Language Variation and Change, **14** (2002), 261–290. Printed in the U.S.A.

DOI: 10.1017/S0954394502143018

Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change

JOAN BYBEE
University of New Mexico

ABSTRACT

The literature on frequency effects in lexical diffusion shows that even phonetically gradual changes that in some cases are destined to be lexically regular show lexical diffusion while they are in progress. Change that is both phonetically and lexically gradual presents a serious challenge to theories with phonemic underlying forms. An alternate exemplar model that can account for lexical variation in phonetic detail is outlined here. This model predicts that the frequency with which words are used in the contexts for change will affect how readily the word undergoes a change in progress. This prediction is tested on data from /t, d/ deletion in American English. Finally, the effect of bound morphemes on the diffusion of a sound change is examined. The data suggest that instances of a bound morpheme can affect the rate of change for that morpheme overall.

Lexical diffusion refers to the way that a sound change affects the lexicon. If sound change is lexically abrupt, all the words of a language are affected by the sound change at the same rate. If a sound change is lexically gradual, individual words undergo the change at different rates or different times. Whether sound changes exhibit gradual or abrupt lexical diffusion is a topic of some recent concern. One early contribution to this debate by Schuchardt (1885) is the observation that high-frequency words are affected by sound change earlier and to a greater extent than low-frequency words. In this article, after discussing the recent literature on lexical diffusion, I document this tendency in a variety of sound change types, discuss briefly the opposite tendency (i.e., for low-frequency words to be affected by a change first), and propose explanations for both types of change. I argue that phonetically conditioned changes that affect high-frequency words before low-frequency words are best accounted for in an exemplar model of phonological representation that allows for change to be both phonetically and lexically gradual. Next I show that a word's contexts of use also affect the rate of change. Words that occur more often in the context for change change more rapidly than those that occur less often in that context. The data for this exploration are from English /t, d/ deletion and involve contracted negation, words end-

For help analyzing the data on /t, d/ deletion, I would like to thank Greg Thomson and Keri Holley; for coding and analyzing the data on Spanish [ð] deletion, I would like to thank Dawn Nordquist, Rena Torres-Cacoullos, and Gabriel Waters. Thanks also for the suggestions of the anonymous reviewers for the journal. Parts of this article were published in Bybee (2002).

ing in unstressed *-nt*, and the regular past tense morpheme. The contexts of use for bound morphemes are examined, using the data from the deletion of [ð] in New Mexican Spanish, to show that sound changes can also progress more rapidly in high frequency morphemes. The role of morphemes in lexical diffusion is examined further by comparing American English /t, d/ deletion, in which a change is retarded in a high frequency morpheme, to Spanish [ð] deletion. It is proposed that the contexts of use determine the rate at which a word or morpheme undergoes a sound change.

This article only addresses word frequency as a conditioning factor in phonological variation and assumes that the variation remaining after frequency is considered is due to other well-established social and linguistic factors. My goal is to demonstrate that word frequency and context frequency are factors that can affect variation and should be taken into account in future studies of phonological variation and change.

REGULAR SOUND CHANGE OR LEXICAL DIFFUSION?

The hypothesis that sound change is lexically regular seems well supported by the facts of change. When we observe that two languages or dialects exhibit a phonological difference, it is very likely that this difference is regular across all the words that have the appropriate phonetic environment. This observation is fundamental to the comparative method; the establishment of genetic relations and the reconstruction of protolanguages are based on the premise that sound change affects all words equally. Schuchardt (1885) was one of the detractors from this position. When he observed sound change in progress, he noted that all words did not change at the same rate, and that the differences were not due to “dialect mixture,” as was often claimed by the Neogrammarians, who supported the regularity position.

A major challenge to the regularity position in the 20th century is expressed in the work of Wang and his colleagues (Wang, 1969, 1977; Wang & Cheng, 1977), which documented changes that seem to occur word by word over a long period of time. While some of these changes result in lexical regularity, Wang and his colleagues also identified changes that appear to be arrested after affecting only part of the lexicon.

Labov (1981, 1994) also dealt with the issue, availing himself of the data from his numerous studies of sound change in progress. He proposed two types of sound change: “regular sound change” is gradual, phonetically motivated, without lexical or grammatical conditioning, and not influenced by social awareness, whereas “lexical diffusion” change, such as the phenomena studied by Wang, is “the result of the abrupt substitution of one phoneme for another in words that contain that phoneme” (Labov, 1994:542). According to Labov, this type of change occurs most often “in the late stages of internal change that has been differentiated by lexical and grammatical conditioning” (ibid.). Labov went so far as to propose that certain changes, such as the deletion of glides and schwa, would be regular changes, while the deletion of obstruents would show lexical diffusion.

A number of researchers have challenged this position. Phillips (1984) argued that even low-level sound changes exhibit gradual lexical diffusion. Similarly, Oliveira (1991) argued that it is likely that gradual lexical diffusion occurs even in changes that turn out to be regular. Krishnamurti (1998) demonstrated that the change of $s > h > \emptyset$ in Gondi exhibits gradual lexical diffusion, but still goes through to completion in some dialects. In this review I present evidence to show that even gradual, phonetically conditioned change exhibits gradual lexical diffusion, though it is perhaps of a more subtle nature than the lexical diffusion studied by Wang and Labov. The lexical diffusion examined here for reductive phonetic change is highly conditioned by word frequency.

Hooper (1976) identified a lexical diffusion paradox. Reductive sound change tends to affect high frequency words before low-frequency words, but analogical leveling or regularization tends to affect low-frequency words before high-frequency words. Working from this observation, Phillips (1984, 2001) studied a number of changes that move in each direction and attempted to refine a hypothesis that predicts the direction of lexical diffusion. In this article I focus on the first direction of change, the characteristic of reductive sound change, with the intent of demonstrating that word frequency affects even changes that may have seemed regular to the Neogrammarians. The second type of change, which affects low-frequency words earlier than high-frequency words, is discussed only briefly.

FREQUENCY EFFECTS ON REGULAR SOUND CHANGE

Sound changes that are complete can be identified as regular or not, depending upon whether they affected all lexical items existing at the time of the change. Ongoing changes cannot be designated as regular or not, since they are not complete. However, one can reference the typical characteristics of a change to project whether it will be regular or not. That is, a phonetically gradual change with clear phonetic conditioning falls into Labov’s first type, and thus we can project its regularity. Here I document lexical diffusion from high-frequency to low-frequency words in ongoing changes that can be expected to be regular, as well as in certain reductive changes that may never be complete because of the nature of the lexicon.[1]

American English /t, d/ *deletion*

Consider the deletion of final /t/ and /d/ in American English, which occurs most commonly in words ending in a consonant plus /t/ or /d/, such as *just*, *perfect*, *child*, or *grand*. This much-studied variable process has been shown to be affected by the preceding and following consonant, with more deletion in a consonant environment; by grammatical status, with less deletion if the /t/ or /d/ is the regular past tense; and by social and age factors, with more deletion among younger, lower socioeconomic class speakers (Labov, 1972; Neu, 1980).

TABLE 1. *Rate of /t, d/ deletion for entire corpus by word frequency*[a]

	Number of Words with High Frequency	Number of Words with Low Frequency
Retention	752 (45.6%)	262 (65.7%)
Deletion	898 (54.4%)	137 (34.3%)

$\chi^2 = 41.67$, $p < .001$
[a]High-frequency words occurred 35 or more times per million in Francis and Kučera (1982); low-frequency words occurred less than 35 times per million.

My study (Bybee, 2000b) of the deletion of /t/ and /d/, using a corpus of phonological variation in Chicano English speakers in Los Angeles (Santa Ana, 1991), focused on lexical frequency as a factor.[2] Using approximately 2,000 tokens of final /t/ and /d/ after a consonant, as transcribed by Santa Ana, and referencing word frequency from Francis and Kučera (1982), I found that deletion occurred more in high-frequency words. Table 1 demonstrates this effect with a cut-off point of 35 per million. This number was chosen because I was also interested in whether a frequency effect could be found among regular past tense verbs, and this figure was the median for such forms in Francis and Kučera (1982).

Similar results were obtained by Jurafsky, Bell, Gregory, and Raymond (2001) using 2,042 monosyllabic content words ending in /t/ or /d/ from the Switchboard corpus, a corpus of telephone conversations between monolingual American English speakers.[3] They found a strong effect of word frequency ($p < .0001$). Another study, using function and content words as well as polysyllabic and monosyllabic words from the Switchboard corpus, found an even higher level of significance for the association of word frequency with final /t/ and /d/ deletion ($p < .00005$) (Gregory, Raymond, Bell, Fosler-Lussier, & Jurafsky, 1999).

Will final /t/ and /d/ deletion after a consonant turn out to be a regular sound change? There is certainly precedent for such a change being regular in the end, especially in certain phonetic contexts. Final consonants in Latin were deleted completely in Spanish and other Romance languages. Latin final /m/ was deleted in all words of two syllables or more (e.g., Latin *quindecim* > Spanish *quince* 'fifteen', *caballum* > *caballo* 'horse', *novem* > *nueve* 'nine'). Latin /t/ was also deleted (*caput* > *cabo* 'end', *amat* > *ama* '3s loves'), as were other consonants (Menéndez-Pidal, 1968). In English, erosion has been working on final consonants for some time. The deletion of /b/ and /g/ after a homorganic nasal, as in *bomb* and *gang*, was completely regular and left English speakers virtually unable to produce final [mb] or [ng] clusters. Final /nd/ could certainly follow and delete regularly as well.

Can the deletion of an obstruent be phonetically gradual? Labov (1994) listed obstruent deletion under "lexical diffusion" changes, presumably because he considered obstruent deletion to involve the phonetically abrupt loss of a phoneme.

TABLE 2. *Rate of deletion for regular past tense compared to all other words of comparable frequency (403 or less)*

	Past Tense	All Other Words
Retention	218 (77.3%)	831 (54.0%)
Deletion	64 (22.6%)	709 (46.0%)

$\chi^2 = 53.2$, $p < .0001$

However, there is evidence that the reduction of final /t/ and /d/ can be gradual. A final /t/ or /d/ may vary in length, and this length variation occurs under the same conditions and in the same direction as the deletion variation. In a laboratory experiment, Losiewicz (1992) found that monomorphemic /t/ or /d/ is shorter in duration than regular past tense /t/ or /d/. As mentioned before, monomorphemic /t/ or /d/ is also more likely to delete. Figures from Bybee (2000b) confirm this, as shown in Table 2. In addition, Losiewicz found that a final past tense /t/ or /d/ is longer in low-frequency verbs than in high-frequency verbs. Her finding for length variation among past tense verbs paralleled the deletion data presented in Bybee (2000b), as shown in Table 3. Regular past tense /t/ and /d/ are more likely to delete in high-frequency verbs than in low-frequency verbs.

The data on this obstruent deletion process, then, suggests both lexical and phonetic gradualness. It thus cannot be said that obstruent deletion is the abrupt deletion of a phoneme. In fact, these data are problematic for any version of phonemic theory. A model that can accommodate these data is presented later in the article.

Spanish [ð] *deletion*

Another good candidate for a phonetically gradual change in progress that exhibits lexical diffusion and that could turn out to be regular is the deletion of intervocalic /d/ or [ð] in many dialects of Spanish. D'Introno and Sosa (1979) regarded the variants ranging from [d] to [ð] to ∅ as a continuum, affirming that the reduction is gradual. My own study of approximately 1,000 tokens of intervocalic /d/ in Spanish spoken by native New Mexican speakers used samples of Spanish-dominant speakers from the corpus of interviews for the Survey of the Spanish of New Mexico and Southern Colorado (Bills, 1997). All medial instances of /d/ were transcribed as present or absent. The results revealed that the rate of deletion is higher among high-frequency words. The frequency count used for Table 4 is taken from the Corpus oral de referencia del español contemporáneo (COREC), a 1.1 million-word corpus of lectures, discussions, and conversations of middle-class Spaniards (Marcos-Marín, 1992). Past participle tokens were removed because the data showed that past participles have a higher rate of deletion than other items. The cut-off point between high and low frequency was

TABLE 3. *Effects of word frequency on* /t, d/ *deletion in regular past tense verbs (non-prevocalic only)*

	High Frequency	Low Frequency
Retention	67 (60.4%)	47 (81.0%)
Deletion	44 (39.6%)	11 (19.0%)

$\chi^2 = 5.00313$, $p < .05$

TABLE 4. *Rate of deletion according to token frequency for all non–past participle tokens in the New Mexican corpus, using the* COREC *as a measure of frequency*

	Low-Frequency Words (0–99)	High-Frequency Words (100+)	Total
Retention	243 (91.4%)	287 (78.6%)	530 (84.0%)
Deletion	23 (8.6%)	78 (21.4%)	101 (16.0%)
Total	266	365	631

$\chi^2 = 17.3$, $p < .001$

arbitrarily chosen to make the number of tokens in the high and low groups approximately equal.

Again we have evidence that obstruent deletion can diffuse gradually through the lexicon, affecting high-frequency words earlier than low-frequency words. We do not know if this sound change will turn out to be completely regular. The earlier deletion of Latin /d/ between vowels left some residue (Menéndez-Pidal, 1968). Thus Latin *credit* gives Spanish *cree* '3s believes', *foedu* > *feo* 'ugly', *pedes* > *pies* 'feet', *fide* > *fe* 'faith', *audire* > *oir* 'to hear', *limpidu* > *limpio* 'clean', and so on, but a few words maintain the /d/, as in *sudare* > *sudar* 'to sweat', *vadu* > *vado* 'ford', *crudu* > *crudo* 'raw', *nidu* > *nido* 'nest', *nudu* > *(des)nudo* 'naked'. However, all of these words are attested without the *d* in Old Spanish, which led Penny (1991) to suggest that the forms with the *d* are influenced by the Latin spelling. Others have argued that the deletion was constrained by whether the resulting vowel combination obeyed the phonotactics of Spanish at the time (see Pensado, 1984). Thus, while the earlier change was not completely regular, it was comparable to the present deletion in being phonetically gradual, and it affected a large majority of words with intervocalic Latin *d*. I suspect the Neogrammarians would have counted it as "regular."

Vowel shifts

Labov (1994) searched his data on vowel shifts in American English for evidence of lexical diffusion. What he was looking for were cases in which words with the

same phonetic environment for the vowel had distinctly different vowel realizations, as opposed to small differences that could be part of a phonetically gradual continuum. The vowel changes reviewed in his Chapter 16 showed detailed phonetic conditioning, gradualness, and lexical regularity. Labov even tested to see if homonyms could split given differences in their token frequency (1994:460–465). This test, on words with very similar phonetic environments, such as *two*, *too*, *to*, *do*, *through*, turned up no lexical diffusion by word frequency. The test, however, was not definitive since all of the words used occurred three or more times in the interview and thus must be considered high frequency. In addition, taking the very small set of words necessary to get comparable phonetic environments yielded too few words to discover any general word frequency effect in these gradual changes. Indeed, it might be that for vowel shifts the phonetic environment is generally more powerful than any other effects because there are fewer words in each phonetic environment.

However, some lexical diffusion was found in Labov's data on vowel shifts. The most famous and still unexplained case concerns the raising of short [æ] in Philadelphia, which affects the adjectives ending in [d]: *mad*, *glad*, *bad*, but not *sad*. This same shift provides some evidence for lexical diffusion by frequency. Labov (1994:506–507) noted that, when word-initial short [æ] "occurs before a voiceless fricative, only the more common, monosyllabic words are tensed: tense *ass* and *ask*; lax *ascot*, *aspirin*, *astronauts*, *aspect*, *athletic*, *after*, *African*, *Afghan*."

Similarly, in Moonwomon's (1992) study of the centralization of /æ/ in San Francisco English, she found that this vowel is more centralized in the environment before a fricative than before a nonfricative; it is also more centralized after [l]. The most commonly used word with this pair of phonetic environments is *class*. *Class* shows more centralization than the other words with these two environments (e.g., *glass*, *laugh*). Moonwomon also studied the fronting of /ɑ/ in the same speakers. Here a following /t/ or /d/ conditions more fronting than other consonants. Of the words in the corpus ending in final /t/, *got* occurred most frequently. Moonwomon found that the fronting in *got* is significantly more advanced than in other words ending in alveolars, such as *not*, *god*, *body*, *forgot*, *pot*.

It appears, then, that there is some evidence that high-frequency words undergo vowel shifts before low-frequency words. As previously mentioned, it may be more difficult to discern frequency effects in vowel shifts because of the effects of the preceding and following environments, which narrow each phonetic class to a small number of words.

Vowel reduction and deletion

In addition to consonant reduction, another type of change that shows robust word frequency effects is vowel reduction and deletion. Fidelholtz (1975:200–201) demonstrated that the essential difference between words that reduce a prestress vowel, such as *astronomy*, *mistake*, *abstain*, and phonetically similar words that do not, such as *gastronomy*, *mistook*, *abstemious*, is word frequency.

Van Bergem (1995) found that reduction of a prestress vowel in Dutch also is highly conditioned by frequency. The high-frequency words *minuut* 'minute', *vakantie* 'vacation', *patat* 'chips' are more likely to have a schwa in the first syllable than the phonetically similar low-frequency words *miniem* 'marginal', *vakante* 'vacant', *patent* 'patent'.

Deletion of reduced vowels may also be conditioned by word frequency. Hooper (1976) asked native speakers of American English for their judgments on whether a post-stress schwa was usually, sometimes, or rarely deleted in words such as *every*, *memory*, and *family*.[4] The results showed that subjects exercised some phonological constraints on schwa deletion, but aside from these, the contexts in which deletion occurred were highly influenced by word frequency. That is, the subjects judged deletion to be more likely in *nursery* than in *cursory*, in *memory* than in *mammary*, in *scenery* than in *chicanery*, and so on.

These deletions appear to be phonetically gradual in that the variants range from those in which a schwa is followed by a resonant, in these cases [r], to those in which the resonant is syllabic, to those in which the resonant is not syllabic (i.e., all syllabicity has been lost). Such changes, then, are both phonetically gradual and lexically gradual. However, they may never fall into the category of regular sound change. The reason is that, as new words enter the language or as low-frequency words become more frequent, there will always be new schwas developing in the context for this deletion process. Similarly, the vowel reductions discussed by Fidelholtz and van Bergem may never be complete. As new full vowels come into unstressed position there will be new material for the reduction to work on.

Factors causing more reduction in high-frequency words

The cases documented so far indicate that high-frequency words tend to change before low-frequency words when the change is the deletion of stops (English /t, d/ deletion), the deletion of fricatives (Spanish ð deletion), some vowel shifts (Labov, 1994; Moonwomon, 1992), the reduction of vowels to schwa (in both Dutch and English), and the deletion of schwa (in American English). One might therefore predict that, in general, reductive changes tend to occur earlier and to a greater extent in words and phrases of high frequency.

If we take linguistic behavior to be highly practiced neuromotor activity (Anderson, 1993; Boyland, 1996; Haiman, 1994), then we can view reductive sound change as the result of the automation of linguistic production. It is well known that repeated neuromotor patterns become more efficient as they are practiced; transitions are smoothed by the anticipatory overlap of gestures, and unnecessary or extreme gestures decrease in magnitude or are omitted. Recent theories of articulatory change point to precisely these two types of changes—increase in overlap of gestures and decrease in magnitude—as describing all changes that occur in casual speech (Browman & Goldstein, 1992) or in sound change (Mowrey & Pagliuca, 1995).

If casual speech processes and reductive sound change are the natural result of the automation of linguistic productions, then it follows that such change will be more advanced in productions that are more highly practiced (i.e., in high-frequency words and phrases). In fact, reductive sound change may be just the most salient aspect of an overall reduction of high-frequency words. Using the Switchboard corpus, Jurafsky et al. (2001) found that, for 1,412 tokens of monosyllabic content words ending in /t/ or /d/, high-frequency words (at the 95th percentile) were 18% shorter than low-frequency words (at the 5th percentile). That is, the entire articulatory span of high-frequency words may be reduced compared to low-frequency words, a phenomenon that could give rise to some noticeable and some not so noticeable articulatory changes.

Other findings by Jurafsky and colleagues indicate that articulatory reduction cannot run rampant, but may be constrained by predictability in discourse. Using the same set of tokens, Jurafsky et al. (2001) found that the predictability of the word given the following word affects duration. That is, the first words in frequently occurring word pairs, such as *Grand Canyon*, *grand piano*, or *Burt Reynolds*, were shorter than words used in less predictable contexts. Gregory et al. (1999) found that, among 4,695 monosyllables ending in /t/ or /d/ from the Switchboard corpus, semantic relatedness to the discourse topic affected word duration: words related to the discourse topic were shorter than words that were not. In addition, in Gregory et al.'s study, words tended to be shorter if they were repeated in the same discourse (see also Fowler & Housum, 1987). In both studies the findings are explained with reference to predictability: predictable words are more reduced. A fuller interpretation of these results might run as follows. The tendency for articulatory reduction due to increased automation is always present and shows itself more prominently in highly practiced, frequent words. Reduction can be inhibited by the speaker's sensitivity to the predictability of words in the context. If the speaker knows that the word will be easily accessed in the context, because it or related words have already been activated, the reductive automating processes will be allowed to advance. If the word is less predictable in discourse, the speaker is likely to suppress the reductive processes and to give the word a more complete articulation.

CHANGES AFFECTING LOW-FREQUENCY WORDS FIRST

As previously mentioned, Hooper (1976) noted a lexical diffusion paradox: sound change seems to affect high-frequency words first, but analogical change affects low-frequency words first. The first tendency has already been documented. The second tendency is evident in the fact that low-frequency verbs, such as *weep*/*wept*, *leap*/*leapt*, *creep*/*crept*, are regularizing, while high-frequency verbs with the same pattern show no such tendency: that is, *keep*/*kept*, *sleep*/*slept*, *leave*/*left* show no evidence of regularizing. Hooper (1976) argued that changes affecting high-frequency words first have their source in the automation of production, whereas changes affecting low-frequency words first are due to imperfect learning. In the latter category are changes that affect words that do not conform to the general patterns of the language. Such exceptional words can be learned and maintained in their exceptional form if they are of high frequency in the input and in general use. However, if their frequency of use is low, they may not be sufficiently available in experience to be acquired and entrenched. Thus they may be subject to changes based on the general patterns of the language.

The example just given above has a clear morphological motivation, but Phillips (1984) showed that some sound changes can proceed from low-frequency to high-frequency words. For instance, the Old English diphthong *eo* monophthongized to a mid front rounded vowel /ö/, with both a long and a short version in the 11th to 12th centuries. In some dialects these front rounded vowels were maintained into the 14th century, but in Lincolnshire they quickly unrounded and merged with /e(:)/. A text written around 1200 AD, the Ormulum, captures this change in progress. The author was interested in spelling reform, and so, rather than regularizing the spelling, he represented the variation using two spellings for the same word (e.g., *deop*, *dep* 'deep') in many cases. Phillips found that, within the class of nouns and verbs, the low-frequency words are more likely to have the spelling that represents the innovation, the unrounded vowel.

If this were a phonetically motivated reduction that facilitates production, we would expect the high-frequency words to change first. Indeed, the frequent adverbs and function words have changed, suggesting that they might be yielding to production pressures, but the fact that nouns and verbs show more change in low-frequency items indicates a different motivation for the change. Phillips proposed that a constraint against front rounded vowels is operating to remove these vowels. Since there were no other front rounded vowels in English at the time, the majority pattern would be for front vowels to be unrounded. The mid front rounded vowels would have to be learned as a special case. Front rounded vowels are difficult to discriminate perceptually, and children acquire them later than unrounded vowels. Gilbert and Wyman (1975) found that French children confused [ö] and [ɛ] more often than any other non-nasal vowels they tested. A possible explanation for the Middle English change, then, is that children correctly acquired the front rounded vowels in high-frequency words that were highly available in the input, but tended toward merger with the unrounded version in words that were less familiar.

Thus there appear to be at least two directions for lexical diffusion with regard to frequency: some changes affect high-frequency words earlier, and some affect low-frequency words earlier. The former changes are phonetically conditioned and gradual, but they may not all turn out to be lexically regular. The evidence cited here suggests that lexical diffusion is much more common than previously supposed. It may be that all sound change diffuses gradually through the lexicon. For this reason we must reject the dichotomy of regular versus lexical diffusion changes and look at the pattern of lexical diffusion for each change. Each pattern of diffusion is associated with a particular source and mechanism for change, which allows us to use the direction of diffusion as a diagnostic for the cause of change (Bybee, 2001; Hooper, 1976; Phillips, 1984, 2001). Changes that affect high-frequency words first are the result of the automation of production, while low-frequency words change first when the change makes the words conform to the stronger patterns of the language. Low-frequency words, with their lesser availability in experience and consequently their weaker representation, are more susceptible to analysis and change on the basis of other forms.

MODELING PHONETIC AND LEXICAL GRADUALNESS

The view of lexical diffusion espoused by both Wang and Labov assumes that a change that diffuses gradually through the lexicon must be phonetically abrupt. This is a necessary assumption if one is to accept a synchronic phonological theory that has phonemic underlying representations. Words can change one by one only if the change is a substitution of phonemes in such a theory. The discovery that sound change can be both phonetically gradual and lexically gradual forces a different view of the mental representation of the phonology of words (Bybee, 2000b; Hooper, 1981). If subphonemic detail or ranges of variation can be associated with particular words, an accurate model of phonological representation must allow phonetic detail in the cognitive representation of words.

A recent proposal is that the cognitive representation of a word can be made up of the set of exemplars that have been experienced by the speaker/hearer. Thus all phonetic variants of a word are stored in memory and organized into a cluster: exemplars that are more similar are closer to one another than to ones that are dissimilar, and exemplars that occur frequently are stronger than less frequent ones (Bybee, 2000a, 2001; Johnson, 1997; Pierrehumbert, 2001). These exemplar clusters, which represent autonomous words, change as experience with language changes. Repeated exemplars within the cluster grow stronger, and less frequently used ones may fade over time, as other memories do.

Changes in the phonetic range of the exemplar cluster may also take place as language is used and new tokens of words are experienced. Thus the range of phonetic variation of a word can gradually change over time, allowing a phonetically gradual sound change to affect different words at different rates. Given a tendency for reduction during production, the phonetic representation of a word will gradually accrue more exemplars that are reduced, and these exemplars will become more likely to be chosen for production, where they may undergo further reduction, gradually moving the words of the language in a consistent direction. The more frequent words will have more chances to undergo online reduction and thus will change more rapidly. The more predictable words (which are usually also the more frequent ones) will have a greater chance of having their reduced version chosen, given the context, and thus will advance the reductive change more rapidly.

The exemplar clusters are embedded in a network of associations among words that map relations of similarity at all levels. Distinct words with similar phonetic properties are associated, as are words with shared semantic features. Bybee

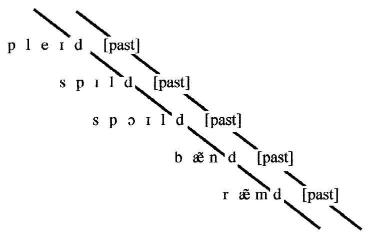

FIGURE 1. Network representation for some English past tense verbs.

(1985, 1988) showed that morphemes and morphological relations in such a network emerge from parallel phonetic and semantic associations, and that schemas or abstractions over relations of similarity can be formulated to account for the regularities and patterns evident in language use. Such a network is illustrated in Figure 1, where the connection lines indicate parallel semantic and phonological relations among instances of the past tense /d/.

Figure 2 shows how complex morphological relations can emerge from the connections based on phonological and semantic similarity. These figures, of course, only show fragments of the network within which these words are embedded, and they do not show all the relations that exist. Other questions about this model with respect to phonology and morphology are answered in Bybee (2001) and Bybee (1985), respectively.

An important property of the exemplar model is the emphasis on words as storage units. Various models have proposed that even multimorphemic words have lexical listing. Vennemann (1974) argued that appropriate constraints on syllable structure can only be applied to whole words, not to morphemes. The common objection to this proposal made in the 1970s was that the human brain does not have sufficient storage capacity for all the words of a language, especially a language with large morphological paradigms. This argument has now been dismissed with the discovery of the huge amount of detail that the brain is capable of recording. Moreover, newer conceptions of the lexicon, not as a list but as a network with tight interconnections, provide the insight that listing two related words, such as *start*, *started*, does not take up as much cognitive space as listing two unrelated words, such as *start*, *flower* (Bybee, 1985). Thus connectionist models (Rumelhart & McClelland, 1986) and analogical models (Eddington, 2000; Skousen, 1989, 1992) have storage of whole words with morphological relations emergent from the categorization involved in storage. In addition, the lexical diffusion data provide evidence that multimorphemic words can have lexical storage. As we saw in Table 3, high-frequency regular past tense English verbs are more likely to have their final /t/ or /d/ deleted than are low-frequency

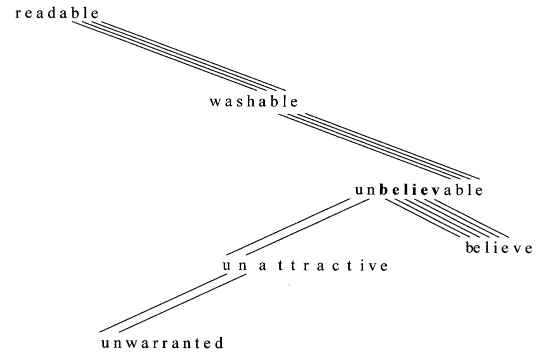

FIGURE 2. Lexical connections supply a morphological analysis of words such as *unbelievable*.

verbs. In order for a frequency effect to accrue to a word, that word must exist in memory storage. Since multimorphemic words evince frequency effects, they must be stored in the lexicon. I return to the issue of how morphological relations are expressed in a network model later in the article.

THE EFFECT OF FREQUENCY OF USE IN CONTEXT

Given that the exemplar model tracks tokens of use and the exemplar cluster changes according to the phonetic shape of these tokens, it follows that if the context of use affects the phonetic shape of a word, the exemplar cluster will change accordingly. The effect of context can be best exemplified in changes that take place around word or morpheme boundaries, where the segment affected by the change is sometimes in the context for the change and sometimes not. Timberlake (1978) called this an alternating environment. Since the exemplar model registers phonetic tokens, the probabilities inherent in the alternating environment affect the shape of the exemplar cluster. For example, the aspiration of syllable-final /s/ in many dialects of Spanish is conditioned by a following consonant, but when that /s/ is word-final, it sometimes precedes a consonant and sometimes a vowel.[5] At first, some tokens will have [s] (those that precede vowels) and some will have [h] (those that precede consonants). These variants become part of the exemplar cluster. Except for words of very high frequency, this type of variation in the phonetic shape of a word seems not to be stable. Rather, over time one variant begins to occur in all environments. In the case of Spanish [s]-aspiration, the [h] variant begins to occur even prevocalically (Bybee, 2000a, 2001). The reason for this is that [h] is the variant for the preconsonantal position, which is the most frequently occurring context in running discourse, as most words in Spanish begin with consonants. The exemplar cluster, then, appears to reorganize itself, with the stronger exemplars being more frequently chosen for use than the less frequent ones despite the context.

Thus, along with the general measure of frequency of use, the relative frequency of the immediate linguistic context of use can also affect the lexical diffusion of a sound change. Even holding frequency constant, a word that occurs more often in the right context for a change will undergo the change more rapidly than a word that occurs less often in the conditioning context. The next section examines the effect of specific contexts of use on American English /t, d/ deletion.

Special contexts for past tense and contracted negation

The deletion of word-final /t/ or /d/ in American English has been studied from a number of different perspectives. Of particular relevance here, Guy (1991a) studied the effects of contextual conditioning on both uniform and alternating environments with respect to /t, d/ deletion. The word-internal context (i.e., the segment preceding the final /t/ or /d/) is uniform or constantly present for each word, whereas the word-external context (i.e., the segment following the word-final /t/ or /d/) is an alternating environment because it varies according to what word follows the /t/ or /d/ in question. Guy found that the word-external context has a lesser effect on the deletion of the /t/ or /d/ than does the word-internal context—a finding predicted by the exemplar model. For example, with a word such as *different*, the /t/ always follows /n/, which presents a favorable condition for deletion, but sometimes *different* precedes a vowel, as in *different one*, and other times it precedes a consonant, as in *different story*. Guy found that the preceding environment was the factor with the strongest effect on deletion, conditioning more consistently than the following environment. As we shall see, the present study reveals the same phenomenon. For example, the *t* in words ending in *-nt* actually deletes more often before vowels than before consonants.

The differential effects of uniform and alternating environments find a natural representation in an exemplar model. As tokens of use are registered for a given word, the preceding word-internal context is constant and provides the conditioning environment in each use. The following word-external context varies, sometimes providing a conditioning context and sometimes not. Thus a word with a preceding favoring environment will have more exemplars with deletion than one without such a preceding favoring environment, other factors being equal.

The English regular past tense /t/ or /d/ has an alternating environment in both the preceding and following contexts. Since past tense occurs on verbs with any final segments, sometimes the past tense follows a vowel (*played*, *tried*), sometimes a resonant (*scored*, *called*), and sometimes an obstruent (*slipped*,

TABLE 5. *Rates of deletion for three word groups and all words before a pause, a vowel, and a consonant*

	Prepausal	Prevocalic	Preconsonantal
n't	16/39 (41%)	31/36 (86%)	254/304 (84%)
-nt	8/21 (38%)	9/13 (69%)	11/25 (44%)
-ed	5/37 (14%)	8/112 (7%)	61/130 (47%)
All words	111/296 (38%)	158/428 (37%)	756/1272 (59%)

rubbed). Thus the preceding context alternates, just as the following word-external context alternates. To the extent that speakers generalize over different instances of the same bound morpheme, the instances with a preceding environment that does not favor deletion may retard the deletion across all items, thus providing part of the explanation for the fact that the past tense /t/ and /d/ delete at a lower rate than other instances of /t/ and /d/. Another important factor for this lower rate of deletion is taken up directly: the environment after the past tense also tends not to favor deletion.

The demonstrated effect that the context has on the deletion rate suggests that a closer look at the contexts in which a word occurs may provide more insight into the rate at which it undergoes a change. For example, in Santa Ana's (1991) corpus, used to study /t, d/ deletion in Bybee (2000b) (henceforth, the LA corpus), 23% of words ending in /t/ or /d/ preceded a vowel. However, in different subsets of words, this rate varied from 9% to 40%. Since the position before a vowel is not conducive to deletion, variations in this rate may affect the rate of deletion for certain sets of words.

We coded approximately 2,000 tokens of words ending in /t/ or /d/ from the LA corpus. Even though the subjects chosen were dominant in English, there might be some question as to the possible effects of Spanish on their speech. Thus the specific results reported on here were checked in the phonemically transcribed portions of the Switchboard corpus, a corpus of telephone conversations among people living in the Dallas, Texas area. As we shall see, all results from the LA corpus were replicated in the Switchboard corpus.

I examined three sets of words: (i) auxiliaries with contracted negation, (ii) a set of words ending in an unstressed syllable with /nt/ (*government*, *different*, *pregnant*), which was intended as a phonologically matched control for the contracted negation, and (iii) the regular past tense forms in the corpus. The specific hypothesis tested was that words that occur more often before vowels, an environment that phonetically favors retention, would exhibit less deletion overall. This hypothesis assumes that the deletion is no longer purely conditioned by the phonetic environment, but that effects of the change are already represented lexically. The evidence for this assumption is the fact that, for two of these groups, the rate of deletion preceding a vowel versus preceding a consonant is not predictable from the phonetics. In fact, Table 5 shows that, for contracted negation,

TABLE 6. *Percentage overall deletion and occurrence before vowels for three sets of words in the LA corpus*

	Overall Deletion[a]	Occurrence Before Vowels[b]
n't	301/379 (79.4%)	36/379 (9.5%)
-nt	28/59 (47.4%)	13/59 (22.0%)
-ed	74/279 (26.5%)	112/279 (40.1%)

[a]For overall deletion, comparison of *-n't* to *-nt* yields $\chi^2 = 27.90025$, $p < .0001$, and comparison of *-ed* to *-nt* yields $\chi^2 = 10.12929$, $p < .001$.
[b]For occurrence before vowels, comparison of *n't* to *-nt* yields $\chi^2 = 8.073901$, $p < .01$, and comparison of *-ed* to *-nt* yields $\chi^2 = 6.757618$, $p < .01$.

the deletion rates are approximately the same in both contexts, and that words ending in *-nt* actually exhibit more deletion before a vowel than before a consonant. Only the past tense *-ed* shows deletion rates more in line with the original phonetic conditioning, a fact I return to later.

The hypothesis, then, is that, as lexicalization proceeds and the exemplar cluster changes, the overall distribution of the exemplars will reflect their distribution in the experience of the speaker at first, and eventually the more frequent exemplars will take over in more and more contexts. Thus, words that occur more frequently before vowels will show less deletion overall than words that occur more frequently before consonants. Table 6 demonstrates that this hypothesis holds for the three sets of words studied in the LA corpus. The overall deletion rates for these word sets are inversely related to their occurrence before vowels in the corpus.

The distributions of these sets of words are strikingly different, as can be observed from Table 7. Auxiliaries with contracted negation occur overwhelmingly before consonants (approximately 80% of the cases in the LA corpus). Of course, their most common site is before a verb and a large majority of verbs begin with consonants. The contracted negation occurs before a pause in only about 10% of the cases. The higher deletion rate for contracted negation, then, is attributable to its distribution in spoken discourse, where it commonly occurs internal to clauses and closely linked to the following verb, which is highly likely to be consonant-initial. Because exemplars without the final /t/ are highly frequent, they come to dominate the exemplar clusters, giving high rates of deletion overall.

Words ending in an unstressed syllable ending in *-nt* were examined in an attempt to find a phonological string that matched the contracted negation but without a morphological function. Of course, these words do not provide an exact phonological match to contracted negation in most cases. So although *student* ends in the same consonants as *didn't*, words such as *pregnant*, *government* do

TABLE 7. *Distribution of word-final /t/ and /d/ before pauses, vowels, and consonants*

	Prepausal	Prevocalic	Preconsonantal	*N*
n't	10%	10%	80%	379
-nt	36%	22%	42%	59
-ed	13%	40%	47%	279
All words	15%	21%	64%	1,996

not provide exact matches to any contracted negatives. However, these words do constitute a set of nouns and adjectives with which to compare the contractions and the past tense in terms of distribution. As Table 7 shows, these nouns and adjectives have a very different distribution from the auxiliary and contracted negation complex. They occur prepausally about 36% of the time. Their prevocalic occurrence of 22% is almost the same as the overall average for the corpus. They occur preconsonantally in 42% of the cases. The lower number of preconsonantal cases and the higher number of prevocalic cases amount to fewer deleted cases in the exemplar cluster and thus a lower overall rate of deletion than the contracted negation, as shown in Table 6. Note again that the lower rate of deletion is not attributable to a higher percentage of prevocalic cases, which would favor retention, because in fact there is not more deletion preconsonantally than prevocalically.

Regular past tense verbs show a distribution that is even more fascinating. Their occurrence before vowels is extremely high: 40% versus an overall figure for the whole corpus of only 21%. Preconsonantal occurrences are at 47%, and prepausal are at 13%. Thus the exemplar cluster for past tense verbs contains many more pre-vocalic occurrences than the other groups and thus would have many more occurrences with the /t/ or /d/ present. In effect, the deletion has not progressed as far for the past tense word group as for the other groups.

Table 6, then, reveals a strong correspondence between the rate of occurrence before vowels and the overall deletion rate for the word group. These facts point clearly to the proposition that sound change occurs in real time as words are used, and that its effects are registered in memory, producing a gradual change in lexical items based on the speaker's actual experience with them. It also reinforces what the data on frequency effects demonstrates, as shown earlier in this article. Words that more frequently occur in the context favoring a change undergo the change at a faster rate than those that occur less frequently in the appropriate context. These findings strongly suggest that phonetic change occurs online, in production, as language is used.

To replicate the findings of the study of the LA corpus, the same word sets were examined in the Switchboard corpus. While the rates of deletion of final /t/ and /d/ were generally somewhat higher in the Switchboard corpus, the correspondence between the overall rate of deletion and the occurrence in prevocalic

TABLE 8. *Percentage overall deletion and occurrence before vowels for three sets of words in the Switchboard corpus*

	Overall Deletion[a]	Occurrence Before Vowels[b]
n't	423/473 (89%)	64/473 (14%)
-nt	104/144 (72%)	40/144 (28%)
-ed	166/492 (34%)	170/492 (36%)

[a]Differences in rates of deletion are significant at the $p < .0001$ level.
[b]Differences in rates of occurrence before vowels are significant at the $p < .01$ level.

position held as it did in the LA corpus, as demonstrated in Table 8. Despite differences in the two corpora in terms of regional and ethnic heritage of the speakers, of the situation, and of potentially different coding criteria, the correspondence between the data from the two corpora is remarkably similar. The differences shown in each column of Table 8 are statistically significant; the rates of deletion in the three word sets are significant at the $p < .0001$ level and in the occurrences before vowels at the $p < .01$ level.

One additional question concerning these results must be addressed: is it possible that the different rates of deletion are due primarily to the token frequency of the words in each set rather than to the contexts in which they occur? After all, the contracted negation set contains the highly frequent words *don't*, *didn't*, which may be responsible for the higher deletion rates of *-n't* words versus *-nt* words. This is less an issue for the regular past tense, since the earlier test to see if these words deleted less than the general corpus had restricted the general corpus words to those with a frequency of less than 403 per million in the Francis and Kučera count (see Table 2). To control for frequency for the comparison of *-n't* words with *-nt* words, the frequency ranges of both groups in the corpora were compared. Both groups contained words with only one occurrence in each corpus, while the highest frequency word in the *-nt* set was *different* with 18 occurrences in the LA corpus and 23 in the Switchboard corpus. This amounted to a total of 41 occurrences in the two corpora. For this test, the data from the two corpora were pooled, and all *-n't* words with a frequency of greater than 41 in the two corpora were removed. The affected words were *don't*, *didn't*, *can't*. The differences between the overall deletion rates and the occurrence before vowels were tested again, and the difference in deletion rate was found to be significant at the $p < .01$ level ($\chi^2 = 7.987$), while the occurrence before vowels was significant at the $p < .05$ level ($\chi^2 = 4.158$); see Table 9 for details. Thus, we see that it is not simply a word's frequency that determines the rate at which it undergoes a sound change, but rather that word's frequency in the context that conditions the sound change.

The data just examined now raises two other questions. First, why do regular past tense verbs occur so much more frequently before vowels? Second, to what extent does a token of a bound morpheme in one word, such as the contracted negation and the regular past tense, have an effect on other instances of the same morpheme in other words?

TABLE 9. *Percentage overall deletion and occurrence before vowels for* -n't *and* -nt *words in the LA corpus and the Switchboard corpus (excluding* -n't *words with a frequency greater than 41)*

	Overall Deletion[a]	Occurrence Before Vowels[b]
-n't	181/223 (81.1%)	40/223 (17.9%)
-nt	132/203 (65.0%)	58/203 (28.6%)

[a]For overall deletion, $\chi^2 = 14.20575$, $p < .001$.
[b]For occurrence before vowels, $\chi^2 = 6.784548$, $p < .01$.

The distribution of regular past tense verbs

A closer examination of the context of use of regular past tense shows that the high rate of prevocalic instances arises from the frequent combination of verbs with particles or prepositions, many of which begin in vowels. In the LA corpus 62 (55%) of the vocalic environments following a regular past tense verb were prepositions or particles. The examples were: *kicked out*, *lived in*, *lived on*, *looked at*, *moved in*, *opened up*, *picked up*, *picked on*, *pulled away*, *pulled out*, *raised in*, *stepped in*, *talked about*, *walked into*, *walked out*, *worked in*, *worked at*. Another 23 (20%) were vowel-initial pronouns, such as *it*, *us*, *him*, *her*, *them* with the first consonant deleted. Another 10 (9%) of these cases consisted of *and* in addition to the commonly used phrases *looked alike*, *changed a lot*, *learned a lot*. And 7 (6%) of the prevocalic contexts had the indefinite article after the verb. In sum, about 90% of the vowel-initial items that follow the past tense verb form what Erman and Warren (2000) called a prefabricated unit with the verb. That is, these sequences are conventionalized units that are very likely stored and processed together.[6]

This chunking and the establishment of semi-autonomous representations protect these verbs from the word-level generalization that leads to deletion outside of the appropriate phonetic contexts. It also suggests that the relevant units for the diffusion of the change very likely involve specific verb + particle and verb + pronoun sequences, such as *looked at*, *picked up*, *liked it*, and so on. These sequences present a uniform environment that does not favor the deletion of the /t/ or /d/. This would explain the very low rate of deletion (7%) of /t/ or /d/ when it is a regular past tense occurring before a vowel.[7]

Other accounts of the past tense phenomena

The lower rate of deletion in the regular past tense has been observed in all studies of English /t, d/ deletion and various explanations have been proposed to account for it. One idea is that, since the past tense /t/ and /d/ have the specific grammatical function of marking past time reference, the deletion of the consonant would entail loss of that information from the utterance. While it may be that in certain contexts this information is uniquely marked in the suffix and that a speaker might make an effort to suppress reduction in such cases, in most cases the information conveyed by this consonant is redundant. Furthermore, we know from the history of many European languages that final consonants with morphological functions often fall to deletion processes. For instance, singular/plural verb agreement in English was lost with the deletion of final unstressed /n/, and the distinction between Latin first and third singular verb agreement fell with the reduction and loss of final /t/ and /m/. In addition, detailed studies of the interaction of functional load with reduction and deletion processes, such as Poplack (1980a, 1980b) and Labov (1994, Chapter 19), have revealed that other factors are more powerful than the preservation of morphological information, and that the predictions of the functional account are often not borne out in the data on variation and change.

Guy's (1991a, 1991b) account attributed the deletion rates for /t/ and /d/ to the structure of the lexical and morphological components of the grammar, adopting the theory of lexical phonology (Kiparsky, 1982). In his account, /t, d/ deletion can apply at a fixed rate at any level of the grammar. Thus some words have a chance to undergo the variable deletion two or three times. Monomorphemic words have a chance to undergo deletion three times, giving them the highest rate of deletion. Semi-weak past tense verbs (those verbs with both a vowel change and a final /t/ or /d/, like *leave*, *left*; *tell*, *told*) are formed at Level 1 and thus have two chances to undergo deletion. Regular past tense is formed at Level 2 and thus has only one chance to undergo deletion. Since each exposure to the variable deletion rule increases the rate of deletion, Guy's model provides an account for the fact that monomorphemic words have a higher rate of deletion than semi-weak verbs, which in turn have a higher rate of deletion than the regular past tense verbs.

Guy's analysis coincides with the exemplar account in two ways. First, it models the cumulative effect of the variable rule applying multiple times. So does the exemplar model, which predicts that the more often a word is exposed to online reduction, the greater rate of deletion it evinces. Second, Guy's analysis recognizes that even a variable rule can interact in significant ways with the lexicon and the morphology. However, his analysis is unable to account for the major role played by word frequency at every level of the lexicon and morphology, and it misses the fact that the behavior of the semi-weak verbs can be accounted for simply by noting their high frequency with respect to the average regular verb. As mentioned earlier, Bybee (2000b) and Gregory et al. (1999) showed that, for all words, there is a significant effect of frequency on final /t, d/ deletion, such that high-frequency words undergo more deletion than low-frequency words. In addition, Bybee (2000b) showed that there is an effect of frequency on regular past tense words occurring in preconsonantal or prepausal position (see Table 3). Lexical phonology has no means of tracking the token frequency of words. Even if words in the lexicon could accrue frequency, the regular past tense verbs are not

TABLE 10. *Semi-weak past tense forms from the LA corpus ordered by frequency, with ties ordered by Francis and Kučera frequency*

Total Tokens	Verb	Tokens with Deletions
32	told	22 (68%)
9	felt	5 (55%)
8	left	2 (25%)
6	kept	4 (66%)
4	sent	1 (25%)
4	built	0 (0%)
3	held	0 (0%)
3	heard	0 (0%)
2	slept	1 (50%)
2	lent	0 (0%)
1	found	0 (0%)
1	lost	0 (0%)
1	meant	0 (0%)

Spearman rank order correlation: $p = .696$, significant at the .01 level (two-tailed or one-tailed)

in the lexicon, but are formed at Level 2, leaving no means of tracking their token frequency. As for the semi-weak verbs, many of them, such as *told*, *felt*, *left*, *kept*, are of very high frequency, which alone can account for their higher rate of deletion. A striking confirmation of the role of frequency is that, even within the semi-weak verbs, the high-frequency verbs undergo deletion much more than the low-frequency verbs, as shown in Table 10, taken from Bybee (2000b).

A final argument against the lexical phonology account of variable rule application refers to data to be presented later in the article, where it is shown that Spanish intervocalic [ð] is deleted more in the past participle morpheme than in lexical morphemes. This is counter to the predictions made by lexical phonology since the past participle in question is perfectly regular and would be formed at Level 2. According to the logic of Guy's model, it should undergo deletion fewer times than instances of lexical morphemes do. As I shall argue, the different rates of change for suffixes are not due to where they appear in the derivation, but to how consistently they occur in the environment for change in language use.

A summary of the usage-based account

The model that can handle the facts just described must allow for gradual changes in the phonetic representations of words, and it must also allow for the expression of certain relations within and among exemplar clusters. As already demonstrated, reductive change occurs in production as already automated sequences of linguistic elements are further reduced. These online reductions feed back into memory representations, since the language user's accumulated experience is represented in memory. Words are represented as clusters of exemplars, and the relative weight of exemplars with different patterns may change over time as reduction proceeds. If the distribution of words in actual discourse contexts differs, the rate at which their exemplar clusters change, and thus the rate at which they undergo a change, may differ.

Exemplar clusters contain information about the contexts in which different exemplars are used. In some cases, this information is lost, and exemplars spread from the more frequent context to less frequent ones. This has apparently occurred in the contracted negation and in *-nt* words, where the rate of deletion before vowels and consonants shows no phonetic effect. In other cases, contextual information is explicit and strong and allows the contextual exemplars to be maintained. This appears to be the case in prefabricated, conventionalized units, such as verb + particle and verb + pronoun sequences, where the maintenance of the regular past tense /t/ or /d/ is high. Note that there is almost no deletion within such units (only 7%), whereas deletion of past tense /t/ or /d/ before consonants correlates with the frequency of the verb, as shown in Table 3. This suggests that the /t/ or /d/ in these sequences is behaving more as a word-internal consonant than as a final one.

Thus, the lexical representation for a verb such as *live* may also contain the more specific entries *lived at*, *lived in*, *lived for* with information about the phonetic shape of these sequences. Figure 3 shows how a portion of the complex network involving verbs and particles might appear. The parallel semantic and phonetic categorization creates exemplar clusters at various levels: all instances of *live* are categorized in a cluster, all instances of *lived* are categorized together, as are all instances of *lived in*. A cross-cutting categorization would put instances of *in* together to the extent that some semantic affinity could be identified among its various uses (e.g., *moved in*, *played in*, and so on). Similarly, it would follow that instances of the regular past tense might be categorized together, with some latitude allowed for the fact that the allomorphs, /t/, /d/, and /ɨd/, differ from each other in voicing and syllabicity.

THE EFFECT OF BOUND MORPHEMES ON CHANGES IN WORDS

Alternating environments

So far we have examined two groups of words that were defined by their suffixes, the regular past tense and contracted negation. An issue we have not yet addressed is the extent to which an affix in one word can have an effect on the same affix in another word. In the model adopted for this study, words can be independent lexical units, but bound morphemes occur only within words and thus are always part of larger units (i.e., they have no independent existence). However, one way that words are categorized in the lexicon is through their shared parts. If these shared parts have both semantic and phonological similarity, they will be identified as morphemes. Thus the regular past tense occurrences in different

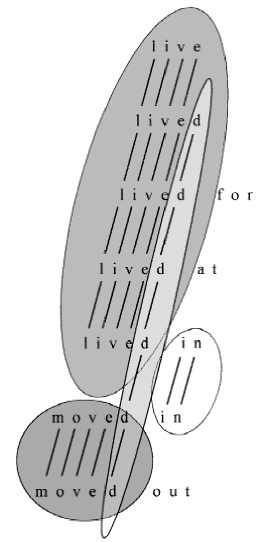

FIGURE 3. A complex network of verbs and particles contains parallel semantic and phonetic categorizations.

words would be represented as in Figure 4, where a dotted line indicates only partial phonological similarity.

Another way of conceptualizing the unity of a suffix, such as regular past in English, would be to consider that it also constitutes an exemplar cluster (as shown in the circled portion of Figure 4). Because the regular past suffix occurs before a vowel about 35% to 40% of the time, the full /t/ or /d/ exemplars remain strong, especially those in conventionalized sequences such as *looked at*, *opened up*. Thus the overall rate of deletion is slowed in regular past tense, even before a consonant or pause.

The evidence that one instance of a past tense can affect another, however, cannot be taken as evidence for an independent unit, the past tense suffix. Remember that preconsonantal and prepausal instances of regular past tense show a frequency effect: namely, high-frequency past tense verbs in these contexts are more likely to undergo deletion than low-frequency ones, as shown in Table 3.

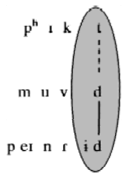

FIGURE 4. Regular past tense morpheme in the words *picked*, *moved*, *painted*.

Thus, deletion of the past tense /t/ or /d/ is never independent of the lexical verb it occurs with.

Similar cases of alternating environments slowing down the progress of a change were presented in Timberlake (1978), and examples of alternating environments blocking a change completely were examined in Bybee (2001).

Uniform environments

The effect of exemplars of bound morphemes upon one another can also be seen in uniform environments, specifically in ones that always uniformly present the context for change. High-frequency morphemes that contain within them the context for the change can undergo the change more rapidly. A case in point is the deletion of intervocalic [ð] in Spanish dialects, which has been in progress for several centuries. This [ð] derives from Latin *t*, which is voiced in Romance; in Spanish it is variably pronounced as a fricative intervocalically or deleted. As mentioned before, a study of the variable deletion of intervocalic [ð] in New Mexican Spanish shows a frequency effect for lexical instances (see Table 4). This effect was also tested on a much larger corpus of peninsular Spanish (COREC). In this 1.1-million-word corpus of spoken language from a variety of situations, deleted elements were noted in the orthographic transcription.[8] We counted all instances of deleted and retained intervocalic *d*. Table 11 shows that the rate of deletion is much lower than for the New Mexican corpus, but that the frequency effect still holds.[9]

In addition to lexical instances of [ð], this sound occurs intervocalically in past participles of all conjugation classes, as in *hablado* ‘spoken’, *comido* ‘eaten’, *vivido* ‘lived’. The rate of deletion for the first conjugation suffix (58%) is higher than the rate for all non–past participle words (17%). This is shown in Table 12 for the New Mexican corpus and in Table 13 for the COREC. The figures here include past participles used either as verb forms, such as the various forms of the perfect, or as adjectives.[10] The second and third conjugation forms of the past participle, *-ido*, also delete at a higher rate than other words with *-id-*, with a deletion rate of 29% ($\chi^2 = 11.345$, $p < .001$). In the COREC, with its lower rate

TABLE 11. *Rate of deletion in high- and low-frequency non–past participles from the COREC*

	Low (0–226)	High (239–2642)	Total
Retention	20,560 (97.6%)	19,171 (91.3%)	39,731 (94.4%)
Deletion	508 (2.4%)	1,830 (8.7%)	2,338 (5.6%)
Total	21,068	21,001	42,069

$\chi^2 = 796.1$, $p < .0001$

TABLE 12. *Deletion rate of* [ð] *in first conjugation past participles versus all other words in the New Mexican corpus*

	First Conjugation Past Participles	All Other Words
Retention	38 (42%)	507 (83%)
Deletion	53 (58%)	101 (17%)

$\chi^2 = 79.9$, $p < .0001$

TABLE 13. *Deletion rate of* [ð] *in first conjugation past participles versus all other words in the COREC*

	First Conjugation Past Participles	All Other Words
Retention	4,903 (78.3%)	39,731 (94.4%)
Deletion	1,355 (21.7%)	2,338 (5.6%)

$\chi^2 = 1999.5$, $p < .0001$

of deletion, the only second or third conjugation past participle that shows deletion is *sido*, the past participle of *ser* ‘to be’.

The apparent explanation for the higher rate of deletion in the past participle morpheme is its high frequency. It appears more than any other single morpheme, lexical or grammatical, with etymological /d/ in it. These data provide evidence for the representation of the past participle suffix (or suffixes) in exemplar clusters. However, again we have evidence that the suffix is not an independent unit, because the rate of deletion in the suffix is higher with high-frequency verbs. The effects of frequency on the deletion rate for the first conjugation past participle are shown in Table 14 for the COREC data. The New Mexican corpus shows a trend in the same direction, which, because of the small number of words coded, does not achieve statistical significance (see Bybee, 2001).

TABLE 14. *Rate of deletion by frequency for first conjugation past participles from the COREC corpus*

	Low (0–33)	High (34–499)	Total
Retention	2,631 (85.1%)	2,275 (71.7%)	4,906 (78.4%)
Deletion	459 (14.9%)	896 (28.3%)	1,355 (21.6%)
Total	3,090	3,171	6,261

$\chi^2 = 165.8$, $p < .0001$

The frequency effect within past participles demonstrates that the suffix is not an independent unit, but exists as a part of a whole word. However, the higher rate of deletion in this suffix suggests some relations among exemplars of the suffix, as indicated in the model with connecting lines for exemplar clusters for parts of words.

Another interesting issue that these data address is the segmentation of the past participle morpheme. Pedagogically, one can teach the past participles as *-ado* for first conjugation and *-ido* for second and third conjugation. However, the first vowel should be associated with the verb stem, as it is predictable from the lexical classification of the verb. This leaves *-do* for the perfect tenses and *-do*, *-da*, *-dos*, *-das* in adjectival uses, depending on the number and gender of the noun. Thus the only constant for the past participle in general is the *-d-*. The difference in deletion rates in *-ado* versus *-ido*, however, suggests that the vowel is segmented with the suffix, and that each two-syllable unit is undergoing change at its own rate.

Conclusions concerning bound morphemes within words

The data discussed here have shown that bound morphemes can have an effect on the diffusion of a sound change, suggesting some association within the lexicon among instances of the same morpheme in different words. We have seen that the English past tense changes more slowly, while the Spanish past participle and, incidentally, the contracted negation in English change more quickly. The reason for these differences can be directly related to variation in the contexts of use of each of the morphemes. The importance of the morpheme in these cases does not overshadow the importance of the words the morpheme appears in. The frequency of the word containing the morpheme, in both the English past tense and the Spanish past participle, contributes to the rate of deletion, strongly pointing to lexical representations for multimorphemic words, which can accumulate their own frequency values or lexical strength.

THE EFFECTS OF EXTREMELY HIGH FREQUENCY

One final effect of context and high frequency of use should be mentioned. In cases of extremely high frequency of words within certain phrases, a new exemplar cluster can be formed. Bybee and Scheibman (1999) studied the reduction of *don't* in conversation. We found that the most common and the most extreme reductions of *don't* occurred in the contexts in which *don't* occurred most often, namely, when *I* was the subject and the following verb was *know*. In fact, the vowel in *don't* was reduced to schwa only when the subject was *I* and the following verbs were *know*, *think*, *have*, *want*, *like*, *mean*, *feel*. The only other context for reduction to this extent was in the phrase *why don't you*, when used as a suggestion. This special reduction suggests that these phrases are separate from the independent representation of *don't*, and that they form exemplar clusters of their own, which allow the reduction of *don't* to progress at a faster rate in these phrases than in other uses. A phrase will form an exemplar cluster of its own when it is used frequently; this frequency of use will also allow it to register phonological, semantic, and pragmatic changes.

Another robust example of the development of variants of words in frequent contexts is liaison in French, where, for example, the definite articles take on special forms depending upon the initial segment of the following noun, as in *le chien* 'the dog' versus *l'ami* 'the friend' and *le*(*s*) *chiens* 'the dogs' versus *le*[*z*] *amis* 'the friends'. While other final consonants were being deleted, liaison consonants remained within highly frequent contexts that were either within a grammatical construction or within a frequent, fixed phrase. As expected, liaison is being lost in a progression from the least frequent to the most frequent contexts (Bybee, 2001).

CONSEQUENCES FOR A USAGE-BASED THEORY

The study of the diffusion of sound change in the lexicon contributes to a better understanding of the nature and causes of sound change. Changes that affect high-frequency words first are a result of the automation of production, the normal overlap and reduction of articulatory gestures that comes with fluency (Browman & Goldstein, 1992; Mowrey & Pagliuca, 1995). The strong directionality of such changes indicates that they are not the result of random variation, but that they stem from reduction processes resulting from repetition and the normal automation of motor activity. If a sound change does not proceed from the most frequent to the least frequent words, then we should seek its explanation in some other mechanisms of change.

Moreover, I have proposed a model in which variation and change are not external to the lexicon and grammar but inherent to it (Pierrehumbert, 1994). Sound change is not rule addition—something that happens at a superficial level without any effect on the deeper reaches of grammar. Rather, lexical representations are affected from the very beginnings of the change. Indeed, they supply an ongoing record of the change since they track the details of the phonetic tokens experienced. Further evidence for sound change having an immediate impact on representation is the fact that sound changes are never reversed or undone (Bybee, 2001; Cole & Hualde, 1998). The morphological structure of words also plays a role from the initial stages of a change, but less because morphemes have some special status with respect to change and more because of the contexts in which they appear. Alternating contexts retard change, while uniform ones allow change to hurry ahead.

Effects of frequency and context demonstrate that language use has an effect on mental representations. In this view, representations and the grammatical structure that emerges from them are based on experience with language. New linguistic experiences are categorized in terms of already stored representations, adding to the exemplar clusters already present and, at times, changing them gradually. Various levels of abstraction emerge as exemplars are categorized by phonological and semantic similarity—morphemes, words, phrases, and constructions can all be seen as the result of the categorization of linguistic experiences.

NOTES

1. There might be some question as to whether the variation discussed is sound change in progress or stable variation. In either case, word frequency must be taken into account along with other factors in describing the variation.

2. The English recorded in this corpus is a contact variety, as is the Spanish discussed in connection with [ð] deletion. However, the data from these corpora are checked against monolingual varieties in both cases, and the results with respect to the effects of frequency are very much the same, strengthening the evidence for the approach of this work.

3. The datasets were drawn from 38,000 words that were phonetically hand-transcribed (Greenberg, 2000; Greenberg, Ellis, & Hollenback, 1996).

4. The weakness of this experiment is that it relies on self-report, which must always be viewed with caution.

5. In the remainder of this article, I will use square brackets and slash lines in more or less traditional ways, even though the model invoked here does not have an explicit level of phonemic representation. The slash lines designate a broad range of variants, and the square brackets a narrower, contextually determined range of variants.

6. That English verbs often occur in collocations with other words is also discussed in Hopper (1991, 1997).

7. There were only 6 deletions of final past tense /d/ before a vowel in the LA corpus, 4 with a particle or preposition (*changed over*, *involved in* (2), *involved into*) and 2 with a pronoun (*called him*, *changed her*). In all of these cases except *called him*, the /d/ ends a complex consonant cluster.

8. The reliability of the transcription is not known; however, it does represent an independent source.

9. The cut-off between high and low frequency was chosen so as to produce two groups of tokens of approximately equal size. The same criterion was used for the data in Table 14.

10. Most of the etymological instances of past participles used as nouns seem to be lexicalized and were not counted as past participles (*marido* 'husband', *estado* 'state', *comida* 'meal').

REFERENCES

Anderson, John R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Bills, Garland D. (1997). New Mexican Spanish: Demise of the earliest European variety in the United States. *American Speech* 72:154–171.

Boyland, Joyce Tang. (1996). *Morphosyntactic change in progress*: *A psycholinguistic approach*. Doctoral dissertation, University of California at Berkeley.

Browman, Catherine P., & Goldstein, Louis M. (1992). Articulatory phonology: An overview. *Phonetica* 49:155–180.

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

______ (1988). Morphology as lexical organization. In M. Hammond & M. Noonan (eds.), *Theoretical morphology*. San Diego: Academic. 119–141.

______ (2000a). Lexicalization of sound change and alternating environments. In M. Broe & J. Pierrehumbert (eds.), *Laboratory V: Language acquisition and the lexicon*. Cambridge: Cambridge University Press. 250–268.

______ (2000b). The phonology of the lexicon: Evidence from lexical diffusion. In M. Barlow & S. Kemmer (eds.), *Usage-based models of language*. Stanford: CSLI. 65–85.

______ (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

______ (2002). Lexical diffusion in regular sound change. In David Restle & Dietmar Zaefferer (eds.), *Sounds and systems. Studies in structure and change. A Festschrift for Theo Vennemann* (Trends in Linguistics). Berlin: Mouton de Gruyter. 59–74.

Bybee, Joan, & Scheibman, Joanne. (1999). The effect of usage on degree of constituency: The reduction of *don't* in American English. *Linguistics* 37:575–596.

Cole, Jennifer, & Hualde, José Ignacio. (1998). The object of lexical acquisition: A UR-free model. *Chicago Linguistic Society* 34:447–458.

D'Introno, Francisco, & Sosa, Juan Manuel. (1979). *Anuario de la Escuela de Letras*. Caracas: Universidad Central de Venezuela. 33–61.

Eddington, David. (2000). Stress assignment in Spanish within the analogical modeling of language. *Language* 76:92–109.

Erman, Britt, & Warren, Beatrice. (2000). The idiom principle and the open choice principle. *Text* 20:29–62.

Fidelholtz, James. (1975). Word frequency and vowel reduction in English. *Chicago Linguistic Society* 11:200–213.

Fowler, Carol A., & Housum, Jonathan. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26:489–504.

Francis, W. Nelson, & Kučera, Henry. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Gilbert, John H., & Wyman, Virginia J. (1975). Discrimination learning of nasalized and non-nasalized vowels by five-, six- and seven-year-old children. *Phonetica* 31:65–80.

Greenberg, S. (2000). Switchboard transcription project. LVCSR Summer Workshop, Johns Hopkins University. http://www.icsi.berkeley.edu/real/stp/.

Greenberg, S., Ellis, D., & Hollenback, J. (1996). Insights into spoken language gleaned from phonetic transcriptions of the Switchboard corpus. *ICSLP-96*, S24-7, Philadelphia, PA. 524–527.

Gregory, Michelle, Raymond, William D., Bell, Alan, Fosler-Lussier, Eric, & Jurafsky, Daniel. (1999). The effects of collocational strength and contextual predictability in lexical production. *Chicago Linguistic Society* 35:151–166.

Guy, Gregory (1980). Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.), *Locating language in time and space*. New York: Academic. 1–36.

______ (1991a). Contextual conditioning in variable lexical phonology. *Language Variation and Change* 3:223–239.

______ (1991b). Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3:1–22.

Haiman, John. (1994). Ritualization and the development of language. In W. Pagliuca (ed.), *Perspectives on grammaticalization*. Amsterdam: Benjamins. 3–28.

Hooper, Joan Bybee. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie (ed.), *Current progress in historical linguistics*. Amsterdam: North Holland. 96–105.

______ (1981). The empirical determination of phonological representations. In T. Myers, J. Laver, & J. M. Anderson (eds.), *The cognitive representation of speech*. Amsterdam: North Holland. 347–357.

Hopper, Paul J. (1991). Dispersed verbal predicates in vernacular written narrative. *Berkeley Linguistics Society* 17:402–413.

______ (1997). Dispersed verbal predicates in vernacular written narrative. In A. Kamio (ed.), *Directions in functional linguistics*. Amsterdam: Benjamins. 1–18.

Johnson, Keith. (1997). Speech perception without speaker normalization. In K. Johnson & J. W. Mullennix (eds.), *Talker variability in speech processing*. San Diego: Academic. 145–165.

Jurafsky, Daniel, Bell, Alan, Gregory, Michelle, & Raymond, William D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins. 229–254.

Kiparsky, Paul. (1982). Lexical phonology and morphology. In I. S. Yang (ed.), *Linguistics in the morning calm*. Seoul: Hanshin. 3–91.

Krishnamurti, Bh. (1998). Regularity of sound change through lexical diffusion: A study of *s* > *h* > ∅ in Gondi dialects. *Language Variation and Change* 10:193–220.
Labov, William. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
______ (1981). Resolving the neogrammarian controversy. *Language* 57:267–308.
______ (1994). *Principles of linguistic change: Internal factors*. Oxford: Blackwell.
Losiewicz, B. L. (1992). *The effect of frequency on linguistic morphology*. Doctoral dissertation, University of Texas at Austin.
Marcos-Marín, Francisco. (1992). Corpus oral de referencia del español contemporáneo. Universidad Autónoma de Madrid. http://elvira.lllf.uam.es./docs_es/corpus/corpus.html.
Menéndez-Pidal, Ramón. (1968 [1904]). *Manual de gramática histórica española*. Madrid: Espasa-Calpe.
Moonwomon, Birch. (1992). *The mechanism of lexical diffusion*. Paper presented at the Annual Meeting of the Linguistic Society of America, Philadelphia.
Mowrey, Richard, & Pagliuca, William. (1995). The reductive character of articulatory evolution. *Rivista di Linguistica* 7:37–124.
Neu, Helene. (1980). Ranking of constraints on /t, d/ deletion in American English: A statistical analysis. In W. Labov (ed.), *Locating language in time and space*. New York: Academic. 37–54.
Oliveira, Marco Antonio de. (1991). The neogrammarian controversy revisited. *International Journal of the Sociology of Language* 89:93–105.
Penny, Ralph. (1991). *A history of the Spanish language*. Cambridge: Cambridge University Press.
Pensado, Carmen. (1984). *Cronología relativa del Castellano*. Salamanca: Ediciones Universidad de Salamanca.
Phillips, Betty S. (1984). Word frequency and the actuation of sound change. *Language* 60:320–342.
______ (2001). Lexical diffusion, lexical frequency, and lexical analysis. In J. Bybee & P. Hopper (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins. 123–136.
Pierrehumbert, Janet. (1994). Knowledge of variation. *Papers from the Parasession on Variation. 30th Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 232–256.
______ (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins. 137–157.
Poplack, Shana. (1980a). Deletion and disambiguation in Puerto Rican Spanish. *Language* 56:371–385.
______ (1980b). The notion of the plural in Puerto Rican Spanish: Competing constraints on (s) deletion. In W. Labov (ed.), *Locating language in time and space*. New York: Academic. 55–67.
Rumelhart, D., & McClelland, J. (1986). On learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In J. McClelland, D. Rumelhart, & PDP Research Group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press. 216–271.
Santa Ana, Otto. (1991). *Phonetic simplification processes in English of the barrio: A cross-generational sociolinguistic study of the Chicanos of Los Angeles*. Doctoral dissertation, University of Pennsylvania.
Schuchardt, Hugo. (1885). *Über die Lautgesetze: gegen die Junggrammatiker*. Berlin: Oppenheim. (Reprinted and translated in T. Vennemann & T. Wilbur. [1972]. *Schuchardt, the Neogrammarians and the Transformational Theory of Phonological Change*. Frankfurt: Athenaum).
Skousen, Royal. (1989). *Analogical modeling of language*. Dordrecht: Kluwer Academic.
______ (1992). *Analogy and structure*. Dordrecht: Kluwer Academic.
Timberlake, Alan. (1978). Uniform and alternating environments in phonological change. *Folia Slavica* 2:312–328.
Van Bergem, Dick. (1995). *Acoustic and lexical vowel reduction* (Studies in Language and Language Use 16). Amsterdam: IFOTT.
Vennemann, Theo. (1974). Words and syllables in natural generative phonology. *Papers from the Parasession on Natural Phonology. 10th Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 346–374.
Wang, William S.-Y. (1969). Competing changes as a cause of residue. *Language* 45:9–25.
______ (ed.). (1977). *The lexicon in phonological change*. The Hague: Mouton.
Wang, William S.-Y., & Cheng, C.-C. (1977). Implementation of phonological change: The Shaungfeng Chinese case. In W. Wang (ed.), *The lexicon in phonological change*. The Hague: Mouton. 86–100.