Chapter 7

WITH JEAN E. NEWMAN

Are Stem Changes as Natural as Affixes?

1. One processing type or two?

Two types of processing have been proposed to account for linguistic behavior associated with morphologically complex words. A traditional generative account of affixation and other morphological processes uses symbolic rules located in a module separate from the lexicon and acts on lexically represented morphemes to create words (Chomsky and Halle 1968). Various recent versions of this model treat morphology as rule derived, analogous to syntax (e.g., Saddock 1985). The other types of processing proposed for morphology are various versions of pattern network associators. Bybee and Slobin (1982) and Bybee and Moder (1983) presented evidence that English irregular past-tense verbs are stored in the lexicon and organized into networks centered around prototypes. Bybee (1985, 1988) proposed that even regular morphological patterns can be accounted for in the same way and that a separate module for morphological derivation was not necessary. The fact that some morphological processes are productive and others are not is a reflex of various factors, but primarily of the type frequency of the particular process, that is, the number of lexical items that participate in the process. A similar claim has been made in the literature on connectionist models that rely exclusively on building up networks of associated patterns based on learned items (Rumelhart and McClelland 1986; MacWhinney et al. 1989). Another recent proposal is that both types of processing are available and utilized and a strict distinction exists between irregular morphological formations, which are stored in the lexicon and organized into networks, and regular morphological processes, which are symbolic rules located in a separate

module of the grammar (Pinker 1991; Marcus et al. 1992; Prasada and Pinker 1993). In this proposal, "regular" processes are defined as the productive ones, that is, the morphological processes that are most easily extended to new items.

It is perhaps not accidental that in English and many other languages (including other Germanic languages and Romance languages) the regular processes are affixal (and often agglutinative in their structure) while the irregular ones often involve changes in the stem or a high degree of fusion between stem and affix. It is possible that the substance of the process—stem change versus affixation—is also associated with the processing type. Since affixation bears a greater resemblance to syntactic concatenation than stem change does, it might be more appropriately handled by symbolic rules and stem changes might be more appropriately handled by lexical connections.

A prominent claim made in the theory of natural morphology, as conceived of by Dressler and others, is that among morphological processes, affixation is more "natural" than internal stem changes. In this theory the criteria for "more natural" include the following: (1) more common in the world's languages; (2) more common within a single language; (3) easier to acquire, and thus acquired earlier; (4) later lost in aphasia; and (5) easier to process in general (Dressler 1985). If affixation generally correlates with regularity, its greater "naturalness" could be explained by a difference in processing type: the conclusion would be that processing by symbolic rules is to be preferred over processing by lexical pattern networks. Marcus et al. (1993) describe symbolic rules as concatenating, that is, affixing, rules, which suggests an association of affixation with regularity and stem change with irregularity.

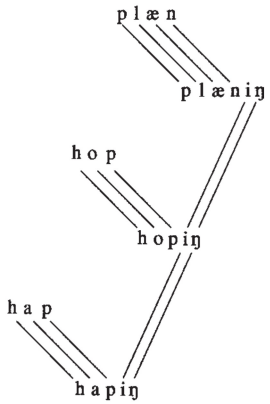
In contrast, a model in which all morphology is handled in the lexicon via associative pattern networks does not distinguish affixation from stem change. Affixes are found in the lexicon attached to words and are analyzed via lexical connections, just as stem changes are. For instance, in the model proposed by Bybee (1985, 1988) a regular suffix is represented by the fragment in (la) and the semiproductive pattern whose prototypical member is *string*, *strung* is partially represented as in (lb) (adapted from Bybee 1988).

If this model is correct, there should be no preference for affixation over stem change from the point of view of acquisition or processing.

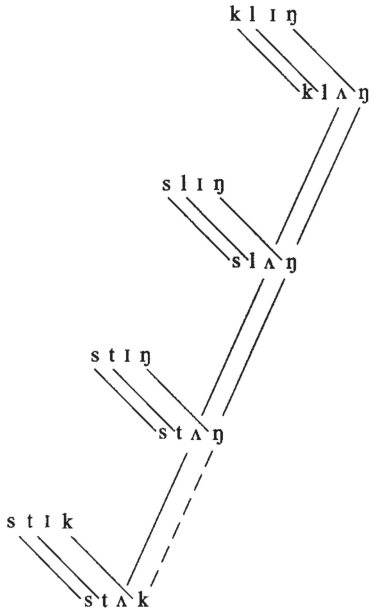
However, such differences *are* found in the acquisition process. Clark and Berman (1987) report that children acquiring Hebrew master morphology requiring suffixation before that requiring stem changes. They point to two reasons for this: stem changes obscure the shape of the lexeme more than affixes do, and stem changes tend to be restricted to lexically arbitrary classes of stems. In the research reported on in this article we attempted to manipulate these two factors in order to determine whether stem changes per se are more difficult to acquire than affixes. Our hypothesis is that stem changes themselves are no more difficult than affixes to store and process, but it is the fact that stem changes, in all reported cases, are lexically arbitrary and of low type frequency that makes them more difficult to master.

In order to find out whether it is stem change itself or other factors that delay acquisition, it was necessary to experiment on artificial languages in which stem change could be made independent of lexical arbitrariness and type frequency. Before describing this experiment and its results, another matter needs to be examined.

(1) a.



b.



If we are claiming that affixation is no easier to acquire and process than stem changes, we need to explain why it is that affixation is so much more common both across the languages of the world and within individual languages. The answer to this question lies in the diachronic phenomenon of grammaticization.

2. The diachronic perspective

Grammaticization is the process by which lexical morphemes or phrases develop into grammatical morphemes and continue their development, often resulting in the development of new affixes. Now that recent research has examined this process in a wide variety of languages (Heine and Reh 1984; Bybee 1985; Heine, Claudi, and Hünnemeyer 1991; Traugott and Heine 1991; Bybee, Perkins, and Pagliuca 1994),

it is clear that the primary sources of grammatical affixes are free words or phrases. (Although there is no direct evidence available on the source of Semitic stem changes, they could have originally been conditioned by the presence of affixes that have long since disappeared. See table 7.1 below for an example of how this occurs.) It is also clear that the grammaticization process is ongoing at all times in all languages and that new grammatical morphemes are constantly arising, some of which are destined to become affixes. Some accessible examples are the development of the English regular past-tense suffix *-ed* (OE *-ede*) from the preterite form of the verb *don* 'to do', *dyde*, which in Proto-Germanic followed the main verb, leading to its development as a suffix. In derivational morphology, the productive affix *-ly*, which appears both on nouns to form adjectives and on adjectives to form adverbs, developed from the OE word *liç* meaning 'body'. Its frequent occurrence in compounds led to its development into an affix.

The development of internal stem changes is a related phenomenon and also occurs during the grammaticization process. However, it is much less common, and more important, the development of stem changes depends entirely upon the prior existence of affixes. Stem changes tend to develop very late in the grammaticization process. The longer an affix has been attached to a stem and the more the combination is used, the greater will be the phonological fusion between the affix and the stem. Sometimes the stem conditions changes in the affix (as when the *-d* of the English past tense is devoiced following a voiceless consonant). At other times, the affix will condition a change in the stem. In Middle English, long vowels were shortened before certain stop clusters so that the past-tense suffix in words such as *slept*, *kept*, *left* caused shortening of the stem vowel, giving them a different vowel from their base forms *sleep*, *keep*, and *leave*. If an affix conditions a change in the stem and then is itself deleted, the stem change becomes the signal for the morphological category. For instance, an old Germanic noun plural suffix was *-i*, and it appeared on nouns such as *mūs* 'mouse', *gōs* 'goose', and *fōt* 'foot'. This suffix, which consisted of a high front vowel, caused the vowel of the stem also to become a front vowel, yielding the long front rounded vowels [ü] and [ö]. At about the same time, this suffix vowel reduced to schwa. Later the schwa deleted entirely and the front rounded vowels became unrounded (to [i] and [u]). Then these long vowels of the singular and the plural were subject to the Great Vowel Shift of early Modem English, giving us the pairs *mouse*, *mice*; *goose*, *geese*; and *foot*, *feet* (see table 7.1). The entire process by which a stem change arose under the influence of an affix that was later lost

TABLE 7.1. The Development of an English Stem Change for Plural

	Singular	*Plural*
Proto-Germanic	*mūs	*mūsi
Proto-Germanic	*mūs	*mȳsə
Old English	mūs	mȳs [müs]
Middle English	mūs	miese [mi:s]
Modem English	mouse [maus]	mice [mais]

took many centuries. Since grammaticization is a continuous process and new grammatical morphemes are constantly arising and supplanting older ones, most affixes do not last long enough to produce stem changes, and even if they do, they are usually being replaced by some other affix, leaving the stem change only as residue in highly frequent items. Thus grammaticization explains why affixes are common, but stem changes are rare.

Another historical dimension of our argument involves grammaticization as well. As we have mentioned, during grammaticization a once free grammatical morpheme can become an affix. Becoming an affix involves a number of changes:

1. The morpheme in question must become fixed in its position. In particular, it must have a grammatically determined position with regard to a lexical stem. (If it is fixed with respect to a phrase, it will be a clitic and it will not become an affix.)
2. No open-class items can come between the grammatical morpheme and the lexical morpheme.
3. When these two criteria are met, the grammatical and lexical morpheme combination may be reanalyzed as a single unit, that is, a word. This reanalysis depends upon the combination constituting a coherent conceptual unit (Bybee 1985).
4. Phonological reduction in the grammaticizing morpheme is ongoing at all times (loss of stress or high tone, reduction and loss of consonants and vowels), forcing it to become more dependent on surrounding material.
5. Phonological fusion between stem and affix takes place.

A well-documented instance of this type of change is the development of the future tense in Romance languages such as Spanish and French. A periphrastic construction in Latin consisting of an inflected auxiliary *habere* 'to have' and an infinitive yielded a meaning of obligation or predestination:

(2) amare habeo
love + inf aux + 1s
'I have to love, I am to love'

The auxiliary reduces phonologically and comes consistently to appear after the infinitive (where previously it could occur in various places in the clause). In Old Spanish we find the construction indicating future:

(3) amar he
love + inf aux + 1s
'I will/shall love'

The auxiliary is written separately from the infinitive because at this stage other morphemes could come between the two, for instance, the object pronoun:

(4) amar lo he
love + inf him aux + ls
'I will love him'

Later this possibility disappears and the auxiliary becomes an actual suffix to the verb:

(5) lo amaré
'I will love him'

In this process the grammaticizing morpheme undergoes phonological reduction (e.g., from *habeo*, to *he*, to *é*), its position becomes fixed, it fuses with the verb, and the whole construction takes on a more abstract, grammatical meaning.

If all regular morphology were handled in a module with symbolic rules that resemble syntactic rules there would be no motivation for the process of affixation ever to take place. Free grammatical morphemes inserted by rule could simply continue to be inserted by rule. However, if affixation is the lexical representation of two units (morphemes) as one unit (a word), then we can understand the affixation process as being the result of the two units occurring together frequently enough and constituting a coherent enough conceptual unit that speakers remember and store them together as a single unit in their mental lexicon. As we said earlier, if affixes are represented lexically, then there is no qualitative difference between affixes and stem changes.

3. The experiment

Because in all languages affixes develop historically more quickly and more easily than stem-internal changes, it is normally the case that affixes are used in the productive patterns and have high type frequency, while stem changes, if they are used at all, occur in lexically arbitrary classes and have low type frequency but high token frequency. It is also often the case (as in English past tense) that there are several types of stem changes, each one restricted to an arbitrary class of verbs. In order to determine whether or not there is an acquisition or processing advantage to affixes, it is necessary to create artificial languages in which the relevant variables can be manipulated independently of one another. Artificial languages have been used in psycholinguistic studies to explore other linguistic variables such as the acquisition of syntax (e.g., Green 1979; Mori and Moeser 1983). Such artificially constructed languages have the advantage that they can be used to study the linguistic strategies that speakers bring to their task without examining the properties of a specific, existing language.

3.1. Background

A study by Bybee, Fox, and Klinck (1987) serves as a point of departure for the current experiment. In the earlier study Bybee and co-authors asked whether affixes were

more "natural" than stem changes. Their operational definition of *natural* was the frequency with which subjects generalized either stem changes or suffixes to novel words in an artificial language–learning task. Twenty subjects were presented with sixteen singular and plural noun pairs in an artificial language. The novel words were accompanied by pictures of their referents in order to facilitate the learning of these forms. All of the singular forms were single syllables with a CVC form. Half of the stimuli (eight pairs) used a suffix schwa in the plural, whereas the other eight pairs had an internal vowel change to schwa to indicate the plural form. Both the suffix and the stem had equal type frequency, and both were equally regular, in the sense that there was no allomorphy to learn.

Subjects underwent three learning trials in which they were presented the singular form and were asked to recall the plural; by the last trial, performance was at 90% accuracy. In the final, critical phase, subjects were subsequently tested on a generalization task. They were presented with six novel words in singular form and were asked to supply what they thought would be an appropriate plural form.

The results indicated that there was no preference for affixation over the stem change. Of the 120 possible responses in the experiment (20 subjects × 6 novel nouns), 63 were affixes (52.5%). When the results were examined by subjects, there was also no clear tendency to choose affixes more frequently than stem changes. The results showed that eight subjects had more affix than stem-change responses, six subjects had an equal number of affixes and stem changes, and six subjects had more stem-change than affix responses.

These data, then, show that there is no obvious advantage for affixation over stem change to indicate the plural and therefore suggest that the former is not a more "natural" process. Accepting these results, however, means that we are accepting the null hypothesis, which is definitely problematic. Furthermore, it may be that Bybee and co-authors did not detect a difference between the two types of plural markers for a different reason that was masked by the design of the study. The experiment did not systematically address the possibility that type frequency, not the type of morphological process, determines the generalization of a suffix or a stem change. In this experiment, the stem change and the suffix were regular, that is, there was only one form for each, and both stem changes and suffixes had equal type frequency. Because the researchers held regularity constant, we cannot determine whether the equivalent generalization of the two types of plural markers was the result of their being equally "natural" or of the fact that they had equal type frequency, which is not usually the case with natural languages where suffixes typically have high type frequency and stem changes typically have low type frequency. The following experiment addresses this issue by using the same basic paradigm but manipulating both type frequency and type of morphological process while keeping token frequency constant in all conditions.

The following variables were controlled:

1. Affixation versus stem change. In each condition, half of the nouns form their plural by the addition of a suffix and half by a change in the stem vowel.

2. Phonetic simplicity (Clark and Berman 1987) is the degree of change between two morphological forms. We recognize that a derivation involving multiple differences between two forms, such as *bring*, *brought*, will be more difficult to acquire than derivations involving only one, such as *string*, *strung*. Therefore, the stem changes used in the experiment involved changing only the vowel of a monosyllabic word.
3. Lexical arbitrariness. In each condition, half of the nouns used required a stem change to form the plural and half used a suffix. The nouns were constructed so that there were no phonological or semantic reasons for the choice. In two of the conditions stem changes took on four different patterns, again lexically arbitrary, and in two conditions the suffix had four arbitrary allomorphs.
4. Type frequency. Type frequency is related in this experiment to lexical arbitrariness. The more variants of the stem change or suffix there are, the lower the type frequency. The conditions in which there are four variants with low type frequency will be called the *irregular* conditions, and the conditions in which there is only one variant with consequent higher type frequency will be called the *regular* condition.
5. Token frequency. Paradigms with irregularities that also have low type frequency tend to have high token frequency. In our experiment we hold token frequency constant by presenting each noun and its plural the same number of times.

There were two independent variables in this experiment: *suffix type* (regular or irregular) and *stem-change type* (regular or irregular), which were crossed with each other in a 2 × 2 design. Each subject received one of four possible combinations of conditions: (1) suffix regular, stem change regular; (2) suffix irregular, stem change irregular; (3) suffix irregular, stem change regular; and (4) suffix regular, stem change irregular. Therefore, all subjects received both types of plurals, suffixes and stem changes, but what varied between conditions was the regularity of the plural for the two morphological processes. This design allows us to assess the effects of regularity on the ease of acquisition and the generalization of suffixes and stem changes while at the same time determining whether there is an overall preference for one morphological process over the other.

Subjects were randomly assigned to one of the four combinations of conditions, resulting in ten subjects in each group. The dependent variables in the generalization phase were percent choices of plural forms (suffix or stem change). Analyses on the learning data used percent correct.

The predictions of the "naturalness" hypothesis are clear: subjects should produce suffixes more frequently than stem changes in the generalization phase, regardless of the type frequencies of the two forms. If type frequency is an important constraint, however, then we expect to see the rate of generalization vary: high type frequency in the input condition could result in more generalization of the form as a plural marker, with the reverse true for plural forms with low type frequency. Of course the picture could be more complex, such that one morphological process might

be favored under conditions of high type frequency and the other under conditions of low type frequency. Finally, given our argument that stem changes are as "natural" as suffixes in marking the plural, we expect the overall rate of stem changes in the experiment to be equivalent to that of suffixes.

4. Method

4.1. Subjects

Forty students at the University of New Mexico served as subjects in this experiment. They were members of introductory linguistics and psycholinguistics courses who were given extra course credit for a written description of their participation in the experiment. In order for their data to be included in the experiment, subjects had to be native speakers of English (defined as their primary language, acquired before the age of 6 years).

The use of English speakers is justified by the fact that English has precisely the pattern that we are hypothesizing will *not* show up in our results. Thus the use of English speakers biases the experiment against the result we expect to obtain.

4.2. Materials

Two sets of materials were constructed for this experiment, corresponding to the two phases, learning and generalization. Sixteen novel words were generated for the learning phase (see table 7.2). These novel words were all of the form CVC and were constructed with a variety of consonants and ensured that vowels were used with equal frequency (see table 7.3).

The novel words were presented to subjects as the "singular" form of a word in a novel language, with the referent depicted in a line drawing. Subjects were also presented with a "plural" form of the same word, accompanied by an appropriate line drawing. The subject's task was to learn the singular and plural forms, such that when presented with the singular forms and pictures followed by the pictures for the plural they could provide the plural forms. They were also given a generalization task at the end of the experiment in which they were presented with a new set of ten novel words in the singular and asked to supply, for each, what they thought would be an appropriate plural form.

The stimuli were constructed in the following manner. The set of words was divided in half, with eight words being assigned to the affix condition and the other

TABLE 7.2. Example Stimuli for Stem-Change and Affixation Conditions

Singular		*Plural*	*Type*	*Picture*
faut	→	fət	stem change	girl(s)
dit	→	ditə	suffix	flower(s)

TABLE 7.3. Stimuli Used in the Learning Phase of the Experiment

Plural form	*Regular*	*Irregular*
Stem change		
waiš	wəš	wəš
zib	zəb	zəb
raik	rək	ruk
sauf	səf	suf
vaum	vəm	voum
faut	fət	fout
ǰik	ǰək	ǰɪk
naid	nəd	nɪd
Affix		
hauk	haukə	haukə
mir	mirə	mirə
sail	sailə	sailəm
nauz	nauzə	nauzəm
ǰaib	ǰaibə	ǰaibni
dit	ditə	ditni
vaud	vaudə	vaudət
faiš	faišə	faušət

eight to the stem-change condition. The other variable was regularity: in the case of regular suffixes and stem changes, there was only one form of the plural; thus each of the eight words (in either the stem-change or affix condition, as appropriate) took the same form. For the *regular* conditions this meant either the addition of [ə] as a suffix or the replacement of the vowel with [ə] for the stem change, for all eight instances of each condition (see table 7.2). By using schwa for both the stem-change and suffix condition, we controlled the vowel such that any differences in the two conditions could not be attributed to the specific vowel used. While we think it unlikely that the specific vowel that is used for the regular conditions would make much difference, future studies could replicate the study using a different vowel or vowels.

In the *irregular* suffix and stem-change conditions, four different plural forms were used, resulting in two novel words in each set of eight whose plural was signaled in the same manner. This contrasts with the regular conditions in which, within each of the eight-item sets, the plural form was marked by the same form. For the irregular suffix condition the four forms used for the affix were [ə] and three other suffixes: [əm], [ət], and [ni]. The irregular plural stem changed to [ə], [u], [ɪ], and [ou]. Care was taken to ensure that the resulting forms did not already exist in English, and that the choice of a particular stem change or suffix in the irregular condition did not correlate with any phonological or semantic properties of the stem (see table 7.3).

Each novel word and its plural form were recorded by a female native English speaker using a Bell and Howell Language Master. The Language Master is an

audiotape recorder that reads from, and records onto, thin cardboard cards that have a strip of audiotape attached to the bottom. Pictures of the referents (both single and multiple forms) were copied onto the card so that in the initial phase subjects could both hear the singular and plural forms of each word and see their referents. Because each item was recorded on a separate card, it was possible to shuffle the stimulus set and present each subject with a unique order of stimuli.

Ten additional novel words were generated for use in the generalization phase of the experiment:

(6) Novel items used in the generalization phase of the experiment
biš
waub
zaud
šaim
lauǰ
kaiǰ
tis
zaif
vin
raus

These were constructed in the same manner as the stimuli for the learning phase. In all cases, for learning and generalization trials both, the materials were carefully constructed so that there were no consistent cues from the structure of the CVC that indicated whether it would take a stem change or suffix in the plural.

4.3. Procedure

4.3.1. Learning phase

Subjects were told that they would be presented with singular and plural forms of words from a language that they did not know. They were instructed to learn the novel words so that they could, when given the singular form, recall the correct plural form.

The set of sixteen words was randomized by shuffling the audio cards and was then divided into four separate sets of four items each. On the initial trial, subjects heard the four items in both the singular and plural form, with their appropriate pictures. They were then presented with the singular form only and asked to generate the plural form. The responses were recorded (both by the experimenter and on audiocassette) and corrected, if wrong. The same four singular items were presented again for two more recall trials of four items, and on each trial the subject was again asked to provide the plural form. Corrections were provided. Following the third iteration of the first four-item set, the subject was presented with the second set of four items (in both singular and plural form), and the same procedure, with three recall trials, was followed. The initial phase of the learning task was repeated for the remaining two sets of four items in the list. For purposes of analy-

sis, each of the three recall trials is referred to as a practice list (numbered from 1 to 3), and the results will be reported across all sixteen words for each list.

Following the initial phase, subjects were presented with the full list of sixteen singular items and instructed to generate the correct plural for each. They were given feedback and corrected where appropriate. The final learning trial consisted of the same sixteen singular items, with instructions to generate the plural form. In this last trial of the learning phase no feedback was given.

4.3.2. Generalization phase

In the last phase of the experiment, subjects were given, one at a time, the new list of ten novel words, accompanied by pictures of their referents. They were told that these were in singular form and, following the presentation of each, were asked to supply what they considered to be appropriate plural forms. Obviously there were no correct answers since these were new items, and subjects were not given feedback on their responses. Following this phase subjects were thanked and debriefed.

4.4. Results

4.4.1. Learning phase

Analysis of the learning data was conducted on the mean percent correct recall across the conditions. In order to observe the pattern of acquisition, the recall data were examined separately for the three practice lists and the two final, full lists. The results revealed that the four conditions differed in initial learnability (see figure 7.1);

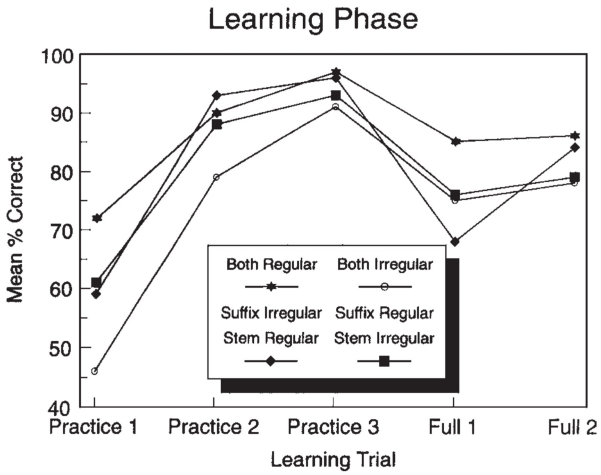


FIGURE 7.1. Correct performance by conditions in the learning phase.

the regular conditions led to higher initial learnability, whereas the irregular conditions led to the lowest initial performance. The effect of regularity was present for both suffixes and stem changes as shown by an analysis of the percent correct data for practice list 1. There was a significant main effect of suffix regularity, $F(1,36) = 5.941$, $p < .02$, MSe = 2065; and a marginally significant main effect of stem-change regularity, $F(1,36) = 4.054$, $p < .052$, MSe = 1409. The interaction of suffix regularity × stem-change regularity was not significant, $F < 1$.

Pairwise comparisons revealed that the source of the initial difference was a significant difference between condition 1 (suffix and stem change both regular) and condition 2 (suffix and stem change both irregular), such that condition 2 led to considerably fewer correct responses (condition 2: 46%, condition 1: 72%), two-tailed $t(18) = 3.31$, $p < .004$. Condition 1 was also marginally different from condition 3, suffix irregular, stem regular (59%): two-tailed $t(18) = 2.03$, $p < .058$. None of the other pairwise comparisons between conditions for practice list 1 approached significance.

Performance on the remaining practice lists quickly reached near ceiling levels. There was a marginally significant effect of stem-change regularity on practice list 2, $F(1,36) = 4.020$, $p < .053$, MSe = 612. No other effects approached significance for either practice list 2 or 3. Mean percent correct on practice list 3, collapsed across conditions, was 94%.

Analyses were also conducted on the percent correct scores for the two full lists. The analyses on the data for full list 1 revealed that there was a marginally significant interaction between suffix and stem-change regularity, $F(1,36) = 3.239$, $p < .08$, MSe = 712, and a marginally significant effect of suffix regularity, $F(1,36) = 3.732$, $p < .061$, MSe = 821. Although neither effect achieved the conventional level for rejecting the null hypothesis, both effects are strong and approaching significance, so that they should not be dismissed. The locus of the interaction can be seen in figure 7.1, where irregular suffixes in the context of a regular stem change (i.e., condition 3: suffix irregular, stem regular) lowered relative performance considerably more than did irregular stem changes in the context of regular suffixes (i.e., condition 4: suffix regular, stem irregular). The suffix-regularity effect is shown in the graph by the fact that the mean of the two suffix-regular conditions (both regular, and suffix regular, stem irregular) is higher than the mean of the two suffix irregular conditions (both irregular, and suffix irregular, stem regular).

Performance on full list 2 was quite high (82%) though not quite at the levels achieved by the end of practice list 3, probably because subjects had to successfully perform a paired-associate task for the entire sixteen items in the list, rather than recalling each set of four items separately. Importantly for our purposes, there were no significant differences in the rate of correct responses across the four conditions: suffix regularity × stem-change regularity, $F < 1$; suffix regularity, $F < 1$; stem-change regularity, $F(1,36) = 1.656$, $p < .206$, MSe = 389.

Thus the learning data indicate that there is no inherent difference in difficulty in terms of acquiring stem changes rather than suffixes. Although there were initial differences in percent correct as a function of regularity for both suffixes and stem changes, these quickly disappeared.

4.4.2. Generalization data

The rate of suffix generalization was 48.75%, whereas the rate of stem-change generalization across subjects in all four conditions was actually 1.5% higher, at 50.25%. Two other types of responses were noted in the generalization task—the use of an English suffix (.25%) and no change to the singular form (.75%). Combined, these two unusual response types totaled only 1% of the generalization data and were not included in the statistical analysis. Thus the overall results of this experiment look very much like those of the Bybee, Fox, and Klinck study, showing that stem changes and suffixes appear to be equally learnable and thus equally "natural." This study, however, differs in that we manipulated regularity as well as morphological form and are thus able to explore whether the apparently equivalent rates of use of the two types of morphological processes hold when regularity varies.

Figure 7.2 presents the percentages for the use of suffixes and stem changes in the generalization phase as a function of stem-change and suffix regularity. The pattern of results shows that the regularity of the stem-change condition during the learning phase had no effect on the use of either suffixes or stem changes in the generalization phase, but that suffix regularity did. Specifically, the rate of suffix use was considerably higher than stem-change use when the suffix input in the learning phase was regular, but lower than stem-change use when the suffix input was irregular. The asymmetry of this pattern is clearly illustrated in figure 7.3, which plots difference scores formed by subtracting the percentage of stem-change choices

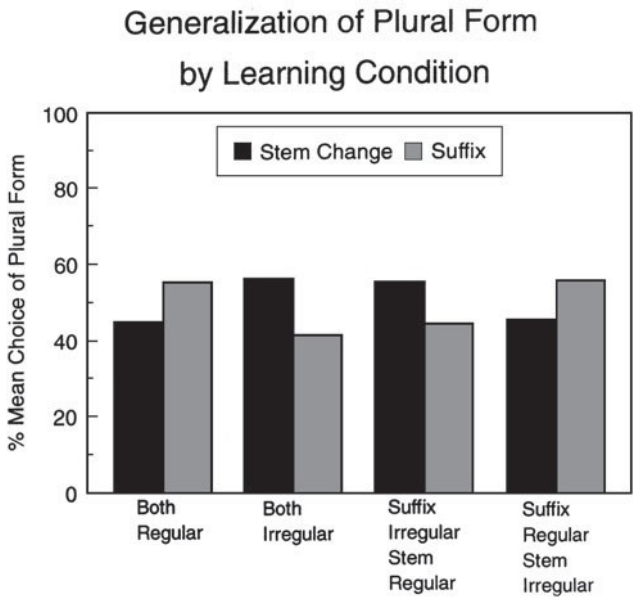


FIGURE 7.2. Generalization of suffixes and stem changes by learning conditions.

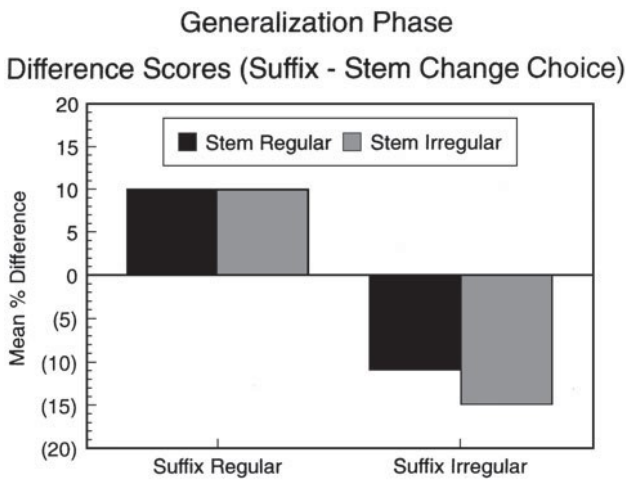


FIGURE 7.3. Generalization difference scores (suffix–stem change) as a function of regularity.

from the percentage of suffix choices. If suffix and stem-change generalization are equivalent, then the difference scores will be zero, regardless of input condition. Figure 7.3 clearly demonstrates that there were more suffix choices than stem-change choices when the suffix was regular (positive difference scores) and more stem-change than suffix choices when the suffix was irregular (negative difference scores).

Because the two dependent measures, percent suffix choice and percent stem-change choice, sum to 100%, thus ensuring that knowledge of one score unambiguously determines knowledge of the other, it would not be appropriate to conduct analyses of variance on both dependent measures. Thus a 2 (suffix regularity) × 2 (stem-change regularity) analysis of variance was conducted on the mean percent suffix choices. There was a main effect of suffix regularity, $F(1,36) = 5.009$, $p < .031$, MSe = 1562. Neither the main effect of stem-change regularity nor the interaction between suffix and stem-change regularity was significant, both Fs < 1. Thus the initial observation that there was no overall difference between the rates of suffix and stem-change generalization must be qualified by the finding that suffix regularity, but not stem-change regularity, determines the rate of suffix and stem-change generalization.

The pattern of results clearly shows that subjects' acquisition of the plural is affected by suffix regularity, but it also demonstrates that acquisition of a stem change to mark the plural is not inherently more difficult than learning to add a suffix as the plural marker. In the two conditions that contained irregular suffixes, subjects chose the stem change more frequently as a plural marker during the generalization phase (condition 2, suffix and stem irregular: 56%; and condition 3, suffix irregular, stem regular: 55%). This level is equal to that of suffix choice in the two regular suffix conditions (condition 1, suffix and stem regular: 55%; and condition 4, suffix regu-

lar, stem irregular: 55%), again showing that it is regularity of the suffix, not difficulty of the stem change, that determines the use of affixation versus stem change. Were affixation a more "natural" process overall, then we would expect to see a much greater advantage for suffixes under the input conditions that most favor their use, namely, the two regular suffix conditions. Indeed, we would expect the greatest use of suffixes in condition 4, in which the stem changes are irregular, but the suffixes are regular; but as we noted, that condition did not result in any higher suffix use than did condition 1, where both the suffix and the stem change were regular.

5. Discussion

Based on a theory in which all morphology is treated in basically the same manner, we hypothesized that stem changes as morphological markers could be learned as easily as affixes. The reason that stem changes seem difficult to acquire in languages such as English, German, and Hebrew is that most stem-change patterns are lexically arbitrary and have very low type frequency. By creating an artificial language and asking subjects to learn it we could distinguish the effects of stem change from the effects of type frequency. The results showed no overall advantage of suffixes over stem changes, but an effect of irregularity or type frequency of the suffixes was revealed. In the learning phase, irregularity in both affix and stem-change conditions led to initial difficulties. In the generalization phase, the overall rate of stem changes was as high as, in fact slightly exceeded, the rate of suffix usage. In this phase also, we see the effects of regularity: regularity of suffixes led to greater use of suffixes, while irregularity of suffixes led to greater use of stem changes.

Thus claims that suffixes (and affixes in general) are somehow psychologically easier to encode may be misguided. Our data suggest instead that regularity, which for diachronic reasons happens to be correlated with affixation rather than stem change, affects the ease of generalization of the suffix. It is also interesting to note that under the right circumstances, such as the two cases of irregular suffixes set up by our artificial language, stem changes can be acquired as easily as, or even more easily than, suffixes.

Our results also show the effects of type frequency on determining generalizability. In the two mixed conditions, one in which suffixes were regular and stem changes irregular and the other in which suffixes were irregular and stem changes regular, the regular morphological process was generalized more frequently than the irregular one. Our interpretation of these results is that the subjects were basically involved in a rote-learning task, setting up lexical representations for this fragment of an artificial language. Where half of the singular/plural pairs they learned had the same morphological process, this process was strengthened. In generalizing, the subjects were accessing the stored units and relations and could form new plurals on the basis of any stored pattern, but they were more likely to use the pattern reinforced by more separate entries.

For the conditions in which both processes were regular or both were irregular, we expected that responses would be approximately evenly divided among suffixes and stem changes. It is logically possible that subjects might have chosen a single

suffix or stem change and generalized it throughout, but it appears that in producing new forms the subjects were in general replicating the distribution of forms in the input. Thus in the all-regular and all-irregular conditions, subjects gave responses of both types. But here we found an interesting effect of regularity. When both processes were regular, subjects used significantly more suffixes, but when they were both irregular, subjects chose significantly more stem changes. Taken together, then, there is no overall suffixation preference, but we did find a difference that needs to be explained.

This result may be due to the subjects' prior experience with natural language(s), in which, as we have said, there is a strong tendency for stem change to be associated with irregularity and for affixes to be associated with regularity. Thus in the very difficult condition in which all plurals were irregular, the subjects might have been more comfortable with the stem-change solution, since stem changes tend to be irregular in English (and in all languages). Conversely, in the all-regular condition, the suffix has an edge since regularity tends to be associated with suffixes.

Because of their experience, English speakers should have been predisposed to use suffixes. The data clearly show that subjects did not, overall, prefer suffixes to stem changes. Since the subjects did generalize suffixes to the novel forms when one of the two input conditions was regular, one possible explanation of this result is that they were speakers of English. The best way to avoid this inherent bias would be to use as subjects speakers of an isolating language such as Mandarin Chinese. In this case we would predict effects of regularity for both suffixes and stem changes, but once again, no overall advantage for one form of morphological expression over the other. If there were a language whose pattern was opposite to that of English, rather than neutral, we would expect to find the opposite results, that is, an effect of stem-change regularity but not suffix regularity, and again, no overall advantage for either form. It is clear that the persistent refrain in our predictions, and in fact our findings, is that there should be no overall advantage for suffixes over stem changes as, in our view, neither form is more "natural" than the other.

In a model in which all morphological relations are treated as lexical connections, there is no particular advantage to affixation over internal changes. Even in such a model, it would still be reasonable to suppose that morphemes that consist of continuous phonological material, rather than interrupted material (as in *sing*, *sang*, *sung*, where the root is *s-ng*), have an acquisitional or processing advantage (Dressler 1985), yet our experiments turned up no evidence of such a tendency. The suggested association of concatenation with regularity and stem change with irregularity (Marcus et al. 1993) was not supported in our experiment, where we demonstrated that the English speakers behave as though these are separate variables.

On a more general theoretical level, we set out to demonstrate that some aspects of the structure of natural language are explainable diachronically but not synchronically. All human languages are the result of long, gradual processes of development. Only certain avenues of development are available, and these lead to certain structures. The fact that some structures are more common than others in the languages of the world does not necessarily mean that the more common structures are optimal from a psycholinguistic point of view. It may only mean that these structures arise more easily and more frequently than others. Since the primary sources of

affixes are words and stem changes can arise under conditioning from affixes, the fact that affixes are more widespread and common than stem changes has a straightforward diachronic explanation.

Note

We would like to thank Laurent Thomin, Patricia Escarraz, Sally Weller, and Valerie Daniel for their assistance in running subjects in this study. Correspondence address: Department of Linguistics, University of New Mexico, Albuquerque, New Mexico 87131.

References

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

———. 1988. Morphology as lexical organization. In *Theoretical morphology*, ed. Michael Hammond and Mickey Noonan, 119–141. San Diego, Calif.: Academic Press.

Bybee, Joan L., Patsy Fox, and Patrick Klinck. 1987. Are affixes more natural than stem changes? Paper given at the "Naturalness and Iconicity: Explanation in Historical Morphology" workshop at the VIIIth International Conference on Historical Linguistics, Lille, France, September.

Bybee, Joan L., and Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59, 251–270. (Reprinted in this volume as chapter 6.)

Bybee, Joan L., Revere D. Perkins, and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Bybee, Joan L., and Dan I. Slobin. 1982. Rules and schemas in the development and use of English past tense. *Language* 58, 265–289. (Reprinted in this volume as chapter 5.)

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Clark, Eve V., and Ruth A.Berman. 1987. Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child Language* 14, 547–567.

Dressier, Wolfgang U. 1985. *Morphonology: The dynamics of derivation*. Ann Arbor, Mich.: Karoma.

Green, T. R. G. 1979. The necessity of syntax markers: Two experiments with artificial languages. *Journal of Verbal Learning and Verbal Behavior* 18, 481–496.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Heine, Bernd, and Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

MacWhinney, Brian, Jared Leinbach, Roman Taraban, and Janet McDonald. 1989. Language learning: Cues or rules? *Journal of Memory and Language* 28, 255–277.

Marcus, Gary F., Ursula Brinkmann, Harald Clahsen, Richard Wiese, Andreas Woest, and Steven Pinker. 1993. *German inflection: The exception that proves the rule*. Center for Cognitive Science, Occasional Paper 47. Cambridge, Mass.: MIT.

Marcus, Gary F., Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen, and Fe Xu. 1992. *Overregularization in language acquisition*. Monographs of the Society for Research in Child Development 57, 4, Serial No. 228. Chicago: University of Chicago Press.

Mori, Kazuo, and Shannon D. Moeser. 1983. The role of syntax markers and semantic referents in learning an artificial language. *Journal of Verbal Learning and Verbal Behavior* 22, 701–718.

Pinker, Steven. 1991. Rules of language. *Science* 253, 530–535.
Prasada, Sandeep, and Steven Pinker. 1993. Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes* 8, 1–51.
Rumelhart, David E., and James L. McClelland. 1986. On learning the past tenses of English verbs. Implicit rules or parallel distributed processing? In *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 2: *Psychological and biological models*, ed. James L. McClelland, David E. Rumelhart, and the PDP Research Group, 216–271. Cambridge, Mass.: MIT Press.
Saddock, Jerrold. 1985. Autolexical syntax: A theory of parallel grammatical representations. *Natural Language and Linguistic Theory* 3, 379–439.
Traugott, Elizabeth C., and Bernd Heine (eds.). 1991. *Approaches to grammaticalization*, vols. 1 and 2. Amsterdam: John Benjamins.