Journal of Memory and Language **40,** 41–61 (1999)
Article ID jmla.1998.2607, available online at http://www.idealibrary.com on IDEAL®

Frequency Effects and the Representational Status of Regular Inflections

Maria Alegre and Peter Gordon

University of Pittsburgh

There have been many recent proposals concerning the nature of representations for inflectional morphology. One set of proposals addresses the question of whether there is decomposition of morphological structure in lexical access, whether complex forms are accessed as whole words, or if there is a competition between these two access modes. Another set of proposals addresses the question of whether inflected forms are generated by rule-based systems by connectionist type associative networks or if there is a dual system dissociating rule-based regular inflections from association-based irregular inflections. A central question is whether there are whole-word representations for regularly inflected forms. A series of five lexical decision experiments addressed this question by looking at whole-word frequency effects across a range of frequency values with constant stem-cluster frequencies. Frequency effects were only found for inflected forms above a threshold of about 6 per million, whereas such effects were found for morphologically simple controls in all frequency ranges. We discuss these data in the context of two kinds of dual models and in relation to competition models proposed within the connectionist literature.

Two surprisingly distinct lines of research in psycholinguistics have addressed the psychological status of inflectional morphology. One line of research comes from the field of language acquisition and has concentrated on the contrast between regular and irregular inflectional forms. The second line of research comes out of the adult lexical access literature and has concentrated on how morphologically complex forms are represented and how such representations are accessed.

In comparing these two literatures, one finds that both contain what are called "dual models," but the meanings of the term differ in the two literatures. Within the acquisition literature, the notion of a dual model arose from an evaluation of claims made within the connectionist community that rule-like behavior could be emulated within a parallel distributed processing (PDP) system that contained no representation of such rules (Rumelhart & McClelland, 1986). Challenging these claims, Pinker (1991) proposed that rule systems are required for regular inflectional morphology, whereas irregular inflectional morphology exhibits characteristics of an associative/connectionist network.

Within the adult psycholinguistics literature, the term "dual model" refers to the access system for complex morphological forms, both inflectional and derivational. Proponents of these dual models assume that access to complex forms can occur either through whole-word representations or through morphological decomposition (Anshen & Aronoff, 1988; Baayen, 1991; Burani & Caramazza, 1987; Burani & Laudanna, 1992; Caramazza, Laudanna, & Romani, 1988; Caramazza, Miceli, Silveri, & Laudanna, 1985; Chialant & Caramazza, 1995; Frauenfelder & Schreuder, 1991; Laudanna, Badecker, & Caramazza, 1989, 1992; Schreuder & Baayen, 1995; Stemberger & MacWhinney, 1988). In contrast, single-route models propose access only through whole-word representations (Butterworth, 1983; Cole, Beauvillain, & Segui, 1989; Cutler, 1983; Emmorey, 1989; Lukatela, Gligorijevic & Kostic, 1980; Rubin, Becker, & Freeman, 1979; Segui & Zubizarreta,

The authors thank David Plaut, Charles Perfetti, Alan Juffs, David Allbritton, and Carl Johnson for helpful comments, and Carey Ryan for advice on statistical issues. The present study was completed in partial fulfillment for the degree of Doctor of Philosophy at the University of Pittsburgh by Maria Alegre.

Address correspondence and reprint requests to Peter Gordon, Department of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260. E-mail: peter+@pitt.edu.

0749-596X/99 $30.00

1985). The present study attempts to bring together some of the issues raised in these two lines of inquiry and to test predictions regarding the representations of regularly inflected forms.

Beginning with the language acquisition literature, Pinker (1991) proposed that forms with regular inflections (e.g. *walked*) are handled by rule, whereas irregular inflections (e.g. *went*) are handled by associative memory. To make such a claim, it is critical to have a clear set of criteria for distinguishing between rule-based processes and generalizations based on associative principles. According to Pinker, a word-formation rule is an abstract operation that concatenates an affix with a variable standing for the stem (Marcus, Brinkmann, Clahsen, Wiese, Woest, & Pinker, 1993; Pinker, 1991; Pinker & Prince, 1991). Rule-based processes are thus default generalizations and apply freely to all items of the right category unless application is blocked by a competing irregular form (e.g., *went* blocks the generation of **goed*).

On the other hand, irregulars are stored as memorized pairs with their stems and their morphological relations. These are linked in an associative structure with certain connectionistlike properties (Pinker & Prince, 1991). The irregular system allows some degree of generalization. This is accomplished by exploiting patterns of similarity between irregular forms. The more similar a stem is to other frequent stems known to undergo irregular inflection, the more likely it is to also undergo such a process. For example, when asked to provide the past tense form for a nonce verb such as *spling*, people are very likely to produce *splang*—in analogy to *spring–sprang, ring–rang, sing–sang, drink–drank*, and so on (Bybee & Moder, 1983; Prasada & Pinker, 1993).

Productive generalizations based on resemblance to families of stored forms are known as "gang effects" (Stemberger & MacWhinney, 1988). The generalization of irregular patterns is highly sensitive to gang effects in both adults (Bybee & Moder, 1983; Prasada & Pinker, 1993) and children (Bybee & Slobin, 1982; Marcus, Pinker, Ullman, Hollander, Rosen, & Xu, 1993; Pinker & Prince, 1988; Xu & Pinker, 1992). On the other hand, gang effects appear to play no role in the productive generation of regular inflections. For example, gangs of regular forms sharing phonological properties do not facilitate production of nonce items sharing those properties, nor do they enhance the acceptability of such nonce forms (Prasada & Pinker, 1993; Marcus et al., 1993).

These facts notwithstanding, the contrast between "rule-based" and "stored" is far from clear. For example, is it the case that only irregular inflections like *sang* are represented in memory, while regularly inflected forms like *walked* are never represented in memory? Prasada and Pinker suggest two possible versions of the dual model. In the strong version, there are no whole-word representations of regularly inflected forms. In the weak version, some storage of regularly inflected forms is possible. Prasada and Pinker state that ". . . prior storage of regulars is possible, and thereby might offer mild analogical assistance to the generalization of the regular inflection to similar forms, but generalization never depends on prior storage of a similar form . . ." (Prasada & Pinker, 1993, p. 9). If the weak version of the dual model is to remain consistent with the lack of gang effects for regular inflections, then it is fundamental that storage be limited to a relatively small number of items.

One way to test the two versions of Pinker's model is to look at frequency effects. If regularly inflected forms are not represented as whole words, then the base form must be accessed every time an inflected word is processed. Thus, access speed should depend on the cluster frequency of the base form, defined as the aggregated frequencies of the base and all its inflectional variants (e.g., for *walk*, the summed frequencies of *walk, walks, walked,* and *walking*). On the other hand, if whole-word representations are available for regularly inflected forms, then the frequency of the inflected form itself should determine access time.

Turning to the lexical access literature, we find a family of dual models. These models share the assumption that access to morphologically complex forms (inflectional and derivational) can occur through either a whole-word representation or through a decompositional route. These two routes are claimed to work in parallel and compete with one another (Anshen & Aronoff, 1988; Baayen, 1991; Burani & Caramazza, 1987; Burani & Laudanna, 1992; Caramazza et al., 1985; 1988; Chialant & Caramazza, 1995; Frauenfelder & Schreuder, 1991; Laudanna et al., 1989, 1992; Schreuder & Baayen, 1995; Stemberger & MacWhinney, 1988).

The models differ considerably in how explicit they are and in how they specify the nature of the competition between access routes. However, there is agreement that the whole-word access system works faster than the morpheme access system and will win out in most cases. An exception to this is low-frequency complex forms, especially those where the constituent morphemes are high in frequency. For example, the complex form *cleans* is very low in whole-word frequency, but the constituents *clean* and -*s* are each very high in frequency. Such forms are predicted to be accessed through decomposition rather than the whole-word route. In general, whole-word frequency effects should be obtained for items in the higher end of the frequency distribution but not for items in the lower end. The threshold for defining “high” versus “low” frequency has not been specified in any of these models and remains an empirical matter to be addressed in the present studies.

Dual models, which assume the existence of rule systems, are challenged by connectionist approaches, which provide alternative perspectives on many of these issues. In connectionist systems, network nodes represent minimal phonetic and/or semantic units. Weighted connections between nodes capture patterns of regularity in the input. Much of the debate within the connectionist literature concerns whether regular and irregular processes can be handled within a single system using associative networks. This discussion addresses the regular–irregular distinction in both morphology (Bybee, 1995; MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1991; Rumelhart & McClelland, 1986; Stemberger, 1995) and spelling–sound correspondences (Plaut, McClelland, & Seidenberg, 1996; Seidenberg & McClelland, 1989). While these models contain uniform representations, they produce outputs that distinguish between regular and irregular processes by exploiting their different distributional properties.

In the case of morphology, Bybee (1995) proposed that complex forms are stored in an associative network where recurring phonological and semantic patterns are represented as links between units. Morphological relations are defined by parallel phonological and semantic connections found in multiple sets of items. Morphemes do not have any independent status or representation and are considered to be epiphenomenal. For example, the word *started* is connected to the word *starting* through the overlapping sequence *start* and to the word *walked* through the common ending *ed*. Such multiple connections allow for rule-like behavior to emerge as generalizations over properties of stored items.

In Bybee’s model, activation levels are the result of two factors: lexical strength and lexical connections. Lexical strength is determined by the frequency of the lexical item. Lexical connections are the pattern of weights associated with the interconnections between related items, as illustrated above. For high-frequency words, lexical strength would be high and would swamp activation due to lexical connections. Low-frequency words would be weak in lexical strength and thus rely on lexical connections for access.

Like the dual access model, Bybee’s model predicts that whole-word frequency effects should be found only above a certain frequency threshold. However, in Bybee’s model, the strength of the frequency effect depends on the composition of the cluster. The number of item types that compose the stem cluster (i.e., stem and inflected forms) determines the strength of lexical connections for that cluster. Therefore, a cluster containing many inflected forms will have strong lexical connections, which would compete strongly against the individual lexical strength of any inflected form within that cluster. When lexical connections are stronger than lexical strength, then whole-word frequency effects should be weak to nonexistent. Conversely, when a cluster contains only a small number of inflectional types, the connection strengths are weak and whole-word frequency effects should be stronger. In other words, Bybee's model predicts an interaction between the number of inflectional types within a cluster and whole-word frequency effects.

This aspect of the model can be tested by comparing nouns and verbs. Noun paradigms are inherently weak, as only two forms, singular and plural, are usually observed. On the other hand, regular verbs usually have four different forms (e.g., *play, plays, playing* and *played*). Thus, for any given cluster frequency value, whole-word frequency effects should be more readily obtained for nouns than for verbs. The opposite pattern should hold for cluster frequency effects. We will test these predictions against the data found in the following studies.

Stemberger's (1995) connectionist model differs from Bybee's model in several respects. In his model, Stemberger attempted to account for the lack of frequency effects for regularly inflected forms compared to strong frequency effects for irregular inflections (Stemberger & MacWhinney, 1988; Prasada, Pinker, & Snyder, 1990). In the case of regular inflections, lexical strength, based on word frequency, faces strong competition from lexical connections to all other words sharing the same ending (e.g., all words ending in *-ed*). Given that this group is extremely large, then the strength of these associations should swamp any activation due to the frequency of the individual inflected word. Rather than predicting a threshold for finding a frequency effect for regular inflections, Stemberger's model suggests that no whole-word frequency effects should be found.

What does current research tell us about the representation of regular inflections? Unfortunately, the results are ambiguous. One set of studies has used priming to examine the relation between base forms and inflected forms (Fowler, Napps, & Feldman, 1985; Hernon & Hall, 1979; Napps, 1989; Stanners, Neisser, Napps, & Fowler, 1987). Several of these studies have shown that regularly inflected forms prime their base forms as strongly as the base forms themselves (identity priming). However, such results only show that base forms are activated in the process of accessing regularly inflected forms. Such results might be expected under both a morpheme access system and a whole-word access system, if one assumes that whole-word representations of related forms cluster together in the lexicon (Lukatela et al., 1980; Butterworth, 1983).

Other studies have manipulated frequency as a way to investigate representational and access issues. Prasada et al. (1990) provided evidence against the existence of whole-word representations for regularly inflected forms. They used a production task in which participants were presented with the base form of a verb (e.g., *walk*) and were asked to produce the corresponding past tense form (*walked*). The items were constructed in pairs matched for cluster frequency but differing in whole-word frequency. For example, the verbs *jump* and *boil* are equally frequent, whereas *jumped* is much more common than *boiled.* If the past tense forms are accessed through the base form, then no differences in reaction times are expected. On the other hand, if regularly inflected words are represented as such, and there is no on-line composition, then *jumped* should be produced faster than *boiled.* Prasada et al. found no differences in RT for the regularly inflected words. However, when irregular forms were tested in the same fashion, high-frequency past tense forms were produced significantly faster than low-frequency past tense forms. These results suggest that regular inflections are put together on line, while irregular inflections are stored as whole words.

One problem with the Prasada et al. (1990) study is their method for computing cluster frequency. They included the frequencies of the stem plus the inflectional forms other than the past tense. This is rather unconventional since the cluster frequency normally includes all inflectional forms. Presumably the reason for this decision was that they were comparing regular and irregular inflections. Irregular inflections are not normally included in the cluster frequency measure because they are not likely to be accessed through the stem. However, by also excluding the regular past tense inflections from the cluster frequency computation, Prasada et al. had items where the word frequency of the past tense form was sometimes higher than its associated stem–cluster frequency. If regularly inflected forms are accessed through their stems, then they should also contribute to the activation levels associated with the frequency effect. Another problem in interpreting the Prasada et al. (1990) study is the nature of the task. Since the participants read the stem form before producing the past tense form, it is not surprising that they would use a compositional route by adding the past tense morphology. It could be the case that whole-word representations are available for regularly inflected forms, but are simply not used under these circumstances.

Other studies in this area have employed the lexical decision procedure and present a rather different picture. Taft (1979) and Burani, Salmaso, and Caramazza (1984) have looked at whole-word frequency effects in lexical access for regular inflections. These studies tested pairs of regularly inflected items where one word in each pair was high in whole-word frequency and the other was very low, but both were matched in cluster frequency (stem frequency plus all inflections). Both studies obtained a significant whole-word frequency effect, suggesting that regular inflections are stored as whole words, in contrast to Prasada et al. (1990).

Once again, there are problems with the design of these studies. In Taft (1979), 9 out of 20 low-frequency items involved stems that were denominal verbs (e.g., *classed*, *numbered*, *sized*), while none of the high-frequency items had derived stems. If denominal verbs are psychologically derived, then this requires two processes (zero derivation plus inflection), which would increase reaction times for many of the low-frequency items. Alternatively, if denominal verbs are stored separately from their related noun forms, then Taft's measure of cluster frequency was inaccurate because it included the frequency of the noun base form as well as the verb form, and the noun form tended to be the major contributor to the cluster frequency count. In either case, RTs for inflected forms based on denominal verbs should be slow, regardless of their representational status.

Burani, et al., (1984) failed to find an effect of word frequency using Taft's procedure in Italian. They did find an effect when they modified the design by testing pairs of words that shared the same stem but had different inflections (e.g., *sent-ire* versus *sent-ivi*). This optimally controls for cluster frequency issues. However, the frequency of the inflected word is likely to be correlated with the frequency of the affix itself. To use English examples, if we were to test a low-frequency inflection like *cleans* versus a high-frequency relative such as *cleaned*, the frequency of the affixes themselves (i.e., *-s* vs *-ed*) would be confounded: past tense morphology is much more frequent than third person *-s*. Thus, even if both *cleans* and *cleaned* were to be accessed through decomposition, attaching *-ed* might take less time than attaching *-s*. Once again, differences in RT might be expected whether inflected words are stored as single units or not.

A common problem in previous studies is that frequency has been treated as a categorical variable (high versus low). This may overestimate or underestimate the existence of whole-word frequency effects, depending on which part of the frequency distribution is being sampled from. Furthermore, this method cannot detect discontinuities in frequency effects that are predicted by dual access models. In the present paper, we will describe several experiments that treat frequency as a continuous variable in an attempt to provide an unambiguous picture of the representational status of regular inflections.

EXPERIMENT 1

Participants

Thirty-one undergraduate students from the Psychology subject pool at the University of Pittsburgh participated in the experiment. In this and subsequent experiments, all participants completed the task and all were native speakers of English.

Materials

Sixty-six experimental items were selected to: (a) be regularly inflected, (b) have similar stem–cluster frequencies, and (c) be evenly distributed in terms of log word frequency. Cluster frequency ranged from 99 to 122 (log 2 to 2.08) based on Francis and Kučera (1982). This range is small but not zero and so cluster frequency is still included as a variable in the analysis. Word frequency varied from 0 to 101 (see Appendix). Items were selected so that word length, syntactic category (noun vs verb), and cluster frequency had very low correlations among themselves and with respect to the critical experimental variable: whole-word frequency (min = .03; max = .2). Filler items included 66 simple words (not inflected); 66 "inflected" nonwords, where an inflectional ending was attached to a nonce base form (e.g., **pramed*); and 66 simple nonwords.

Procedure

Participants performed a lexical decision task on a PC using the Micro Experimental Lab (MEL) software (Schneider, 1988). After reading the instructions they began with 40 practice trials. They were then tested on the experimental and filler items, which were presented in a different random order for each subject. Participants were allowed 3 1-min breaks during the experiment.

On each trial, an asterisk appeared in the middle of the screen for 1.5 s and then an item was displayed, centered around the asterisk position. Participants were instructed to press the "1" key each time they saw a word, and the "2" key each time they saw a nonword. The presentation of the item was terminated by the subject's response. If participants took longer than 1.5 s to respond, the screen showed a message: "too slow." Data for such trials were discarded. This amounted to 3.3% of the responses in Experiment 1. Also, participants were informed if their responses were incorrect with the message: "wrong answer."

Results and Discussion

For the item analyses we performed standard multiple regressions. That is, we collapsed data across subjects and regressed RT on five independent variables: length, category, neighborhood frequency, cluster frequency, and word frequency.

For the subject analyses, a straightforward MANOVA could not be used because we are dealing with continuous variables. The present study requires us to capture within-subject effects using regression analyses. For this, a separate regression equation was estimated for each subject in which RTs for that subject were regressed on the five independent variables in our model. This leads to a slope estimate for each independent variable from each participant. Then we computed the mean slope for each independent variable across subjects and tested whether those mean slopes were significantly different from zero (see Ryan, Sherman, & Judd, 1994 for an example of this type of statistical analysis). Such tests tell whether each of the independent variables has a significant impact on RT, assessing those effects within subjects. Degrees of freedom are based on the number of subjects. These statistical analyses were also used for the accuracy data and were repeated in all of the studies reported in this paper.

Because some of the experimental items did not occur in the Francis and Kučera word count, their zero frequency was a problem for the log transformation. We dealt with this in two ways. One way was to use the log of the word frequency + 1. The second way was to postulate a real frequency between zero and one for these items, which we set at 0.5, yielding a log word frequency of -0.3. In most cases, the two methods did not produce differences in significance. Unless there was a difference, we report only results based on log (word frequency + 1).

The present analyses initially included neighborhood frequency as a variable. Grainger, O'Regan, and Segui (1989) and Grainger (1990) have shown that word frequency effects can sometimes be confounded with frequency of orthographic neighbors (i.e., items sharing all but one letter of the target word). However, in none of the analyses in this paper did this variable show any effects and so it was dropped from the model.

The only variable that was significant by subjects and items was word length [$F1(1,30) =$

TABLE 1

Summary of RT Results for Experiment 1

	Mean slope	*SD*	*F*	*p*
A. Subject analysis				
Category	−1.78	13.61	0.51	.48
Word length	9.0	8.85	31.02	<.001
Log-cluster frequency	−4.7	640.7	0	.97
Log-word Frequency	−8.15	19.72	5.12	.03
	Slope	*SE*	*F*	*p*
B. Item analysis				
Category	−2.97	3.94	0.57	.45
Word length	9.00	2.37	14.42	<.001
Log-cluster frequency	−46.94	141.6	0.11	.74
Log-word frequency	−9.59	7.36	1.7	.2

31.02, $p < .0001$; $F2(1,65) = 14.42$, $p < .001$; Min $F'(1,97) = 9.84$, $p < .005$]. Each additional letter contributed about 9 ms to RT. The effect of word frequency was significant only by subjects [$F1(1,30) = 5.12$, $p < .05$; $F2(1,65) = 1.7$, $p = .19$]. Slope coefficients are given in Table 1, and Fig. 1 is a scatter plot of RT against log-word frequency. Accuracy was at ceiling (98%) and so there were no significant effects.

The present results provide weak evidence for the existence of whole-word representations for regularly inflected forms within the frequency range tested (0 to 101). The interpretation of the present results will become clearer when items in a different frequency range are considered in the next experiment.

EXPERIMENT 2

This experiment sought to examine whole-word frequency effects for a different stem–cluster frequency value of around 50 rather than 100 for the first experiment.

Participants

Thirty undergraduate students from the same population as in Experiment 1 participated.

Materials

Sixty-eight regularly inflected words were selected whose stem–cluster frequencies ranged from 49 to 60 (log 1.70 to 1.78). Word frequency varied from 0 to 44 (see Appendix). All other aspects of the materials were identical to those in Experiment 1. Filler items included 68 simple words, 68 inflected nonwords, and 68 simple nonwords.

Procedure

The procedure was the same as in Experiment 1.

Results and Discussion

The results are displayed in Table 2. As in Experiment 1, there was a significant effect of word length on RT [$F1(1,29) = 55.6$, $p < .0001$; $F2(1,67) = 14.1$, $p < .0001$; Min $F'(1,96) = 11.23$, $p < .005$]. Each additional letter contributed 10 ms to RT. Unlike Experiment 1, log-word frequency was highly significant for both subject and item analyses [$F1(1,29) = 33.5$, $p < .0001$; $F2(1,67) = 10.9$, $p < .005$; Min $F'(1,96) = 8.24$, $p < .01$]. Figure 2 shows a scatter plot of word frequency against RT. One unit increase in log-word frequency corresponded to a 32-ms decrease in RT. Accuracy was again at ceiling (97%) yielding no significant effects.

A major assumption of dual access models is that for low-frequency inflected forms, the compositional route should win out over the whole-word access route. As a consequence, whole-

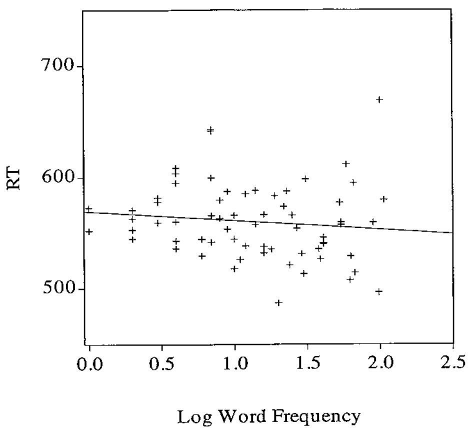

FIG. 1. Reaction times for regularly inflected words in Experiment 1.

TABLE 2

Summary of RT Results for Experiment 2

	Mean slope	*SD*	*F*	*p*
A. Subject analysis				
Category	2.8	13.4	1.27	.27
Word length	9.92	7.17	55.55	<.001
Log-cluster frequency	42.46	398.3	0.60	.44
Log-word frequency	−32.17	29.94	33.48	<.001
	Slope	*SE*	*F*	*p*
B. Item analysis				
Category	2.72	4.53	0.36	.55
Word length	9.86	2.63	14.08	<.001
Log-cluster frequency	45.75	145.9	0.10	.75
Log-word frequency	−32.14	9.72	10.93	.002

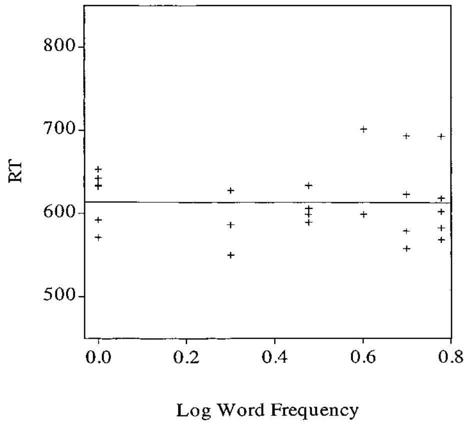

FIG. 3. Reaction times for whole-word frequency values in the 0 to 6 range for Experiment 2.

word frequency effects are not predicted to be found within the lower end of the frequency distribution. To test this prediction, we did a median split on the frequency values, which divided them at the frequency value of 6 per million. Items with word frequencies below and above the median were separately reanalyzed. There were 30 items in the 0 to 6 range. For these items, length was highly significant [$F1(1,29) = 32.19$, $p < .0001$; $F2(1,29) = 8.0$, $p < .01$; Min $F'(1,44) = 6.41$, $p < .05$], but log-word frequency was no longer significant [$F1(1,29) = .28$, $p = .6$; $F2(1,29) = .07$, $p = .8$]—see Fig. 3. Again, we found no effects on accuracy, which remained high at 97%.

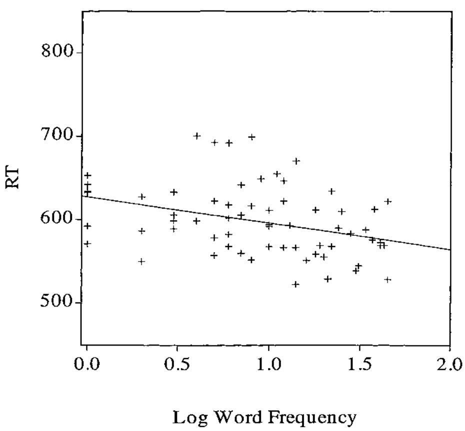

FIG. 2. Reaction times for regularly inflected words in Experiment 2.

There were 38 items with frequencies above the median split in the 7 to 50 range. This represents a log frequency spread of 0.8 (0.9 to 1.7), which was slightly lower than the 0.84 spread in the 0 to 6 range. Log-word frequency was highly significant [$F1(1,29) = 16.7$, $p < .0005$; $F2(1,37) = 6$, $p < .005$; Min $F'(1,60) = 4.41$, $p < .05$]. The estimated slope coefficient for log-word frequency was −44. In other words, a unit increase in log-word frequency corresponded to a decrease of 44 ms in RT.

This reanalysis of the data supports the prediction of dual access models that inflected forms in the lower frequency range should not show whole-word frequency effects. The lack of frequency effects is unlikely to be the result of a decrease in the number of items tested. There was a similar decrease in the number of items with frequencies above 6 per million, yet frequency effects remained in that reanalysis. Another possibility is that the lack of frequency effects in the 0 to 6 range was due to range compression. We will address that issue in the next experiment.

Reanalysis of Experiment 1 Data

How do the present data compare to those for Experiment 1 where we found only weak effects of word frequency? Items in Experiment 1 covered a broader frequency range (0 to 101) than Experiment 2 (0 to 44), so there could have been a floor effect above a certain frequency range for items in Experiment 1. Within the higher frequency ranges, it generally takes a larger increase in frequency to make a difference in RT and the distribution of inflected forms at the higher frequency ranges is relatively sparse.

We decided to reexamine the data from Experiment 1 by looking at the frequency range comparable to Experiment 2. There were 54 items in the 0 to 50 range in Experiment 1. Log frequency was significant by subjects [$F1(1,30) = 15.74$, $p < .0005$] and marginally significant by items [$F2(1,53) = 3.35$, $p = .07$]. Item analyses showed significance when the alternative log word frequency measure was used ($p < .05$). The estimated slope coefficient for log word frequency was -15 for both subjects and items. This suggests that the data from the two experiments are comparable within the overlapping frequency range of 0 to 50. Once again, accuracy remained close to ceiling at 98%.

In Experiment 2, the word frequency effect on RT disappeared in the lower end of the frequency distribution. For a comparable analysis, we examined the 0 to 6 frequency range for Experiment 1. There were 23 inflected items in this range. Log word frequency was not only nonsignificant, but was in opposite directions for subject and item analyses [$F1(1,30) = .44$, $p = .51$; $F2(1,22) = 1.81$, $p = .19$; slope coefficient $= -11.5$ by subjects but 34.2 by items]. Accuracy remained at 98%.

Overall, the results of Experiments 1 and 2 seem to support the conclusion that there is a reliable word frequency effect for regularly inflected words, which indicates the availability of whole-word representations. However, when we performed a median split of the items based on their frequency, the word frequency effect appeared to be restricted to words with a frequency of 7 or more. In the lower end of the frequency distribution, we found no indication of whole-word representations.

There were three problems with these experiments. First, the 0 to 6 range was established when the data had already been collected and examined. It is important to replicate the lack of a frequency effect on RT within this range in a prospective study, with new subjects and additional items. Second, it is possible that frequency counts in this range are simply not accurate. When frequency values are very small, it may be the case that the difference between one value and another might be more related to sampling and chance than to real underlying differences in frequency. Finally, the smaller range of frequency values might make it harder to get a significant correlation with RT. Therefore, we needed to investigate whether word frequency effects could be observed within the lower frequency ranges when uninflected items are considered.

EXPERIMENT 3

Participants

Thirty undergraduate students from the same population as Experiments 1 and 2 participated.

Materials

Fifty regularly inflected words were used whose cluster frequencies were in the 49 to 60 range (log 1.70 to 1.78). Word frequency varied only from 0 to 6. Fifty monomorphemic adjectives were used as control items. Adjectives do not take inflectional endings and thus there is no distinction between word and cluster frequency. Like the inflected items, the adjectives also varied in word frequency from 0 to 6. Filler items included 30 uninflected nouns and verbs, 50 inflected nonwords, and 80 uninflected nonwords.

Procedure

The procedure was the same as in the previous experiments.

Results and discussion

The results from Experiment 3 are summarized in Table 3. For inflected forms, word length was again significant [$F1(1,29) = 52.52$, $p < .0001$; $F2(1,49) = 16.63$, $p < .0005$; Min $F'(1,74) = 12.63$, $p < .001$]. There was also an effect of category whereby RTs were faster for

TABLE 3

Summary of RT Results for Experiment 3

	Inflections				Adjectives			
	Mean slope	*SD*	*F*	*p*	Mean slope	*SD*	*F*	*p*
A. Subject analysis								
Category	−25.89	36.38	14.7	<.001	—	—	—	—
Word length	13.25	9.84	52.5	<.001	20.77	18.64	36.0	<.001
Log-cluster freq.	−104	568	0.97	.33	—	—	—	—
Log-word freq.	5.61	57.8	0.27	.60	−57.48	68.89	20.2	<.001

	Inflections				Adjectives			
	Slope	*SE*	*F*	*p*	Slope	*SE*	*F*	*p*
B. Item analysis								
Category	−27.07	13.28	4.16	.04	—	—	—	—
Word length	12.95	3.18	16.63	<.001	23.5	6.17	14.49	<.001
Log-cluster freq.	−94.01	195	.23	.63	—	—	—	—
Log-word freq.	8.15	22.99	.13	.72	−79.58	29.07	7.5	.009

verbs than for nouns. This effect was significant by subjects and items and was marginal by Min F' [$F1(1,29) = 14.69$, $p < .001$, $F2(1,49) = 4.16$, $p < .05$; Min $F'(1,73) = 3.24$, $p = .08$]. Since we did not see a category effect in any of the other studies reported here, it is possible that the presence of the adjectives somehow made category differences more salient. However, the reasons for such an effect are unclear.

Log-word frequency for inflected forms was not only nonsignificant but the effect was in the opposite direction than would be predicted, with positive slopes for both subjects and items [$F1(1,29) = 0.27$, $p = .6$; $F2(1,49) = 0.13$, $p = .72$]—see Table 3 and Fig. 4. Therefore the data for inflected forms replicate those of the previous studies.

When the same analyses were performed on the adjectives, word length was again significant [$F1(1,29) = 36.0$, $p < .0001$; $F2(1,49) = 14.49$, $p < .0005$; Min $F'(1,77) = 10.33$, $p < .005$). Unlike the inflected items, the effect of log word frequency was significant for the adjectives in this frequency range [$F1(1,29) = 20.19$, $p < .0005$; $F2(1,49) = 7.5$, $p < .01$; Min $F'(1,76) = 5.47$, $p < .05$]—see Fig. 5. To directly compare adjectives with inflected items, we performed a paired t test by subjects comparing the slope coefficients for log-word frequency for inflected forms versus adjectives. The difference was significant [mean slope for inflected forms = 5.61, for adjectives = −57.48, $t(29) = 3.57$, $p < .005$].

The results were similar for the accuracy data. For the inflected forms, only length had a

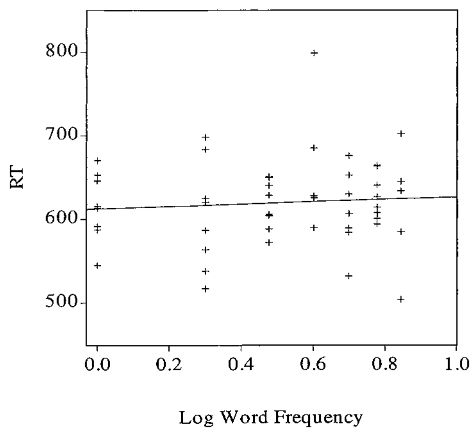

FIG. 4. Reaction times for regularly inflected items in Experiment 3.

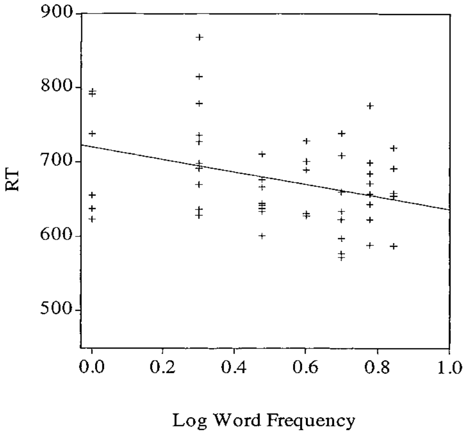

FIG. 5. Reaction times for uninflected adjectives in Experiment 3.

significant effect [$F1(1,29) = 8.54$, $p < .01$; $F2(1,49) = 6.46$, $p < .05$; Min $F'(1,79) = 3.68$, $p = .052$]. Word frequency was not significant [$F1(1,29) = 1.15$, $p = .29$; $F2(1,49) = .61$, $p = .44$]. Accuracy remained high at 96%. For the adjectives, word frequency was significant [$F1(1,29) = 71$, $p < .0001$; $F2(1,49) = 6.38$, $p < .05$; Min $F'(1,59) = 5.85$, $p < .05$], while length was significant by subjects but not by items [$F1(1,29) = 20.55$, $p < .0001$; $F2(1,49) = 1.84$, $p = .18$]. Average accuracy for the adjectives was 79%. The differences in accuracy rates for adjectives and inflected forms probably reflects the fact that inflected forms could be accessed compositionally through the higher frequency base form, whereas the adjectives could not.

A paired t test by subjects comparing the slope coefficients for log word frequency for inflected forms versus adjectives was significant for the accuracy data as well [Mean slope for inflected forms $= -0.015$, for adjectives $= 0.285$, $t(29) = -9.05$, $p < .0001$].

In summary, we replicated the lack of a word frequency effect for inflected forms in the 0 to 6 frequency range. However, a word frequency effect was found in this range for uninflected adjectives. This result suggests that the lack of frequency effects for inflected forms reflects a real unavailability of whole-word representations rather than sampling limitations due to unrepresentativeness of items within a small, low frequency range.

It is possible that the frequency effects for adjectives might be affected by the clusters formed by derivational forms. Although adjectives are not inflected, some can take derivational affixes such as *-er, -est, -ly, -en, -ness,* and *-ity.* We calculated derivational cluster frequencies based on this family of derivational forms for the adjectives. In addition, we calculated similar derivational cluster frequencies for the inflected forms. It is possible that, if derivational cluster frequency is correlated with item frequency, then either the derivational cluster frequency or a combination of the item and cluster frequency might be pushing the adjectives to exhibit a frequency effect. However, the log-transformed derivational cluster frequency was only moderately correlated with log word frequency ($r = 0.14$ for the inflected items; $r = 0.15$ for the adjectives). Effects of log-derivational cluster frequency for RT were not significant for inflected items [$F1(1,29) = .30$, $p = .59$; $F2(1,49) = .06$, $p = .81$], although they were significant for adjectives [$F1(1,29) = 27.86$, $p < .0001$; $F2(1,49) = 6.91$, $p < .05$; Min $F'(1,44) = 5.54$, $p < .05$]. This suggests that the family cluster formed by adding derivational morphology can support access for adjectives.

We reran all of the analyses with derivational cluster frequency included as a variable. The results were not essentially affected by this variable. In particular, the word frequency effect for adjectives remained [$F1(1,29) = 17.66$, $p < .0005$; $F2(1,49) = 6.06$, $p < .05$; Min $F'(1,48) = 4.51$, $p < .05$], and there was still no effect for inflected forms.

EXPERIMENTS 4 AND 5

The previous experiment provided evidence that frequency effects occur in the lower frequency ranges for uninflected forms, but not for regularly inflected forms. In the next two studies, we sought to replicate and extend this result by examining inflected forms and adjectives separately. This would allow us to negate any possible contamination between the two types of items. In addition, by using a larger number

TABLE 4

Summary of RT Results for Experiments 4 and 5

	Inflections				Adjectives			
	Mean slope	*SD*	*F*	*p*	Mean slope	*SD*	*F*	*p*
A. Subject analysis								
Category	−.53	18.06	.02	.88	—	—	—	—
Word length	13.16	13.02	26.56	<.001	12.84	9.37	46.9	<.001
Log-cluster freq.	−237	314	14.8	.006	—	—	—	—
Log-word freq.	−23.27	29.58	16.1	<.001	−91.61	48.52	89.1	<.001

	Inflections				Adjectives			
	Slope	*SE*	*F*	*p*	Slope	*SE*	*F*	*p*
B. Item analysis								
Category	−.02	4.78	.00	.99	—	—	—	—
Word length	13.12	3.04	18.63	<.001	13.48	3.21	17.64	<.001
Log-cluster freq.	−259	117	4.84	.03	—	—	—	—
Log-word freq.	−21.45	12.92	2.76	.10	−96.25	13.97	13.97	<.001

of items for each type, we could increase the power of the tests.

We were also concerned that the choice of the 0 to 6 frequency range was relatively arbitrary based on the median split of the frequency range for Experiment 2. We were interested to see if the same pattern of results occurred for a wider low-frequency range of 0 to 24. Although this is also arbitrary, it does allow for further exploration. It is possible to increase the potential number of inflected forms if the cluster frequency is kept relatively low. This is because most low-frequency items belong to low-frequency clusters. We therefore chose frequency cluster values of 25 to 31 for the next experiment. This also allowed us to examine the effects of word frequency for a new cluster value.

Participants

Undergraduate students from the same population as previous experiments participated. Twenty-seven subjects were tested in Experiment 4 and 26 subjects were tested in Experiment 5.

Materials

For Experiment 4, 94 regularly inflected words were selected with cluster frequencies in the 25 to 31 range (log 1.40 to 1.49). Word frequency varied from 0 to 24. Filler items included 94 uninflected nouns and verbs, 94 inflected nonwords, and 94 uninflected nonwords.

In Experiment 5, 94 monomorphemic adjectives were selected with word frequencies in the 0 to 24 range. Filler items included 94 uninflected nouns and verbs and 188 simple nonwords.

Procedure

The procedure was the same as in the previous experiments.

Results and Discussion

The results of Experiments 4 and 5 are displayed in Table 4. For the inflected items in Experiment 4, word length had a significant effect on RT [$F1(1,26) = 26.56$, $p < .0001$; $F2(1,93) = 18.63$, $p < .0001$; Min $F'(1,100) = 10.95$, $p < .005$]. Unlike the other experiments, there was a cluster frequency effect [$F1(1,26) = 14.8$, $p < .001$; $F2(1,93) = 4.84$, $p < .05$; Min $F'(1,121) = 3.65$, $p = .06$]. Although we tried to keep the range of cluster frequency to a minimum, there was enough of a spread in the present experiment to create an effect. The effect for word frequency was significant by subjects [$F1(1,26) = 16.1$, $p < .0005$], but not by items [$F2(1,93) = 2.76$, $p = .10$]—see Fig. 6. The frequency effect suggested by these data indicate that the 0 to 24 frequency range might exceed the cut-off point for whole-word representations of inflected forms.

When the same analyses were performed on the adjectives in Experiment 5, both length and word frequency were significant [length: $F1(1,25) = 46.93$, $p < .0001$; $F2(1,93) = 17.64$, $p < .0001$; Min $F'(1,118) = 12.82$, $p < .001$; word frequency: $F1(1,25) = 89.11$, $p < .0001$; $F2(1,93) = 47.45$, $p < .0001$; Min $F'(1,109) = 30.96$, $p < .0001$]—see Figure 7.

To directly compare the adjectives with the inflected items, we performed a t test by subjects comparing the slope coefficients for log word frequency for inflected forms versus adjectives. The difference was significant [mean slope for inflected forms = -23.27, for adjectives = -91.61, $t(51) = -6.22$, $p < .0001$].

For the accuracy measures, there was a significant effect of word frequency for the adjectives in Experiment 5 [$F1(1,25) = 84.58$, $p < .0001$; $F2(1,93) = 44.79$, $p < .0001$; Min

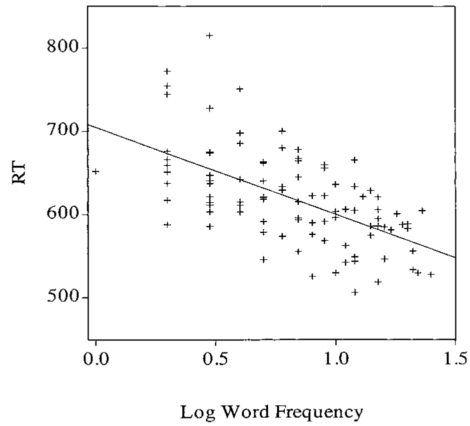

FIG. 7. Reaction times for uninflected adjectives in Experiment 5.

$F'(1,109) = 29.28$, $p < .0001$], but not for the regularly inflected words in Experiment 4 [$F1(1,26) = 2.7$, $p = .11$; $F2(1,93) = 2.41$, $p = .12$]. Average accuracy for the inflected items was 95% while for the adjectives it was 90%.

To summarize, by increasing the frequency range of the inflected forms to 24, we appear to have gone beyond the threshold for whole-word representations. To be sure that the present data are compatible with the previous findings, we looked at items in the 0 to 6 word frequency range in Experiments 4 and 5. There were 51 inflected items in Experiment 4 and 51 adjectives in Experiment 5 within this range. For the inflected items, length was the only variable with a significant effect on RT [for length: $F1(1,26) = 19.12$, $p < .0001$; $F2(1,50) = 13.27$, $p < .0001$; Min $F'(1,77) = 7.83$, $p < .01$; for word frequency: $F1(1,26) = .32$, $p = .58$; $F2(1,50) = .34$, $p = .56$).

Analyses on accuracy for inflected items showed a significant effect of word frequency by subjects and a marginal effect by items [$F1(1,26) = 4.47$, $p < .05$; $F2(1,50) = 3.14$, $p = .08$]. These results are unexpected since they seem to contradict the data from reaction times and the findings from previous experiments where accuracy showed no significant effects for inflected forms. Furthermore, we found no significant frequency effect for accuracy when we looked at the 0 to 24 range from the same experiment. If there were a true un-

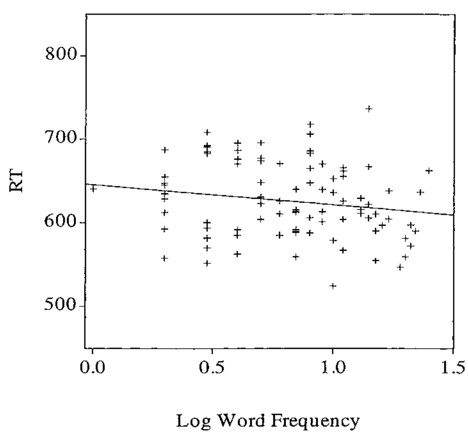

FIG. 6. Reaction times for regularly inflected items in Experiment 4.

derlying frequency effect, then one would expect it to be greater when the frequency range is extended. This was the case with RTs for all forms and for accuracy with adjectives. In addition, average accuracy remained at 95% in the 0 to 6 range. This trend therefore makes little sense in the context of the other findings. In other ways, the present results are generally consistent with the results of previous experiments. Taken together, they suggest that whole-word representations are not available in the 0 to 6 frequency range, but do begin to be available at frequencies above that range.

For the adjectives, length had a significant effect on RT within the 0 to 6 range [$F1(1,25) = 24.83, p < .0001$; $F2(1,50) = 11.4, p < .005$; Min $F'(1.77) = 7.81, p < .01$]. Word frequency was significant by subjects and marginal by items when log (word frequency + 1) was used [$F1(1,25) = 4.35, p < .05$; $F2(1,50) = 3.69, p = .06$]. This effect did reach significance when the alternative log method was used [$F1(1,25) = 7.85, p < .01$; $F2(1,50) = 8.58, p < .005$; Min $F'(1,67) = 4.1, p < .05$]. For the accuracy data, word frequency was also significant [$F1(1,25) = 33.88, p < .0001$; $F2(1,50) = 6.93, p < .05$; Min $F'(1,68) = 5.75, p < .05$]. Average accuracy went down to 80% when only adjectives in the 0 to 6 range were considered.

As in Experiment 3, it was important to ensure that word frequency effects for the adjectives were not an artifact of derivational cluster frequency effects. Once again, we reran all of the analyses with derivational cluster frequency included as a variable. This variable was not significant for either the inflections or the adjectives and word frequency results remained unaffected.

CLUSTER FREQUENCY EFFECTS

Our approach in designing the present set of experiments has been to hold cluster frequency constant and to look at effects due to variations in word frequency. However, because we examined different cluster frequencies across experiments, we can ask whether the base form is activated when a regularly inflected form is accessed.

Taft (1979) and Burani et al. (1984) both carried out lexical decision experiments in which pairs of regularly inflected words were matched for word frequency but differed considerably in cluster frequency. RTs were significantly higher for low cluster frequency items than for high cluster frequency items. In order to do a comparable analysis with the present data, we selected a range of word frequencies that was common across experiments and where the cluster frequency differed significantly.

TABLE 5

Cluster Frequency Effects Comparing Experiments 1 and 2

	Cluster frequency	Mean word frequency	RT (MS)
Experiment 1	100	12.61	560
Experiment 2	50	12.63	598

Experiments 1 and 2 included several items with word frequencies in the 0 to 50 range, while the cluster frequency was around 100 in Experiment 1 and 50 in Experiment 2. Mean RTs for the items in the 0 to 50 word frequency range for each experiment are presented in Table 5. A t test revealed that RTs for items in Experiment 1 were significantly shorter than those for Experiment 2 [$t(121) = 5.6, p < .0001$]. Presumably, this difference reflects the activation of the base form, replicating the results of Taft and Burani et al.

As a further test, we can add the results of Experiment 4 by limiting the analysis to items with word frequencies in the 0 to 24 range. Mean RTs for these items in Experiments 1, 2, and 4 are presented in Table 6. These data show that RTs decreased with increases in cluster frequency across experiments, as predicted if base forms are activated when accessing regularly inflected forms. Differences between Experiments 1 and 2 were significant [$t\ (101) = 5.24, p < .0001$], as were those between Experiments 2 and 4 [$t\ (150) = 3.11, p < .005$].

Another interesting test of cluster frequency effects is in the comparison of inflected forms, which are members of clusters, and adjectives, which lack inflectional clusters. If clusters fa-

TABLE 6

Cluster Frequency Effects Comparing Experiments 1, 2, and 4

	Cluster frequency	Mean word frequency	RT
Experiment 1	100	8.29	563
Experiment 2	50	8.02	603
Experiment 4	25	7.3	626

cilitate lexical access, then we would expect RTs to be lower for inflected forms than for adjectives. This was the case in Experiment 3, where the mean RT for the inflected forms was 619 ms compared to 677 ms for the adjectives. However, this trend did not hold in the comparison of Experiments 4 and 5 (mean RT for inflections: 626 ms, for adjectives: 619 ms). This may be due to the fact that in Experiment 3, the cluster frequency was 50 and in Experiment 4 it was 25. In addition, the item frequency range in Experiment 3 was 0 to 6 and in Experiments 4 and 5 it was 0 to 24. Thus, the cluster frequency effect would be weaker in Experiment 4 than in Experiment 3.

TESTING PREDICTIONS FROM CONNECTIONIST MODELS

In the Introduction to this article, we discussed the connectionist accounts of Bybee (1995) and Stemberger (1995) with regard to whole-word frequency effects for regularly inflected words. Both accounts focused on the competition between lexical strength and lexical connections, but differed in the form of this competition. The present results are consistent with one aspect of Bybee's model, which predicted whole-word frequency effects only above a certain frequency threshold. Stemberger (1995) argued that whole-word frequency effects for regular inflections should not be obtained due to the competition from the very strong interlexical connections to other items sharing inflectional endings. This model is not supported by the present data. However, it is likely that Stemberger's model could also accommodate the pattern of results we found by lessening the role of interlexical connections relative to the lexical strength of the inflected items.

Unfortunately, it is difficult to test predictions from these models based on simple whole-word frequency effects. This is because these models depend on postulated relative contributions of lexical strength and lexical connections, which can be altered to fit the data. It is also difficult to differentiate them from dual models since they are designed to account for the same kind of data outcomes. Therefore, we should look for predictions that are specific to the connectionist models. One strong prediction of Bybee's model, noted earlier, is that nouns should show stronger effects of whole-word frequency than verbs, and verbs should show stronger effects of cluster frequency than nouns.

We tested whether there were any interactions between category and whole-word frequency for Experiments 1, 2, 3, and 4. None of the interactions was significant regardless of whether all of the items or only the low-frequency items (0 to 6) were considered ($p > .27$). Similarly, the category by cluster interactions for Experiments 1, 2, and 4 were also not significant ($p > .39$). Therefore, neither of these analyses based on category differences supports the specific predictions derived from Bybee's model.

It is possible that the current data might be explained by a more straightforward competition model in which cluster frequency competes with whole-word frequency (David Plaut, personal communication). In our studies, whole-word frequency varied, which would promote whole-word frequency effects. On the other hand, cluster frequency was kept constant, which would promote similar reaction times across items. On this account, the threshold for finding whole-word frequency would be the point at which the competing strengths of these two factors are about equal. The model also accounts for the strong whole-word frequency effects found for adjectives. Because adjectives lack inflectional clusters there would be no competition to promote a flat frequency effect for low-frequency items.

This model makes the distinct prediction that, as cluster frequency decreases, so should the threshold value for finding a whole-word frequency effect for inflected items. We can approach a test of this prediction by looking at identical word frequency ranges across experiments. Experiments 1, 2, and 4 examined cluster frequencies of 100, 50, and 25, respectively. As cluster frequency decreases, the whole-word frequency effect should become stronger as the threshold shifts downward.

Selecting items in the 0 to 24 range from Experiments 1, 2, and 4, we examined the interaction between whole-word frequency effects and cluster frequency across these experiments. We ran a single test combining data from all three experiments and regressed RT on category, length, log word frequency, and two contrast coded variables that compared cluster 25 versus clusters 50 plus 100 (contrast 1) and cluster 50 versus cluster 100 (contrast 2). In addition, we looked at two interaction terms between log word frequency and our two contrast coded variables. The results replicated main effects for length, log word frequency, and cluster frequency. However, the log word frequency by cluster interactions were not significant [contrast 1 $\times$ log word frequency: $F(1,193) = 0.02$; $p = .89$; contrast 2 $\times$ log word frequency: $F(1,193) = 1.14$; $p = .29$].

We also looked at the data from each experiment noting whether the slope coefficients for whole-word frequency increased as cluster frequency increased across experiments. Contrary to the predictions of the model, there was no linear relation between the cluster frequency and the size of the whole-word frequency effect (Experiment 1: slope $= -10.67$; Experiment 2: slope $= -31.09$, Experiment 4: slope $= -21.44$). In other words, the cluster frequency of 50 showed the strongest whole-word frequency effects, followed by the cluster of 25, then 100. Bearing in mind that the present experiments were not directly set up to test this model, at present we cannot endorse the model based on the analysis of our results.

GENERAL DISCUSSION

Our results show that whole-word frequency effects can be reliably obtained for regularly inflected forms, although they disappear at a word frequency threshold of about 6 per million. This indicates that whole-word representations are available for regularly inflected forms with frequencies above that value, otherwise they require compositional representations based on morphological structure. Frequency effects did not disappear for simple words (adjectives) below any frequency threshold. This presumably reflects the fact that simple words must have whole-word representations throughout the frequency range.

We began this paper with a discussion of two kinds of dual models. In one case, the dual model proposed that morphologically complex words could be accessed either through a whole-word representation or through a decompositional route. The two access routes were said to be in competition with one another, and word frequency determined which access route would win out. For high-frequency inflected forms, whole-word representations are strong and therefore provide the primary access route. As frequency decreases, the compositional route becomes more reliable and whole-word access no longer occurs. The threshold for frequency effects found in the present studies clearly supports this dual model.

The second kind of dual model that we considered was that proposed by Pinker and his colleagues (Pinker, 1991). This model proposed that regular and irregular inflections are dissociated at many levels. Regular inflections were said to show evidence of rule-based processes whereas irregular inflections show evidence of being stored in associative memory. We noted that there was a strong and a weak form of Pinker's model. The strong version claimed that regularly inflected forms are never represented as whole words; one should find no frequency effects in any range. The present results do not support this position since we did find frequency effects above the threshold of 6 per million. The weaker version of Pinker's model allows that there could be some whole-word representations for regularly inflected forms, but that such representations are not necessary. This version appears to provide a better fit to the present data.

However, endorsing the weak version of Pinker's dual model raises important questions with respect to the concept of storage. If storage leads to connectionistlike associations between items, then such items should show gang effects (Stemberger and MacWhinney, 1988). Prasada and Pinker (1993) found that, unlike irregular inflections, regularly inflected forms show no gang effects. Since the number of regularly inflected forms above threshold for frequency effects is considerable, it is paradoxical that they do not show some evidence of associative processes in the form of gang effects.

There are two possible solutions to this apparent paradox. First, it is likely that the two tests address different issues. On the one hand, we have a question about the nature of our representations: do we keep whole-word representations in memory for complex words? Manipulations of frequency are designed to address this question. A very different issue is how morphological patterns generalize. Prasada and Pinker's results suggest that regular inflections generalize on the basis of abstract rules, while irregular inflections generalize on the basis of lexical associations (i.e., gang effects). Generalizations based on gang effects require whole-word representations; they exploit similarities at the level of the entire word. Generalizations based on rules are independent of the existence of whole-word representations, but by no means exclude them. In fact, from a processing perspective, it may be advantageous or even necessary to store whole-word representations for regularly inflected forms when they occur frequently in the language. As Sandra (1994) pointed out, on-line composition of morphologically complex words reduces the need for storage space but at the cost of greater processing demands. While whole-word representations may not be necessary for regularly inflected items, the limited nature of our on-line computations paired with our enormous resources for language memory may make them convenient.

If all of the above is true, then this would seem to imply that there is lexical storage for regular and irregular inflections but that associative mechanisms arise only in the irregular vocabulary. Why should associative memory act only on irregular morphology? One possibility is that gang effects and whole-word frequency effects tap into different levels of processing. In other words, these two tests for storage may address different kinds of representations, and only one of those exhibits associative properties.

Dual models for lexical access propose that whole-word representations exist in the access system (Anshen & Aronoff, 1988; Baayen, 1991; Burani & Caramazza, 1987; Burani & Laudanna, 1992; Caramazza et al., 1985; 1988; Chialant & Caramazza, 1995; Frauenfelder & Schreuder, 1991; Laudanna et al., 1989, 1992; Schreuder & Baayen, 1995; Stemberger & MacWhinney, 1988). Marslen-Wilson, Tyler, Waksler, and Older (1994) have proposed that access representations are defined by the modality of the stimulus input (e.g., visual, auditory). It is here that frequency effects are assumed to arise. The modality-specific access representations are said to feed into a more abstract, modality-neutral, internal lexicon. It is possible that the lexicon, rather than the access system, is the locus of associative mechanisms giving rise to such things as gang effects.

With regard to connectionist models, we noted that a threshold for whole-word frequency effects could be predicted by certain kinds of connectionist models. However, when more specific predictions were tested, the connectionist models did not fare so well. On the other hand, it was only the connectionist models that were specific enough to allow such predictions, and one should not therefore take such failures as an endorsement of alternative models that make no such specific predictions. In addition, the kinds of analyses we carried out to test these models were not within the main design of the study and so further evaluation is warranted.

To summarize, we have shown that there are indeed whole-word representations available for regularly inflected items. At the same time, we have demonstrated that this is not true for items in the lower end of the frequency distribution. In other words, there is a critical role for combinational principles in the system, given that some items can only be accessed through morphological decomposition. The results highlight the need to distinguish between frequency effects and gang effects and the different conceptual issues associated with these two phenomena. The distinction between access representation and central representation is particularly useful in this regard.

APPENDIX

Experimental Items

Experiment 1

Word	W Freq	Cluster Freq	Word	W Freq	Cluster Freq
destroys	0	104	squares	13	121
sevens	0	114	supplying	13	103
suns	1	117	cared	15	108
behaviors	1	100	oils	15	111
finishes	1	120	selecting	15	112
captains	1	99	balls	17	122
incomes	1	110	refers	18	104
eats	2	122	parks	19	111
listens	2	122	traditions	21	115
settles	2	105	successes	22	116
naming	3	109	reflects	23	107
announces	3	116	tended	24	104
faiths	3	110	sessions	26	106
discovers	3	122	responses	28	105
justices	3	104	trips	29	109
regarding	3	100	instances	30	112
suffers	5	110	ladies	37	122
saves	5	121	smiling	38	122
compares	6	114	pages	40	102
councils	6	115	regions	40	119
denies	6	109	shared	40	105
saints	6	103	shoulders	52	112
proposing	6	110	motors	53	108
forgetting	7	119	pushed	53	102
gases	7	107	wondered	58	119
observes	8	120	weapons	61	103
mouths	8	113	arrived	62	108
fixing	9	109	yards	65	100
dinners	9	100	fingers	66	106
depended	9	106	published	89	108
seas	10	104	parents	97	113
improves	11	121	facilities	100	111
treating	11	122	interested	101	106

Experiment 2

Word	W Freq	Cluster Freq	Word	W Freq	Cluster Freq
constructs	0	56	advertised	9	49
convinces	0	57	vacations	9	54
televisions	0	51	strangers	9	50
tuesdays	0	59	retains	9	49
installs	0	49	confronting	10	55
mixes	0	56	afforded	11	58
appointing	1	50	rejects	11	58
cleans	1	58	constitutes	11	54
jumps	1	58	restaurants	12	53

Word	W Freq	Cluster Freq	Word	W Freq	Cluster Freq
adjusting	2	60	enabling	13	57
bays	2	60	printing	13	53
operas	2	49	chickens	13	49
pouring	2	49	sisters	15	55
solves	2	49	pockets	17	59
bibles	2	60	implied	17	53
majorities	3	60	favored	18	49
uncles	3	58	victims	19	50
absences	3	56	secrets	20	52
interiors	4	49	shadows	20	54
promotes	4	60	bullets	21	49
chests	4	57	financing	21	55
chemicals	4	57	substances	23	56
accounted	5	49	prospects	24	49
tendencies	5	54	critics	27	53
ignoring	5	57	stars	29	58
prefers	5	60	dancing	30	59
tying	5	50	cooling	33	59
dramas	6	49	advertising	36	49
republics	6	49	authorized	37	49
roots	6	53	approved	40	56
implying	7	53	composed	40	50
responds	7	54	neighbors	42	59
tragedies	7	56	shoes	44	58
liberties	8	56	acres	44	54

Experiment 3: Inflections

Word	W Freq	Cluster Freq	Word	W Freq	Cluster Freq
televisions	0	51	appointing	1	50
constructs	0	56	dictionaries	1	59
convinces	0	57	leans	1	60
installs	0	49	pouring	1	49
isolates	0	49	advertises	1	49
mixes	0	56	jumps	1	58
surrounds	0	53	adjusting	2	60
tuesdays	0	59	operas	2	49
preferring	1	60	retires	2	54
cleans	1	58	solves	2	49
bays	2	60	chests	4	57
bibles	2	60	concludes	4	60
composing	2	50	ignores	5	57
commits	2	50	affords	5	58
majorities	3	60	reminding	5	57
uncles	3	58	tendencies	5	54
absences	3	56	confronts	5	55
constituting	3	54	tying	5	50
creations	3	50	doctrines	5	57
interiors	4	49	authorizing	5	49

Word	W Freq	Cluster Freq	Word	W Freq	Cluster Freq
promotes	4	60	possessing	6	54
rejecting	4	58	republics	6	49
diameters	4	53	responding	6	54
defends	4	56	roots	6	53
forts	4	55	dramas	6	49

Word	W Freq	Cluster Freq	Word	W Freq	Cluster Freq
neglects	1	28	denotes	7	25
scaring	1	26	persists	7	25
bellies	2	25	protesting	7	29
charms	2	25	borrowing	7	31
journeys	2	31	statues	8	25
dilemmas	2	27	crystals	8	31
gears	2	28	monuments	8	29
shirts	2	29	clarified	8	25
parades	2	26	bears	9	25
regimes	2	25	charts	9	30
compels	2	25	rails	9	25
arresting	2	27	borders	10	30
reverses	2	27	clues	10	25
exciting	2	28	champions	10	31
academies	3	28	departing	10	28
bowls	3	26	enduring	10	31
sheriffs	3	28	assessing	10	25
crafts	3	25	blades	12	26
cooperated	3	28	debts	12	25
exerts	3	29	pretending	12	27
arousing	3	30	bubbles	13	25
aunts	4	27	deputies	13	27
circuits	4	27	drifting	13	27
pistols	4	31	breathing	13	31
rituals	4	27	flames	14	27
disasters	4	30	organs	14	26
empires	4	26	minerals	14	26
conveys	4	27	potatoes	15	30
inducing	4	29	vegetables	16	26
squeezing	4	30	devised	16	25
baths	5	31	tanks	18	30
cafes	5	25	toes	19	26
garages	5	25	documents	19	30
praying	5	30	boots	20	30
linking	5	25	residents	20	28
storms	6	31	educated	21	31
cottages	6	25	dedicated	22	25
cycles	6	30	norms	24	31

Experiment 3: Adjectives

Word	WFreq	Word	WFreq
macabre	0	candid	3
nimble	0	mundane	3
supple	0	lavish	3
somber	0	transient	3
brusque	0	diagonal	4
abstruse	0	docile	4
avid	1	innate	4
insipid	1	inverse	4
affable	1	numb	4
convex	1	salient	4
minuscule	1	languid	4
obstinate	1	plump	4
opulent	1	anterior	5
rotund	1	fertile	5
treble	1	mediocre	5
succinct	1	fluent	5
agile	2	lethal	5
arrogant	2	aberrant	5
covert	2	obsolete	5
deft	2	sparse	5
obscene	2	cordial	6
profuse	2	clumsy	6
tacit	2	meager	6
insolent	2	opaque	6
exquisite	3	placid	6

Experiment 4

Word	W Freq	Cluster Freq	Word	W Freq	Cluster Freq
coped	0	30	fortunes	6	29
cancers	1	25	gangs	6	27
funerals	1	31	veins	6	31
spheres	1	26	pretended	6	27
tributes	1	25	yelling	6	31
cellars	1	25	vectors	7	26
climates	1	27	motels	7	31
perceiving	1	29	cabins	7	30
explores	1	29	negotiated	7	25

Experiment 5

Word	W Freq	Word	W Freq
nimble	0	futile	6
astute	1	rude	6
brash	1	meager	6
concise	1	opaque	6
oblique	1	placid	6
clandestine	1	blunt	7
supine	1	bizarre	7
convex	1	brisk	7
rotund	1	polite	7

Word	W Freq	Word	W Freq
avid	1	dismal	8
affable	1	pregnant	8
affluent	2	sane	8
arid	2	serene	8
deluxe	2	fierce	8
morose	2	latent	9
prolific	2	potent	9
terse	2	compact	9
deft	2	coarse	10
tacit	2	tense	10
obscene	2	weird	10
profuse	2	adverse	11
agile	2	eloquent	11
covert	2	prone	11
bland	3	strict	11
erudite	3	tender	11
lenient	3	shy	11
sordid	3	solemn	12
candid	3	diverse	13
mundane	3	acute	13
lavish	3	transparent	13
adept	4	dumb	14
tame	4	elegant	14
turbulent	4	intact	14
lucid	4	ripe	14
docile	4	shallow	14
innate	4	sincere	15
inverse	4	loud	15
numb	4	sober	16
plump	4	absurd	17
austere	5	abrupt	18
hoarse	5	naive	19
inert	5	savage	19
timid	5	nude	20
torrid	5	smart	20
cordial	6	bold	21
clumsy	6	moderate	22
desolate	6	stupid	24

REFERENCES

Anshen, F., & Aronoff, M. (1988). Producing morphologically complex words. *Linguistics,* **26,** 641–655.

Baayen, H. (1991). Discussion: On frequency, transparency and productivity. In G. E. Booij & J. van Marle (Eds.), *Yearbook of morphology* (pp. 181–208). Dordrecht: Foris.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes,* **2,** 217–227.

Burani, C., & Laudanna, A. (1992). Units of representation for derived words. In R. Frost & L. Katz (Eds.), *Orthography, phonology, and meaning* (pp. 361–376). Amsterdam: North–Holland.

Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. *Visible Language,* **4,** 348–358.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp. 257–294). London: Academic Press.

Bybee, J. L. (1995). Diachronic and typological properties of morphology and their implications for representation. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 225–246). Hillsdale, NJ: Erlbaum.

Bybee, J. L., & Moder, C. L. (1983). Rules and schemes in the development and use of the English past tense. *Language,* **59,** 251–270.

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

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition,* **28,** 297–332.

Caramazza, A., Miceli, G., Silveri, M. C., & Laudanna, A. (1985). Reading mechanisms and the organization of the lexicon: Evidence from acquired dyslexia. *Cognitive Neuropsychology,* **2,** 81–114.

Chialant, D., & Caramazza, A. (1995). Where is morphology and how is it processed? The case of written word recognition. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 55–78). Hillsdale, NJ: Erlbaum.

Cole, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language,* **28,** 1–13.

Cutler, A. (1983). Lexical complexity and sentence processing. In G. B. Flores d'Arcais and R. J. Jarvella (Eds.), *The process of language understanding* (pp. 43–79). New York: Wiley.

Emmorey, K. D. (1989). Auditory morphological priming in the lexicon. *Language and Cognitive Processes,* **4,** 73–92.

Fowler, C. A., Napps, S. E., & Feldman, L. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition,* **13**(3), 241–255.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston, MA: Houghton, Mifflin.

Frauenfelder, U. H., & Schreuder, R. (1991). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In G. E. Booij & J. van Marle (Eds.), *Yearbook of morphology* (pp. 165–183). Dordrecht: Foris.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language,* **29,** 228–244.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing units in visual word recognition: The neighborhood frequency effect. *Perception and Psychophysics,* **45,** 189–195.

Laudanna, A., Badecker, W., & Caramazza, A. (1989). Priming homographic stems. *Journal of Memory and Language,* **28,** 531–546.

Laudanna, A., Badecker, W., & Caramazza, A. (1992). Processing inflectional and derivational morphology. *Journal of Memory and Language,* **31,** 333–348.

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

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition,* **40,** 121–157.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., Woest, A., & Pinker, S. (1993). *German inflection: The exception that proves the rule.* MIT Center for Cognitive Science, Occasional paper No. 47.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in Language Acquisition. In *Monographs of the Society for Research in Child Development* (Vol. 57, No. 228). Chicago University Press.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review,* **101,** 3–33.

Napps, S. E. (1989). Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory and Cognition,* **17**(6), 729–739.

Napps, S. E., & Fowler, C. A. (1987). Formal relationships among words and the organization of the mental lexicon. *Journal of Psycholinguistic Research,* **16,** 257–272.

Pinker, S. (1991). Rules of language. *Science,* **253,** 530–535.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a Parallel Distributed Processing model of language acquisition. *Cognition,* **28,** 73–193.

Pinker, S., & Prince, A. (1991). Regular and irregular morphology and the psychological status of rules of grammar. *Proceedings of the 1991 Meeting of the Berkeley Linguistics Society.*

Plaut, D. C., McClelland, J. L., & Seidenberg, M. S. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review,* **103,** 56–115.

Plunkett, K., & Marchman, V. (1991). U-shaped Learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition,* **38,** 43–102.

Prasada, S., & Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes,* **8,** 1–56.

Prasada, S., Pinker, S., & Snyder, W. (1990). Some evidence that irregular forms are retrieved from memory but regular forms are rule generated. Poster presented at the Psychronomic Society meeting, November, 1990.

Rubin, G. S., Becker, C. A., & Freeman, R. H. (1979). Morphological structure and its effect on visual word recognition. *Journal of Verbal Learning and Verbal Behavior,* **18,** 757–767.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing* (Vol. 2, pp. 216–271). Cambridge, MA: MIT Press.

Sandra, D. (1994). The morphology of the mental lexicon: Internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes,* **9,** 227–269.

Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM-PC compatibles. *Behavior Research Methods, Instrumentation and Computers,* **20,** 206–217.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–156). Hillsdale, NJ: Erlbaum.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review,* **96,** 523–568.

Segui, J., & Zubizarreta, M. L. (1985). Mental representation of morphologically complex words and lexical access. *Linguistics,* **23,** 759–774.

Stemberger, J. (1995). Phonological and lexical constraints on morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 247–270). Hillsdale, NJ: Erlbaum.

Stemberger, J. P., & MacWhinney, B. (1988). Are inflected forms stored in the lexicon? In M. Hammond & M. Noonan (Eds.), *Theoretical morphology: Approaches in modern linguistics* (pp. 101–116). San Diego, CA: Academic Press.

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

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition,* **7,** 263–272.

Xu, F., & Pinker, S. (1992). Weird past tense forms. Paper presented at the Seventeenth annual Boston University Conference on Language Development, October 23–25.

(Received June 12, 1997)
(Revision received August 10, 1998)