587

JOURNAL OF MEMORY AND LANGUAGE 26, 419–452 (1987)

Lexical Organization and Welsh Consonant Mutations

S. BOYCE, C. P. BROWMAN, AND L. GOLDSTEIN

Haskins Laboratories, New Haven, Connecticut and Yale University

The role of phonological form in the lexical organization of morphologically related words is investigated using consonant mutations in Welsh. Consonant mutations are a regular form of nonaffixal morphology, in which the initial consonant of a word changes depending on its syntactic context; for instance, the word *pont* "bridge" may appear as *bont* or *font*. In English, it has been shown that affixal variants, such as POUR–POURED, prime each other strongly, while nonaffixal variants, such as HUNG–HANG, show weak or nonexistent priming (S. T. Kempley & J. Morton (1982) *British Journal of Psychology*, 73, 441–454; R. F. Stanners, J. J. Neiser, W. P. Hernon, & R. Hall (1979) *Journal of Verbal Learning and Verbal Behavior*, 18, 399–412). Using the task of auditory repetition priming, we show that mutation is similar to affixing in English in that mutated variants prime each other. We further show that abstract morphological categories, rather than identity of phonological form, are required to organize the Welsh lexicon, thus suggesting that current phonologically based lexical models need to be revised. An alternative model utilizing an underspecified autosegmental representation is proposed.

Recent studies (e.g., Kempley & Morton, 1982) have shown that morphologically related words such as POUR and POURED share a common lexical representation. In this paper, we investigate the extent to which such sharing is constrained by the similarity in form between the related words. In particular, we ask whether such sharing is possible when the related words differ in a more complex way than the simple suffixation seen in English. To this end, we make use of an interesting feature of Welsh, the fact that initial consonants in Welsh words undergo systematic changes as a function of their syntactic and lexical context. These consonant changes, or *mutations* as they are known, are the focus of our study.

Words can be defined as those unique intersections (associations) of semantic and phonological material that can stand independently as minimal utterances. Thus, BOIL, GRAPHIC, and DOG are each a different combination of meaning and phonological form. But what of sets such as POUR, POURING, POURED, etc? These also fulfill the definition as different words, yet there is something intuitively unsatisfying in saying that their relationship is analogous to that between, for example, BOIL, GRAPHIC, and DOG. Traditional linguistic analysis (and common sense) address this issue by pointing out that each of these words has a smaller pattern in common, POUR, which is itself a unique intersection of meaning and phonological shape. It is in this sense that POUR, POURED, etc. can be alternately described as the same word in different forms, or, in the terms of the "minimal utterance" definition of a word, as different words with the common semantic/phonological pattern POUR. This common se-

This work was supported by NIH Grant HD-01994. We thank Dr. Frank Gooding, of the University of North Wales, Bangor, for making the facilities of the Linguistics Department available, and Carolyn Iorwerth, of the Department of Extra-Mural Welsh Instruction, for help in finding subjects. We also thank Patrice Beddor, David Fay, Len Katz, and an anonymous reviewer for helpful criticisms. Requests for reprints should be addressed to Catherine P. Browman, Haskins Laboratories, 270 Crown St., New Haven, CT 06511.

0749-596X/87 $3.00

mantic/phonological pattern is referred to as the *root morpheme*, or *base*. In addition, linguistic theory posits *grammatical morphemes* that capture the grammatical function and phonological form that distinguish the morphological variants POUR, POURING, POURED from each other—for example, the forms -ING and -ED. In this paper we will refer to the entire set of morphologically related forms as a *morphological complex*, and words that belong to the same morphological complex as *morphological variants*. For inflectional processes (such as those involving -ING and -ED), we usually think of the morphological variants as forms of the same abstract "word" (or "lexeme," cf. Matthews, 1974), in the sense that we would expect to find only a single entry in a dictionary, corresponding to the whole morphological complex.

It has often been proposed that this special relationship between morphological variants is reflected in the organization of the lexicon; that is, words are generated or accessed according to morphological grouping (e.g., Lukatela, Gligorijević, Kostić, & Turvey, 1980; MacKay, 1979; Murrell & Morton, 1974; Stanners, Neiser, Hernon, & Hall, 1979; van der Molen & Morton, 1979). Assuming that such is the case, there are logically two ways by which these groups or access routes might be organized in the lexicon: according to the phonological form of the likeness between variants or according to their functional (semantic or syntactic) relationship. Putting this somewhat less categorically, we can ask whether there are any constraints on the kind and degree of phonological likeness, or the kind and degree of functional relationship, which in and of themselves exclude certain words from participating in such morphologically based groups or access routes. Evidence for such constraints on the possible groupings, or access routes, would obviously provide useful limitations on possible models of lexical organization.

It happens that in the most frequent and familiar English case (as in POUR, POURING, etc.) functional and formal analyses converge, thus making it difficult, if not impossible, to distinguish the two. In this "regular" English system, it is possible to decompose the phonological form of inflected words into two independent, physically continuous surface forms, the base and the suffix. These forms bear a one-to-one relation to the root meaning and grammatical tense marker. There are, however, more peripheral examples in English where formal and functional analyses diverge, in particular, in the case of "irregular" inflected nouns and verbs, whose formal relationship is often relatively arbitrary (SHAKE–SHOOK, CREEP–CREPT) but whose functional relationship is straightforwardly parallel to that of regular forms (PRESENT–PAST TENSE). In addition, there are cases of derived forms in English where formal relationships are transparent, but where functional relationships are not completely predictable (e.g., FORM–FORMAL; LYRIC–LYRICAL). These irregular examples have been the focus of a good deal of recent research (e.g., Fowler, Napps, & Feldman, 1985; Kempley & Morton, 1982; Stanners et al., 1979), and we will return to a discussion of this work, below.

Models of lexical access have tended to echo the pattern of the English regular paradigm, picturing the lexicon as based on physically continuous, separable, shared units. In one well-known model, the logogen theory of Morton (Kempley & Morton, 1982; Morton, 1969, 1970), the lexicon (or at least the *input* lexicon, cf. Morton, 1979) is a set of receptors or recognition units, each of which fires upon detecting the presence of a particular stimulus in the incoming signal. For a word such as POURED there would be two such receptors, corresponding to the base morpheme POUR and the affix -ED. The base morpheme POUR is considered to be a shared receptor, or logogen, which fires alike for each variant of the morphological complex containing POUR: POURING, POURED, etc. The logogen POUR is only excited by other variants and not by accidentally similar auditory or orthographic sequences. Moreover, because receptors can only fire for discrete, continuous units, morphological variants whose surface forms do not share this same unit cannot be grouped together, and would be represented by separate logogens (Kempley & Morton, 1982, p. 443). Presumably, this would be a problem both for pairs such as SHAKE–SHOOK, whose formal relation is irregular, and DECIDE–DECISION whose relationship is regular but requires the application of vowel and consonant-altering phonological rules. Similar notions of shared entry, and the separation of affix from base, have been proposed by Taft and Forster (1975) and Manelis and Tharp (1977), among others. In all these models, because the lexicon is organized by shared phonologically continuous units, morphological variants that do not share such a unit cannot be grouped together, or share recognition units.

Not all models of the lexicon, of course, are constrained in this way. Even if the lexicon is viewed as listing only base forms, it is possible to generate the variants by the application of stored rules (e.g., MacKay, 1979), where these rules can be either concatenative (e.g., "make a past by adding ed") or form-changing (e.g., "make a past by changing the vowel"). Thus, this type of model allows a dictionary base entry to be shared among morphological variants that do not share a continuous surface sequence corresponding to the base. However, perceptual experiments have not, in general, supported this possibility. As discussed below, various experiments have suggested (e.g., Kempley & Morton, 1982; Stanners et al., 1979) that such sharing does not occur in the case of English irregular forms (e.g., SHAKE–SHOOK). Therefore, lexical models based on these experiments have assumed that morphological variants can share a lexical entry only if the base occurs in unmodified form in each variant.

These shared-entry approaches to lexical access are plausible in English and the more commonly studied European languages, where the primary device for encoding morphological variation is affixation. Because the exceptions to this format —primarily irregular nouns and verbs of the SHAKE–SHOOK variety—are few and idiosyncratic, it has been possible to assume that the lexicon treats them as associated, but not as sharing a lexical entry. There are many languages, however, in which there are regular morphological processes that are more complex than simple concatenation. The variants related by such processes often violate the constraint of shared, continuous base forms. Some languages, such as Tagalog, violate this requirement by inserting grammatical morphemes into the body of a base morpheme, as in /sulat/ "to write," /sumulat/ "wrote," /sinulat/ "was written," from the combination of the base form /sulat/ with infixes /-um-/ and /-in-/. Many languages challenge both continuity and separability by modifying the base morpheme internally. In Semitic languages, for instance, roots are varied by vowel substitution and prosodic adjustment while maintaining a common consonantal core. In Hebrew /sapar/ "barber" and /siper/ "he cuts," the contrast between inserted vowels /a–a/ and /i–e/ differentiates the noun and the verb (cf. Bentin, Bargai, & Katz, 1984). For another example, compare Arabic /daras/–/daaris/–/darraas/–/diraasat/, to their English equivalents "study" (verb), "studying," "student," "study" (noun), where variations in placement and doubling of the vowels /i/ and /a/ and the consonant /r/ differentiate large sets of related words (cf. McCarthy, 1981). As a third example, consider Indo-European, which had a regular process of vowel change from /e/ to /o/ in root forms that made (among other things) past from present verb forms. (German strong verb alternations, and our present-day English irregular forms SIT–SAT, are fossilized remnants of this once regular system.) In each of these cases, the nonaffixal form of morphological variation is known to be, or have been, a regular and productive process in the language.

The existence of such processes (and the complex paradigms found in "inflecting" languages like Latin and Russian) has led some linguists to question the general utility of the base plus affix morpheme model (e.g., Anderson, 1982; Matthews, 1974); it also calls into question the universal applicability of the base plus affix model of lexical access. How, then, will the lexicon be organized in those languages whose regular morphological processes do *not* involve separable, phonologically continuous morphemes? If continuity and separability are essential to lexical organization, related morphemes in these languages will behave like separate words. Conversely, if the related morphemes in these languages behave like regular, affixed forms, then it would seem that lexical structure can be adapted to accommodate the various kinds of regular grammatical alternations found in the world's languages.

The issue of shared lexical representations, and in particular the problems posed by irregular forms, has been primarily investigated using variants of the repetition priming paradigm. This technique relies on the fact that if a word is presented twice to a subject it is recognized more easily, and faster, on the second repetition. This has been demonstrated using the lexical decision task (e.g., Forbach, Stanners, & Hochhaus, 1974), and recognition of words in noise (e.g., Murrell & Morton, 1974). In its application to morphological investigation, two different morphologically related words are presented instead of repeating exactly the same word twice; the extent of facilitative influence from the first presentation (prime) to the second (target) is measured. In the logogen model, facilitation is explained by positing an activation threshold for each logogen; activation then lowers its threshold so that the logogen is more easily activated for some following period of time. Priming between morphologically related forms is explained by the effect of previous activation on the logogen they have in common. As noted above, this shared logogen necessarily corresponds to the base morpheme, that is, the continuous string of phonological elements they hold in common. All shared-entry models echo this concept (e.g., Manelis & Tharp, 1977). In terms of priming, therefore, any word that contains a base morpheme such as POURED or POURING should be as effective a prime for the word POUR as POUR itself would be.

Inflectionally affixed forms such as POURED have been shown to successfully prime a base form POUR. The magnitude of such priming is consistently less than, but statistically equivalent to, that for POUR–POUR. This has been shown for lexical decision tasks (Fowler et al., 1985; Stanners et al., 1979) and for auditory word recognition in noise (Kempley & Morton, 1982). That is, identical variants prime themselves best, and regular, inflectionally related variants prime each other nearly as well. This result for the regular variants has been referred to as "full" priming and has been interpreted as evidence for shared entries between morphologically related forms. However, the situation is different for irregular inflections, such as irregular strong verbs (WOVE–WEAVE) (Kempley & Morton, 1982; Stanners et al., 1979) and suppletive (WORSE–BAD) forms (Kempley & Morton, 1982). Kempley and Morton found no priming between irregularly related inflectional forms, while Stanners et al. found "partial" priming, that is, a significant effect, but one that was significantly less than that for regularly related forms. Both papers interpreted these results to mean that irregular variants do not have enough phonological material in common to share a single lexical representation and are therefore represented as separate entities.

As Kempley and Morton point out, the failure of irregular inflectional variants to prime each other also argues against an alternative interpretation of the repetition priming effects. This alternative attributes the priming to the semantic relations between the prime and the target. Since semantically related (but morphologically unrelated) words can prime each other (e.g., Meyer & Schvaneveldt, 1976, for lexical decision), priming by inflectionally related words might be completely attributed to the semantic overlap between them, rather than to their sharing a common lexical entry or recognition unit. However, the semantic relations are equally tight in the case of irregular inflections, and thus, the failure to obtain priming with them argues against the strictly semantic alternative. In addition, Henderson, Wallis, and Knight (1984) have shown that priming between inflectionally related forms is significantly greater than priming between words related only semantically, when the two types are controlled for semantic closeness. Thus, repetition priming between inflectionally related forms seems to be revealing an independent layer of organization in the lexicon that is morphological in nature.

Results of priming between derivationally related variants presents a somewhat more confused picture. Stanners et al. (1979) looked at priming using derived adjectives (e.g., SELECTIVE–SELECT) and derived nouns (e.g., DECISION–DECIDE). Results showed a significant decrease in priming from derived words to their bases compared to the priming of base to itself. On the basis of this pattern, Stanners et al. suggest a compromise model of the lexicon, in which inflectionally related variants share entries while the reduced priming observed between irregularly and derivationally related variants is modeled as resulting from associations between separate entries. This conclusion is further supported by the fact that the degree of change in (phonological or orthographic) form between the derivation and the base did not seem to influence the degree of priming. This would be predicted if the connection between such items were solely semantic. Thus, Stanners et al. conclude that shared, separable phonological units entail shared entries in the lexicon for regular base and affixed variants, but not for words whose formal or functional relationship is irregular (as derivational relations may be).

The results of Fowler et al. (1985) call into question the strong form of Stanners et al's (1979) conclusion. They replicated the above experiments, but also ran conditions with a longer lag between prime and target, and found that with this longer lag, full priming could be observed from derived forms to their bases. (They attribute this to reduction of "episodic" aspects of priming.) While this result casts some doubt on the ability to decide whether priming in any particular case is the result of shared entries or associative connections, it is still the case that the derived forms behave differently than regular inflected forms. That is, there exist conditions under which priming among derived forms is not full, while regular inflections show full priming under the same conditions. Stanners et al. attribute the difference between derivation and inflection to the fact that derivational processes in English are less *regular* than derivational ones. Note that regularity is a complex concept; even when no change in base form is involved, English derivations may be less regular in two ways: a particular derivational affix may attach only to a subclass of eligible items (e.g., -IVE cannot be added to all verbs, or all transitive verbs), or there may be unpredictable changes in meaning associated with the derivations (as in FORM–FORMAL). Inflectionally related variants, in contrast, must have regularly, that is predictably, related meaning (Bybee, 1985). Stanners et al. argue that shared entries are involved only if two conditions are met: (1) there is a continuous, shared phonological sequence between the forms and (2) the relationship between the forms is regular.

It should be clear by now, however, that in English, the factors regularity and shared form are (partially) confounded. There are few (if any) cases in which a completely regular morphological process involves a change in form. For this reason, we chose to examine a language in which a change in form is associated with a completely regular process.

Welsh

Welsh is an Indo-European language, one of the few members of the Celtic branch to have survived into modern times. Other members of this group are Irish, Scots Gaelic, and Breton. Welsh is spoken by approximately half a million people in the western portion of the British Isles (1981 census figures). The following discussion and examples of Welsh were culled from several sources: primarily Jones (1977), a grammar of colloquial Welsh, Awbery (1973, 1986), and Willis (1986), the latter three references being linguistic analyses of the mutation system in modern Welsh using native speakers. Each of these primarily treats the South Wales dialects. Information on the Northern dialects comes from the above works and Fynes-Clinton (1913).

In common with other Celtic languages, Welsh shows a phenomenon known as "initial mutation," in which the initial consonant of a word changes as a function of its lexical and syntactic context. There are three classes of such changes, traditionally called the SOFT, ASPIRATE, and NASAL mutations; of these, we will be concerned with only two, the SOFT and the ASPIRATE mutations. The SOFT mutation changes initial /p,t,k/ into /b,d,g/ and initial /b,d/ into /v,ð/ (/ð/ is the *th* sound in English *breathe*), while initial /g/ is deleted. (The SOFT mutation also affects other initial consonants, specifically /m/, the voiceless aspirated trill /r̥/, and the voiceless lateral fricative /ɬ/; however, we will be concerned only with /p,t,k,b,d,g/ in this paper.) The ASPIRATE mutation changes the initial consonants /p,t,k/ into /f,θ,x/ (/θ/ is the *th* sound in English *thing* and /x/ is the *ch* sound in German *Bach*); it does not affect initial /b,d,g/. These changes are summarized in Table 1. Thus for example, a word whose dictionary form is /pont/ "bridge" will appear in a SOFT mutation context as /bont/ and in an ASPIRATE mutation context as /font/. A word whose dictionary form is /brawd/ "brother" appears as /vrawd/ in a SOFT mutation context but stays as unchanged /brawd/ in an ASPIRATE mutation context.

It is important to note that, in terms of the phonology of Welsh, all the changes we are concerned with are phonemic, that is, they involve change from one phoneme to another (Awbery, 1986). (The phonemic consonants of Welsh are listed in Table 2.) While in theory this could result in neutralization of the difference between minimal pairs of words, such as /pis/ "peas" and /bis/ "finger," in practice this happens very rarely, partly because many phonemes (e.g., /v/, /θ/, and /ð/) appear word-initially in base forms primarily as the result of borrowing (Awbery, 1973), in which case the different phonological structure of borrowed words makes neutralization unlikely. In addition, the fact that the mutated and citation, or base, forms appear in very different linguistic contexts makes confusion unlikely.

Certain consonants are not subject to mutation. These nonmutating consonants include the consonants /f,θ,v,s,ʃ,r,l,h/ and

TABLE 1
Initial Consonant Changes for Mutating Words in Welsh

Citation	Aspirate mutation	Soft mutation
p	f	b
t	θ	d
k	x	g
b	b	v
d	d	ð
g	g	—

TABLE 2
PHONEMIC CONSONANTS OF WELSH (FROM AWBERY, 1986)

	Bilabial	Labiodental	Dental	Alveolar	Palatoalveolar	Palatal	Velar	Uvular	Glottal
Voiceless stops	p			t			k		
Voiced stops	b			d			g		
Voiceless fricatives		f	θ	s,ɬ	ʃ			X	h
Voiced fricatives		v	ð						
Nasals	m			n			ŋ		
Liquids				l,r					
Glides	w					j			

Note. Awbery lists the phoneme /X/ as being uvular rather than velar. This is probably a matter of dialectal variation, as Jones (1926) lists it as velar. In any case, velar and uvular place of articulation are not contrastive in Welsh phonology.

the glides /w,j/ as well as the nasal /n/ (/ŋ/ occurs initially only as a result of the nasal mutation). Words beginning with these consonants do not change form, regardless of surrounding context; thus, /frind/ "friend" remains /frind/ in all circumstances.

The mutation classes are differentiated from each other not only by the particular phonological changes they engender, but also by the different sets of linguistic circumstances that trigger them. The contexts that trigger mutation are predominantly lexical (i.e., occurring with a particular word), but also can be syntactic (i.e., occurring in a certain type of grammatical phrase or syntactic structure). The contexts have no obvious phonological, structural, or semantic commonality in modern Welsh. As a historical note, however, membership in the class of mutation-triggering contexts, which seems so unmotivated in the modern language, stems from developments in the 5th and 6th centuries, when early phonological rules affected the pronunciation of word-initial consonants according to the final sound of the preceding word; later, the preceding vowels and consonants ceased to be pronounced, but the custom of mutating initial consonants was preserved and even extended (Willis, 1986, p. 42).

Examples of lexical contexts in which consonants are obliged to undergo the SOFT mutation include grammatical function words such as /i/ "to," as in /i vaŋgor/ "to (the city of) Bangor"; and /pan/ "when, at the time," as in /pan ðaiθ/ "when he came" (from /daiθ/ "he/she/it came"). In these cases, the preceding word itself is the trigger, that is, synonyms and homonyms do not have the same effect. The vast majority of the SOFT mutation contexts are of this type, as are all of the aspirate mutation contexts (Awbery, 1973): for example, consonants undergo the ASPIRATE mutation after the word /a/ "and," as in /te a xofi/ "tea and coffee"; and after the word /tri/

"three," as in /tri xant/ "three hundred" (from /kant/ "hundred").

The SOFT mutation is also used in syntactic contexts. For example, it occurs in adjectives following a feminine singular noun, as in /ə vasged bert/ "the pretty basket" (from /basged/ "basket" and /pert/ "pretty"); in words used vocatively, as in /bore da, blant/ "good morning, children" (from /bore/ "good," /da/ "morning," and /plant/ "children"); in a direct object following a verb (with or without intervening material), as in /gweloð Tom gi/ "Tom saw a dog" (from /gweld/ "see" and /ki/ "dog"); and in the subject of a sentence when an adverb or other material is inserted between the verb and subject (Welsh has verb–subject–object sentence order). In some cases, the application of mutation by itself carries functional or semantic load; for instance, according to Jones (1977, p. 335) "the Soft Mutation is the sole interrogative marker in spoken Welsh," as in /welest ti ve/ "did you see him?" (from /gweld/ "see" and /ti/, /ve/ "you," "him").

In spite of the current arbitrariness of the triggering contexts, mutation in Welsh is a regular and productive process, in the sense that the mutations apply (with minor exceptions) to all phonologically eligible words in the language. Moreover, new words and borrowings are mutated just as entrenched Welsh words would be. It is also obligatory: use of the mutations cannot be varied for stylistic reasons and native speakers are identified by the consistency and ease with which they use mutations. Indeed, it is as pervasive a presence in Welsh sentences as, for instance, the use of the suffix /-s/ in plurals and the present tense is in English: a typical Welsh sentence will contain at least one instance of mutation and often more. All changes in initial consonants due to mutation are reflected in the writing system, that is, *pont* is written *bont* in a mutation context.

The Welsh mutations present some interesting opportunities for testing the hypotheses about lexical organization discussed above. First, the relationship among the forms /pont/, /bont/, and /font/ is not an instance of affixation, analogous to English POUR–POURED, but rather an instance of internal modification analogous to the English HANG–HUNG case. (Although the three share a continuous unit, ONT, this unit is not meaningful by itself.) Thus, if a shared, continuous base form is required in order for words to share a lexical entry or recognition unit, then /bont/ or /font/ would be expected to prime /pont/ no more than HUNG primes HANG (i.e., partial priming in lexical decision, Stanners et al., 1979; and no priming in auditory word recognition, Kempley & Morton, 1982). On the other hand, if the key factor in lexical organization is paradigmatic regularity, then the regular Welsh variants /pont/ and /bont/ would be expected to prime one another as much as the regular English variants POUR and POURED prime each other.

The particular pattern of phonological changes in Welsh mutations provides us with Welsh-internal controls on the role of form change in priming. In particular, words beginning with voiced stops show no change in ASPIRATE mutation, although words beginning with voiceless stops do show such changes. Thus, we can compare, for example, how well variants in the ASPIRATE mutation prime BASE forms, when this difference is or is not accompanied by an actual change in phonological form. This kind of control is important, since without it, the effect of form change could only be assessed by comparing results across experiments involving different languages.

The primary goal of Experiments 1 and 2 is to examine the role of form change in priming using the Welsh mutations. Robust, or full, priming would indicate that sharing a lexical entry or recognition unit does not depend on shared, continuous surface sequences, as long as the variation is regular. In addition, the experiments were designed to allow evaluation of alternative lexical organizations for morphologically related words. One alternative is the model that Stanners et al. (1979) propose for derivations and irregular inflections in English. In this model, these irregularly related forms do *not* share lexical entries, but rather constitute separate, closely associated entries. In addition, Stanners et al. hypothesize that processing of irregular inflections (and derivations) requires activation of the entry for base form in memory, although this activation is less strong than if the base and related form shared the same entry (as they do for regular inflections). This hypothesis was motivated by their results showing that derivations and irregular inflections require a longer lexical decision time (when unprimed) than their associated base forms. The Stanners et al. model predicts an asymmetry in priming: since processing irregular inflections and derivations requires the activation of the base forms, we would expect measurable priming of the bases by such variants. Conversely, activation of the variants is presumably not required in order to process the base (otherwise there would be no accounting for the differences in unprimed decision time), and thus we would expect bases to prime these variants much less strongly. Experiment 1 was additionally designed, therefore, to look for evidence of this symmetry, by examining both priming of mutations by base forms and priming of base forms by mutations. Such asymmetries, particularly if accompanied by overall weak priming effects, could be taken as evidence for a "separate entries" organization of the Welsh mutations.

Experiment 2 partially replicates Experiment 1, and is also designed to look for evidence of the "satellite" model of lexical organization proposed by Lukatela et al. (1980). This model, developed for Serbo-Croatian inflections, assigns a separate lexical entry to each member of an inflectional paradigm, but ties each one to a single core entry (the nominative singular form). Since the satellites are not related to each other, but only to the nucleus, we would expect less priming between two satellite entries than between the nucleus and a satellite. To the extent that this kind of organization is appropriate for Welsh mutations, we would expect the base form to be the nucleus, since it is the most common and representative of the possible variants: it is the form any Welsh speaker will give in response to the question "What's the word for *X*?" and the only form that can occur in isolation. Experiment 2 will allow us to test the predictions made by this kind of lexical organization by comparing priming between two mutations with priming between mutation and base forms.

Experiment 3 is designed to compare the overall recognition rate for words in different mutations. Of particular interest here is the fact that Welsh mutations specifically effect changes in the word initial segment. There is substantial evidence that beginnings of words have some special status in the lexicon (Browman, 1978; Fay & Cutler, 1977) and in word recognition (Cole & Jakimik, 1980; Marslen-Wilson & Welsh, 1978). Theories of word recognition model the importance of the initial portion of the word by assuming that recognition involves creating a "short list" of potential recognition candidates on the basis of the word-initial segments. Such a short list is either seen as being created on the fly by independent word recognition units operating in parallel (cf. Marslen-Wilson, 1984), or as inherent in how the lexical entries are organized in a mental dictionary (cf. Forster, 1976; Taft, 1979). For Welsh, such an organization might be expected to cause difficulties, particularly if the evidence from Experiments 1 and 2 points to a single lexical entry or recognition unit for all mutation variants. If there is a single entry for the Welsh forms /pont/, /bont/, and /font/, in which "short list" would it be found—the /p/-initial, /b/-initial, or /f/-initial lists? If it were entered with the /p/-initial forms (appropriate to its base), then this would suggest some added complexity in the process of recognizing the mutated forms. The recognition units for mutating words would have to be sensitive to the interaction of phonological form and syntactic environment. Experiment 3 looks for evidence of such complexity by comparing recognition of mutating words with control words that do not mutate.

General Method

An unusual aspect of the study described here is the fact that all three experiments were run simultaneously on the same set of speakers. This was done to make maximal use of a limited amount of time in Wales and a potentially limited number of Welsh-speaking subjects. Those design features and procedures the three experiments have in common are described here, while the aspects of design and methodology unique to each experiment are described in the separate experimental sections.

The experimental paradigm used was a variation of the auditory priming methodology found in Kempley and Morton (1982). In our experiments, as in theirs, subjects were first required to attend to a list of priming words, presented under optimal listening conditions, and then to identify, using written responses, a list of target words presented in noise. Each target word was associated with a particular prime-target condition, depending on the specific experimental design. Typical prime-target conditions include priming of a target by itself (the "identical" form in prime and target lists), priming of the target by the same word in a different form ("related" or "different" priming) and no priming ("none"). Subjects were divided into eight groups, with each group receiving a given target word in a different prime-target condition. The use of eight subject groups was dictated by the structure of Experiment 2, which compared eight prime-target conditions. In Experiments 1 and 3, both of which logically required only six subject groups, certain prime-target conditions were duplicated in order to fill out the stimulus tapes for subject groups 7 and 8.

For each experiment, the experimental treatments were distributed among the eight subject groups and the selected lexical items using a balanced fractional factorial design based on Latin Squares. For example, given eight prime-target conditions and eight words, subject group I received word A in condition 1, word B in condition 2, word C in condition 3, and so on, while subject group II had word B in condition 1, word C in condition 2, word D in condition 3, etc. Thus, each word occurred in a different experimental treatment for each group of subjects. The number of words used in any experiment was always an even multiple of the number of prime-target conditions defined by the experiment.

Stimulus Construction

Words. There were 72 words altogether, of which 24 were used in Experiment 1, 32 in Experiment 2, and 16 in Experiment 3. All except eight nonmutating words (used in Experiment 3) are regularly subject to SOFT and ASPIRATE mutations in Standard Welsh. The eight nonmutating words are unambiguously nonmutating; that is, there are no corresponding mutating words that share the same form with the nonmutating words in some mutation. Criteria for word selection were that the word be a monosyllabic or bisyllabic masculine noun with initial stress and end in a closed syllable. The number of words with a particular initial consonant in the base form (for mutating words /p,t,k,b,d,g/; for nonmutating words /f,x/) was balanced in each experiment. Equal numbers of monosyllabic and bisyllabic words were chosen. We tried to limit the list to frequently used words for commonplace objects or concepts, and to avoid words with limited areal (dialectal) distribution, or unstable gender identification, by checking with a native speaker. The complete set of words is listed in Appendix I.

Mutation frames. Although base (dictionary) forms may occur in isolation, Welsh speakers find mutated forms removed from their triggering context extremely strange. Accordingly, in the experiment all forms were always presented embedded in an appropriate syntactic frame. The SOFT and ASPIRATE mutation contexts consisted of the preceding possessive pronouns /ei/ "his" and /ei/ "her," plus a disambiguating postposition commonly used in colloquial speech. For the SOFT mutation, this was /o/, meaning "he" or "him." For the ASPIRATE mutation, this was /hi/, meaning "she" or "her." The BASE frame consisted of the preceding definite article /ə/ plus a postposition /əma/ meaning "there," which when combined have a colloquial meaning "that." Since the definite article is a base form context for masculine (but not feminine) nouns, our experiments used only singular number masculine nouns. For example, presentation of the word /pen/ "head" for the SOFT mutation occurred in the phrase /ei ben o/ "his head," for the ASPIRATE mutation in the phrase /ei fen hi/ "her head," and for the BASE form in the phrase /ə pen əma/ "that head."

Recording. Each word was recorded three times in each of these three frames by a native speaker of North Wales, in a sound-treated booth. The speaker was instructed to produce the words with identical intonation, speaking rate, and loudness (as much as possible), and was constantly monitored for signs of fatigue or change in these variables. To avoid encouraging systematic intonational differences between BASE and MUTATION forms, phrases were recorded in a randomized list. The most intelligible and normal-sounding example of each phrase, as determined by a second Welsh speaker, was then digitized at 20 Khz using the PCM system at Haskins Laboratories.

Noise levels. Each target stimulus was presented in noise. The noise employed was signal correlated, that is, noise manufactured from the corresponding stimulus by reversing the polarity of digitized signal points in a random pattern. Unlike white noise, which is known to differentially degrade recognition of different segment types, noise made by this technique, which matches the amplitude envelope of the noise to that of the signal, may affect different segments and words more nearly equally.

Addition of a given level of noise affects the recognizability of different words to different degrees, particularly as a function of word frequency (Howes, 1957; Pollack, 1963; Rubenstein & Pollack, 1963). In addition, individual subject differences in hearing or word recognition strategy may also affect word recognition under noisy conditions. Thus in analyzing the outcome of recognition tests in noise, it is desirable to control to some extent the effects of individual word and subject variation. In this study, we drew on previous experience with Welsh pilot studies to determine an anchor noise level at which subjects averaged 50% recognition. This anchor level differed for the monosyllabic and bisyllabic stimulus words. In addition, we attempted to calibrate noise levels for individuals subjects. To this end, we administered a noise-level sensitivity pretest to each subject before commencing the experiment proper. The pretest session was designed to choose a noise level for each subject that would yield an average 40% recognition rate. (Details of the process by which noise levels were chosen can be found in Appendix II.)

Stimulus orders. For each subject group, the target words (in the appropriate mutation frame) for all three experiments were combined into a single target list; similarly, the priming words (in their appropriate frames) were combined in a single priming list. Extra lead-in items (also in appropriate frames) were then added to the priming and target lists. Both the priming list and the target list were randomized separately from each other, and separately for each subject group, using a block randomization procedure that guaranteed that each quartile of each list contained the same number of words from each experiment. In line with Kempley and Morton's (1982) procedure, the priming list was repeated twice in succession on the test tape. The priming list was differently randomized and contained different lead-in items for each presentation.

Thus, a test session for any one subject included, in order, (1) the noise level pretest described above, (2) the priming items assigned to his or her subject group, presented twice, and (3) the target list for that group. Eight separate test tapes were made, one for each group of seven subjects. One channel of each tape held the noise-free recordings of pretest, priming, and target stimuli; the second channel carried synchronized correlated-noise images of the pretest and target stimuli, at an amplitude equal to the amplitude of the words. (For the priming session, this channel carried only silence.) Noise and signal channels were then mixed (using a circuit that allowed variable attenuation of the two channels) at a signal-to-noise ratio determined by the subject's response on the pretest, as detailed in Appendix II.

Procedure

Subjects were instructed, at the beginning of the test session, to expect words to occur in only one of the three phrases described above: /ə_əma/, /ei_o/, and /ei_hi/. For the pretest, subjects were instructed to write down the entire phrase as they heard it. For the two presentations of the priming list, subjects were told that they would hear a number of different mutated and base-form nouns embedded within one of the three frames, and were requested to provide a written estimate of the frequency with which the central noun in the phrase was used in everyday conversation, on a scale of 0 to 100. They were provided with an answer sheet with appropriate scales. This task ensured the subjects' attention to each word on the priming lists, and provided two estimates of token frequency for each lexical item in the experimental corpus. The estimates of frequency are reported in Appendix III, and further discussed in the relevant experiments; in general, while there were frequency differences among the lexical items, these were distributed evenly among the various conditions, with the primary exception of slightly higher frequency judgments for items in their base form than in their mutated forms (base mean = 59, mutation mean = 52). Four seconds of silence intervened between the presentation of each item in the priming lists.

Approximately 5 minutes after the second presentation of the priming list, the subjects were presented with the list of target items. In the interim, they were instructed that they were now to write down the entire phrase as they heard it, just as in the pretest case. After five lead-in items (different from the priming lead-ins but identical for each subject group), the list of target items was played, with 8 s between each phrase.

Each subject was tested individually in a quiet room. The total priming list for any subject group (from all three experiments plus lead-ins) consisted of 53 items, the total target list of 77 items. The entire test session for any one subject lasted approximately 35 minutes.

Subjects

Subjects were 56 native speakers of Welsh. Forty-seven subjects were students between the ages of 17 and 35 (45 at the University of North Wales in Bangor, 2 at Cambridge University in Cambridge, England), while 9 were older members of the Welsh speaking community in Cambridge, England. All had early schooling in Welsh, and many had been educated in Welsh up to university level. All were fully bilingual in Welsh and English. Subjects were assigned to groups in a random fashion.

Experiment 1

As discussed in the introduction, the primary goal of Experiment 1 was to examine whether full priming is found between base and mutation forms of Welsh nouns. An affirmative answer would indicate that a shared, continuous phonological base form is not required for the variants to share a lexical entry. A second goal of the experiment was to test for possible asymmetries in the direction of priming. As argued above, the Stanners et al. (1979) model for irregular morphological relations predicts that mutations should prime base forms much more strongly than base forms prime mutations. Experiment 1 was therefore designed to show any differential effects of priming between BASE and MUTATED forms.

The design for Experiment 1 included three type of primes: IDENTICAL (prime and target were the same lexical item in the same mutation), DIFFERENT (prime and target were different forms of the same lexical item, one a base, the other a mutation), and NONE (no prime at all). In addition, for each prime type, targets were of two types: MUTATION and BASE. This yielded the six prime-target conditions shown in Figure 1, exemplified using the lexical item /pen/ ("head"). The six conditions were replicated for each of two mutation classes: the SOFT and ASPIRATE mutations. A set of 12 words (6 monosyllabic, 6 bisyllabic; each group of 6 including one each of word-initial /p,t,k,b,d,g/) was assigned to each of these replications. The distribution of the 12 words across the six prime-target conditions varied over six subject groups according to the modified Latin Square design described earlier. In order to conform to the overall study design of eight subject groups, the prime-target patterns for subject groups 1 and 2 were repeated for subject groups 7 and 8.

TARGET	PRIME: IDENTICAL	PRIME: DIFFERENT	PRIME: NONE
BASE	ə pen əmə [ə pen əmə]	ei ben o [ə pen əmə]	-- [ə pen əmə]
MUTATION	ei ben o [ei ben o]	ə pen əmə [ei ben o]	-- [ei ben o]

FIG. 1. Example of experimental design for Experiment 1. Within each box, the top item is the prime, the bottom item is the target.

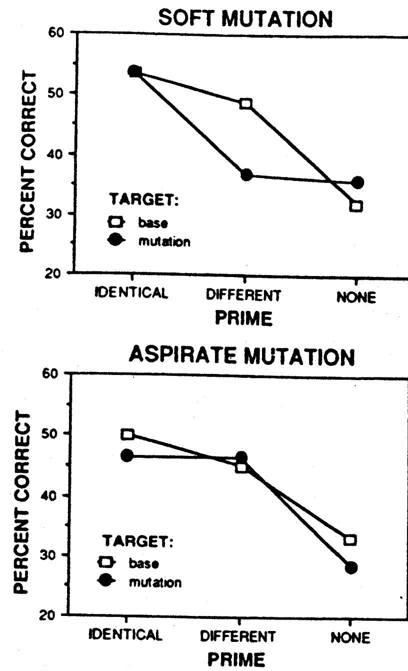

FIG. 2. Percentage of correct recognition for each prime condition (Experiment 1).

Results

Subject responses were scored as correct if the entire stimulus, including mutation and surrounding frame, was correctly transcribed. Results are shown graphically in Figure 2. The top panel of the figure shows the percentage correct recognition for the six prime-target conditions for the SOFT mutation replication, the bottom panel shows the same for the ASPIRATE mutation. The figures show that IDENTICAL priming yields the highest recognition rate, followed closely by DIFFERENT priming, which in turn shows better recognition than the NONE condition.

Two separate analyses of variance were performed (in BMDP4V), one using the variability of individual subjects and one using item means to form the error term. Factors were SYLLABLE (ONE vs TWO), MUTATION (SOFT vs ASPIRATE), TARGET (BASE vs MUTATION form), and PRIME (IDENTICAL vs DIFFERENT vs NONE). Those factors involving different sets of words (SYLLABLE and MUTATION) were treated as within-group, or repeated measures, factors in the subject analyses and as between-group, or grouping, factors in the item analyses. There were no systematic differences between subject groups. Note that only data from subject groups 1 through 6 are reported here (the results did not differ when data from subject groups 7 and 8 replaced data from groups 1 and 2 in the analyses).

Only the main effect for PRIME was significant in both the subjects and items analyses: subjects, $F(2,82) = 16.78, p < .0001$; items, $F(2,40) = 12.53, p < 0.001$. Planned comparisons between the IDENTICAL and DIFFERENT conditions (the test for full vs partial priming) showed the two to be marginally statistically different in the subjects analysis, $F(1,41) = 6.70, p = .013$, but not in the items analysis, $F(1,20) = 3.30, p = .084$. However, the DIFFERENT condition was reliably different from the NONE condition in both analyses: subjects, $F(1,41) = 11.49, p < 0.005$; items, $F(1,20) = 10.55, p < .005$. The effect of SYLLABLE was significant in the subjects analysis, but not the items: subjects, $F(1,41) = 23.65, p < .0001$; items, $F(1,20) = 1.71, p = .205$. The recognition rate was higher for monosyllables (monosyllabic mean = 49%, bisyllabic mean = 36%). This may indicate that the expected recognition advantage of bisyllabic words was overcompensated by the difference in noise levels; it might also reflect frequency differences between the monosyllables and bisyllables, since these differences are in the right direction, although not significant. No other main effects, nor any interactions, were significant. Note that the lack of a TARGET effect indicates that the small (but significant) difference in estimated frequency between base and mutated forms was not reflected in the recognition accuracy.

In Kempley and Morton (1982), a distinction was drawn between the results as scored by strict criteria and the results as scored by more lenient criteria in which a response containing the correct lexical item was considered correct, regardless of whether it was the correct morphological variant. In their experiment, this amounted to a distinction between "whole word" scoring (transcription of the stimulus as given) and "morphemic" scoring (responding with some variant of the given stimulus). In our experiment, there were several possibilities for "morphemic scoring," ranging through (1) responding with the right frame but wrong mutation; (2) responding with the wrong frame but right mutation; and (3) responding with the wrong frame and mutation. Informal analysis of the incidence of these response types revealed no systematic pattern; in addition, statistical reanalysis of the data obtained by scoring all three as correct responses showed the same results as the strict scoring procedure reported above.

Discussion

The most important result in this experiment is the significant effect of priming between BASE and MUTATED forms. Clearly, unlike Kempley and Morton's (1982) result for English, priming does occur between morphological variants that are not eligible to share a logogen. This, in turn, supports the view that the kind of intimate lexical sharing that is indexed by repetition priming does not require that the forms share a unique, continuous phonological sequence. Rather, the lexical organization among variant forms of a given lexical item in Welsh seems to be more akin to the organization of regular paradigms in English than to the organization of irregular English pairs.

This conclusion is also compatible with another major result of the experiment, namely, the failure to find any asymmetries in priming. The lack of a significant effect for TARGET and the lack of any interactions between PRIME and TARGET indicates that BASE and MUTATED forms are equally effective both as primes and as targets. Note that this result would not be expected under a theory of lexical structure that assumes that the variant forms have distinct lexical entries and that one of the forms is always accessed through the other, dominant form (a principal entries hypothesis). Rather, the current results argue for symmetrical lexical access.

It might be objected that this interpretation is not valid because it is based on partial rather than full priming, while most arguments for shared entries are based on full priming between related forms (cf. Stanners et al., 1979). However, we should point out that the difference between the IDENTICAL and DIFFERENT conditions only just approaches significance in our results, even in a powerful, a priori test. Further, it has been noted (Fowler et al., 1985) that full priming results in the literature have tended to show numerically, if not statistically, weaker priming for the DIFFERENT condition. That is, priming between related forms may be consistently less than priming between identical forms, but at a level that only sometimes reaches statistical significance.

It should be noted that, even if this difference between conditions is real, the differences are not necessarily due to the *phonological* differences between the prime and target words. The prime and target in the DIFFERENT condition also differ in their morphosyntactic (and semantic) mutation frames. The question of form vs frame can be pursued, in a preliminary way, by breaking down the data for the ASPIRATE mutation further. Since words whose BASE forms begin with voiced /b,d,g/ do not undergo changes in the aspirate mutation, but words beginning with voiceless /p,t,k/ do undergo change to the corresponding fricatives, comparison of these two sets of stimuli could shed some light on the importance of phonological form change in contributing to the difference between IDENTICAL and DIFFERENT conditions. In fact, words with voiceless base forms show a decrement from 45.2% in the IDENTICAL condition to 39.2% in the DIFFERENT condition, while words with voiced base forms show 51.2% in the IDENTICAL condition and 52.3% in the DIFFERENT condition. While this difference suggests that phonological form plays a role, the design of Experiment 1 did not allow this difference to be properly evaluated statistically. In Experiment 2, therefore, the underlying voicing status of the target was explicitly set up as a factor in the design.

EXPERIMENT 2

Experiment 2 extends the findings of Experiment 1 to additional types of prime-target pairings, and attempts to explore more carefully the role of phonetic factors and form changes in priming. In Experiment 1, the issue of symmetrical vs asymmetrical relations among forms of a given lexical item was tested using BASE–MUTATION and MUTATION–BASE prime-target pairs. Results suggested that the relations between the forms is symmetrical. It is possible, however, that whatever form of organization links the base form with the different mutation forms (and vice versa) does not operate between mutated forms. This would be the case, for example, in a "satellite" form of organization (e.g., Lukatela et al., 1980). In such a model, the base form would still be the central element of a cluster whose peripheral members remain unconnected with each other. In Experiment 2, therefore, priming between related forms was further explored by including conditions in which an item in one mutation form (SOFT or ASPIRATE) served as a prime for that item in the other mutation form.

The evidence for symmetrical relations in Experiment 1 lies in the absence of a significant TARGET effect, or of any significant interactions involving TARGET. However, as is clear in Figure 2, for the SOFT mutation there is a difference between BASE and MUTATION targets in the DIFFERENT priming condition (49 vs 37%), while for the ASPIRATE mutation, such a difference is not found (45 vs 46%). It is not clear whether this difference between the mutations is reliable (as noted above, interactions were not significant), or how to interpret it, if it should turn out to be reliable. It could reflect slight but genuine differences in lexical structure between the two mutations (i.e., relations are more symmetrical in the case of ASPIRATE mutation). Alternatively, it might be due to the fact that the SOFT and ASPIRATE mutation data come from separate sets of lexical items in Experiment 1. Or it may be the case that phonetic differences in the discriminability of consonants produced by the two mutations are, in fact, obscuring an underlying similarity between the two mutation classes. Experiment 2 was designed so that the same lexical items occurred in both mutation classes; and the experiment was set up to allow the phonetic effects of the target on recognition to be evaluated.

The experiment was designed as follows. All targets were presented in mutation form (either SOFT or ASPIRATE mutations), and they were primed under one of four conditions: IDENTICAL (same mutation form used for prime as for target), BASE (BASE form used as prime), OTHER (ASPIRATE form used as prime if target was SOFT and vice versa), and NONE.

The targets themselves were divided into four groups by two additional experimental factors. Half of the targets used in the experiment, which we will refer to as BITEMS, had base forms beginning with one of the voiced stop consonants (/b,d,g/), while the other half of the targets (PITEMS) had base forms beginning with one of the voiceless stops (/p,t,k/). Note (with the aid of Table 1) that the PITEMS undergo phonetic changes in both the SOFT and ASPIRATE mutations, while the BITEMS undergo changes only in the SOFT mutation.

The second factor on which targets varied was the mutation form in which the target was presented. However, rather than code the mutation form as SOFT vs ASPIRATE, the target was coded according to the *phonetic* effect of the mutation, because of our interest in factoring out possible phonetic effects on recognition. Looking again at the consonant changes wrought by the two mutations (illustrated in Table 1), we see that for words beginning with stop consonants, there is one mutation context in which the phonetic feature of manner of articulation is changed from stop to fricative, and one mutation context in which the stop remains a stop. For BITEMS, this change from stop to fricative is induced by the SOFT mutation, while for PITEMS it is the ASPIRATE mutation that produces this change. Thus, the target is characterized according to the MANNER of its initial consonant—STOP vs FRIC. Note that vowel-initial targets derived from /g/-initial base forms (by the SOFT mutation that derives /v/ from /b/ and /ð/ from /d/) are characterized as FRIC. This was done for the sake of symmetry in experimental design.

The four PRIME conditions were crossed with the four target conditions (2 BP × 2 MANNER), giving a total of 16 prime-target conditions. These are illustrated in Table 3, using the items /pen/ "head" and /beic/ "bicycle" as examples. Two sets of eight monosyllabic words were chosen, one PITEM set containing two words each beginning with /p/ and /k/ plus four words beginning with /t/, and one BITEM set containing two words each beginning with /b/ and /g/ plus four words be-

TABLE 3
DESIGN OF EXPERIMENT 2

		Identical	Base	Other	None
BITEMS	STOP	ei beic hi ei beic hi	ə beic əma ei beic hi	ei veic o ei beic hi	— ei beic hi
	FRIC	ei veic o ei veic o	ə beic əma ei veic o	ei beic hi ei veic o	— ei veic o
PITEMS	STOP	ei ben o ei ben o	ə pen əma ei ben o	ei fen hi ei ben o	— ei ben o
	FRIC	ei fen hi ei fen hi	ə pen əma ei fen hi	ei ben o ei fen hi	— ei fen hi

Note. Within each group, the top item is the prime, the bottom is the target.

ginning with /d/. Each set was distributed over the eight prime-target combinations appropriate to that item and over the eight subject groups as described above in Experiment 1. Each word appeared in each of the eight prime-target conditions across different subject groups, and no subject heard the same words in more than one prime-target condition. Also as in Experiment 1, this structure was duplicated with a set of bisyllabic words.

Results

Subjects' responses were scored by strict and lenient ("morphemic") criteria, and repeated measures analyses of variance were carried out by subject and by item as described in Experiment 1. As before, results were the same by either method of scoring; only the results scored by strict criteria are reported below. Figure 3 shows the percent recognition for each of the 16 prime-target conditions. Lines connect the points for a given target type—BITEMS are connected by solid lines, PITEMS by textured lines. STOPS are represented by solid circles, and FRICS by open squares.

Factors for the analysis of variance included MANNER (STOP vs FRIC), BP (BITEM vs PITEM), PRIME (IDENTICAL, BASE, OTHER, and NONE), and SYLLABLE (ONE vs TWO). As in Experiment 1, PRIME was the only factor that was significant in both subject and item analyses: subjects, $F(3,165) = 15.12$, $p < .0001$; items, $F(3,84) = 13.82$, $p < .0001$. As expected, recognition was highest in the IDENTICAL condition and lowest in the unprimed (NONE) condition, while priming between morphological variants (the BASE- and OTHER-mutation priming conditions) showed intermediate, and approximately equal, levels. Contrasts showed no significant difference between these two intermediate conditions: subjects, $F(1,55) = .36$, $p = .55$; items, $F(1,28) = .21$, $p = .65$. To test for full vs partial priming, we contrasted the IDENTICAL

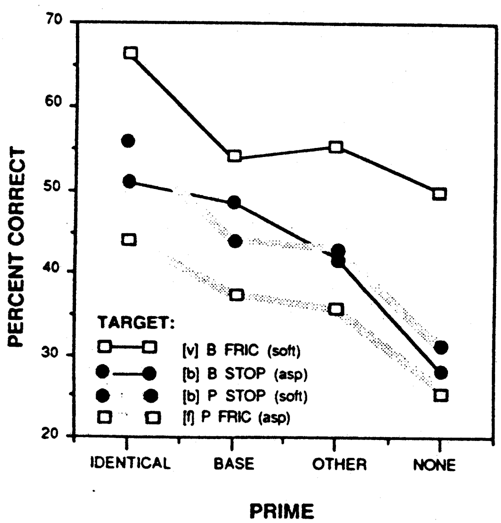

FIG. 3. Percentage of correct recognition for each prime condition (Experiment 2).

condition against the combined BASE and OTHER (non-IDENTICAL) conditions. This contrast was significant in both analyses: subjects, $F(1,55) = 10.16$, $p < .005$; items, $F(1,28) = 11.68$, $p < .005$. Finally, these latter two conditions contrasted significantly with the NONE condition: subjects, $F(1,55) = 19.36$, $p < .0001$; items, $F(1,28) = 25.87$, $p < .0001$.

Returning to the main analysis, overall differences in the recognizability of targets, as a function of the target factors (BP and MANNER), are visible in the separation between means corresponding to the four categories of targets (shown with connected lines in Figure 3). While MANNER itself (FRIC vs STOP) was not a significant main effect, there was a significant effect for BP (BITEMS vs PITEMS) in the subjects analysis, but not the items analysis, as well as a significant interaction of BP × MANNER (BP: subjects, $F(1,55) = 20.66$, $p < .0001$; items, $F(1,28) = 2.37$, $p = .13$; BP × MANNER: subjects, $F(1,55) = 19.46$, $p < .0001$; items, $F(1,28) = 3.8$, $p = .06$). The pattern of results visible in Figure 3, that is, the wide separation of lines for the voiced and voiceless FRIC targets and the clustering of lines for STOP targets midway between them, suggests that the significance of the BP factor could be entirely attributable to the MANNER × BP interaction. This would mean that the data can be explained phonetically by positing a recognition advantage for voiced fricative targets, relative to the stop baseline recognition, and a similar disadvantage for voiceless fricative targets. This interpretation is supported by the fact that an analysis of simple main effects shows that the BP factor is significant for FRIC targets, but not for STOP targets (FRIC: $F(1,55) = 30.92$, $p < .0001$; STOP: $F(1,55) = .17$, $p = .69$; analysis done for subjects only, since the overall effects were only significant there). Alternatively, it is possible that rather than being a function of the phonetic factors the data reflect a basic difference in recognizability between the two mutations. (Note this is not reflected in the estimated frequencies of the items in the two mutations, which are not significantly different: 54 vs 52.) Reanalysis of the data by the factor MUTATION (targets derived by the SOFT vs targets derived by the ASPIRATE mutation) in fact revealed a significant difference between the two. All other results, including a significant BP effect, were the same in this reanalysis; there was, however, no BP × MUTATION interaction. From this reanalysis, it would appear that the two mutations have different baseline recognition rates, and that there is also some difference between recognition rates for BITEMS and PITEMS. At this point, there is insufficient evidence to decide between the alternative analyses. Experiment 3 will, however, be useful in this regard. In any case, none of these issues appear to affect priming; relationships between priming conditions are consistent for all categories of targets.

As in Experiment 1, the SYLLABLE effect was significant in the subjects but not items analysis: subjects, $F(1,55) = 7.97$, $p < .01$; items, $F(1,28) = .95$, $p = .34$. The monosyllabic recognition mean was 41%, the bisyllabic mean was 47%. No other effects or interactions reached significance. There was a very small, nonsignificant difference, in the right direction, in the estimated frequencies of the monosyllables and bisyllables.

Experiment 2 was also set up to allow us to compare priming in BASE prime-target pairs that involve an actual change in initial consonant with those BASE prime-target pairs that do not actually show any consonant change. The relevant comparisons here involve BITEMS vs PITEMS for the STOP manner, in which both BITEMS and PITEMS have targets that are mutations (BITEMS are ASP, PITEMS are SOFT) and that begin with /b,d,g/. In the IDENTICAL condition both sets of items are primed by targets beginning with /b,d,g/. However, in the BASE condition, primes for the BITEMS have the same initial consonants as the targets (e.g., /ə beic əma/ priming /ei beic hi/ in Table 3), while for PITEMS, the primes differ by having voiceless initial stops (e.g., /ə pen əma/ priming /ei ben o/ in Table 3). If form change per se between target and prime was a major determinant of strength of priming, then we would expect recognition to decrease from IDENTICAL to BASE conditions in the case of the PITEMS, but we would expect much less decrement in the case of the BITEMS. The points in Figure 3 are indeed in the right direction. However, a separate analysis of variance performed on these four cells alone revealed no interaction, and indeed, no significant main effect of PRIME, that is, IDENTICAL vs BASE (PRIME: subjects, $F(1,55) = 1.91, p = .172$; items, $F(1,28) = 1.64, p = .211$; PRIME X BP: subjects, $F(1,55) = 1.04, p = .312$; items, $F(1,28) = .64, p = .43$).

Discussion

The apparent equivalence of the BASE and OTHER conditions, like the lack of asymmetries found in Experiment 1, argues against a kind of "satellite" lexical organization for Welsh mutations in which all the variant forms are related to the base, but not to each other. Once again, the results from Welsh suggest that the relationship among mutation variants is quite close, and there is no evidence to compel us to the view that each mutation variant must have its own independent lexical entry.

In Experiment 2, the difference between IDENTICAL and non-IDENTICAL (BASE plus OTHER) priming conditions reached significance (corroborating a trend found in Experiment 1). This suggests that priming between mutation variants, or between base and mutations, is only partial, rather than full. In this sense, the results for Welsh differ from those found for regular morphological processes in English. However, the priming difference does *not* seem to stem from the differences in phonological form between the lexical item in different mutations. As the separate smaller ANOVA on the BASE condition showed, indications of a difference in pattern for those lexical items that do show form changes (PITEMS) and those that do not show form changes (BITEMS) were not statistically significant. Thus, the fact that priming in the BASE and OTHER conditions is not full must be attributed to other differences between the conditions—for example, to syntactic or semantic differences between the mutation frames, or to abstract membership of an item in a particular mutation class. The fact that priming between items that show form change can be, statistically, as strong as between items that do not show such a change argues against defining logogens, or any notion of lexical units, strictly in terms of the criterion of unique shared physical sequences.

Experiment 3

In Experiment 3, we turn away from examining prime-target relationships, and instead look for differences in average recognition rates for different morphological variants. There was a suggestion of such differences in Experiment 2, where one interpretation of the results of the target manipulations was that there are differences in recognizability between the ASPIRATE and SOFT mutations. As discussed in the introduction, differences in recognizability among mutation variants are potentially quite interesting because of the important role initial segments play in lexical access (e.g., Cole & Jakimik, 1980; Taft, 1979). Since the evidence from Experiments 1 and 2 points to mutation and base variants sharing a single lexical entry, this single entry would have to be accessed by inputs varying in their initial consonants. If this kind of flexibility were problematic for the word recognition system, then we would expect to find some evidence of this difficulty in different recognition rates for the different morphological variants. In addition, we might expect mutating lexical items to be more difficult to recognize than words whose initial consonant is invariant across mutation environments (e.g., Welsh words whose base forms begin with fricatives).

In this experiment, only the unprimed (NONE) and IDENTICAL prime-target conditions were employed. One comparison of interest to us was the contrast between the behavior of the BASE, ASPIRATE, and SOFT forms. A second comparison of interest was the difference between MUTATING and NONMUTATING words, both in terms of overall recognition rate and also in terms of similarity of behavior in the different mutation contexts. If the MUTATING and NONMUTATING words display the same pattern of behavior in the different mutation contexts, any difference in recognizability between the morphological variants can not be attributed to difficulties associated with a variable initial consonant. Rather, it must be attributed to some aspect of the different mutation contexts themselves.

Thus, given two prime conditions and three morphological variants, there were 6 prime-target conditions tested: BASE primed by BASE, SOFT MUTATION primed by SOFT MUTATION, ASPIRATE MUTATION primed by ASPIRATE MUTATION, and all three forms unprimed (NONE). Two sets of eight items were employed, MUTATING and NONMUTATING. The NONMUTATING words were chosen to begin with either /f/ or /x/, segments that are also the reflex of an initial /p/ or /k/ in the ASPIRATE mutating context. The mutating words began with /p,k,b,g/. The items were distributed across the six prime-target conditions so that for each subject, two items were presented in each of the BASE conditions (BASE–BASE and BASE–NONE), and one item in each of the other four conditions. This allowed us to take advantage of all eight subject groups. As in the other experiments, the distribution of items over conditions and subject groups was varied according to a modified Latin Square.

Results

Figure 4 shows the percentage correct recognition for each of the prime-target conditions, separately for the MUTATING and NONMUTATING words. Again, repeated measures analyses of variance were carried out by subject and by item as described in Experiment 1. The results reported below reflect strict scoring criteria; as in the two previous experiments, "morphemic" scoring results were parallel. Factors were WORDTYPE (MUTATING vs NONMUTATING), TARGET (BASE, SOFT MUTATION, or ASPIRATE MUTATION), and PRIME (IDENTICAL vs NONE). The BASE level of the TARGET factor was represented by the mean (per subject or per item) of the two duplicated BASE–BASE and NONE–BASE prime conditions. (Separate analyses of the data using one or the other set of conditions showed parallel results.)

As expected, the IDENTICAL priming condition showed a significantly higher recognition rate than the NONE condition: subjects, $F(1,55) = 57.39$, $p < .0001$; items, $F(1,14) = 38.33$, $p < .0001$. The significant main effect for TARGET (subjects,

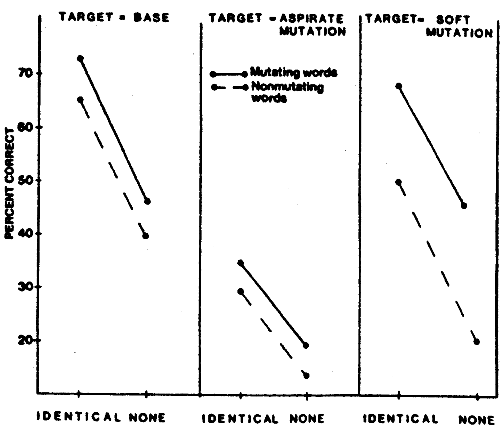

FIG. 4. Percentage of correct recognition for each prime condition (Experiment 3).

$F(2,110) = 40.03$, $p < .0001$; items, $F(2,28) = 30.62$, $p < .0001$) was primarily due to the lower recognizability of the ASPIRATE mutation. The difference between SOFT and ASPIRATE mutations was significant in a planned comparison: subjects, $F(1,55) = 29.43$, $p < .0001$; items, $F(1,28) = 34.51$, $p < .0001$; however, the BASE and SOFT mutations also differed marginally in recognition: subjects, $F(1,55) = 7.81$, $p < .01$; items, $F(1,14) = 6.21$, $p = .0259$. The TARGET effect was similar and consistent for both MUTATING and NONMUTATING words, as shown in the nonsignificant TARGET × WORDTYPE interaction. WORDTYPE was significant in the subject analysis only: subjects, $F(1,55) = 21.54$, $p < .0001$; items, $F(1,14) = 1.24$, $p = .28$. No interactions were significant.

Discussion

The comparison of recognizability for NONMUTATING and MUTATING words fails to support the notion that variation in the initial consonant of a word produces difficulties in recognition. Specifically, it is not the case that NONMUTATING words were better recognized than MUTATING words, as might have been expected if variability in initial consonants decreases recognition. In fact, the NONMUTATING words were recognized (nonsignificantly) less well than the MUTATING words. This might be due to their lower estimated frequency; however, the frequency difference is the same as that noted between the monosyllables and bisyllables in Experiment 1, which had only a marginal effect on recognition accuracy.

The behavior of the mutations also fails to support the role of initial consonant variability in recognition difficulties. While overall recognizability is greater for the SOFT MUTATION and BASE than for the ASPIRATE MUTATION, this pattern holds true for both MUTATING and NONMUTATING words. Since the NONMUTATING words have the same initial consonants in BASE, SOFT MUTATION, and ASPIRATE MUTATION contexts, their adherence to this pattern indicates that any explanation must make reference to what the MUTATING and NONMUTATING words have in common: the frame in which the word occurs. It is not clear from this experiment, however, which properties of the frame are relevant. For example, we would like to know whether the effect is specific to the morphemes for "his" /o/, and "her" /hi/, or whether the same pattern would be obtained with other contexts for the ASPIRATE or SOFT mutations. In the former case, the effect would be specific to a particular morphological category; in the latter, it would be specific to an abstract mutation class (abstract in the sense that it doesn't always result in any phonological alterations).

Suggestions of the superiority of the SOFT MUTATION and BASE forms over the ASPIRATE MUTATION can be found in Experiments 1 and 2, as well. For Experiment 1, if we examine the ASPIRATE MUTATION results in Figure 2, we see that for both the IDENTICAL and NONE conditions, the BASE targets are superior to MUTATION targets. However, for the SOFT MUTATION results, this is not the case. In the IDENTICAL condition, BASE and MUTATION are equivalent, while for the NONE condition, the MUTATION form is actually better recognized. Thus, we see in Experiment 1 the same basic pattern that we have documented in Experiment 3.

As mentioned above, in Experiment 2 reanalysis of the data by mutation category showed a significantly higher recognition rate for the SOFT MUTATION targets than for the ASPIRATE MUTATION targets. In analyzing those results, we could not choose between an analysis based on inherent differences between the two mutations and an analysis based on acoustic phonetic properties. In Experiment 2 a large portion of the difference between targets derived by SOFT and ASPIRATE MUTATIONS could be ascribed to the higher rate of recognition associated with initial voiced fricative consonants, and the much lower rate associated with initial voiceless fricatives. However, the NON-MUTATING words in Experiment 3 provide the crucial evidence—they show that there are systematic differences between SOFT and ASPIRATE MUTATION contexts even when there are no phonetic differences between the words in these contexts. This suggests that it would make more sense to interpret the results of Experiment 2 in the same way. This interpretation would be supported by other considerations as well. The alternative analysis rests on the assumption that voiced fricatives are more recognizable than stops, which are, in turn, more recognizable than voiceless fricatives. However, phonetic studies (at least for English) have tended to find a rather different order of intelligibility (e.g., Goldstein, 1977; Miller & Nicely, 1955)—stops, followed by voiceless fricatives, followed by voiced fricatives.

It is interesting to note that the apparent advantage of the SOFT over the ASPIRATE MUTATION is in accord with the linguistic facts. Awbery (1986) has shown that in many areas of Wales the phonetic changes associated with the SOFT MUTATION are replacing ASPIRATE MUTATION changes in ASPIRATE mutating environments. It is possible that what we see here is a reflection of the decreasing probability, over the language as a whole, that the ASPIRATE MUTATION will be used. Before making this kind of conclusion, however, it would be important to know whether the effect generalizes to other mutation contexts, or is specific to "his" and "her." It is unlikely to be a reflection of a specific weakness in the "her" context, since this context is one of the strongest holdouts of the ASPIRATE MUTATION (Jones, 1977).

As mentioned above, unequal recognizability among morphological categories represents inherent differences in probability of access as a function of mutation class or of morphological environment. It is hard to know how to fit such a notion into current models of lexical access. The notion of word frequency, of course, can be used to account for differences among lexical items in probability of access: in the logogen model, for instance, high frequency words fire more readily because they have a higher state of resting activation. However, within the logogen model there is no mechanism available to register frequency of use for categories like mutation class (should that turn out to be the relevant domain of the effect) as abstract entities. A possible account would be that provided by the "satellite" model of Lukatela et al. (1980), which attempts to account for differences in access *time* between nominative and nonnominative forms in Serbo-Croatian. However, as we have seen, other aspects of the satellite model do not fit the Welsh situation very well: for example, the fact that in the model all inflected forms are separate entries connected only to the nominative (or basic) form, is not supported by the results of Experiment 2.

General Discussion

The three experiments described all point to a single generalization about the organization of mutation variants in Welsh speakers' lexicons—namely, that the particular phonological relationships among the forms related by mutation processes seem to have little relevance for their relationships during lexical access. Experiments 1 and 2 showed that repetition priming is possible between mutation variants, and that this priming is equally robust between variants with and without change in phonological form. Further, Experiment 3 showed that there are differences among mutation variants in how readily they are recognized, but that such differences do not require corresponding differences in phonological form. Thus, it seems clear that the human cognitive capacity underlying lexical organization and access does not place severe limits on how phonological form can vary in lexical relations. From the Welsh results, in particular, we learn that morphological variants, in order to prime each other (1) do not need to be concatenative (affixal); (2) do not need to share an unmodified base form; and (3) do not need to share an initial consonant. Thus, any theory of the lexicon that attempts to model human lexical retrieval, as opposed to language-specific behavior, must be equally appropriate for all the variations in phonological form that occur in natural language. As discussed in previous sections, neither the strong form of the logogen model nor the form of the shared entries hypothesis that requires an unmodified base form is adequate for this purpose.

A further problem for these models, and for the satellite model as well, stems from the fact that, in Experiments 1 and 2, priming was found to be symmetrical between variants. It appears that, for Welsh at least, all the variants are symmetrically represented in the lexicon. On the other hand, the fact that Welsh variants prime each other at all indicates that they must have some lexical structure in common, that is, they cannot be wholly independent entries. These findings, together with the differing recognizability of the mutations, mean that we need a model of lexical organization that does not entail asymmetries in priming among morphological variants, but that does accommodate asymmetries in ease of access to the variants.

With these points in mind, we feel that an adequate model of the lexicon should have two independent dimensions: a lexical dimension that accesses a single entry for each set of three morphophonological variants, and a morphological dimension that accesses the morphophonological classes of BASE, SOFT, or ASPIRATE mutations. It is important to emphasize that in this view of the lexicon, the dimension containing the mutation classes is autonomous. That is, the mutation class category can not be conflated with its phonological effects, because, as we have seen, the priming and recognition behaviors of the mutations are the same regardless of the presence or absence of phonological change in the first consonant. In particular, the nonmutating items, which do not show the phonological effects of the mutation classes, nevertheless show the recognizability differences for the different mutations. It should be noted that recent linguistic analyses have also emphasized the autonomy of the morphological level, both with respect to phonological structure (e.g., Anderson, 1982) and with respect to syntactic structure (Sadock, 1985).

The two proposed independent dimensions account for the dual aspects of priming symmetry and access asymmetry in the following way. First, priming a specific morphophonological variant entails recognizing two kinds of information: the specific lexical entry common to all the morphophonological variants, and the specific mutation class. The common lexical entry accounts for the fact that all the morphophonological variants prime each other equally, while the fact that both the lexical entry and the mutation class are primed accounts for the differences between priming with the identical form and priming with other variants. Further, the independence of the mutation class dimension permits differential access of the mutations in a manner analogous to the effect of frequency of occurrence on recognizability of lexical items. By modeling the mutation classes as separate entries on this independent dimension, it is possible to associate different degrees of recognizability with each mutation class, just as lexical items with different frequencies of occurrence are associated with differing amounts of recognizability. In fact, the differing ease of access (i.e., recognizability) of the different mutations may be a direct reflection of the frequency of the mutation, where the relevant index of frequency is the number of different lexical and syntactic environments in which the particular mutation is used (Boyce, 1983). Since the SOFT mutation is used in a much wider variety of environments than the ASPIRATE, it would have a higher frequency of occurrence, which, as we have seen, is associated with higher recognizability. Thus, the two dimensions allow recognizability to be affected both by the frequency of the individual lexical item and by the frequency of the particular mutation class.

This model of lexical organization could also be employed for languages with more conventional kinds of morphology—for example, the regular plural and past tense affixes in English. However, it should be recalled that there is one difference between the present results for Welsh and those for regular inflectional morphology in English. English does not show significant differences between priming with identical forms and priming with other variants. If we assume English has the same kind of organization that we are proposing for Welsh, then the difference must somehow reside in the susceptibility of the morphological classes themselves (mutation classes vs past tense) to priming.

Thus far, the lexical organization we have proposed involves abstract lexical and mutation dimensions. To complete our proposed lexical representation, we need to relate these dimensions to the actual phonological form of words by specifying the internal structure of the lexical items. This internal structure must incorporate symmetrical relations among all the morphological variants, in order to accommodate the symmetry of priming we observed. That is, as noted above, a satellite model such as that proposed for Serbo-Croatian is not appropriate for Welsh. The minimal representation that meets the symmetry requirement is an "allomorphic" approach, in which all the morphological variants of a lexical item are fully specified as to their phonological form, and are linked together into a single lexical item either directly, or via connections to a single abstract node.

However, the use of fully specified allomorphs fails to capture, in the lexical structure, any phonological regularities that occur across the morphological variants. These regularities are most directly captured by "shared entry" models in which the phonological information common to all the variants is specified in a single shared phonological node, and the information unique to each variant is specified separately for each variant. For English and other inflectional languages, of course, this corresponds to the shared entry models of the logogen type, or (more generally) a concatenative, base plus affix approach. While a concatenative approach is not appropriate for Welsh, the basic insight of the shared entry models, that of analyzing phonological information into common and unique components, can be retained simply by relaxing the sequentiality and continuity constraints implicit in concatenative affixal morphology.

Recent work in *autosegmental phonology and morphology* has focused on precisely this issue: namely, the type of models appropriate for describing nonconcatenative morphological processes (e.g., Lieber, 1984; Marantz, 1982; McCarthy, 1981, 1984). Rather than restricting phonological representation to strictly sequential units, autosegmental analyses decompose phonological form into parallel and autonomous simultaneously occurring units called *tiers*. Each tier provides a sequence of specifications for some subset of phonological features. For example, in the Semitic case discussed in the introduction, McCarthy (1981) has proposed that consonant and vowel "melodies" are represented on separate tiers. Lexical items are distinguished on the basis of consonantal sequences on the consonant tier, while different inflectional morphological categories are represented by different sequences of vowels on the vowel tier. The units on the two tiers are linked to positions on a third, common tier (the skeleton), and in this way the actual intercalation of consonants and vowels is represented.

In the Semitic case, each tier contains information about whole segments—consonants and vowels. Other morphological analyses have been suggested in which one of the tiers contains only a subset of the features required to completely specify a segment. Thus, Lieber (1984) has proposed a linguistic analysis of the morphological–phonological relations for a language (Fula) that shows patterns similar to those found in Welsh. Fula nouns may show as many as three different initial consonants as a function of morphological categories such as singular, plural, and diminutive. The actual forms for a given lexical item all share the same place of articulation, but may differ in that one is a continuant, one is a stop, and one is a prenasalized stop (e.g., the stem meaning "free man" will have the variants /rim/, /dim/, and /ndim/). In Lieber's analysis, the morphological variants are decomposed into a stem whose initial consonant is unspecified for the features [continuant] and [nasal], and a separate tier that contains values for these features corresponding to the particular morphological categories. The combination of these two partially specified representations is seen as parallel to the concatenation of stem plus affix in the more common sequential case, in the sense that both separate the common and unique information.

It would be attractive to apply a similar analysis to Welsh. As in the Fula case, the morphological variants would be decomposed into a single underspecified stem whose initial consonant is specified only for the place of articulation, and a separate tier with the phonological information unique to each mutation. Although Lieber (1983) has attempted such an analysis, she has done so only for the SOFT mutation, and it is unclear how the analysis could be extended to account for the ASPIRATE mutation as well. The problem is that the mutations in Welsh do not have consistent phonological characteristics, as they do in Fula. That is, it is impossible to associate the mutations with any unique set of phonological features. While it is the case that the BASE, SOFT, and ASPIRATE mutations differ either in voicing or in degree of oral closure, these differences are not associated one-to-one with any single mutation. Thus, if the BASE form contains a voiceless stop, it will differ from the SOFT mutation in terms of voicing (e.g., /p/ vs /b/), and from the ASPIRATE mutation in terms of degree of closure (e.g., /p/ vs /f/). But if the BASE form contains a voiced stop, it will be the SOFT mutation, rather than the ASPIRATE, from which it differs in terms of degree of closure (e.g., /b/ vs /v/); and the BASE and ASPIRATE mutations do not differ at all. The difficulties posed by these relationships have been clear to those who have tried to describe them by means of more traditional generative phonological rules. Awbery (1973), for example, notes that two completely unrelated rules are required to derive the SOFT mutation forms from the BASE forms (one for voiceless stops and liquids, and another for voiced stops), while the ASPIRATE mutation requires a rule that, in terms of its phonological effects (changing a stop to a fricative), ought to be collapsible with one of the SOFT mutation rules. But because the two stop-to-fricative rules apply in different morphological and syntactic environments, they cannot be collapsed. Thus, the phonological differences between these mutations are impossible to characterize in a simple and general way, at least using ordinary phonetic features.

However, it is possible to unify the descriptions of the phonological effects of the various mutations by considering them to be dynamic processes, as proposed by Griffen (1985). Griffen suggests that the mutations can be considered to be dynamic processes operating on a single scale of strength, where voiced fricatives are the weakest (designated by the value 1), voiced stops next (with a value of 2), voiceless stops next (3), and voiceless fricatives the strongest (4). (Acoustically, the strength scale corresponds to an increasing ratio of high-to-low frequency energy, according to

Griffen, 1975.) For the labial place of articulation, this results in the following scale: (1) /v/, (2) /b/, (3) /p/, (4) /f/. In Griffen's analysis, the membership of each lexical item in the mutation system is defined in terms of the place of articulation of the beginning portion, and the strength of the BASE form. Starting from this classification of lexical items, the mutations are simply related in terms of their degrees of strength: the SOFT mutation is one degree weaker than the BASE form, and the ASPIRATE mutation one degree stronger. Thus, a lexical item whose BASE form begins with /p/ is defined as a labial with strength 3. Its corresponding SOFT mutation form would then have strength 2 (/b/) and the ASPIRATE mutation strength 4 (/f/). Similarly, a lexical item whose BASE form begins with /b/ is a labial with strength 2; its SOFT mutation form has strength 1 (/v/). The ASPIRATE mutation here has the same strength as the BASE, resulting in a gap in the application of the strength relations such that the ASPIRATE mutation only differs from the BASE if the BASE has strength 3.

While the notion of the strength scale has gaps in its application, it nevertheless provides a description of the phonological changes associated with the mutation classes that is simple enough to be used in an autosegmental account. This is true not only for the similarities and differences among the mutations and segments immediately relevant to this paper, but also, as Griffen (1985) shows, for the other mutations and segments in Welsh. And the strength scale is not simply an arbitrary scale. As noted above, it may reflect the acoustics of the segments. Moreover, increase and decrease in strength has been observed to be a frequently occurring historical process. This is true for Welsh, where the modern mutations are the fossilized remains of an earlier process that was quite phonologically regular (Jones, 1931). And changes in the strength of consonants have been observed for a number of languages, especially those in the Indo-European language family (Lass, 1984).

Thus, Griffen's strength scale appears to be well motivated, as well as being well suited to an autosegmental analysis of the phonological structure of the Welsh mutation classes. In such an analysis, the lexical item includes the underspecified phonological representation (with only place of articulation specified in the initial portion) and its link to the appropriate value on the separate strength tier (3 for voiceless BASES, 2 for voiced BASES). In addition, the mutation classes are associated with operations on the strength tier (adding values of 0, +1, and −1, respectively, for the BASE, ASPIRATE, and SOFT mutations). The mutation classes operate on the strength values associated with the lexical item to determine the correct strength for that morphophonological variant, and hence, the complete phonological representation. Certain restrictions are necessary to capture the gaps in the mutations, as well as to handle the nonmutating segments. Thus, the mutation classes will only change lexical strength values of 2 or 3, thereby exempting the nonmutating segments being considered in this paper, whose lexical item will be linked to the value of 4 on the strength tier. And the ASPIRATE mutation is further restricted to operate only on lexical items of strength 3.

Note that the phonological realization of the mutation classes with respect to the strength tier requires an operation, that of additivity. Current notions of the relationships among autosegmental tiers include linking between specific items, blocking of such links, removal of links, and reassignment of links, but no arithmetic operations are posited. Thus, the use of operations such as addition for the combination of two (scalar) feature values represents an extension to autosegmental theory. This extension in fact brings the proposed autosegmental model of lexical organization quite close to a class of models currently being actively developed by psychologists, biologists, and computer scientists—parallel distributed processing or "connectionist" models (cf. Rumelhart & McClelland, 1986a; McClelland & Rumelhart, 1986). The image of a word suggested in autosegmental analyses—that of a set of links among nodes on independent tiers—is quite similar to the connectionist image of a network of interconnections among layers of simple, but dynamic, processing units. Each unit has some state of activation, which is influenced over time by the inputs from connected units.

In spite of this apparent surface convergence between the autosegmental and connectionist approaches, there is nevertheless a difference in underlying assumptions. The autosegmental approach assumes that the rules that combine information from the various tiers, for example, are not simply descriptive, but are also the mechanism by which the system operates. Connectionist models, in contrast, assume that rules serve to describe the behavior of a system but are not the mechanism whereby that behavior is achieved. Thus, as stated by Rumelhart and McClelland (1986b, p. 218), "[t]here is no denying that rules . . . provide a fairly close characterization of . . . performance. . . . We would only suggest that parallel distributed processing models may provide a mechanism sufficient to capture lawful behavior, without requiring the postulation of explicit but inaccessible rules. Put succinctly, our claim is that PDP [Parallel Distributed Processing] models provide an alternative to the explicit but inaccessible rules account of implicit knowledge of rules."

Rumelhart and McClelland (1986b) support the above claim by developing a connectionist model that learns both regular and irregular past tenses of English verbs, simply through repeated presentations of present and past tense forms. The system gradually develops a set of weights associating the nodes corresponding to phonological units of present tense forms with the nodes corresponding to past tense forms. In the course of learning, the system exhibits the various stages shown by children learning past tenses. That is, first a small set of verbs is learned with the correct past tenses, regardless of whether the past tenses are regular or irregular. In the second stage, the regular past tense is overgeneralized so that irregular verbs no longer have the correct past tense. In the final stage, a much larger set of verbs, with both regular and irregular past tenses, is learned correctly. What is interesting about this model in the current context is that even though the system "learns" regularities and can generalize, this knowledge resides only in the connections among the present and past tense pairs. The ability to generalize results from the similarity in the present–past associations for the multitude of regular verbs. That is, regularity consists of the same pattern being repeated in association after association—exactly the same situation as occurs in the Welsh mutation system. The exact nature of the pattern does not matter, only its regular occurrence.

From this point of view, the distinction between an allomorphic approach and an underspecified or autosegmental approach to characterizing lexical items may be more apparent than real. That is, the connectionist model associates fully specified allomorphs; as long as these allomorphs are specified using phonological features, it does not matter how the features associated with the morphological information are distributed throughout the lexical items. They can be completely isolatable in one portion of the lexical item, as an affix; or they can be only partially isolatable, as in the case of the mutations. The connectionist model should handle both these cases with equal ease. A similar conclusion is reached by Bybee (1985), who argues that these alternatives are endpoints on a continuum. For her, the lexicon consists of sets of dynamically linked allomorphs, each with variable lexical strengths (related to their frequency of occurrence: cf. also

Stemberger & MacWhinney, 1986) and variable degree of connection to other morphologically related forms (at least partially as a function of phonological similarity).

This kind of connectionist model accommodates the basic similarities and differences between the English and Welsh morphological systems. Those patterns that recur in each language, regardless of their phonological nature, will result in generalizations based on their regularity. The specific nature of the generalization—affixal vs mutation—will of course differ between the two languages. Nevertheless, certain details of the Rumelhart and McClelland (1986b) connectionist model for English are not appropriate for the facts of Welsh. In particular, the lack of a significant effect on priming due to phonological similarity (i.e., identity) among the mutations suggests that a simple connection via phonological feature nodes is not sufficient to characterize the connections among Welsh morphological variants. It might be possible to partially reconcile the Rumelhart and McClelland model with the Welsh data by positing a phonological identify effect for the first consonant that is so small as to be overwhelmed by the much larger effects due to the rest of the word and the additional differential recognizability of the individual morphological classes. Nevertheless, because the effect of the mutations remains strong regardless of their phonological realization, it is necessary to include an independent morphological level. This might be handled by additional input units or, possibly, by explicit "hidden units." The additional level in the Welsh lexicon might be capable of accounting for the fact, discussed above, that Welsh and English differ when comparing performance in the identical-priming and other-priming conditions. (However, note that a morphological level independent of phonological structure has also been suggested for English—cf. Fowler et al.'s (1985) suggested modification to Dell's (1984) network model.) In general, then, while it remains an appealing possibility that a connectionist model could be constructed that captures both the similarities and differences between English and Welsh, the possibility cannot be satisfactorily confirmed without the actual construction of an explicit model for testing.

To summarize, the two-dimensional model of lexical organization we have proposed, augmented with an autosegmental representation of the associated phonological structures, seems well suited to describing the kinds of lexical relations that develop in the case of regular nonconcatenative morphology like the Welsh mutations. Moreover, connectionist models offer the possibility of understanding the mechanism that gives rise to this kind of system behavior. By investigating languages with different morphophonological structures, it is possible to increase our understanding of the limits on the complexity of the relations among (semi)autonomous linguistic structures (i.e., phonological, morphological, syntactic, semantic). The present paper has provided one attempt to extend our understanding of these limits.

Appendix I: Stimuli

The stimuli are listed for each experiment in Welsh orthography, followed by their English gloss. Note that *w* and *y* represent vowels, pronounced roughly as in FOOL and FIN, *ch* represents a voiceless velar fricative, as in German *Bach*, *f* represents the English voiced labiodental fricative as in VENT, the double *ff* is the English voiceless labiodental fricative as in FOOL, the double *ll* is a voiceless lateral fricative, while the double *dd* is a voiced alveolar fricative as in English THAT.

Monosyllables

Experiment 1	Experiment 2	Experiment 3
beic "bicycle"	bedd "grave"	bwyd "food"
brawd "brother"	bocs "box"	budd "profit"
dial "revenge"	dwrn "fist"	gair "word"
drych "mirror"	dyn "man"	glyn "glen"
gwg "frown"	drwg "evil"	pwys "weight"
gwallt "hair"	darn "piece"	pôt "pot"
pen "head"	gwin "wine"	cwm "valley"
plyg "fold"	glaw "rain"	crys "shirt"
trwyn "nose"	pris "price"	
tân "fire"	pwrs "purse"	
cig "meat"	tolc "dent"	
caws "cheese"	taid "grandfather"	

trên "train"
tir "land"
cais "attempt"
cwsg "sleep"

BISYLLABLES

Experiment 1	Experiment 2
bwrlwm "bubble"	bachgen "boy"
bocsach "boast"	bywyd "life"
dillad "clothing"	diwrnod "day"
deintydd "dentist"	darllun "picture"
gorwel "horizon"	dyddiad "date"
gobaith "hope"	dosbarth "class"
pellter "distance"	golau "light"
priddfain "brick"	gelyn "enemy"
talcen "forehead"	plentyn "child"
tywydd "weather"	peiriant "engine"
colwyn "puppy"	toriad "cut"
ceffyl "horse"	tafod "tongue"
	tebot "teapot"
	tegan "toy"
	canol "center"
	capel "chapel"

NONMUTATING WORDS

Experiment 3
ffŵl "fool"
ffrind "friend"
ffug "fiction"
fflach "flash"
chwant "desire"
chwys "sweat"
chwart "quart"
chwyth "breath"

APPENDIX II: NOISE LEVELS

The pretest consisted of 30 words (none of which were used in the experiment) set in the same BASE, SOFT MUTATION, and ASPIRATE MUTATION frames used in the experiment. The words were played to the subjects through a two-channel amplifier whose second channel could be attenuated in 2-db steps. The first channel (the noise) was mixed with the second channel (the signal words) and the output was directed to the subject via stereo headphones. Beginning with the most intelligible of the signal-to-noise levels, the channel containing the signal words was attenuated by one 2-dB turn of the dial every six items, a total of five different settings ranging in −2-dB steps from +2 to −6 dB for monosyllables. The range for bisyllables was 3 dB lower, from −1 to −9 dB, to control for the higher recognition rate of bisyllables found in pilot studies (see Note 1). The subject's noise level for the target portion of the experiment was that noise level in which the subject had last guessed correctly one or more of the items within a six-item group. The distribution of the different noise levels over subjects, and the effects of noise levels on recognition rates, are given in Note 2.

Note 1. The major problem for the technique of word recognition in noise is the possibility that a too-easy or too-difficult noise level for a particular combination of word and subject will swamp out any experimental effects due to priming for that combination. The ideal solution to this problem, of course, is extensive pretesting of individual subjects and individual words to establish performance under different noise conditions. In such a case, a different noise level can be chosen for each combination of word and subject that maximizes the likelihood that priming can be detected. Unfortunately, such testing is often not practical. Instead, investigators (Kempley & Morton, 1982; Magen & Manuel, 1982) have concentrated on choosing a single noise level for the experimental corpus that is free of ceiling and floor effects for a significant proportion of words. This can be done in several ways. In Kempley and Morton (1982) a pilot group of subjects was used to establish a 40% average recognition rate for the experimental corpus. In Magen and Manuel (an experiment on German) a pilot group of native and nonnative speakers was tested on a sample corpus to locate a 40% average recognition rate. In our case, not being able to test the entire experimental corpus in New Haven due to the lack of native speakers, we drew on previous experience with small Welsh pilot studies on different subjects and different sets of words (some of which occur also in the experimental corpus) to estimate the average effects of various signal-to-noise ratios. The anchor level, that is, the 50% recognition level, was determined to be (approximately) at signal-to-noise equivalence, or a signal-to-noise ratio of 1 dB for monosyllabic words and at −3 dB for bisyllabic words. To control for this apparent general effect of redundancy (due to the increased word length), during the target phase of the experiment bisyllables were presented at a noise level 3 dB greater than that for monosyllables.

Note 2. As mentioned previously, the noise level at which a subject heard the target tape was individually calibrated to each subject by means of a pretest. As it happened, subject performance on the pretest was an indifferent predictor of performance on the target list; in general, the subjects assigned the more difficult signal/noise levels had proportionately lower rates of recognition. Approximately half (30) of the subjects received a noise level of −4 dB for monosyllables and −7 dB for bisyllables, with the others normally distributed over remaining levels. The mean rate of recognition for all subjects, over all three experiments, was 46%. To check for the possibility that ceiling and floor effects influenced the pattern of the data reported here, the statistical analysis of each experiment was repeated using three subjects assigned noise level

−4/−7 (monosyllables/bisyllables) from each subject group. Because for each experiment the results were the same as when the entire subject pool was used, but the reduced number of subjects in a subject group meant a loss of power, the results reported in the text are for the full complement of 56 subjects.

APPENDIX III: ESTIMATED FREQUENCIES OF STIMULI

Item	Mean	Mutation			Experiment	Type
		BASE	SOFT	ASP.		
bedd	57	57	57	56	2	Mono
bocs	64	72	67	54	2	Mono
beic	66	65	—	68	1	Mono
brawd	73	72	74	—	1	Mono
bwyd	80	86	67	80	3	Mono
budd	43	47	47	32	3	Mono
dwrn	53	52	57	50	2	Mono
dyn	69	83	58	68	2	Mono
drwg	64	70	64	57	2	Mono
darn	59	73	53	54	2	Mono
dial	40	46	—	36	1	Mono
drych	45	48	41	—	1	Mono
gwin	58	68	53	52	2	Mono
glaw	68	78	76	48	2	Mono
gwg	25	28	—	21	1	Mono
gwallt	70	68	71	—	1	Mono
gair	65	68	67	59	3	Mono
glyn	34	35	40	25	3	Mono
pris	65	80	55	58	2	Mono
pwrs	64	71	57	65	2	Mono
pen	77	78	—	77	1	Mono
plyg	27	29	23	—	1	Mono
pwys	43	51	29	44	3	Mono
pot	57	67	49	42	3	Mono
tolc	34	50	20	30	2	Mono
taid	68	63	64	76	2	Mono
tir	66	76	64	56	2	Mono
tren	58	68	52	55	2	Mono
trwyn	73	72	—	74	1	Mono
tan	68	73	63	—	1	Mono
cais	45	48	45	42	2	Mono
cwsg	53	58	57	46	2	Mono
cig	69	74	—	65	1	Mono
caws	67	76	61	—	1	Mono
cwm	56	63	43	54	3	Mono
crys	75	72	72	82	3	Mono
bachgen	72	73	74	70	2	Bi
bywyd	70	68	66	76	2	Bi
bwrlwm	29	32	—	28	1	Bi
bocsach	18	19	17	—	1	Bi
diwrnod	74	83	70	68	2	Bi
darllun	61	69	60	56	2	Bi
dyddiad	58	69	48	57	2	Bi
dosbarth	64	70	68	52	2	Bi
dillad	81	83	—	80	1	Bi
deintydd	50	56	46	—	1	Bi
golau	67	74	68	60	2	Bi
gelyn	56	59	56	54	2	Bi

gorwel	35	38	—	31	1	Bi
gobaith	55	52	59	—	1	Bi
plentyn	72	73	75	68	2	Bi
peiriant	55	54	58	52	2	Bi
pellter	48	60	—	40	1	Bi
priddfain	15	15	15	—	1	Bi
toriad	38	48	38	29	2	Bi
tafod	64	62	75	53	2	Bi
tebot	66	65	61	73	2	Bi
tegan	51	50	55	48	2	Bi
talcen	54	53	—	55	1	Bi
tywydd	69	84	59	—	1	Bi
canol	64	67	64	60	2	Bi
capel	65	68	62	65	2	Bi
colwyn	14	15	—	14	1	Bi
ceffyl	61	64	58	—	1	Bi
ffwl	50	66	40	28	3	Nonmutating
ffrind	77	73	78	83	3	Nonmutating
ffug	30	28	26	40	3	Nonmutating
fflach	36	38	38	30	3	Nonmutating
chwant	48	52	49	37	3	Nonmutating
chwys	53	54	58	47	3	Nonmutating
chwart	34	31	45	28	3	Nonmutating
chwyth	24	22	20	32	3	Nonmutating
Mean	55	59	54	52		
SD	16	18	16	17		
Min	14	15	15	14		
Max	81	86	78	83		

MEAN FREQUENCIES BY EXPERIMENT AND TYPE

Experiment 1		Experiment 2		Experiment 3	
Mono	Bi	Mono	Bi	Mono	Nonmutating
58	44	59	62	57	44

Reliability of frequency estimates. The frequency estimates can only be considered to be approximations; the values in the table should be interpreted as an indication of, at best, the nearest decile. This is due both to measurement error and to the nature of the subject responses.

Measurement error for the frequency responses of individual subjects is at least ±5%, due to the experimenters' decision to round subjects' responses to the nearest 5. Additional possible sources of measurement error include lack of double checking during data entry into the computer (unlike the experiment proper, in which responses were double or triple checked, as necessary), and possible inconsistency in rounding decisions between the two individuals entering the responses.

In addition, the subject responses varied in accuracy. For some of the subjects, accuracy is ±20%; these subjects circled only the labeled frequency values on the answer sheets (frequency values of 0, 20, 40, 60, 80, and 100 were explicitly labeled). For other subjects, accuracy is ±10%; these subjects circled the labeled values and also the intermediate bisecting tick marks. A few subjects resorted to effectively binary choices (0 or 100); the rest of the subjects used the intermediate values, as expected.

As a rough measure of frequency, however, we believe these data may be fairly reliable, for several reasons. First, the presentation of the items in different mutation environments serves as a kind of replication; the relatively high correlations among the frequency estimates of the same item in different mutation environments (reported below) suggest that the estimates are at least fairly replicable. Second, Goldstein and Van den Broecke (1977) have shown that estimates of phoneme frequency correlate well with other measures of phoneme frequency, in English. Finally, anecdotal reports from subjects at the time of running the experiments were consistent with the frequency estimates for the low frequency items in the list.

For instance, all of the words with an overall score of 25 or less were spontaneously noticed by at least one subject as being odd or unnatural. *Priddfain* "brick," for example, was remarked on as a word everyone knows from church but that is never used in conversation. (It may be worth noting that the North Wales dialect speaker we consulted in New Haven kept up her Welsh by attending a Welsh chapel.) Subjects also revealed that, for them, the primary denotation of *colwyn* "puppy" was the nearby Colwyn Bay. *Bocsach* "boast" and *gwg* "frown" were apparently somewhat obscure, as several subjects claimed to have never encountered them.

Subject comments also revealed that five more words—*chwant* "desire," *tegan* "toy," *taid* "grandfather," *bachgen* "boy," and *cwm* "valley"—have different frequencies in North and South Wales, *chwant* and *bachgen* being more common in South Wales, and *taid, tegan,* and *cwm* being used predominantly in North Wales. The estimated frequencies for these words therefore reflect the different experiences of North and South Wales dialect speakers among our subjects. Note, however, that we had approximately equal numbers of subjects from North and South Wales. Note also that, although the experiment took place in North Wales, the subjects were all familiar with the literary language, which incorporates aspects of both dialects.

The high frequency words in the list were also consistent with expectation. For instance, words with the highest frequency scores were common words for common objects or concepts: *bwyd* "food," *crys* "shirt," *dillad* "clothing," and *ffrind* "friend." Words such as *brawd* "brother," *dyn* "man," and *diwrnod* "day," which are expected to be frequent, have high scores also.

EXPERIMENT and TYPE effects. An analysis of variance was performed using the mean estimated frequency for each item, where the mean was the average across subjects and mutation classes (as reported in the leftmost numerical column of the above table). The grouping factors of EXPERIMENT (1, 2, or 3) and TYPE (MONO, BI, or NONMUTATING) were tested. Both EXPERIMENT ($p < .02$) and EXPERIMENT × TYPE ($p < .04$) were marginally significant; TYPE was not significant. Tests of simple main effects indicated that the EXPERIMENT effect was confined to a significant difference between the BIsyllables in Experiments 1 and 2, with the BIsyllables in Experiment 1 being significantly lower in frequency than those in Experiment 2 ($p < .006$). No other simple main effects were significant.

Although TYPE was not significant in the overall analysis, nevertheless the NONMUTATING items in Experiment 3 were markedly lower in frequency than the MONOsyllabic items. Since this contrast was an important one, we felt justified in further tests of significance. TYPE remained nonsignificant when tested in Experiment 3 alone. When MONO and BI from all the experiments were grouped together and compared (without EXPERIMENT as a grouping factor) to the NONMUTATING items, thereby increasing the degrees of freedom, the NONMUTATING items were marginally significantly lower than the grouped other items ($p < .04$).

MUTATION effects. Here we used the estimated frequency averaged across subjects (as reported in the second through fourth numerical columns in the above table). The estimated frequencies for the three MUTATIONS, BASE, SOFT, and ASPIRATE, were fairly highly correlated: for BASE with SOFT, $r = .81$; BASE with ASPIRATE, $r = .81$; and SOFT with ASPIRATE, $r = .75$. The analyses of variance detailed below revealed that the items were judged to be significantly higher in frequency when they occurred in the BASE form than when they occurred in either of the other MUTATIONS; there was no significant difference in the estimated frequencies for items occurring in the SOFT and ASPIRATE MUTATIONS.

Using subsets of EXPERIMENT (1, 2, or 3) and TYPE (MONO, BI, or NONMUTATING) as grouping factors and MUTATION (BASE, SOFT, or ASPIRATE) as within-group factors, analyses were run to provide a rough indication of the significance of the differences of the frequency estimates for items in the different mutations. Since not all combinations of the grouping factors occurred in all the experiments, the analyses were run both on subsets of EXPERIMENTS and on subsets of MUTATIONS. No main effects, other than that for MUTATION, and no interactions were significant in any of the analyses reported below.

For EXPERIMENT levels 2 and 3, using all levels of MUTATION, MUTATION was significant ($p < .0001$). Further analyses revealed that the MUTATION effect was confined to the significantly higher estimated frequencies for items occurring in the BASE, compared to either of the other MUTATIONS. Thus, for all levels of EXPERIMENT (1, 2, and 3), items were estimated to be significantly higher in frequency when they occurred in the BASE than in the SOFT MUTATION ($p < .0058$). Again for all levels of EXPERIMENT, items were estimated to be significantly higher in frequency when they occurred in the BASE than in the ASPIRATE MUTATION ($p < .0003$).

For Experiments 2 and 3, frequency estimates in the SOFT and ASPIRATE MUTATIONS were not significantly different. For Experiment 1, the SOFT and ASPIRATE MUTATIONS could not be directly compared since each lexical item occurred with only one of the two MUTATIONS. Thus, the SOFT and ASPIRATE MUTATION frequencies were averaged for each lexical item in Experiments 2 and 3; for Experiment 1, the "average" consisted of the single frequency for whichever of the two MUTATIONS occurred with that lexical item. These average values were contrasted with the frequencies in the BASE for all levels of EXPERIMENT. The frequencies in the BASE were significantly higher than the averaged frequencies of the two MUTATIONS ($p < .0001$).

REFERENCES

ANDERSON, S. R. (1982). Where's morphology? *Linguistic Inquiry*, 13, 571–612.

AWBERY, G. M. (1973). *Initial mutation in a generative grammar of Welsh* (Report No. 4, pp. 28–46.). Leeds: University of Leeds, Phonetics Department.

AWBERY, G. M. (1986). Moves towards a simpler, binary mutation system in Welsh. In H. Andersen (Ed.), *Sandhi phenomena in the languages of Europe*. The Hague: Mouton.

BENTIN, S., BARGAI, N., & KATZ, L. (1984). Orthographic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 353–368.

BOYCE, S. E. (1983). A perceptual analog of change in progress in Welsh. *Haskins Laboratories Status Report on Speech Research*, SR-76, 159–165.

British Office of Population Censuses and Surveys. (1983). *Census 1981: Report for Wales*. London: Her Majesty's Stationery Office.

BROWMAN, C. P. (1978). Tip of the tongue and slips of the ear: Implications for language processing. *UCLA Working Papers in Phonetics*, 42.

BYBEE, J. L. (1985). *Morphology. A study of the relation between form and meaning*. Philadelphia: John Benjamins.

COLE, R. A., & JAKIMIK, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *The perception and production of fluent speech*. Hillsdale, NJ: LEA.

DELL, G. (1984). *A spreading activation theory of retrieval in sentence production* (Cognitive Science Technical Report). Rochester, NY: University of Rochester.

FAY, D., & CUTLER, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry*, 8, 505–520.

FORBACH, G. B., STANNERS, R. F., & HOCHHAUS, L. (1974). Repetition and practice effects in a lexical decision task. *Memory & Cognition*, 2, 337–339.

FORSTER, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland Press.

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

FYNES-CLINTON, O. H. (1913). *The Welsh vocabulary of the Bangor district*. Oxford: Oxford University Press.

GOLDSTEIN, L. M. (1977). Perceptual salience of stressed syllables. *UCLA Working Papers in Phonetics*, 39, 37–60.

GOLDSTEIN, L. M., & VAN DEN BROECKE, M. (1977). Response bias and subjective estimation of consonant frequency. *Journal of the Acoustical Society of America*, 61(Suppl. 1), S65. (Abstract)

GRIFFEN, T. D. (1975). *A new Welsh consonant shift, description and implications*. Doctoral dissertation, University of Florida, Gainesville. (University Microfilms International No. 76-4237.)

GRIFFEN, T. D. (1985). *Aspects of dynamic phonology*. Amsterdam: John Benjamins Publishing Company.

HENDERSON, L., WALLIS, J., & KNIGHT, D. (1984). Morphemic structure and lexical access. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X. Control of language processes*. London: Erlbaum.

HOWES, D. (1957). On the relation between the intelligibility and frequency of occurrence of English words. *Journal of the Acoustical Society of America*, 29, 296–305.

JONES, J. M. (1931). *A Welsh grammar: Historical and comparative. Phonology and accidence*. Oxford: Clarendon Press.

JONES, S. (1926). *A Welsh phonetic reader*. London: University of London Press.

JONES, T. J. R. (1977). *Living Welsh*. Suffolk: Hodder and Stoughton.

KEMPLEY, S. T., & MORTON, J. (1982). The effects of irregularly related words in auditory word recognition. *British Journal of Psychology*, 73, 441–454.

LASS, R. (1984). *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.

LIEBER, R. (1983). New developments in autosegmental morphology: Consonant mutation. In M. Barlow, D. P. Flickinger, & M. T. Wescoat (Eds.), *Proceedings of the West Coast Conference on Formal Linguistics* (No. 2, pp. 165–175). Stanford, CA: Stanford Linguistics Association (Stanford University, Department of Linguistics).

LIEBER, R. (1984). Consonant gradation in Fula: An autosegmental approach. In M. Aronoff & R. T. Oehrle with F. Kelley & B. W. Stephens (Eds.), *Language sound structure* (Chap. 17, pp. 328–345). Cambridge, MA: MIT Press.

LUKATELA, G., GLIGORIJEVIĆ, B., KOSTIĆ, A., & TURVEY, M. T. (1980). Representation of inflected nouns in the internal lexicon. *Memory & Cognition*, 8, 415–423.

MACKAY, D. G. (1979). Lexical insertion, inflection, and derivation: Creative processes in word production. *Journal of Psycholinguistic Research*, 8, 477–498.

MAGEN, H., & MANUEL, S. Y. (1982, December). *Auditory recognition of morphologically related*

German words. Paper presented at the meeting of the Linguistic Society of America.

MANELIS, L., & THARP, D. A. (1977). The processing of affixed words. *Memory & Cognition*, 5, 690–695.

MARANTZ, A. (1982). *Re reduplication*. *Linguistic Inquiry*, 13, 435–482.

MARSLEN-WILSON, W. (1984). Function and process in spoken word recognition: A tutorial review. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X. Control of language processes* (pp. 125–150). London: Erlbaum.

MARSLEN-WILSON, W. D., & WELSH. A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29–63.

MATTHEWS, P. (1974). *Morphology*. Cambridge: Cambridge University Press.

MCCARTHY, J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, 12, 373–418.

MCCARTHY, J. (1984). Prosodic organization in morphology. In M. Aronoff & R. T. Oehrle with F. Kelley & B. W. Stephens (Eds.), *Language sound structure* (pp. 299–317). Cambridge, MA: The MIT Press.

MCCLELLAND, J. L., & RUMELHART, D. E. (1986). A distributed model of human learning and memory. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 2. *Psychological and biological models* (pp. 170–215). Cambridge, MA: The MIT Press.

MEYER, D. E., & SCHVANEVELDT, R. W. (1976). Meaning, memory structure, and mental process. In C. N. Cofer (Ed.), *The structure of human memory*. San Francisco: W. H. Freeman.

MILLER, G. A., & NICELY, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338–352.

MORTON, J. (1969). The interaction of information in word recognition. *Psychological Review*, 76, 165–178.

MORTON, J. (1970). A functional model for memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press.

MORTON, J. (1979). Word recognition. In J. Morton & J. C. Marshall (Eds.), *Psycholinguistics 2: Structure and processes* (pp. 107–156). Cambridge, MA: MIT Press.

MURRELL, G. A., & MORTON, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology*, 102, 963–968.

POLLACK, I. (1963). Reaction times to unknown word sets in noise. *Language and Speech*, 6, 189–195.

RUBENSTEIN, H., & POLLACK, I. (1963). Word predictability and intelligibility. *Journal of Verbal Learning and Verbal Behavior*, 2, 147–158.

RUMELHART, D. E., & MCCLELLAND, J. L. (1986a). PDP models and general issues in cognitive science. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 1. *Foundations* (pp. 110–146). Cambridge, MA: The MIT Press.

RUMELHART, D. E., & MCCLELLAND, J. L. (1986b). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 2. *Psychological and biological models* (pp. 216–271). Cambridge, MA: The MIT Press.

SADOCK, J. M. (1985). Autolexical syntax: A proposal for the treatment of noun incorporation and similar phenomena. *Natural Language & Linguistic Theory*, 3, 379–439.

STANNERS, R. F., NEISER, J. J., HERNON, W. P., & HALL, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, 18, 399–412.

STEMBERGER, J. P., & MACWHINNEY, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition*, 14, 17–26.

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

TAFT, M., & FORSTER, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638–647.

VAN DER MOLEN, H., & MORTON, J. (1979). Remembering plurals: Unit of coding and form of coding during serial recall. *Cognition*, 7, 35–47.

WILLIS, P. (1986). *The initial consonant mutations in Welsh and Breton*. Bloomington, IN: Indiana University Linguistics Club.

(Received August 5, 1986)
(Revision received February 9, 1987)