On the Retrieval and Lexical Structure of Verbs
DONALD G. MACKAY*
University of California, Los Angeles

This study examined the retrieval of regular and irregular past tense verbs. Subjects were presented with present tense verbs (e.g., TEACH) and had to produce the past tense form (TAUGHT) as quickly as possible. Reaction times and errors in this task suggested that preterites such as TAUGHT are not stored as separate and independent lexical units but are formed from the verb stem (TEACH) by means of derivational rules. The form of the errors suggested that different phonological operations for the same stem are to some extent independently specified and apply to distinctive features rather than fully integrated phonemes.

This paper examines the retrieval of past tense verbs in the light of two general views on the organization of the internal lexicon: the Derivational hypothesis (DH) and the Independence hypothesis (IH). The definitive or essential component of DH is that stems and suffixes are separately stored in the brain and at least some derivatives such as AGREEMENT and some inflected forms such as TALK are generated by applying suffix rules to base forms resembling AGREE and TALK. A more extreme possibility under DH is that preterites such as DUG are produced by applying vowel alternation rules to a base resembling the present tense form DIG. The essential and contrasting component of IH is that all words are stored as separate and independent phonological units in the brain and are generated as fully integrated phonological units in natural speech production (by means of rather simple read-out mechanisms). Under IH, words such as AGREEMENT and TALK are stored and produced as fully integrated units. There need be no rules for adding affixes to stems or for altering vowels: preterites such as DUG are stored and produced as units without reference to the stem DIG.

These hypotheses appear in many different guises in models of the internal lexicon and sometimes generate identical empirical predictions depending on the nature of other assumptions of the model. But models of lexical storage and retrieval must incorporate either IH or DH or both and these lexical storage hypotheses carry strong implications for the nature of the semantic or conceptual store. For example, if preterites such as DUG are stored as independent units at the phonological level, a unitary semantic representation seems likely as well. But if preterites are generated by applying alternation rules to a base resembling the present tense form, then the semantic representation for retrieving DUG must be componential, resembling [DIG]+[Past]. Evidence favoring DH constitutes evidence favoring rules for combining concepts or semantic components.

At present there is no conclusive support for DH. Berko (1958) showed that young children added regular preterite markers to nonsense words such as GLING in a past tense context, while adults had much more difficulty deciding on an appropriate response and seemed to be torn between GLINGED and GLANG. Neither result constitutes
conclusive evidence for either regular or irregular preteritization rules: GLINGED or GLANG could have been formed by analogy with phonologically integrated prototypes such as WINGED or RANG. Alternatively, new words such as GLANG may be created using separately stored rules which have nothing to do with the storage and retrieval of already existing words (see Halle, 1973, for example).

Current evidence can readily be handled by either IH or DH. However, there are strong priori arguments in favor of IH for the retrieval of irregular Anglo-Saxon preterites such as DUG. Being relatively few in number and frequent in use, irregular preterites could easily be learned and stored as independent, rote units. And the limited scope of alternation rules for irregular preteritization suggests that acquiring these rules would serve little purpose. For example, the vowel alternation rules for HIDE-HID or RIDERODE seem to provide little simplification of the lexicon, since few other verbs form their preterite in exactly this way. Such considerations suggest that irregular preterites are produced as fully integrated units rather than by means of vowel alternation rules.

We conducted two experiments to test between IH and DH for irregular preterites. In the first experiment, subjects heard present tense verbs such as DIG and had to produce the simple past (DUG) as fast as possible. Production time was the dependent variable, and the following four classes of stimuli were the independent variables. Regular T-suffix verbs, for example, WALK-WALKED; Irregular T-suffix verbs, for example, SINGLE-GLANGED; Regular T-suffix verbs, for example, TAUGHT-TEACH; Irregular T-suffix verbs, for example, HIDED-HIDE.

In the second experiment, we introduced 20 practice verbs into the experimental conditions and controlled for frequency, phonemic, and acoustic factors. The materials were also controlled for phonemic and acoustic factors, for the nature of the word-initial segments, and for ambiguity or homophony. There were, in addition, 20 practice verbs similar to those in the experimental classes described above.

**Experiment I**

**Method**

The subjects were 12 native speakers of English enrolled at UCLA (mean age, 18.7). None reported a history of abnormal hearing or articulatory difficulty in response to a formal questionnaire. The following instructions were read to each subject:

This is an experiment on the transformation of English verbs. You will hear verbs in present tense, e.g., WRITE, and your task will be to transform them into past tense as quickly as possible. For example, if you hear TALK, you would respond TALKED. After you have given your answer you are to write down on your check sheet what you thought the stimulus verb was so that we can determine whether you heard the verb correctly. Are there any questions?

The experimental materials consisted of 20 verbs, shown in Table 1. All were present tense, three segments long, and had CVC structure. The stimuli fell into four conditions: Regular T-suffix verbs, for example, KISS-KIST (N = 5); Irregular T-suffix verbs, for example, TEACH-TEACHED (N = 5); Simple vowel alternation, for example, TEAR-TORE (N = 5); and Complex vowel alternation, for example, HIDE-HIDED (N = 5). Stimuli and responses in the Regular and Irregular T-suffix conditions were controlled for frequency using Carroll, Davies, and Richman (1971), as were those in the Simple and Complex vowel alternation conditions. The materials were also controlled for phonemic and acoustic factors, for the nature of the word-initial segments, and for ambiguity or homophony. There were, in addition, 20 practice verbs similar to those in the experimental classes described above.

**TABLE 1**

<table>
<thead>
<tr>
<th>Complexity Level 1: Regular ending</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. T ending</td>
</tr>
<tr>
<td>chip chipet kick kickt bake baket soakt soakt</td>
</tr>
<tr>
<td>kist kist walk walkt lead leakt fake faket</td>
</tr>
<tr>
<td>2. D ending</td>
</tr>
<tr>
<td>gain gained rob robt rub rubt chill chill</td>
</tr>
<tr>
<td>seal seald rig rigt fall faild love loved</td>
</tr>
<tr>
<td>lag laged time timed</td>
</tr>
<tr>
<td>3. ED ending</td>
</tr>
<tr>
<td>pet petted fit fitted side wade waded</td>
</tr>
<tr>
<td>dot dotted cheat cheated heed weeded weed</td>
</tr>
<tr>
<td>pat patted beat heated raid raider chide chided</td>
</tr>
<tr>
<td>bat batted rate rated need needed fade faded</td>
</tr>
<tr>
<td>pit pitted wait waited nod nodded shade shaded</td>
</tr>
</tbody>
</table>

**Complexity Level 2**

<table>
<thead>
<tr>
<th>Complexity Level 3: Vowel and glide alternation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vowel alternation</td>
</tr>
<tr>
<td>dig dug ring rang wear wore run ran</td>
</tr>
<tr>
<td>tear tore win won sing sang sit sat</td>
</tr>
<tr>
<td>fall fell get got</td>
</tr>
<tr>
<td>2. Consonant alternation</td>
</tr>
<tr>
<td>bend bent make made send sent have had</td>
</tr>
<tr>
<td>lend lent build built</td>
</tr>
<tr>
<td>3. Glide alternation</td>
</tr>
<tr>
<td>wind wound</td>
</tr>
</tbody>
</table>

**Complexity Level 4: Vowel, glide, and consonant alternation**

<table>
<thead>
<tr>
<th>Complexity Level 4: Vowel, glide, and consonant alternation</th>
</tr>
</thead>
<tbody>
<tr>
<td>teach taught keep kept leave left tell told</td>
</tr>
<tr>
<td>mean meant kneel knelt lose lost feel felt</td>
</tr>
</tbody>
</table>
The stimuli were recorded clearly and at normal rate on two tapes in differing random orders, approximately one every 10 sec. Half the subjects heard one tape, and half heard the other. Responses were recorded on a Revox A77, and stimuli were presented over Pioneer stereo earphones (Model SE 30A) using a TEAC Model 4070 (attached to a Model KA2000 Kenwood solid state stereo amplifier). Each stimulus word set off a Model SI Standard electric timer via one channel of a Model 6602A Lafayette voice key. A Monarch Model TM-18 microphone was situated about 3 in. from the subject’s lips so that his responses triggered the voice key and stopped the timer.

**Results**

Since Experiment II was designed to investigate errors in the production and perception of tense, we excluded the infrequent misperceptions ($N = 8$) and response errors ($N = 22$) from detailed analysis in the present study.

The lag time of the voice key was determined by comparison with the printout of a Siemens Oscillograph (tape speed 250 mm/sec) receiving identical inputs. The average lag time (approximately 6 msec) varied with the word-initial characteristics of the stimuli, but did not differ across conditions. The response time for each stimulus trial was determined by subtracting stimulus duration from voice key time, giving stimulus offset to response onset times. Stimulus duration was determined by Oscillograph (paper speed 100 mm/sec). Subtracting stimulus duration from voice key time seemed essential, since stimulus duration varied substantially across conditions and since offset to onset time was the most conservative measure with respect to the hypotheses under test. Less than 0.007 of the responses were lost due to technical difficulties with the voice key.

Correct response times are shown by condition in Table 2. Response times for Irregular T-suffix verbs, for example, TAUGHT,

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Response time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular T-suffix verbs</td>
<td>0.819</td>
</tr>
<tr>
<td>Irregular T-suffix verbs</td>
<td>0.910</td>
</tr>
<tr>
<td>Simple vowel alternation</td>
<td>0.690</td>
</tr>
<tr>
<td>Complex vowel alternation</td>
<td>0.858</td>
</tr>
</tbody>
</table>

*RTs in seconds.

were about 90 msec longer than for Regular T-suffix verbs, for example, TALKT $F(1, 11) = 16.95, p < .001$, and response times for Simple vowel alternation, for example, DUG, were about 160 msec shorter than for Complex vowel alternation, for example, GAVE, $F(1, 11) = 18.63, p < .001$. There were significant subject differences in both analyses ($p < .001$), but no significant subject by condition interactions ($p > .10$). In order to handle words as well as subjects as random effects (cf. Clark, 1973, p.348; Winer, 1971, p. 364), the data were analyzed by word (averaged across subjects) and the Simple vs. Complex vowel alternation conditions were compared by means of a Mann-Whitney test for independent samples, the significant results of which ($U = 3, p < .05; U = 2, p < .05$) allow generalization of the data to new samples of words as well as subjects.

Post-hoc analyses showed an effect of distinctive feature differences between the nuclear vowel of stem and preterite in the three irregular preterite conditions. Using standard distinctive feature analyses (cf. Gleason, 1961), the vowels required a change in either one distinctive feature (e.g., SI-M, or more than one distinctive feature (e.g., TEAR-TORE). Response times for single feature changes were significantly faster than for multiple feature changes ($p < .01$, sign test with subjects as unit of analysis). Shadowing latencies for the 20 stimuli were determined by Oscillograph for two control subjects instructed to repeat the auditorily presented stimuli as rapidly as possible. The onset to onset shadowing latencies did not differ across conditions, suggesting that the main results were not due to low level phonetic factors.

**EXPERIMENT II**

Experiment II was similar to Experiment I, but included many more subjects and stimuli, so as to study the nature of errors in perception and production of tense. The stimuli were not designed for reaction time analyses, since it was impossible with such a large corpus ($N = 174$) to control for relevant factors such as frequency, acoustic duration, phonological length, and ambiguity or homophony (cf. Appendix for further discussion). The main independent variable was Complexity level, a measure of the number of phonological differences between stem and preterite. Irregular verbs requiring no phonological changes, for example, PUT-PUT, HIT-HIT, were designated Complexity Level 0. Regular verbs, for example, TALK-TALKT, being phonologically more complex than Level 0 verbs but lexically less complex than Level 2, 3, or 4 verbs (see Discussion), were designated Complexity Level 1. Irregular verbs requiring a single change—either vowel alternation or glide alternation, for example, TALK-TALKT, or glide alternation, e.g., FIND-FOUNDE, were designated Complexity Level 2. Complexity Level 3 verbs required vowel alternation along with addition, deletion, or substitution of a glide, for example, GIVE-GAVE (glide addition); LEAD-LED (glide deletion); and WADE-WADED (glide substitution). And Complexity Level 4 verbs required glide, vowel, and consonant alternation, for example, TEACH-TAUGHT.

The main question was whether the errors would provide further support for derivational processes. A secondary question was whether irregular verbs were stored with the feature [+Irregular] to block the application of regular rules. Other questions concerned the nature of phonological rules: do alternation rules apply to distinctive features rather than fully integrated phonemic units? Are various alternation rules for the same stem to some extent independent of one another, for example, is the vowel alternation rule for TEACH-TAUGHT independent of the consonant alternation rule? Finally, are alternation rules ordered or applied in sequence as postulated by Chomsky and Halie (1968)?

**Method**

The subjects were 24 undergraduates from the UCLA subject pool (mean age 21.3). Instructions were identical to those in Experiment I, except that subjects were informed that some stimuli had to be transformed from past to present tense. The stimuli are shown in Table 1. There were 58 practice verbs drawn from the experimental classes described below. Half the verbs were present (e.g., BEND) and half were past (e.g., BENT). Most (95%) of the present tense verbs were three segments long and had CVC structure. About half were irregular ($N = 84$) and half regular ($N = 80$). Suffixes for regular verbs were -T, for example, TALK-TALKT ($N = 20$); -D, for example, LOVE-LOVED ($N = 20$); -ED, with stem-final -T, for example, WAIT-WAITED ($N = 20$); and -ED with stem-final -D, for example, WADE-WADED ($N = 20$). There were 10 verbs of Complexity Level 0, for example, CUT-CUT; 34 of Complexity Level 2, for example, DIG-DUG (vowel alternation), BEND-BENT (consonant alternation); FIND-FOUND (glide alternation); 18 of Complexity Level 3, for example, GIVE-GAVE; and 22 of Complexity Level 4, for example, TEACH-TAUGHT. Stimuli were recorded in different random orders on four tapes with random assignment of subjects to tapes.

**Results**

Errors fell into two mutually exclusive classes, discussed below: misperceptions and response errors.

---

**TABLE 2**

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Response time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular T-suffix verbs</td>
<td>0.819</td>
</tr>
<tr>
<td>Irregular T-suffix verbs</td>
<td>0.910</td>
</tr>
<tr>
<td>Simple vowel alternation</td>
<td>0.690</td>
</tr>
<tr>
<td>Complex vowel alternation</td>
<td>0.858</td>
</tr>
</tbody>
</table>
Misperceptions.

Misperceptions were scored whenever the word written on the check sheet failed to match the intended stimulus (obvious misspellings ignored). Of 96 misperceptions, 88 seemed to represent genuine mishearings, and eight were the result of homophony (two different lexical items having identical sound, e.g., WORE and WAR). Since the stimuli were uncontrolled for homophony, it is perhaps not surprising that subjects sometimes comprehended a verb in an unintended homophonic sense (e.g., WINED for WIND, FINED for FIND, BILLED for BUILD).

But homophonic misperceptions displayed an unexpected systematicity connected with the distinction between regular versus irregular verbs. Only irregular verbs were perceived as unintended homophones. And the perceived homophones (e.g., FINED) invariably had lower frequency of usage than the intended stimuli (e.g., FIND). However, regular verbs never were perceived homophonously, despite ample opportunities; for example, WADE was never taken as WEIGHED, or NEED as KNEED.

The remaining misperceptions were classified as mishearings (N = 88) and included vowel substitutions, for example, FILL for FAIL (N = 7); consonant substitutions, for example, RAPED for RAKED (N = 76); and consonant additions, for example, HEAPT for HEAT (N = 5). Substitute and target consonants generally differed by only a single distinctive feature, usually place of articulation, for example, BOUGHT for GOT (N = 53).

Mishearings did not vary with Complexity Level and were not due to poor quality recording of certain words, since any one word was never misheard more than once per tape. Subjects invariably made no response when stimuli were misheard as nonverbs (N = 4), but always carried out an appropriate transformation for stimuli misheard as other verbs (N = 92), for example, GOT misheard as BOUGHT was transformed to BUY.

Response errors. Response errors were scored whenever subjects perceived the stimulus correctly but responded incorrectly. Two major classes of response errors, regularizations (N = 104) versus nonregularizations (i.e., all other errors, N = 349), showed basically different relations with derivational complexity and are treated separately below.

Nonregularizations. Nonregularizations, discussed in detail below, included mistransformations (N = 201) and nontransformations (N = 47) of vowels and consonants, misinflections (N = 64), backformations (N = 6), and stutterings (N = 31). Nonregularizations were, in general, more common for past than present tense stimuli (x^2 = 14.87, p < .001), and more common for irregular than regular verbs, (x^2 = 11.02, p < .001), and increased in frequency as a systematic function of Complexity Level, as shown in Fig. 1.

Mistransformations. Mistransformations (N = 201) were scored whenever vowels or consonants underwent either partial or inappropriate alternation, defined below. In cases of partial alternation (N = 83), some but not all of the appropriate derivational steps were carried out. An example is CATCH-CAT (instead of CAUGHT), where consonant alternation occurred, but not vowel alternation. A more complex example of partial alternation was MAKE-MAYG (instead of MADE): the /G/ has the wrong place of articulation but the right voicing, as if one distinctive feature had undergone alternation, but not the other. A second type of mistransformation was inappropriate alternation. In inappropriate alternations (N = 118), subjects misapplied alternation rules appropriate to another class of lexical items. An example is RIDE-RID (rather than ROED), where alternation rules appropriate for other lexical items, for example, HIDE-HID, seem to have been misapplied. In no case did an inappropriate feature alternation occur which would not have been appropriate for some other item in the internal lexicon. Curiously, however, inappropriate vowel alternations often resulted in real words (N = 62), which were sometimes related in meaning to the stimulus verb, for example, PAT-PET (instead of PATTED). Whether this phenomenon was an artifact associated with the use of short CVC words is at present unknown.

Inappropriate consonant alternations were rare (N = 6), and always occurred in conjunction with vowel alternation in irregular verbs, for example, DIG-GUG (rather than DUG), where the vowel underwent the appropriate (backing) alternation, and the initial consonant underwent an analogous but inappropriate (backing) alternation. Unlike inappropriate vowel alternations, inappropriate consonant alternations may reflect inadvertent side effects of vowel alternation rather than misapplication of otherwise possible consonant alternation rules.

Nontransformations. Nontransformations were scored whenever subjects reiterated a stimulus without making any change whatsoever (excluding Complexity Level 0 verbs, where reiteration is appropriate). Typical examples are BUILD-BUILD (instead of BUILT) and RUN-RUN (instead of RAN). Like other nonregularization errors, nontransformations were more frequent for irregular than regular verbs, and increased systematically with Complexity Level (as in Figure 1). But unlike other nonregularizations, these errors were significantly more common for present than past tense stimuli, x^2 = 4.27, p < .05.

Backformations. Backformations were scored whenever a word-final -T or -D was dropped, leaving a nonviable verb stem, for example, LEFT-LEF (instead of LEAVE), MEANT-MEN (instead of MEAN). Backformations always involved preterites of Complexity Level 4 (as in the above examples), and were not due to a general phonological strategy of dropping stem-final dentals, since there were no errors of the form SIT-SI (instead of SAT), HEEED-HEE (instead of HEEDED) or HIDE-HI (instead of HID). Backformations in the present study, therefore, differ from the superficially similar outputs of children (who, out of ignorance, might produce HEE as the stem of HEED) and from historical backformations, for example, the change from PEAS (singular) to PEA in the history of English.

Misinflections. Misinflections were scored whenever subjects produced inappropriate inflectional forms, for example, TAKE-TAKEN (instead of TOOK), DIG-DIGS (instead of DUG), and RIDE-RIDING (instead of ROE), and apparently reflect temporary amnesia for the instructions to produce simple-past or present tense forms. But not just any such errors are misinflections. Derivational suffixes such as -ER (e.g., RIDE-RIDER) or suffixes appropriate to other syntactic categories (e.g., RIDE-RIDELY) were never mistakenly added, but only verb inflection suffixes as in the above examples.

Stuttering. Thirty-one errors closely resembled misarticulations characteristic of pathological stuttering. The subjects either repeated an initial consonant (e.g., SANG-S-SING), prolonged an initial consonant (e.g., SANG-S-ING), or made false starts (e.g., ROSE-RO-RISE). False starts resembled error corrections (N = 276, e.g., SUNK, I mean SANK) except that only...
part of the word was uttered prior to correction.

These findings are consistent with the hypothesis that stuttering in “nonstutterers” reflects a process of error correction. False starts are directly explained under an error correction hypothesis, and repetitions and prolongations increase with derivational complexity and occurred mainly with past rather than present tense stimuli, and mainly with irregular rather than regular verbs. But unlike false starts, prolongations and repetitions must reflect anticipatory corrections of programming errors which never become manifest in the surface output.

Regularizations. Regularizations were scored whenever subjects mistakenly added regular suffixes to irregular verbs, for example, MAKE-MAKET, GIVE-GIVED, CATCH-CATCHHT, HIDE-HIDED, WEAP-WEAPT. As in the above examples, the regularizations always followed appropriate rules for regular suffixation, for example, -T was never added to stems ending in voiced segments. And regularizations were significantly more common for present than past tense stimuli ($X^2 = 95.92$, $p < .001$). In fact, there were no instances of “unirregular” (addition of regular markers to irregular preterites, e.g., WALKT-WALKTED) and only two instances of “semiregularization” (addition of regular markers to irregular preterites, e.g., G0T-GOTED and DIG-DUGED).

The probability of regularization per subject per verb is shown as a function of derivational complexity in Fig. 1. As can be seen there, the data show an interesting difference between Level 0 vs. other irregular verbs. Post-hoc analyses indicated that regularizations were significantly more frequent for Complexity Level 0 verbs than Levels 2 to 4, $X^2 = 15.89$, $p < .001$, between which there was no significant difference ($p > .10$, same test). It should be noted however, that 16 regularizations of Level 4 verbs (e.g., KEEP-KEEPT) are ambiguous, since they could have been counted as partial alternations, but were not.

**General Discussion**

Taken as a whole, the data supported DH rather than IH as the mechanism for retrieving preterites at the phonological and semantic levels. Consider first the time data of Experiment I. As predicted under DH, response times were faster for Regular than Irregular T-suffix verbs. This difference clearly fits the hypothesis that producing irregular preterites such as KEPT requires additional rules not necessary for producing regular preterites such as WALKT.

This finding is probably not due to an inherent difference between regular versus irregular regularization rules, in view of the longer response times for Complex versus Simple vowel alternation. This latter difference provides the strongest support for DH as opposed to IH, and rules out the hypothesis that suffixes such as -T or -ED are necessary for triggering the acquisition and application of vowel alternation rules. The fact that reaction time increased as a function of the number of distinctive feature changes further supports DH and suggests that vowel alternation rules apply to distinctive features rather than fully integrated phonemes.

Error data from Experiment II also supported DH, but were for the most part difficult to explain under IH. For example, IH cannot readily handle regularizations (e.g., FALL-FALLED) or backformations (LEFT-LEFTF), since forms such as FALLED and LEFT do not appear in the internal lexicon and in all likelihood were never before encountered by the subjects. The IH also fails to explain why regularizations occurred mainly with present tense Level 4 verbs. Equally problematic for IH are mistransformation errors such as MAKE-MAYG, since MAYG is clearly not stored in the internal lexicon. And although IH is not incompatible with instances of nontransformation (e.g., RUN-RUN) or misinflections (e.g., DIG-DIGS instead of DUG), IH fails to explain why nontransformations were more common for present than past tense stimuli, or why misinflections only involved inflectional affixes and not, say, nominalization affixes (e.g., RIDE-RIDER).

However, DH readily explains both the nature and distributional characteristics of these errors. Consider first the partial alternations and backformations. Partial alternations such as CAUGHT-CAT (rather than CATCH) are readily explained under DH as due to partial or incomplete application of independently specified derivational rules. Partial alternations such as MAKE-MAYG (instead of MADE) are similarly explained under DH, and further support the hypothesis that alternation rules apply not to phonemes as fully integrated units, but to distinctive features (see also Chomsky & Halle, 1968).

Backformations (e.g., KEPT-KEEP) also can be explained under DH as a special type of partial alternation: due to time pressure, one of the derivational steps was retraced (giving SOL or LEF), but not the other (to give SELL or LEAVE). Viewed as partial alternations, backformations clearly support the suggestion that Level 4 preterites such as SOLD are derived from a base resembling SELL by at least two steps: a nearly regular suffix rule (giving SELL + D) and a complex vowel alternation rule (giving SOLD). Together, backformations and partial alternations suggest that different phonological operations for the same stem are, to some extent, independently specified. These errors are also consistent with an Incomplete Specification hypothesis. Under this hypothesis, alternation rules apply sequentially, so that partial alternations and backformations reflect incomplete or partial specification of the sequence of phonological operations (due to haste or other factors). However, the data provide no conclusive support for the assumption that alternation rules are ordered.

Consider now the misinflections, nontransformations, mistransformations and regularization errors, which are readily captured under DH as results of rule misapplication. Mistransformations such as RIDE-RID (rather than RODE) are readily explained as misapplications of alternation rules appropriate for other lexical items. And under DH, the misinflections (such as DIG-DIGS instead of DUG) reflect application of inappropriate verb inflection rules. Thus, not just any suffix was added in misinflection errors because the subjects were applying inflectional rather than derivational rules, such as, say, nominalization. As would be expected under this hypothesis, subjects instructed to nominalize verbs (e.g., DECADE-DECIDES) often added inappropriate nominalization suffixes (e.g., DECIDE-MENT), but not inappropriate in-<tional suffixes (e.g., DECIDES) (MacKay, 1974).

Nontransformations (e.g., RUN-RUN) may also reflect rule misapplication whereby subjects misapplied to other irregular verbs the null alternation rule appropriate only for Level 0 verbs such as HIT-HIT. Under this hypothesis, subjects producing nontransformations coded stimuli such as RUN as present tense and irregular, but simply applied the wrong alternation rule for forming the preterite, an hypothesis which explains why nontransformations were more frequent for present than past tense verbs, and more frequent for irregular than regular verbs.

Regularizations are also readily explained under DH as due to misapplication of rules for regular preteritization, but it remains to explain why Level 0 verbs (e.g., PUT, HIT) were much more susceptible to regularization than other irregular verbs. The high frequency of regularization for Level 0 verbs was probably not due to the fact that Level 0 verbs always end in a dental consonant (T or D),
since subjects did not regularize other dental-final verbs (e.g., TAUGHT, Sought, or READ) with higher than average probability. Rather, subjects making these errors must simply have failed to recognize that Level 0 verbs were irregular, which suggests the interesting possibility that Level 0 verbs may be stored without lexical features such as [+Irregular], and are, in fact, the most primitive verbs at the lexical as well as phonological levels. That is, speakers may produce Level 0 verbs in present tense form unless situational or deictic context (e.g., a temporal adverb, such as YESTERDAY) specifies otherwise. These contextual features would serve to block application of subject agreement rules, giving HE HIT HIM YESTERDAY, rather than HE HITS HIM YESTERDAY. Under this as yet untested hypothesis neither the feature [+Past] nor [+Irregular] are necessary for producing Level 0 preterites.

In summary, the error data are in harmony with the time data and can be explained under DH as due to either misapplication of inappropriate alternation rules or incomplete application of appropriate alternation rules. Like the production time data, the errors also support the hypothesis that derivational rules apply to distinctive features rather than fully integrated phonemes and suggest that vowel and consonant rules for the same stem are applied independently.

We turn now to alternative explanations. Limited aspects of the present data are open to alternative explanations, to be designated the Lexical Marker hypothesis and the Search Procedure hypothesis. Under the Lexical Marker hypothesis, regular preteritization requires fewer lexical markers than irregular preteritization. That is, in addition to the feature [+Past], irregular preteritization requires activation of the feature [+Irregular] for triggering retrieval of the irregular form and for blocking application of the regular rules. Activation of the extra feature [+Irregular] may thus account for the longer response times for irregular preteritization.

Under the second alternative, the Search Procedure hypothesis, regular preterites follow DH and are produced by rule, but irregular preterites follow JH and are produced as independent phonological units, which can only be located or retrieved by complex and time-consuming search procedures. The longer response times for irregular versus regular preteritization may thus reflect the additional time required to locate the lexically stored irregular preterites.

But neither the Lexical Marker hypothesis nor the Search Procedure hypothesis is capable of explaining the longer response times for complex vs. simple vowel alternations or the nature and distributional characteristics of the errors and so must be discarded as general explanations of the present findings.

Consider now the limitations of the present data. The data leave unanswered the question of whether alternations are ordered or apply in sequence and whether irregular verbs are stored with the feature [+Irregular] to block the application of regular rules. The data also fail to delimit the scope of DH. For example, it seems likely that suppletive forms such as GO-WENT are learned by rote and produced as independent units at the phonological level, since suppletive alternation rules seem so unnatural and complex as to be unlearnable. Suppletive preterites thus provide an interesting domain for testing the Search Procedure hypothesis. The Search Procedure hypothesis postulates longer retrieval times for lexically stored units, so that reaction times for suppletive preteritization, for example, GO-WENT, should be longer than for rule-generated preteritization for example, KEEP-KEPT.

Another limitation of the present study is that DH and JH are not necessarily mutually exclusive: Words may be stored both as independent units and as morphological components with rules for vowel and consonant alternation. And given more than one memory system for words, it may be possible to retrieve irregular preterites as independent units from some other memory system not accessed in the present task. Moreover, the present data only indicate that speakers can use derivational rules for forming preterites. Whether speakers normally use such rules in producing sentences is clearly an open question, although speech error data (MacKay, 1974; Fromkin, 1973) suggest that these rules do play a role in natural speech—at least for some speakers some of the time. However, neither error nor time data tell us much about the form of these rules, except that they probably apply to distinctive features rather than phonemes as integrated units. The data also bypass the question of why irregular alternation rules are learned in the first place. As pointed out in the introduction, straightforward economic criteria cannot explain the acquisition of irregular preteritization rules. But the possibility of more complex economic factors deserves exploration. If lexical economies encompasses the entire internal lexicon rather than individual subsections, then irregular preteritization rules may be more economic than originally supposed, since other derivational paradigms require identical rules. For example, the vowel alternations for SIT-SAT, HIDE-HID, and LEAD-LED also appear in adjective (PARADIGM-PARADIGMATIC), nominal (BREATHE-BREATHE), and semantic derivations (e.g., FLIP-FLAP-FLOP, CLINK-CLANK-CLUNK). We cannot, at present, reject the possibility that learning rules for irregular preteritization may be simpler than storing and producing as units all the items requiring these rules.

APPENDIX: DERIVATIONAL COMPLEXITY AND WORD FREQUENCY

A number of problems were encountered in determining the word frequency of materials for Experiment II. As these problems appeared to have general relevance for experiments on morphology, as well as theoretical significance in their own right, they are discussed in detail below.

The first set of problems concerned the relative frequency of corresponding present and past tense forms. Determining the frequency of present versus past tense forms appeared important for two major reasons. First, controls for word frequency seemed necessary since numerous studies have shown that word frequency is an important variable in recognition and production of words (see, for example, Solomon & Postman, 1952). Second, relative frequency of present versus past tense forms was of interest in its own right, since Greenberg (1966) postulated that unmarked forms such as present tense TALK are universally more frequent than marked forms such as past tense TALKT. But contrary to Greenberg's hypothesis, Svartvik (1966) reported that pasts outnumbered presents in English. However, Svartvik's data seemed ten to question, as his analyses were based on a limited number of sources (two novels and a biology text), which may not be representative of usage within the language as a whole. Moreover, there were vast differences between his sources: pasts were only more frequent than presents for the novels (cf. Table 3), while the opposite was true for the biology text, which presents were overwhelmingly more frequent than pasts (98% presents vs. 2% pasts). Such source-dependent variability suggests that frequency analyses based on written materials must consider large and representative samples from many sources.

Large and representative word counts are of course readily available. For an example, the Standard Frequency Index (SFI) of Carroll, Davies, and Richman (1971) seem highly representative of written English, being based on 10,043 samples from 6162 books in 22 fields. As an added advantage, the SFI provides an adjustment for the dis-
Since pasts such as TALK are longer than excused when either the stem or preterite analysis would confound length and frequency in a manner favoring Greenberg’s hypothesis, with more than one form class. A second problem concerned the length (in phonemes) would distort the relative frequency of present to overcome possible sources of bias. For example, present tense forms such as FALL and RUN are ambiguous as to noun versus verb, whereas their corresponding preterites, FELL and RAAN, are not. Since such ambiguity could distort the relative frequency of present versus past tense forms, verb pairs were excluded when either the stem or preterite appeared in Webster’s Collegiate Dictionary with more than one form class. A second problem concerned the length (in phonemes) of stems and preterites: a simple frequency analysis would confound length and frequency in a manner favoring Greenberg’s hypothesis, since pasts such as TALKT are longer than presents such as TALK and length generally correlates negatively with frequency (Zipf, 1965). As a control for length, the verbs were divided into three categories: (1) regular, for example, WALK-WALKT (with preterites longer than present tense stems); (2) same length irregular, for example, HIDE-HID (with stems and preterites of equal length, diphthongs counted as single phonemes); (3) different length irregular, for example, KEEP-KEPT (with preterites longer than stems). The verb stems were all either one or two syllables long.

The analyses called for several controls to overcome possible sources of bias. For example, present tense forms such as FALL and RUN are ambiguous as to noun versus verb, whereas their corresponding preterites, FELL and RAAN, are not. Since such ambiguity could distort the relative frequency of present versus past tense forms, verb pairs were excluded when either the stem or preterite appeared in Webster’s Collegiate Dictionary with more than one form class. A second problem concerned the length (in phonemes) of stems and preterites: a simple frequency analysis would confound length and frequency in a manner favoring Greenberg’s hypothesis, since pasts such as TALKT are longer than present tense forms such as TALK and length generally correlates negatively with frequency (Zipf, 1965). As a control for length, the verbs were divided into three categories: (1) regular, for example, WALK-WALKT (with preterites longer than present tense stems); (2) same length irregular, for example, HIDE-HID (with stems and preterites of equal length, diphthongs counted as single phonemes); (3) different length irregular, for example, KEEP-KEPT (with preterites longer than stems). The verb stems were all either one or two syllables long.

The data in Table 3 show the frequency of present versus past tense form translated into percentages. The present tense frequency included inflected forms such as HIDES as well as the uninfl.ected form HID. Taken together, present tense forms were significantly more common than their corresponding preterites (p < .001, sign test with past vs. present tense forms as variables). More importantly, present tense forms were more frequent than preterites for irregular verbs with equally long present and past tense forms (p < .001, same test). This finding therefore appears consistent with Greenberg’s hypothesis that present tense forms are more common than their corresponding preterites, independent of phonological length.

But the data still contain an important source of bias, since frequency of present tense forms includes occurrences of infinitives, which are inherently neither present nor past tense, for example, the infinitive TO POUR in a sentence such as SHE WANTED TO POUR THE DRINKS. To estimate the importance of this source of bias, we determined the relative frequency of infinitives as opposed to present tense use of verbs for five 1000-word samples from six sources (three novels and three scientific texts). The results suggested that relative frequency of present tense forms was overestimated by approximately 12% in Table 3. We therefore recomputed relative frequency of present versus past tense forms, but again present tense irregular forms were significantly more frequent than their corresponding same length preterites (p < .01, same test). These findings clearly call into question explanations of experimental data based on the assumption that preterites are more frequent than present tense forms (e.g., Clark & Stafford, 1969). However, the present findings do not explain why present tense forms are more frequent. One possible explanation may lie in the variety of semantic functions or uses made of present as compared to past tense forms (cf. Harris & Brewer, 1973). An alternative explanation can be phrased in terms of the underlying operations required to produce present versus past tense forms. That is, we may produce preterites such as DUG or WALKT by applying past tense rules to bases resembling the present tense form. This being the case, present tense forms may be more frequent because they are simpler or easier to produce than their corresponding preterites.

A second, somewhat independent, set of problems concerned the relationship between frequency and derivational complexity of irregular preterites. Our analyses indicated a curious intersection between age, frequency of use, and derivational complexity of 21 irregular preterites. These verb forms had CVC structure in either present or past tense form (or both), so as to control for phonemic length and syllabic complexity, and were obtained from Siobin (1971, p. 2.). The preterites fell into four categories of derivational complexity: Level 0 (e.g., PUT-PUT); Level 2 (e.g., FALL-FELL); Level 3 (e.g., WAKE-WOKE); and Level 4 (e.g., LOSE-LOST). General adult usage of these preterites was determined from Kuéera and Francis (1967), and is shown as a function of derivational complexity in Figure 2 (right ordinate). As can be seen there, general adult usage did not vary with derivational complexity: there were no appreciable differences in frequency between Complexity Levels. But usage varied systematically with derivational complexity in the speech of children and of adults speaking to children. The average relative frequency with which 24 children (aged 1.5 to 4.0) used these preterites (correctly or incorrectly) was determined from Siobin (1971, p. 221), and is shown as a function of derivational complexity in Fig. 2 (left ordinate). As can be seen there, frequency decreased with derivational complexity in the speech of these children. Adults talking to children showed a similar relation-
ship between frequency and derivational complexity for these preterites. The data were again obtained from Slobin (1971, p. 221), who recorded and tabulated the frequency of these preterites from 64 hours of speech to children (aged 1.5 to 4.0) by mothers and investigators under conditions of natural dialogue. The results are shown in Fig. 2 (right ordinate), and, as can be seen there, frequency of use by adults talking to children also decreased systematically with derivational complexity. These findings should be considered as tentative, since only a small number of verbs was examined (five per complexity category, on the average). Nevertheless, the data at hand suggest the interesting possibility that children may find complex derivatives more difficult to learn than simple derivatives, a phenomenon appreciated by adults and reflected in their frequency of use in talking to children.

References


HALLE, M. Prolegomena to a theory of word formation. Linguistic Inquiry, 1973, 4, 3-16.

(Received April 29, 1975)