonemes with irregular forms. There were only 3 errors on these regular verbs out of 550 trials, as compared to 26 errors on the List 1 verbs out of 750 trials. This difference is significant ($\chi^2(1) = 12.42$, $p < .001$). The number of shared phonemes does appear to have a great effect on error rate for regular verbs. These results are predicted by gang effects, and we have been able to devise no other explanation to account for them. (The schema hypothesis of Bybee and Slobin, 1982; and the prototype hypothesis of Bybee and Moder, 1983, cannot account for the differences in the rate of no-marking errors between the different groups of verbs, at least as currently described. See Stemberger, 1983; Stemberger and MacWhinney, 1986b, for additional arguments against these approaches.)

These data provide additional evidence that irregular inflected forms are stored in the lexicon, as the theory predicts. However, it is not particularly surprising that irregular forms are stored in the lexicon. It is of greater interest to determine whether all regularly inflected forms are stored in the lexicon. We must thus examine regular verbs that resemble regularly inflected forms. If regularly inflected forms are in general in the lexicon, then a gang effect should be present. If few regularly inflected forms are in the lexicon, then there should be no gang effect. It is relatively difficult to examine this question using past tense forms, since there are so many irregular past tense forms. We thus shift to examining third person singular present tense forms, where there are almost no irregular forms at all. Any obtained gang effects can then be due only to regularly inflected forms being stored in the lexicon. We contrast base forms that closely resemble present tense forms (as *gaze* resembles *plays*) and base forms that do not (as with *buzz* and *cross*). Gang effects predict that base forms that resemble actual inflected forms should show a greater rate of no-marking errors than forms that do not.

Study 4: Regular Verbs that Resemble Regularly Inflected Forms

A list of 90 monosyllabic English verbs was constructed. Thirty of these verbs ended in /s/ and /z/ and were designed to address the current hypothesis. The other 60 verbs ended in other segments and were designed to address other hypotheses; they are not discussed further here (but see Stemberger and MacWhinney, 1984, 1986b). Half of these closely resembled regularly inflected present tense forms and half did not. The groups were balanced for length and frequency. For the verbs that resembled actual present tense forms (e.g., *cause, coax, gaze, please*), there was an average of 3.3 present tense forms that shared at least three contiguous phonemes with them, which is comparable to the size of the gangs of irregular past tense forms in Study 3, and there was an average of 19.5 present tense forms that shared at least two contiguous phonemes. In contrast, for those verbs that did not closely resemble actual present tense forms (e.g., *buzz, force, race, toss*), there was an average of 0.3 present tense forms that shared at least three contiguous phonemes with them, and an average of 6.5 that shared at least two contiguous phonemes. These differences in gang size and the number of shared phonemes would lead us to predict a greater rate of no-marking errors on the verbs that resembled actual present tense forms.

The procedure was similar to that in Study 2, except that the frame *is ____ing* was used, and subjects produced the present tense forms of the verbs. There were 60 subjects from the same subject pool as in Study 2.

No-marking error rates on the two groups of verbs were identified and analyzed. the results are shown in (8).

(8)

Item type	Number of errors	Number of trials	Rate
Resemblance	36	900	.040
No resemblance	32	900	.036

The two groups of verbs show essentially identical error rates. Number of shared phonemes and gang sizes thus have no apparent effect. Gang effects are not present. Regular verbs are not penalized for physically resembling regularly inflected forms.

The lack of a gang effect here implies that regularly inflected forms are not in general stored in the lexicon. If they are stored in the lexicon, we would have obtained a gang effect. We conclude that speakers use inflectional rules instead, creating regularly inflected forms on-line during language production by adding an affix directly to the base form.

We have arrived at an apparent contradiction. In Study 2, we demonstrated, via a frequency effect, that at least high-frequency regularly inflected forms are stored in the lexicon in some fashion. In Study 4, we demonstrated, via (the lack of) a gang effect, that regularly inflected forms are not in general stored in the lexicon. This contradiction is only apparent. In point of fact,

the results of the two studies are quite compatible. The results of Study 2 require only that high-frequency regularly inflected forms be stored in the lexicon and are compatible with the hypothesis that low-frequency regularly inflected forms are not stored in the lexicon. The results of Study 4, on the other hand, require that low-frequency regularly inflected forms are not stored in the lexicon but are compatible with the storage of just high-frequency regularly inflected forms. The reason for this is quite simple. There are relatively few high-frequency present tense forms that closely resemble the base forms of regular verbs. How frequent an inflected form must be before it is stored in the lexicon is unknown, but let us suppose a (probably too-low) frequency of 15 occurrences per million words. The only gang of present tense forms that could have an impact on present tense verbs that physically resemble them is made up of four verbs that end in /ouz/ (*knows*, *goes*, *shows*, and *grows*), as no other forms with a frequency over 15 are similar enough phonologically to form a strong gang. We would thus predict that resembling actual present tense forms would in general have no effect on processing, because the present tense forms that regular verbs resemble are in general too infrequent to be stored in the lexicon. We might expect a higher error rate only on verbs that end in /ouz/, since this one gang should be present in the lexicon. There is in fact a nonsignificant difference in that direction in our data, with 25.0% of the errors in the high-resemblance group occurring on such verbs, compared to 18.8% expected by chance. However, the difference is small and nonsignificant, and we did not set up our stimuli to be able to test this suggestion. This prediction should be followed up. There are too few stimuli available to test it with the present tense, but there are enough to test it with perfect forms and (marginally) with plural forms. At any rate, the results of Study 4 are problematical for the gang effect hypothesis if all regularly inflected forms are stored in the lexicon, and are compatible with it if only high-frequency regularly inflected forms are stored in the lexicon.

Discussion and Conclusion

We began by laying out three generic positions taken by researchers on the representations of inflected forms: having them stored fully suppletively, having them stored as morphologically complex lexical items or through lexically conditioned rules, and having no inflected forms stored but creating them on-line via lexical-item-free rules. Our data allow us to say that irregular forms are stored, but do not allow us to say how they are stored. However, our data require storage of high-frequency regularly inflected forms as bimorphemic or with lexically conditioned inflectional rules and require that low-frequency regularly inflected forms not be in the lexicon at all. We prefer the alternative given in (3) for high-frequency regulars, with lexically conditioned rules (see Stemberger, 1985b), but we know of no data that require it.

Given that high-frequency regularly inflected forms are stored, we can consider at what stratum they are stored in a lexical phonology framework. Irregular inflected forms are placed in Stratum 1 by Kiparsky (1982) and Halle and Mohanan (1985), while regularly inflected forms are placed on the final lexical stratum. Since high-frequency forms are explicitly listed in the lexicon, should they be at Stratum 1? Halle and Mohanan implicitly assume that this must be the case, since their analysis of engma in English requires that *longer*, *stronger*, and *younger* were placed in Stratum 1 at a historical period when they were fully regular semantically and phonologically. However, this is not possible for high-frequency past tense forms. Many of our high-frequency forms have long vowels, for example, *need*. If inflection were to take place at Stratum 1, the other rules that are posited for that stratum would give us /nɛd/, just as they give us /spɛd/ as the past tense of *speed*. Thus, regularly inflected forms are lexicalized on the final lexical stratum, not at Stratum 1. (Ramifications for Halle and Mohanan's analysis of engma are left to the reader.)

The storage of high-frequency regularly inflected forms has ramifications for compounding. Kiparsky has noted that irregular plurals can show up as the first part of a compound, but that regular plurals cannot. Sproat (1985) and Anderson (this volume) have suggested that this is because irregular forms are listed in the lexicon but regular forms are not. Our data suggest that this explanation is incorrect, since some regulars are listed. Whatever underlies this phenomenon, it cannot be the simple fact of whether something is stored or not.

We mentioned above that affix shifts (*tell-us-ing*) appear to be a postlexical phenomenon, reflecting a very late problem with the ordering of lexical items in sentences (Garrett, 1980; Stemberger, 1984). Shifts imply that high-frequency regularly inflected forms are treated as two morphemes even at this late point in processing. In lexical phonology, however, it is assumed that all inflected forms are treated as monomorphemic elements postlexically, due to the operation of Bracketing Erasure at the end of each lexical stratum. The shift data imply that there is no operation of Bracketing Erasure in language production, since the status of the affix as an independent morpheme must be preserved in order to account for the data. (Note that there are also problems for Bracketing Erasure in linguistic theory, as certain morphological phenomena such as morphological haplology require that morphological structure be maintained through all lexical strata and into the postlexical domain; see Stemberger, 1981, for relevant data.)

We note one possible problem for our approach as well. We have been assuming that the prototype effects reported by Bybee and Moder (1983) derive

from gang effects, as Bybee (this volume) also assumes. Gang effects imply that the largest effects come with the greatest number of phonemes shared with the greatest number of words, but Bybee and Moder report that the greatest effect in their study was with /ʌŋ / rather than /ʌŋk/. This is counter to our predictions. Further, studies of the phonological processing of word-initial clusters in language production by Stemberger and Treiman (1986) have found that competing single consonants tend to reinforce a cluster that contains both of the consonants involved. By implication, the presence of irregular forms like *rung* and *stuck* should reinforce /ʌŋ k/ (which shares two phonemes with all the forms and contains a cluster with both /ŋ / and /k/) more than it reinforces either /ʌŋ / or /ʌk/. Bybee and Moder's morphological data thus appear to be somewhat problematical, given what we currently know about phonological processing. However, more study is needed on the phonological processing of word-final consonant clusters in language production (which could conceivably differ from word-initial clusters), as well as on these gang effects. We point out this possible difficulty, however, to emphasize the point that morphological processing cannot be studied completely separately from phonological processing. The two interact in many ways, and theories of morphological processing may fail because they are incompatible with what we know about phonological processing.

In conclusion, we have presented data that we think provide an answer to the question of whether inflected forms are stored in the lexicon. The answer is yes and no. Irregular forms appear to be stored in the lexicon, as do high-frequency regularly inflected forms. Low-frequency regularly inflected forms, which include the inflected forms of most English words, are not stored in the lexicon, however. Further study is now needed to determine whether stored inflected forms might be created in language production using lexically conditioned rules.

Acknowledgments

This research was supported by a Sloan Post-Doctoral Fellowship to Joseph Paul Stemberger while he was at CMU, and by NICHHD Grant 1-R01-HD17790 to Brian MacWhinney. The data in Studies 1 and 2 have been reported in another context in Stemberger and MacWhinney (1986a). The data in Studies 3 and 4 have been reported in other contexts in Stemberger and MacWhinney (1984, 1986b).

References

Aronoff, M. (1976). "Word Formation in Generative Grammar." MIT Press, Cambridge, MA.

Atkinson, R.C., and Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. *In* "The Psychology of Learning and Motivation" (K.W. Spence and T.J. Spence, eds.), Vol. 2. Academic Press, New York.

Bradley, D. (1980). Lexical representation of derivational relation. *In* "Juncture" (M. Aronoff and M.L. Kean, eds.). MIT Press, Cambridge, MA.

Butterworth, B. (1983). Lexical representation. *In* "Language Production: Vol. 2. Development, Writing, and Other Language Processes" (B. Butterworth, ed.), (pp. 257–294). Academic Press, London.

Bybee, J., and Moder, C. (1983). Morphological classes as natural categories. *Language* **59**, 251–270.

Bybee, J., and Slobin, D.I. (1982). Rules and schemas in the development and use of the English past tense. *Language* **58**, 265–289.

Francis, W.N., and Kucera, H. (1982). "Frequency Analysis of English Usage: Lexicon and Grammar." Houghton Mifflin, Boston.

Garrett, M. (1980). Levels of processing in sentence production. *In* "Language Production, Vol. 1: Speech and Talk" (B. Butterworth, ed.), (pp. 177–220). Academic Press, London.

Halle, M., and Mohanan, K.P. (1985). Segmental phonology of English. *Linguistic Inquiry* **16**, 57–116.

Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language* **51**, 639–671.

Kempley, S.T., and Morton, J. (1982). The effect of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology* **73**, 441–454.

Kiparsky, P. (1982). From cyclic phonology to lexical phonology. *In* "The Structure of Phonological Representations" (H. van der Hulst and N. Smith, eds.), (Part 1), pp. 130–175. Foris, Dordrecht.

Leben, W., and Robinson, O.W. (1977). Upside-down phonology. *Language* **53**, 1–20.

Lukatela, G., Gligorijevic, B., Kostic, B., and Turvey, M.T. (1980). Representation of inflected nouns in the internal lexicon. *Memory and Cognition* **8**, 415–423.

McClelland, J.L., and Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology* **18**, 1–86.

MacKay, D.G. (1976). On the retrieval and lexical structure of verbs. *Journal of Verbal Learning and Verbal Behavior* **15**, 169–182.

MacKay, D.G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review* **89**, 483–506.

MacWhinney, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development* **43**, whole no. 1.

Peters, A.M. (1983). "The Units of Language Acquisition." Cambridge University Press, Cambridge England.

Rumelhart, D., and McClelland, J. (1982). An interactive activation model of the effect of context on perception, Part 2. *Psychological Review* **89**, 60–94.

Rumelhart, D., and McClelland, J. (1986). On learning the past tenses of English verbs. *In* "Parallel Distributed Processing: Explorations in the Microstructure of Cognition (D. Rumelhart and J. McClelland, eds.), Vol. 1. Bransford, Cambridge, MA.

Sproat, R.W. (1985). *On deriving the lexicon*. Unpublished doctoral dissertation, MIT.

Stanners, R.F., Neiser, J.J., Hernon, W.P., and Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior* **18**, 399–413.

Stemberger, J.P. (1981). Morphological haplology. *Language* **57**, 791–817.

Stemberger, J.P. (1983). Inflectional malapropisms: Form-based errors in English morphology. *Linguistics* **21**, 573–602.

Stemberger, J.P. (1984). Structural errors in normal and agrammatic speech. *Cognitive Neuropsychology* **1**, 281–313.

Stemberger, J.P. (1985a). Bound morpheme loss errors in normal and agrammatic speech: One mechanism or two? *Brain & Language* **25**, 246–256.

Stemberger, J.P. (1985b). An interactive activation model of language production. *In* "Progress in the Psychology of Language" (A. Ellis, ed.), Vol. 1. Erlbaum, London.

Stemberger, J.P., and MacWhinney, B. (1984). Extrasyllabic consonants in CV phonology: An experimental test. *Journal of Phonetics* **12**, 355–366.

Stemberger, J.P., and MacWhinney, B. (1986a). Frequency and the lexical storage of regularly inflected words. *Memory and Cognition* **14**, 17–26.
Stemberger, J.P., and MacWhinney, B. (1986b). Form-oriented inflectional errors in language processing. *Cognitive Psychology* **18**, 329–354.
Stemberger, J.P., and Treiman, R. (1986). The internal structure of word-initial consonant clusters. *Journal of Memory and Language* **25**, 163–180.
Walsh, T., and Parker, F. (1983). The duration of morphemic and nonmorphemic /s/ in English. *Journal of Phonetics* **11**, 201–206.

Part II

Rules and Representations