Bruce Hayes
322 HGS

R. M. CENA

When is a Phonological Generalization Psychologically Real?

WHEN IS A PHONOLOGICAL GENERALIZATION PSYCHOLOGICALLY REAL?

BY

R. M. Cena
UNIVERSITY OF HAWAII

Reproduced by the
Indiana University Linguistics Club
310 Lindley Hall
Bloomington, Indiana 47401

August, 1978

WHEN IS A PHONOLOGICAL GENERALIZATION PSYCHOLOGICALLY REAL?[1]

In recent years, the notion of psychological reality has played an increasingly important role in discussions of the validity of phonological generalizations. When the validity of a generalization is in doubt, the question invariably asked is whether or not it has psychological reality. A psychologically real phonological generalization is characterized as one which applies to an open lexical domain. When a phonological generalization is formulable independently of specific lexical items, the generalization is said to express a true statement about the sound system of the language.

A behavioral interpretation typically applied on this characterization of a psychologically real phonological generalization is that speakers will make use of the generalization in the derivation of novel forms; if it is shown that they in fact do, this demonstration is then considered to be evidence for the psychological reality of the generalization. This interpretation has engendered an experimental paradigm for testing for psychological reality, whose main feature is that subjects combine morphemes into nonsense word forms, about which the generalization being tested makes certain predictions. There is nothing wrong with this interpretation; however, the experimental paradigm that arose from this interpretation cannot guarantee data relevant to the issue of psychological reality, particularly since in most of the studies that made use of this paradigm, the dependent measures consisted of what essentially are acceptability judgments.

One difficulty with experiments that rely on acceptability judgments is that it is not an easy matter to know the basis for the subjects' rejection or acceptance of a form, since there is always the likelihood that the subjects may be analyzing the materials differently than expected. It is true that in planning the experiment, we try to define the experimental problem as narrowly as possible, but with novel forms, it is unlikely that we can anticipate accurately the range of reactions that will be evoked. Irrelevant behavior is thus a threat that cannot be ignored, and the experimental situation must have a control for systematic but irrelevant behavior.

Consider, for example, Anisfeld's (1969) attempt to test for the reality of an intermediate representation. English has a small set of verb-adjective pairs that exhibit the alternations {t, d} ~ s, for example, *permit* - *permissive*, *submit* - *submissive*, *extend* - *extensive*, *decide* - *decisive*. In a generative description of these facts (see Chomsky & Halle 1968:229), a rule is proposed that changes the dentals *t* and *d* to the spirants *s* and *z*, respectively. This rule accounts for the *t~s* alternation but not for the *d~s* alternation.

An additional rule is required to devoice the intermediate segment *z*. Thus, in the *extend* - *extensive* derivation, *exten*[z]*ive* serves as an intermediate form. Anisfeld was interested in the reality of this form.

Synthetic verbs were constructed (e.g. *flamit*, *garlude*), and each verb was paired with *-ive*-adjective forms whereby in one condition the crucial sound was [z]. Subjects were asked to judge the acceptability of the *-ive*-forms as adjective counterparts of the verbs. On the assumption that in derivations of this sort [z] is an intermediate representation of /d/ but not of /t/, Anisfeld predicted that the pair *garlude* - *garluzive*, for example, would receive significantly better acceptability ratings than the pair *flamit* - *flamizive*. The results showed that this was just the case, and the conclusion of psychological reality for the intermediate representation was made.

The conclusion appears to be unwarranted. The fact has not been established that subjects based their judgments of the adjective forms as INTERMEDIATE forms, nor is this deducible from any aspect of the design of the experiment. It is plausible that the subjects treated the supposed intermediate forms as surface forms. Indeed, a reasonable explanation of the results could be based on phonetic distance. The segments *d* and *z* (in *garlude* - *garluzive*) differ in only one feature, which is spirantization, whereas the segments *t* and *z* (in *flamit* - *flamizive*) differ in two features, namely, spirantization and voicing. Moreover, there is evidence that the change from *t* to *z*, i.e. from voiceless to voiced, is less salient. Lackner and Goldstein (1975) have "verbal transformation" data showing that auditory misperceptions with respect to the feature voicing is unidirectional from voiced to voiceless, and therefore a change from *t* to *z* should be less favored. On the basis of these considerations, more favorable ratings should be expected for the pair that exemplifies *d* → *z*, which was precisely the pattern of results obtained. The results are thus explainable in terms unrelated to the issue of the reality of intermediate representations.

I believe that the notion PSYCHOLOGICALLY REAL PHONOLOGICAL GENERALIZATION can be given a much more explicit behavioral interpretation, for which there exist experimental paradigms with built-in controls against systematic but irrelevant behavior. These paradigms are thus able to generate data whose relevance to the issue of psychological reality is less in doubt.

Let us interpret a psychologically real phonological generalization as follows: A phonological generalization that serves as a variable in mental processing is psychologically real. Given this interpretation, the prediction can be made that scores associated with mental activities (perceiving, recognizing, learning, recall, etc.) will vary systematically directly as a function of the reality

or unreality of the generalization being tested. With regards to learning, the prediction is made that if a generalization has psychological reality, it will facilitate the learning of materials that exemplify the generalization, and interfere in the learning of materials that violate the generalization.

Now, in learning paradigms, feedback is employed to guide the subjects to focus attention on just those properties of the stimulus materials that are relevent. Suppose that some of the items to be learned are relatable by the rule being tested, while other items are not so relatable. If a subject makes irrelevant responses, negative feedback will force him to look for other response strategies. On the other hand, if he makes relevant responses, positive feedback keeps him on the right track. If at the termination of the learning session the subject shows better learning of the items relatable by the generalization, we assume, provided that proper control of other factors have been made, that the subject has discovered and made use of the generalization to learn the items, and that learning of the other items, on the other hand, has not been so facilitated since no comparable rules relate the items. Thus, feedback steers the subject to the aspects of the stimulus materials relevant to the experimental hypothesis. Experiments that rely on acceptability judgments do not have this control mechanism, since any and all responses will have to be accepted at face value.

It is within this explication of a psychologically real phonological generalization that the present experiment has been conducted.

ON THE TEMPLATE ROLE OF LINGUISTIC EXPERIENCE

Underlying the above behavioral interpretation of a psychologically real phonological generalization is the general assumption that linguistic experience (or, to use progressively stronger terms, habits, abilities, knowledge, competence) guides linguistic performance, in the natural setting as well as in the experimental laboratory. This general assumption has three components, namely, first, that some aspects of linguistic behavior are rule-governed; second, if there exists a useful generalization in the linguistic data, the mature speaker has discovered and internalized it (Nessly 1974); and third, an internalized generalization will be used in appropriate situations (Nessly 1974). These ideas are fundamental to all linguistic construct validation experiments.

There is sizeable support from classical verbal learning studies for the role of linguistic experience in performance in verbal experiments. In a review of these studies, Cofer (1969) concluded:

> *It is virtually impossible to design units which are independent of the S's language skills and background. The prominence we have given to meaningfulness, familiarity, association value, pronunciability, inter-word associations, sequential organization, category membership, and grammatical structures makes this abundantly clear. Past experience with the language results in skills and habits which interact, in an inexorable way, with contemporary verbal learning tasks.* (p. 364)

The nature of this interaction, Glanzer (1967) suggested, is in the form of proactive interference, where linguistic experience either inhibits or facilitates the learning of the experimental tasks. In the area of psycho-phonological research, there have been some studies that provide evidence for the facilitative or inhibitory effect of linguistic knowledge. A few of these studies will be discussed.

One of the earliest of these experiments is reported in Esper (1925). Esper compared the learnability of three artificial mini-languages that varied in their degree of approximation to English. One of the languages that had to be learned departed from certain syntactic and phonological structural properties of English (a confounding that unfortunately was not controlled for). The task required the subjects to learn the referents of the component words of compound words of the form CV##CCVC. The task essentially reduced to learning to locate the word boundary. Now the CC clusters used in the experiment do not occur in English in syllable-initial position (e.g. *lg*, *gd*, *zg*, *mb*, etc.), and therefore, it was difficult for the subjects, who were native speakers of English, to see that the second component word began with one of those clusters. As expected, subjects had difficulty learning this language.

While the increased difficulty of this language may be the confounded effect of both syntactic and phonological distortions, there is reason to believe that the phonological anomaly contributed a great deal of the errors. Esper pointed out that substitution errors showed that the subjects extended the first word to include the next consonant, resulting in a more natural CVC.CVC syllable division. Esper concluded that "in general, we find a tendency [to] modify the non-English syllable divisions ... in accordance with English speech habits" (p. 38). The phonological distortion apparently made the correct responses difficult to integrate, thereby rendering them less available as response items.

The phonological fusion experiments, initiated by Day (1968), have provided data on phonological mediation in the perception of the temporal order of segments in clusters. The technique requires presentation of dichotic pairs, e.g. *pay/lay*, one word to each ear, and the subject is asked to report what he hears. Day reported that

subjects typically heard *play*. There was fusion when both stimulus words were real words, as in *pay/lay* → *play*; when the stimulus items were nonwords and the fusions were real words, *banket/lanket* → *blanket* (Day 1970); and even when both stimulus and response were nonwords, *gorigin/lorigin* → *glorigin* (Day 1968). Fusion was observed even when presentation of *banket* was delayed by up to 100 msecs (Day 1970). In other words, even when the crucial sounds [b] and [l] were objectively ordered as [lb], subjects heard them as [bl]. This result was duplicated when the dependent measure was judgment of the temporal order of the crucial sounds. In the visual mode, comparable results were reported by Rommetveit (1968:99), in which *shar/shap* were reported to have been seen as *sharp*.

In explaining these misperceptions, Day (1970) suggested that, in the perception of linguistic materials, a "linguistic filter" is operating. The filter rejects *lbanket*, for example, and alters it to the acceptable *blanket*. Day wrote:

> *Before a subject can give a response, the results must be related to past experience with the language, perhaps by way of a linguistic filter or similar device. The filter operates on the basis of the sequential dependencies of phonemes in the language.* (p. 84)

The phonological constraint operating here is the restriction on the order of stop-liquid sequence in word initial position in English, which disallows [lb] but not [bl]. That a linguistic explanation is in order is given some support in another experiment (Day & Cutting 1970), in which the stimuli consisted of nonlinguistic materials. In this experiment, subjects made accurate judgments of the temporal order of signals. Furthermore, when the stimuli were reversible clusters, as for example, [sp] and [ps] in final position, reordering of sounds was not observed. The dichotic pair *lis/lip* was heard as *lisp* when *lis* led, and as *lips* when *lip* led. Phoneme sequence misperceptions did not occur since no phonotactic constraint restricts the ordering of the sounds [s] and [p] in cluster word finally.

The "verbal transformation" experiments have provided another line of support for the template role of natural language habits in the processing of linguistic materials. In the earliest of these experiments, Warren and Gregory (1958) reported a tendency by subjects to hear different words from an unchanging stimulus presented repeatedly over and over. The relevant point here is that subjects' reports show evidence that phonological resegmentations and reorganization of the stimuli are subject to the phonological constraints of the language of the subjects. For example, the sequence *tress - tress - tress* was reported to have been heard

as *stress - stress - stress* (Warren & Warren 1970). Even when subjects reported hearing nonsense words, these words tended to be segment sequences permissible in English (Warren 1961). A similar conclusion was reached by Moorehead (1970), who employed the technique to see if the phonological rules or strategies children use in processing non-English phoneme sequences are different from those used by adult speakers. He wrote: "Whenever a semantic or phonetic change was made, [the subjects] changed the original sequence to English in the most economical way according to the rules of English phonology" (p. 62).

Certain linguistic habits may be conditioned by universal properties of languages. An experiment by Schane, Tranel, and Lane (1974) shows evidence for the psychological reality of a linguistic universal. Schane et al. were interested in whether an artificial mini-language system exemplifying a natural rule would induce more learning than one that exemplifies an unnatural rule, even though the natural rule has limited application in the language of the subjects. An example of a natural rule is one that deletes a consonant before a word beginning with a consonant, as opposed to a rule that deletes a consonant before a vowel. The naturalness of the first type of rule is based on the observation that the preferred segment sequence structure in many languages is CVCVC..., within morphemes and even across morpheme boundaries. A rule of this sort has many examples in French. Thus, *petit + garsõ* is *peti garsõ* (cf. *petit ami*). In English, Schane et al. suggested that the *a ~ an* alternation is the only example that is true in all dialects of English; they then claimed that this natural tendency has very limited application in English.[2] If it is in fact the case that this formal universal of preferred segment sequence has psychological reality, the prediction is that speakers of English, given the chance, would show preference for such a structure.

Schane et al. first taught their subjects four novel nouns as paired associates to English "translations". Two of the novel nouns began with a consonant sound, and the other two began with a vowel sound. In the second stage of the experiment, three novel adjectives, all with final consonant sounds, were paired with each of the novel nouns, following the natural rule (consonant deletes before consonant) for one group of subjects, and the unnatural rule (consonant deletes before a vowel) for the other group of subjects. The 12 phrases were cued by their English translations, and the subjects were to respond with the phonologically correct sequence of adjective and noun. The results showed that "the *S*s learning the unnatural corpus had a strong tendency to give natural responses, whereas the converse was not true. Consequently they made many more errors en route to mastery than their natural counterparts,... It appears that *S*s had implicit knowledge of the natural rule, even though it does not operate to any significant extent in English" (p. 351).

Why should phonological constraints of the sort involved in the above experiments interfere in mental activities? The answer is not hard to see if these constraints are viewed as "expectations", "strategies", or "rules" that subjects have and use in handling linguistic materials. After all, the subjects go into the experimental laboratory with a lifetime of linguistic experience behind them. When they are asked to process materials that appear to them to be organized the way a human language is organized, they invariably employ modes of processing that they have become accustomed to use in their own language.

Psychologists have been concerned with this type of proactive interference for some time now. Interference may be negative, in which case processing of the new materials is hindered, as in the Esper and the fusion experiments. Or it may be positive, in which case learning of the new materials is facilitated (cf. studies on syntactic facilitation, for example, Glanzer 1962; Baker & Prideaux 1975). Interference is explained in terms of such familiar psychological notions as "response similarity", "response mediation", "response competition", and lately, "response restriction" (Bower 1970). When two response systems are sufficiently similar, learning or recall of the correct response is hampered by the fact that something similar to it competes for attention. Of these competing responses, the stronger response is frequently selected; the stronger response is that which is more strongly linked to the stimulus, typically by a mediating rule or generalization.

For example, in the context of the transfer schema A-B, A-C, if B is a type of response associated with linguistic usage, and C is the experimentally prescribed response, B interferes in the learning of C and this interference is a function of the similarity of B to C. In the testing phase, response B may predominate over C if C has not been fully learned. Imperfect learning of C may arise if interference by B is sufficiently strong, in other words, if the linkage between A and B has not been effectively extinguished. Thus (all other things equal), the dominance of the linguistically determined response B is taken as an indirect evidence for the existence of linguistic mediation between A and B. In Bower's (1970) response restriction hypothesis, the predominance of B over C may result if a mediating rule exists that restricts the range of possible responses to A to responses of the type B. Thus, a rhyming rule, for example, that relates stimulus and response, would help the subject to restrict his choices of responses to words that rhyme with the stimulus word.

SUBSTANTIVE REALITY VS. FORMAL REALITY

Following Kiparsky (1968), a distinction is made here between the psychological reality of the substance or content of a rule and the psychological reality of its form. Two rules that differ in form may express the same generalization, and thus account for the same phenomenon. If the predictions turn out to be correct, then the rules have content reality. Such a demonstration, however, does not say which form or which rule has reality. Hence, testing for the reality of a rule involves testing for its form as well as its content; these tasks are, to a certain extent, independent of each other (a point also made by Ohala 1970:5). Testing for substantive validity means testing for knowledge of the generalization underlying the rule, in whatever form the generalization maybe represented, while testing for formal validity amounts to testing for the way this knowledge is represented in terms of the formal language of a particular theory. If a rule is shown to be without content, the rule is thereby rendered invalid; however, general properties of rule form, e.g. disjunctiveness, conjunctiveness, mirror-imageness, etc., have not been tested. Competing rules within a theory may differ crucially only in their form, and it is desirable that the form of a rule be tested for validity as well. However, the question of the substantive reality of a rule takes precedence over the question of its formal validity, since it is pointless to examine the correctness of the form of a rule if the rule in the first place has no substantive basis. The present study is directed to this more basic task.

THE EXPERIMENT

THE AREA

It was decided to apply this behavioral explication of psychological reality in the area of vowel alternation in English. The choice of this area was motivated by the fact that there is no other aspect of English phonology known to us for which there exist diametrically opposite claims about its psychological reality at both the formal and experimental levels. At the formal level, whatever formal differences the versions of the vowel shift rule may have (cf. Chomsky & Halle 1968; W. Wang 1968; Ladefoged 1971; Krohn 1972; Stockwell 1972), there is implicit agreement among those who have formulated the generalization as a phonological rule that what the rule expresses are lexeme-independent properties of the sound system of English, hence, that what is being captured is a psychologically real process. On the other hand, the claim has been made that English vowel alternation is a nonproductive process (M. Wang 1974:43), that it is an irregular and lexeme-

dependent process (Cearley 1974:30), and that the alternations are fossils of a dead process, which are of historical rather than synchronic interest (Maher 1971; Olofsson 1974).

At the experimental level, there have been three major studies in this area, conducted at about the same time, using different methodologies, and producing conflicting results. Steinberg and Krohn (1975) and Ohala (1974) used the word-derivation paradigm and uncovered negative results, whereas Moskowitz (1973) used a learning paradigm--concept formation--and produced positive results (these studies are discussed in some detail in a later section). In the light of these conflicting claims and findings, further investigation is necessary.

THE EXPERIMENTAL PARADIGM

The experimental paradigm used here is PAIRED-ASSOCIATE LEARNING. In paired-associate learning, subjects learn pairs of items. Evidently, if the items are relatable by a generalization or a rule, then these items will be easier to learn than pairs of unrelated items. Thus, if subjects knew the rules of exponentiation, a list of correct exemplars such as $2^2 = 4$, $4^3 = 64$, etc. should be found easier to memorize than a list of incorrect instances of exponentiation. Similarly, if native speakers of English have internalized the vowel alternation generalization, then they should find pairs of made-up words such as *subk*[ay]*pe* - *subk*[I]*pity* easier to memorize than pairs such as *malp*[ay]*de* - *malp*[ɛ]*dity*, since the former but not the latter exemplifies a type of vowel alternation in English, i.e. the alternation in *divine* - *divinity*. Furthermore, knowledge of this alternation should interfere in the learning of the latter pair since, given *malp*[ay]*de*, the normal course of action is to respond with the rule-governed alternate *malp*[I]*dity*, and not with the experiment-prescribed incorrect alternate *malp*[ɛ]*dity*. The manner in which a psychologically real vowel alternation generalization will facilitate the learning of real alternations in a paired-associate learning situation will be explained.

Consider a three-stage model of paired-associate learning (McGuire 1961). First the stimuli are discriminated from one another. Then the responses are learned, a process that is essentially of the recall type. Finally, the stimuli are associated with the appropriate responses. If now a vowel alternation generalization of the sort linguists postulate is part of the speaker's mental grammar, this should provide a link between stimulus and response. The ready availability of such a link simplifies the associative phase, and improvements in learning should result. A second way by which an available mental generalization should facilitate learning is in the minimization of

response learning, or, in Bower's (1970) term, in the restriction of available responses. If indeed a mental vowel alternation generalization exists, which specifies what base vowels are paired with what derived vowels, the choice of responses is severely restricted. In fact, under ideal conditions (i.e., knowledge of the generalization is explicit), the response term is completely recoverable--the subject needs only to add the suffix *-ity* to the stimulus term, and, using the generalization, change the shape of the crucial vowel. Paired-associate learning of real alternations is thereby reduced to stimulus identification. Such a simplification of the learning process is not suggested by alternations that are not part of the language of the subject. The expectation in these cases is that these nonexisting vowel alternations will be harder to learn. The remainder of this paper is concerned with the experiment conducted to test for these speculations.

MATERIALS AND PROCEDURE

Twenty-five alternations were included in the study, of which five were the correct English alternations ay ~ I, iy ~ ɛ, ey ~ æ, aw ~ ʌ, and ow ~ ɔ. The 20 incorrect alternations (incorrect in the *-ity*-paradigm) were formed by pairing each of the diphthongized vowels with the correct alternants of the other diphthongized vowels. For example, ay was paired with ɛ, æ, ʌ, and ɔ; these incorrect alternations served as the control conditions for the correct ay ~ I. The major alternation uw ~ ɔ (*school - scholar*) was not included in the study since both this and the ow ~ ɔ alternations have the same lax vowel. The 25 alternations are shown in Table 1.

Alternations	Stimulus Word-Pairs	English Example
ay ~ I	subk[ay]pe-subk[I]pity	divine - divinity
ay ~ ɛ	malp[ay]de-malp[ɛ]dity	
ay ~ æ	purg[ay]te-purg[æ]tity	
ay ~ ʌ	surf[ay]me-surf[ʌ]mity	
ay ~ ɔ	disr[ay]ne-disr[ɔ]nity	
iy ~ I	surk[iy]te-surk[I]tity	
iy ~ ɛ	malm[iy]de-malm[ɛ]dity	serene - serenity
iy ~ æ	disf[iy]me-disf[æ]mity	
iy ~ ʌ	del[iy]ne-del[ʌ]nity	
iy ~ ɔ	ent[iy]pe-ent[ɔ]pity	
ey ~ I	prel[ey]be-prel[I]bity	
ey ~ ɛ	subn[ey]ze-subn[ɛ]zity	
ey ~ æ	deb[ey]me-deb[æ]mity	profane - profanity
ey ~ ʌ	surt[ey]ce-surt[ʌ]city	
ey ~ ɔ	disd[ey]pe-disd[ɔ]pity	
aw ~ I	misn[aw]d-misn[I]dity	
aw ~ ɛ	surs[aw]n-surs[ɛ]nity	
aw ~ æ	pren[aw]se-pren[æ]sity	
aw ~ ʌ	perc[aw]t-perc[ʌ]tity	profound - profundity
aw ~ ɔ	imp[aw]ze-imp[ɔ]zity	
ow ~ I	end[ow]ne-end[I]nity	
ow ~ ɛ	emp[ow]be-emp[ɛ]bity	
ow ~ æ	perh[ow]de-perh[æ]dity	
ow ~ ʌ	disl[ow]pe-disl[ʌ]pity	
ow ~ ɔ	exg[ow]me-exg[ɔ]mity	verbose - verbosity

Table 1. List of Alternations and Stimulus Materials. The underlined pairs are exemplars of correct English alternations.

The alternations were embedded in nonsense word pairs, e.g. *subk*[ay]*pe* - *subk*[I]*pity*. In an attempt to make the words sound like real English words of Romance origin, Romance prefixes were used as the first syllable. The second syllable, the syllable that contained the crucial sound, had the structure CVC. Three graduate students who were native speakers of English and who knew the purpose of the experiment went through the list to make sure that the words were passably English-sounding, equally pronounceable, and did not elicit strong associative meaning. The list went through a number of revisions before the final list (see Table 1) was deemed acceptable. The words were voiced on audio tape by a female native speaker of English.[3] These recordings were digitized using a PDP 12 computer, and re-recorded on LINC tape. From these master recordings, the words were recorded again on audio tape. This procedure allowed us to eliminate excessive hissing sounds that accompanied certain sibilants, and to achieve uniform amplitude for all signals.

The words were presented to the subject using the study-test procedure. In the study phase, the subject tried to remember the word pairs. The subject heard and repeated both words of each pair, the words separated by a 2-sec gap. All 25 word pairs were heard in a study phase, with a 4-sec interval between word pairs. A short tone signalled the end of a study phase and the start of a test phase. In the test phase, the subject recalled the noun forms. Only the adjective forms were presented, at 5-sec interval. The subject was to give the noun form within the 5-sec gap. A long tone signalled the end of a test phase and the beginning of the next study phase. The order of presentation of word pairs was randomized in each study phase; likewise the order of presentation of the adjective forms was randomized in each test phase. However, since randomization was made during the recording of the materials, there was one fixed ordering for all subjects. The subject was asked to treat the words as uncommon adjective-noun pairs exemplifying the same grammatical relations as the pairs *plausible* - *plausibility*, *possible* - *possibility*, etc. Criterion was set at one error-free trial, or ten trials, whichever was first. Twenty university students who were native speakers of English served as subjects.

RESULTS

A summary of the responses is presented in Table 2, where each entry represents responses of a given vowel type, summed over subjects and across trials for each stimulus and required response vowel type combination. For example, entries in the first row are frequency responses to the stimulus word *subk*[ay]*pe*, with *subk*[I]*pity* as the required response. 154 (of a possible 200)

responses were [I] responses, i.e. responses of *subk*[I]*pity*. There were 31 *subk*[ɛ]*pity* responses to this condition, 6 *subk*[æ]*pity* responses, 3 *subk*[ʌ]*pity* responses, and 2 *subk*[ɔ]*pity* responses (with 4 instances of irrelevant or nonresponses).

Stimulus	Required Response	Actual Response				
		I	ɛ	æ	ʌ	ɔ
ay	I	154	31	6	3	2
	ɛ	83	92	4	4	12
	æ	67	27	70	13	12
	ʌ	32	16	60	83	3
	ɔ	28	18	2	4	131
iy	I	75	68	5	38	2
	ɛ	28	142	4	2	10
	æ	19	25	113	35	4
	ʌ	52	81	4	48	3
	ɔ	67	61	8	7	53
ey	I	99	25	53	4	12
	ɛ	24	45	88	2	31
	æ	6	9	168	4	7
	ʌ	14	12	70	92	8
	ɔ	25	17	27	22	101

(Table 2)

(Cont. Table 2)

Stimulus	Required Response	Actual Response				
		I	ɛ	æ	ʌ	ɔ
	I	64	26	11	34	48
	ɛ	11	84	13	25	61
aw	æ	22	23	77	8	62
	ʌ	38	14	3	100	34
	ɔ	9	15	13	12	135
	I	45	30	7	15	91
	ɛ	25	54	7	12	95
ow	æ	14	12	81	4	72
	ʌ	14	19	2	100	58
	ɔ	2	3	4	4	185

Table 2. Responses Summed Over Subjects and Trials. Entries along the main diagonal in each square are correct test scores. The underlined entries are correct English scores.

ANALYSIS OF THE LEARNING SCORES

Three learnability measures were established. The first was Correct Test scores; these are the entries along the main diagonal in each square of scores in Table 2, out of which a two-way table was constructed (Table 3).

		Required Response				
		I	ɛ	æ	ʌ	ɔ
Stimulus	ay	**154**	92	70	83	131
	iy	75	**142**	113	48	53
	ey	99	45	**168**	92	101
	aw	64	84	77	**100**	135
	ow	45	54	81	100	**185**

Table 3. Correct Test Scores.
The underlined values represent correct English responses. ANOVA results: $F = 12.74$, $p. < .001$, $df = 4,475$.

A two-way analysis of variance of these scores (Stimulus x Required Response) yielded a significant interaction ($F = 12.74$, $p < .001$). This significant interaction effect suggests that the learnability of the alternations was in fact influenced by the type of vowels paired. A simple and direct way of identifying the alternations that caused the greatest amount of learning is to plot certain values associated with the alternations. The values we chose to plot are mean deviation scores, instead of raw means, because mean deviation scores are truer interaction scores since they are corrected for main factor effect. Table 4 contains the Correct Test mean deviation scores, and Figure 1 plots these scores.

		Required Response				
		I	ɛ	æ	ʌ	ɔ
	ay	2.81	-0.08	-2.10	-0.59	-0.01
	iy	-0.14	3.40	1.03	-1.35	-2.92
Stimulus	ey	0.31	-2.18	3.04	0.10	-1.26
	aw	-0.98	0.21	-1.05	0.95	0.88
	ow	-1.98	-1.33	-0.90	0.90	3.33

Table 4. Correct Test Mean Deviation Scores.
Values along the main diagonal represent correct English alternations.

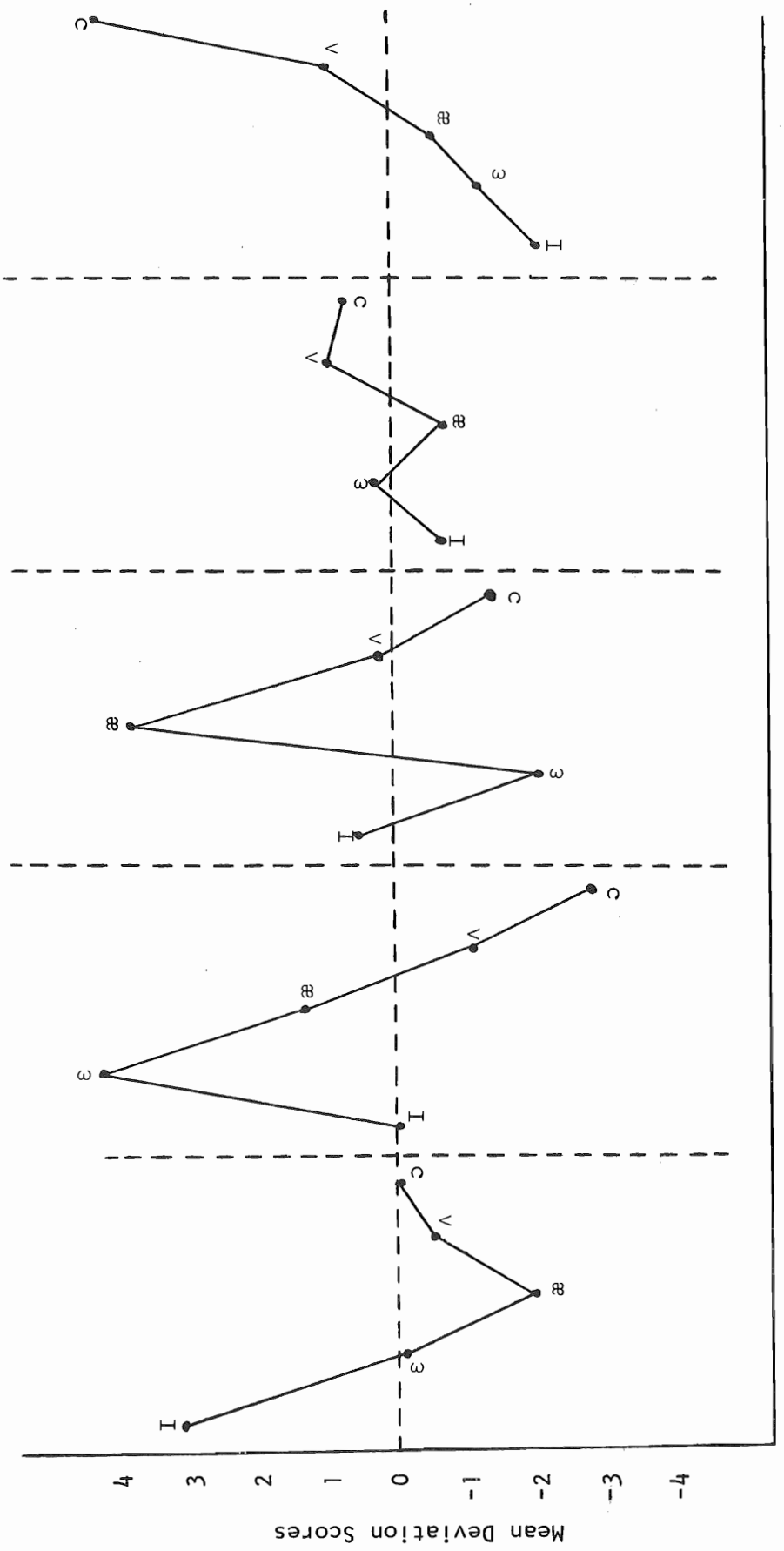

Figure 1. Stimulus - Response Interaction, Correct Test

It can be seen in Figure 1 that for each group of alternations, represented by the frames, the correct English alternation in fact induced the greatest amount of learning in contrast to their respective control alternations. Thus, when the stimulus was [ay], the correct English alternate [I] received a score higher than the controls; similarly for all the other four alternations.

In addition to Correct Test Scores, two other learning measures were established, namely, First Correct Trial score and Last Error Trial score. The First Correct Trial score of an alternation is the trial number in which the required response was first given (since there were ten trials, if none of the responses for all the trials was correct, a value of 11 was assigned). The Last Error Trial score is the trial number in which the last error occurred (if all responses were incorrect, a value of 0 was assigned; if all responses were correct, a value of 11 was assigned). The idea behind these measures of learning is that early learning of a condition, as indexed by the first correct response trial number, and its early mastery, as indexed by the last error trial number, are indications of the superior learnability of the alternation. The expectation was that correct English alternations, in contrast to their respective controls, will be learned and mastered early, and thus, they should have LOWER First Correct Trial and Last Error Trial scores.

The analyses of variance performed on these scores showed that in both analyses, the interaction between stimulus and required response was significant (First Correct Trial $F = 8.44$, $p < .001$; Last Error Trial $F = 8.61$, $p < .001$). These significant interaction effects suggest that early learning and early mastery of the conditions are a function of the type of alternations being learned. Figures 2 and 3 plot the mean deviation scores of the alternations.

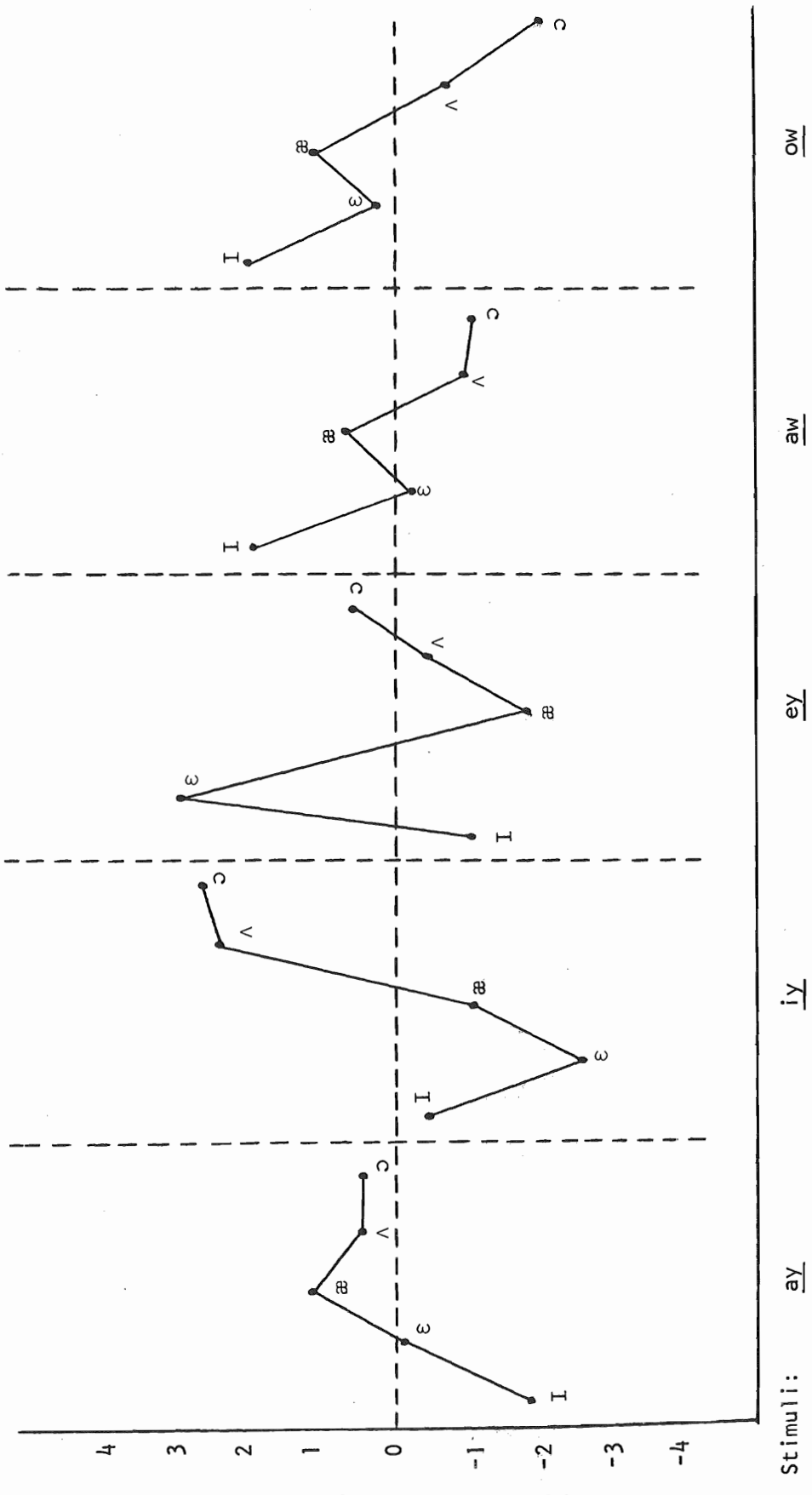

Figure 2. Stimulus - Response Interaction, First Correct Trial

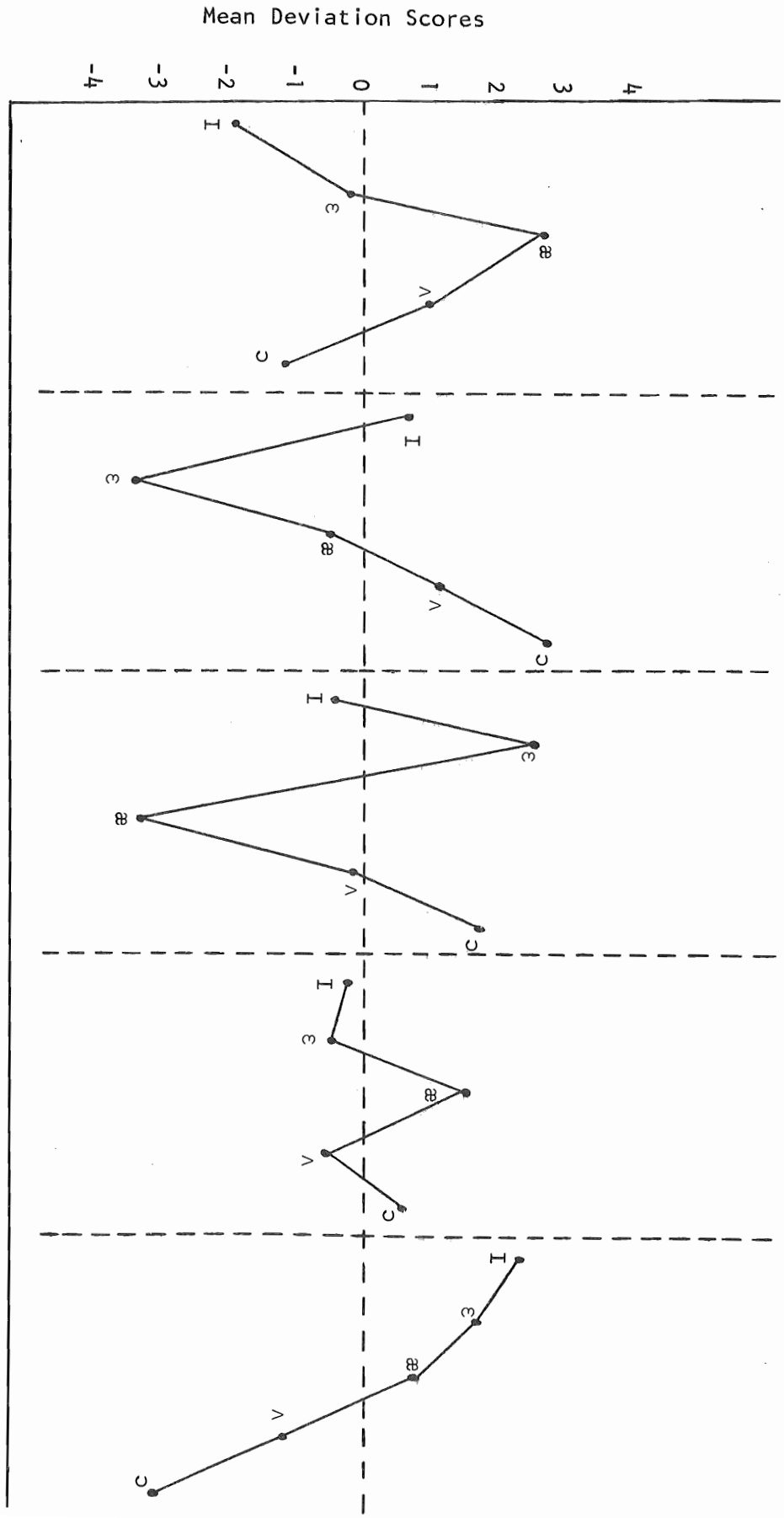

Figure 3. Stimulus - Response Interaction, Last Error Trial

With one exception, the deviation scores of the alternations are LOWER than their respective control conditions in both first correct and last errors analyses, suggesting early learning of the word pairs exemplifying correct English alternations. The one exception was the deviation score of the First Correct Trial score for the fourth alternation, which is slightly higher than the score for the fifth alternation.

Since these scores are all measures of learnability, it makes sense to collapse them into one composite learnability measure. To make the Correct Test Scores compatible with the other two scores, they were transformed into error scores. The three mean deviation scores corresponding to each of the alternations were then averaged to yield a composite learnability score for the alternations. These composite scores are plotted in Figure 4, which shows that the correct English alternations have lower scores--significantly superior learnability--than their respective control conditions.

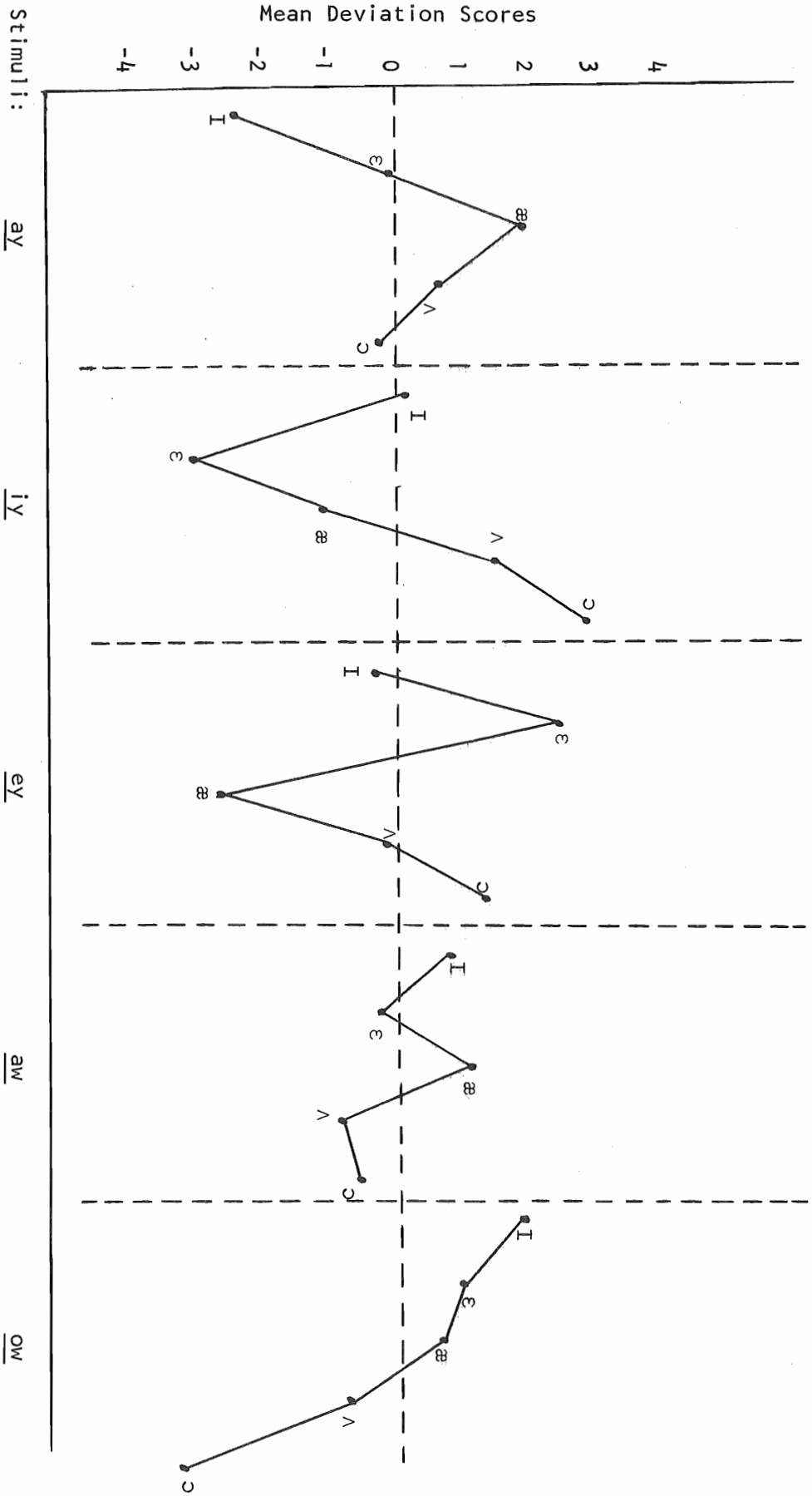

Figure 4. Composite Learnability Measure in Mean Deviation Scores

ANALYSIS OF THE CORRECT ENGLISH INTERFERENCE SCORES.

Error scores represent learning interference. Part of these interference scores consists of correct English responses. Let us look at Table 2 again. The values in the second row are frequency responses to the condition ay ~ ɛ (i.e. *malp*[ay]*de* - *malp*[ɛ]*dity*). This condition required the subjects to memorize an incorrect English alternation. There were 92 instances of successful memorization, and 103 unsuccessful attempts (with 5 irrelevant or nonresponses). Of these errors, a large majority, 83 out of 103, was the correct English alternate [I]. That is, the subjects responded with *malp*[I]*dity* nearly as often as they responded with *malp*[ɛ]*dity*, in spite of the fact that the experimental situation required that they responded with the latter. Thus, there was interference in the learning of the incorrect alternation, and this interference can be attributed largely to knowledge of the correct English alternation.

A first approximation of the extent of this interference can be made by pulling out these English interference scores and transforming them as percentages of the error scores. These values are as follows: ay ~ I = 55%, iy ~ ɛ = 47.75%, ey ~ æ = 53.50%, aw ~ ʌ = 20.75%, ow ~ ɔ = 66.25%. Thus, when the stimulus was [ay], and the required response was not the correct alternate [I], 55% of the error responses were nonetheless the correct response [I]; similarly, when the stimulus was [iy], and the required response was not the correct alternate [ɛ], 47.75% of the error responses were the correct alternate [ɛ], and so forth. These values show that, except for the fouth alternation, at least a third, and in the case of the fifth alternation two-thirds, of the errors are attributable to knowledge of the correct English alternations.

To test for the extent of this interference, a comparison of Correct English scores and Correct Test scores was made. The motivation for this comparison is here outlined: If knowledge of vowel alternation guides performance, memorization of the word pairs that exemplify English alternations should be facilitated and memorization of the word pairs that do not exemplify real English alternations should be impeded. The strongest confirmation of such an expectation is, of course, a response pattern in which the non-English alternations are never learned and only the correct English alternations are learned, and furthermore, that all of the error responses consist of correct English alternants. However, to expect to get this ideal result is to make the unreasonable assumption that there is a rule of language so strong that it cannot be temporarily suppressed in an experimental situation by highly motivated and often ego-involved university students. A moderate expectation was made, which was that interference, if any, would be strongest in the first trials, and decrease as trials progressed. In other words, that subjects

will make many more correct English responses in the absence of learning, but that they will make more correct test responses as learning takes place. Testing for this expectation requires comparison of subjects' Correct English and Correct Test scores across trials.

Analysis of variance of these scores revealed that the interaction between Trial, Stimulus Type, Required Response, and Correct English/Correct Test is significant, with $F = 1.25$, $p < .001$. Figure 5 plots the mean deviation scores.

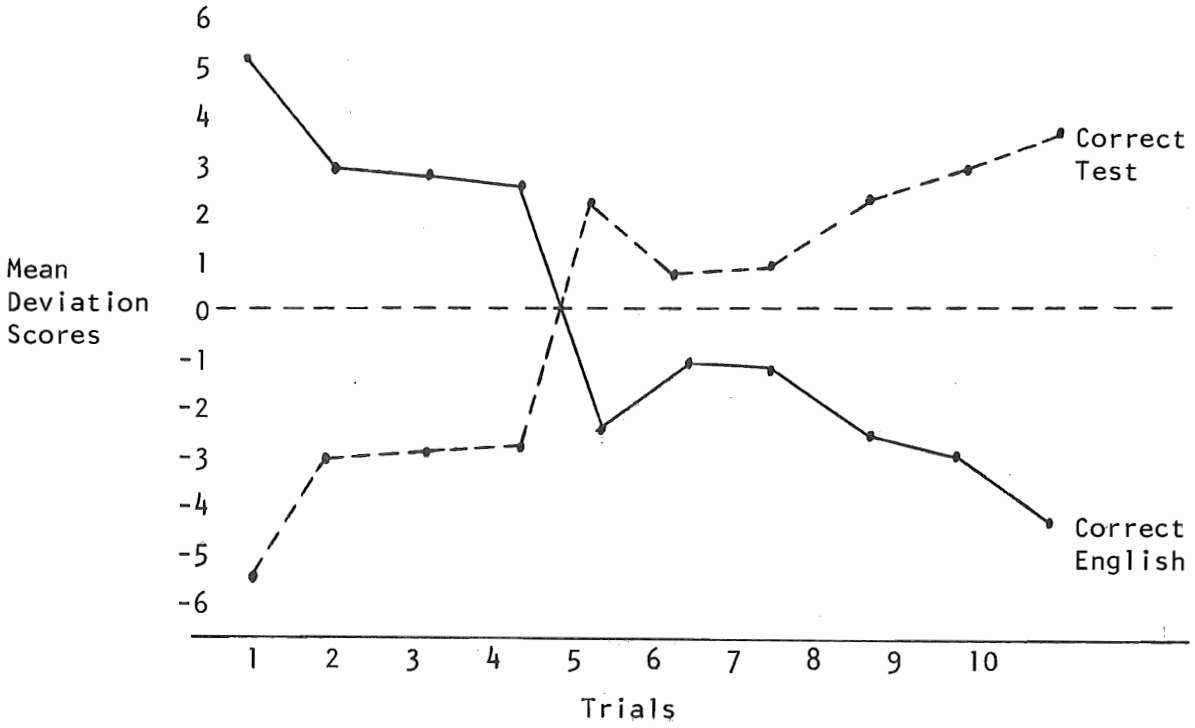

Figure 5. Mean Deviation Scores of the Categories Correct Test and Correct English

(Because the values plotted represent deviation from cell means, and because there are only two categories, the line corresponding to one category is the mirror image of the other line.) The relevant information in this figure is the superiority of Correct English in the first four trials. The values for the first four trials are positive for Correct English and negative for Correct Test. Starting with the fifth trial, the values are reversed, i.e., Correct English scores are below the mean and Correct Test scores are above the mean. This is precisely the kind of results that would suggest the inhibitory effect of the internalized linguistic patterns being tested.

Subjects were making errors in the first four trials in the direction of Correct English alternations; in other words, in the absence of learning, the nature of the responses was largely determined by related linguistic experience, in this case, experience with the correct English alternations.

DISCUSSION

The general findings of the experiment have been that nonsense word pairs exemplifying correct English alternations induced more learning than nonsense word pairs that do not exemplify correct English alternations, and furthermore, that the relative low learnability of the non-English alternations may be attributed to interference due to experience with or knowledge of certain aspects of the phonology of English. The phrase "knowledge of a linguistic generalization" is used to mean possession of a mental representation of the alternations (in whatever modality or form). Or, somewhat weakly, that speakers have become sensitive to the pattern of alternations through experience with the language, and that they CAN notice it in a similar set of data and employ it as a heuristic in, say, a learning situation.[4]

The assumption of previous knowledge of some of the alternations used in the study in effect leads to a bi-partitioning of the stimulus materials into two types: alternations that subjects already 'knew' and alternations that they had to learn for the first time. The subjects thus performed not one but two tasks. The first task was the recall of associations already known, a recall task that was made simple by the cueing function of the test word. The other task was the acquisition of new associations, a task more difficult than simple recall. Acquisition of the new associations was rendered even more difficult in that, in contrast to the usual experimental transfer situation, the old associations were NOT to be extinguished, but were to be retained and used as bases for generating responses. Competition between 'old' and 'new' responses was, therefore, active throughout the experiment. Linguistic knowledge proactively inhibited the learning of the new, experimentally prescribed, linguistic patterns. No linguistic generalization was available as basis for generating the responses to stimulus words that exemplified non-English alternations, hence, the task of remembering these word pairs suffered grossly from a lack of an organizing principle, and the pairs had to be memorized individually.

The findings suggest that knowledge of the pattern of English vowel alternations under study goes beyond the existing lexicon. This point is strongly supported not only by the overall superior learnability of word pairs with English alternations, but also by the fact that, in the early stages of the learning process, subjects made errors in the direction of the alternants predicted by

the generalization. We conclude that vowel alternation in English, because it serves as a variable in learning, is psychologically real. The findings thus confirm the major finding of the Moskowitz (1973) study, and appear to contradict the findings reported in Steinberg and Krohn (1975) and Ohala (1974). An attempt will be made to account for these contradictory findings. But first, a brief summary of the Moskowitz and the Steinberg and Krohn studies.

Steinberg and Krohn collected real English base words which when combined with any of the suffixes, *-ic*, *-ical*, *-ity*, *-ify*, and *-ish* produce "meaningful" but nonexistent forms, for example, *snide* + *ity* → *snidity*. The experimental procedure supplied a verbal context for a particular base and suffix combination, and the subjects, who were university students, were asked to supply the combination. Steinberg and Krohn assumed that a response of the type *sn*[I]*dity* is supportive of vowel alternation, whereas the response *sn*[ay]*dity* (i.e. no change in the quality of the crucial vowel) is to be considered disconfirmatory. Nearly 90% of the valid responses were of the disconfirmatory type. Since the critical vowel sounds remained unchanged in the derived forms, Steinberg and Krohn concluded that the subjects did not "know" vowel shift, and on this basis, they questioned the validity of (Chomsky & Halle's) vowel shift rule.

Moskowitz, on the other hand, used made-up forms as bases. To elicit responses, the concept-formation technique was employed. The subjects were told that to made-up words they were to add *-ity*, and that they were to pronounce the resulting words. One of the nonsense words used was [fayp]. A response of [fIpItiy], which is predicted by the rule, was acknowledged to be correct; any other responses were commented on as erroneous and the expected response was supplied. The concept to be 'formed' was the correct vowel alternant. Two control conditions were set up, both of which exemplified alternations not in consonance with vowel alternation. The subjects were children aged 9 to 12. The results showed that subjects who worked on the correct alternations conditions performed to criterion, whereas only two of the 20 subjects who belonged to the incorrect alternations conditions learned the alternations. In other words, alternants predicted by the generalization were discovered and learned by many more subjects. Moskowitz concluded that vowel shift is part of the subjects' linguistic knowledge.

I will discuss briefly two approaches towards explaining these contradictions. The first approach faults the methodology of the Steinberg and Krohn study; the other approach assumes that both negative and positive results are essentially correct, given the methodologies employed, and that the contradictions point to a possible sharpened characterization of the notion of psychological reality. These approaches are in themselves contradictory, but I believe that a consideration of each has a contribution to make towards a clarification of psychological reality.

Faulting the methodology of a study that disconfirms certain aspects of the formal grammar being tested has been resorted to from time to time by even respectable scholars.[5] As pointed out by Ohala (1970:7), there is something fundamentally undesirable in this practice, for then we shift the burden of proof to negative evidence, and thus we make it difficult to disconfirm a theory, when the more sensible thing to do is to make it difficult to confirm a theory. And yet, because the methodological aspect of psychological experimentation revolves essentially around the problem of quantification, the possibility cannot be discounted that the experimental technology may not be advanced enough (Fromkin 1975), or simply that an error has been made in the selection of the measuring device. If the behavior being tested manifests itself in very fine gradations and if the device is not appropriately calibrated, a report of no difference is made and the conclusion erroneously reached is that the theory being tested has been disconfirmed.

Let me examine the possibility that the measuring device Steinberg and Krohn employed may not have been adequate to do the job it was supposed to do, on two counts: inappropriate stimulus materials and inappropriate experimental paradigm. Let us start with Kiparsky's (1975) criticism of the Steinberg and Krohn materials, which was that some of the stimulus words were "morphologically deviant". The claim is that Germanic bases (*snout*, *overgrown*) were used, in spite of the fact that the *-ity*-paradigm covers only the Romance part of the vocabulary. The second point was that noun bases (*maze*, *ground*) were used, although the paradigm calls for adjective bases. The first point is tenable only if it is the case that native speakers of English, as part of learning their language, internalize a distinction between the Romance and the Germanic parts of the lexicon and of the word derivational processes that these parts of the lexicon can participate in. My feeling along this line is that this is an unlikely event. In any case, this is an empirical question, and Kiparsky has in fact suggested an experiment to test for his claim (1975:197, n.3). As to the second point, if the nounness of some of the bases have affected results, it can only be minimal. In an informal experiment, we asked speakers to frontform *supinity* (an uncommon alternant of *supineness*) from the Latin-based adjective *supine*; the responses we obtained showed no shift.

I think that a significant part of the negative results of the Steinberg and Krohn experiment can be attributed to the use of REAL English words as base forms of the stimuli. The use of made-up forms in productivity tests is motivated by the consideration that with real words, it is uncertain whether the rule being tested has mediated a correct form, or whether a response is simply a memorized form. With synthetic stimuli, the experimenter is reasonably certain that the responses have not been memorized. In the Steinberg

and Krohn study, while a base plus affix combination resulted in a derived form that was a non-existing English item, the CRUCIAL part--the part that contained the critical vowel sound--was a real English word. As far as the relevant materials were concerned, they were not novel to the subjects. These real English words may have been treated not as the abstract stems required by the rule, but as surface forms with their own established pronunciation. The pressure of the regular pronunciation of these words may have been too strong to allow for the expression of vowel alternation, and thus, in the words of Kiparsky (1975), "what the subjects did seems to have been in part unrelated to their knowledge of English grammar and to have taken the form of a mere agglutination of the morphemes provided by the experimenter" (197).

The near unanimity of judgment by Steinberg and Krohn's subjects (90% of the valid responses were "nonshifted" forms) may have been an artifact of the experimental task. The task required the subjects to express their preference between two derived words, for example, *maze + ic* and *maze + ity*. Since both suffixes supposedly trigger vowel shift, the actual choice was of course irrelevant to the main concern of the experiment; the relevant point was how the base vowel would be pronounced. Suppose the subject replied with *mazic*, with no vowel change; now his motivation may be nothing more than the fact that, as he understood it, the task required him to combine morphemes. The experimenter could not but accept this response. Such an acceptance, in effect, was an acknowledgement that the response was correct. The subject would then proceed to simply add suffixes to succeeding base words with no concomitant vowel shift, thinking that this was all there was to it. Thus, the procedure had no adequate control for response bias.

The technique of concept-formation, which Moskowitz employed, has a built-in control for response bias. Subjects who formed words by simply adding the suffix to the base received negative feedback, and hence, they were forced to search for other strategies. Subjects who belonged to the correct alternations conditions "found" the alternation generalization and used it, but subjects who belonged to the incorrect alternations conditions had no comparable generalization available to them, hence, learning of these incorrect alternations was far more inefficient. Similarly, in the paired-associate learning paradigm, feedback serves to control for systematic but irrelevant behavior.

I have suggested above that the negative results of the Steinberg and Krohn study can be attributed to the use of the real English stems and of the word-derivation paradigm; this point applies as well to the comparable Ohala study. We could stop here, and leave matters as they stand. But let me push this point a little further. No doubt someone will point out that the word-derivation paradigm cannot be all that ineffectual in testing for

psychological reality. After all, Berko in 1958 employed a word-derivation technique in testing for the productivity of the English sibilant alternation, with positive results. And Berko's data are essentially judgmental in that when a subject responds with, say, "Two wug[z]", he is in effect saying that he accepts *wug*[z] as the plural form of *wug*.

What seems to have brought about the negative results in the Steinberg and Krohn and in the Ohala studies is the use of the word-derivation paradigm on vowel alternation in English, that is, the combination of paradigm and generalization to test. Now when we address the question of why one psychologically real generalization is more sensitive to a paradigm than another generalization, which generalization another paradigm has shown to be likewise psychologically real, we are led to consider the possibility that the paradigms differ in strength or power and that these two aspects of English phonology do not have equal psychological reality. In other words, that psychological reality is a matter of degree;[6] that is, that certain generalizations are more psychologically salient than others, and certain data elicitation techniques are appropriate for some generalizations but not for others. A gross data-elicitation technique such as word-derivation can expose the productivity of only those generalizations with very high degree of productivity. On the other hand, the demonstration of the reality of a generalization with lesser psychological salience requires the use of data-elicitation techniques with much more sensitivity and power such as a learning paradigm.

As to the question, which generalizations are more psychologically salient, let me make this speculation: Phonetically conditioned processes such as surface phonetic constraints have a higher degree of saliency than nonphonetically conditioned processes. Furthermore, if a process is functionally widespread, it has a better chance of occupying a high place in the scale of productivity or reality. Now sibilant alternation in English has been considered to be part of a more general surface phonetic constraint (see Shibatani 1974); it is employed in a number of paradigms, namely, noun pluralization, possessive formation, *is* and *has* contraction, and in verb morphology. On the other hand, vowel alternation is nonphonetically motivated, since there is no phonetic reason for [ay], for example, to alternate with [I]; also, it occurs in fewer paradigms: The Adjective-Noun+*ity* paradigm and in a few irregular verb forms, for example, *keep - kept*. Even in the absence of frequency counts, it can be safely said that the sibilant alternation is used much more frequently than vowel alternation. Once we treat vowel alternation to exhibit only partial productivity, the contradictory nature of the results of the studies on vowel alternation can be explained in a revealing way. Both the Steinberg and Krohn and the Ohala studies are word-derivational, hence, the negative results; on the other hand, both the Moskowitz and the present study are learnability studies, hence, the positive results.

This scalar view of psychological reality may appear to some readers as a contradiction in terms in that we normally think of something as either real or not real. To think of something as, say, 80% real is difficult--unless we define REALITY as RELEVANCE or SALIENCY. Now the weaker interpretation of "knowledge of a linguistic generalization" as "predisposition towards the generalization" takes on some significance. Under this weaker interpretation, it is not necessary to claim that speakers know in any strong sense the vowel alternation generalization, but only that through experience with the alternations they have developed a predisposition towards them and can appeal to them given the right context (e.g. in coining new words, in nativizing loan words, or in performing in verbal experiments). How salient a generalization is to psychological processing depends on how strongly predisposed the speakers are to this generalization, which may in turn depend on the amount of experience the speakers have had with the facts expressed by the generalization.

In conclusion, the above assessment of experimental studies on vowel alternation in English suggests that knowledge of some linguistic generalizations may be less accessible than others, and one of the ways the more deeply ingrained knowledge may be exposed is to make the task sufficiently demanding such that coping with the task is possible only if recourse to this knowledge is made. For example, the information contained in an incoming stimulus signal may be deliberately set to a level that is not up to threshold so that the subjects, in attempting to supply the 'missing parts', will draw on this hidden knowledge. Or the input information may come in chunks too large for immediate processing so that the subjects, in transforming the input signal to manageable units, will use language-based processes. It is only when we frame our studies of the psychological reality of linguistic generalizations within processes or activities that are observably psychological can we be confident of the validity of our claims about the psychological relevance of these generalizations.

FOOTNOTES

[1]This is an expanded version of a paper read at the Annual Meeting of the Linguistic Society of America, held in Philadelphia on December 28-30, 1976. For invaluable help and suggestions, I wish to thank J. T. Hogan, M. L. Marckworth, B. R. Rochet, W. S-Y. Wang, and especially G. D. Prideaux. J. Gray has helped clarify many points in this version, and I also wish to thank him.

[2]J. Gray (personal communication) takes issue with Schane et al. on the claim of limited application of this preferred structure in English. He has suggested that tense vowels may be treated as a sequence of vowel and consonant-like (glide) segments. Whenever a vowel is laxed preceding a consonant, the glide is deleted (or, whenever a glide is deleted, the preceding vowel is laxed), thus: *th*[ə] *man*, but *th*[iy] *apple*; *t*[ə] *school*, but *t*[uw] *a house*; *beaut*[iy] - *beaut*[ə]*ful*, etc. Gray has pointed out that this treatment suggests that what Schane et al. regard as rare in English is, in fact, something with which English speakers are thoroughly familiar.

[3]My thanks to Ms. Tracy Love-Derwing for the recordings, and to Allan Opperthauser and Richard Stephens for 'taming' the PDP 12 for me.

[4]In a recent paper, Fromkin (1975) made this same point: "The [generative] rules proposed are not performance rules; it does not claim that every speaker or any speaker 'runs through' each rule. The claim is rather that such rules can be learned by speakers--that there are generalizations which are available to be abstracted by speakers" (p. 47).

[5]For example, this quotation from Saporta (1965): "If a test designed to demonstrate behavioral correlates for ... [linguistic notions such as nouns, phonemes, etc.] ... fails to yield the predicted results, one feels obligated to modify the test, not the theory" (as quoted by Ohala 1970:7).

This approach to experimental results, which takes the view that only positive results can be useful, has recently been expounded on by Kiparsky (1975): "Psychological experiments involving production and perception are unfortunately going to constitute evidence in one direction only--a positive result will confirm the psychological reality of the tested grammatical rule, but a negative one does not disconfirm it. This is an unfortunate but normal situation in many sciences. When we dig and find a skull we conclude that the place was inhabited, but when we fail to turn up anything we don't know for sure that the place was uninhabited. When we find a certain pattern of blips with the radiotelescope aimed at some point in the sky, we may have evidence for a pulsar, but when we

fail to find such a pattern, there might be either no pulsar or the signals we are looking for are so weak that they are buried in all kinds of other stellar and terrestrial noise" (p. 203).

[6]Kiparsky (1975) reached this same conclusion. He wrote: "Moreover, what many recent discussions (e.g. Skousen, 1972; Ohala, 1972 [1974]; Krohn, Steinberg, and Kobayashi, 1972) ignore is that productivity is traditionally and correctly viewed as a gradient phenomenon. The strongest sense in which one might speak of a rule being productive is that any new word or formation which meets the structural analysis of the rule MUST undergo it. Another, weaker sense is that in which we speak of a rule as being productive if new words or formations CAN become subject to it, that is, if the scope of the rule is being extended in the language. I shall distinguish the two cases as FULL and PARTIAL productivity, respectively. Thus, in my terms there are degrees of partial productivity depending on how strongly the rule spreads" (p. 195).

Incidentally, Krohn (1972) and Ohala also entertained the view that productivity may be a gradient phenomenon. Ohala wrote: "Some general conclusions that can be drawn from this experiment are that the productivity of a given sound pattern varies depending on the particular sound pattern..." (p. 45). And Krohn wrote: "Since it is clear that the plural suffix rule and the vowel shift rule occupy different positions on a productivity-nonproductivity scale, it is not at all unreasonable to conclude that the former, which is highly productive, is a major rule, while the latter, which is less productive, is a minor rule" (p. 18).

REFERENCES

Anisfeld, M. 1968. Psychological Evidence for an Intermediate Stage in a Morphological Derivation. *Journal of Verbal Learning and Verbal Behavior* 8.191-95.

Baker, W. and G. Prideaux. 1975. Word Category Effects on Paired-Associate Learning: A Critique of Glanzer's Results. Paper presented at the meeting of the Canadian Linguistic Association, Edmonton, Alberta.

Berko, J. 1958. The Child's Learning of English Morphology. *Word* 14.150-57.

Bower, G. 1970. Organization Factors in Memory. *Cognitive Psychology* 1.18-46.

Cearley, A. 1974. The Only Phonological Rule Ordering Principle. A. Bruck, R. Fox, & M. La Galy (eds.), *Papers from the Parasession on Natural Phonology*. Chicago, Illinois: Chicago Linguistic Society.

Chomsky, N. and M. Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Cofer, C. 1969. Verbal Learning. M. Marx (ed.), *Learning: Processes*. New York: The Macmillan Company.

Day, R. 1968. *Fusion in Dichotic Listening*. Unpublished Doctoral Dissertation, Stanford University.

_____. 1970. Temporal Order of Perception of a Reversible Phoneme Cluster. *Haskins Laboratories Status Report on Speech Research* 24.47-56.

Day, R. and J. Cutting. 1970. Levels of Processing in Speech Perception. Abstract of a paper presented at the 10th annual meeting of the Psychonomic Society, San Antonio, Texas.

Esper, E. 1925. A Technique for the Experimental Investigation of Associative Interference in Artificial Linguistic Material. *Language Monographs* 1.

Fromkin, V. 1975. When Does a Test Test a Hypothesis, or, What Counts as Evidence? D. Cohen & J. Wirth (eds.), *Testing Linguistic Hypotheses*. Washington: Hemisphere Publishing Corporation.

Glanzer, M. 1962. Grammatical Category: A Rote Learning and Word Association Analysis. *Journal of Verbal Learning and Verbal Behavior* 1.31-41.

_____. 1967. Psycholinguistics and Verbal Learning. K. Salzinger & S. Salzinger (eds.), *Research in Verbal Behavior and Some Neurophysiological Implications*. New York: Academic Press.

Kiparsky, P. 1968. Linguistic Universals and Linguistic Change. E. Bach & R. Harms (eds.), *Universals in Linguistic Theory*. New York: Holt, Rinehart, & Winston.

_____. 1975. What Are Phonological Theories About? D. Cohen & J. Wirth (eds.), *Testing Linguistic Hypotheses*. Washington: Hemisphere Publishing Corporation.

Krohn, R. 1972. The Vowel Shift Rule and Its Productivity. *Language Sciences* 20.17-18.

Lackner, J. and L. Goldstein. 1975. The Psychological Representation of Speech Sounds. *Quarterly Journal of Experimental Psychology* 27.173-85.

Ladefoged, P. 1971. An Alternative Set of Vowel Shift Rules. *Working Papers in Phonetics* 17. University of California at Los Angeles.

Maher, P. 1971. Etymology and Generative Phonology in Traditional Lexicon. *General Linguistics* 11.71-98.

McGuire, W. 1961. A Multi-Process Model for Paired-Associate Learning. *Journal of Experimental Psychology* 62.335-47.

Moorehead, D. 1970. Processing of Phonological Sequences by Young Children and Adults. *Papers and Reports on Child Language Development* 1. Stanford University.

Moskowitz, B. 1973. On the Status of Vowel Shift in English. *Cognitive Development and the Acquisition of Language*. New York: Academic Press.

Nessly, L. 1974. Experimental Phonology and English Stress. Reproduced by the Indiana University Linguistics Club, Bloomington, Indiana.

Ohala, J. 1970. Aspects of the Control and Production of Speech. *Working Papers in Phonetics* 15. University of California at Los Angeles.

Ohala, J. 1974. On the Design of Phonological Experiments. Ms. University of California at Berkeley.

Olofsson, A. 1974. Latin Rules in English Phonology? *Linguistic Inquiry* 5.320-33.

Rommetveit, R. 1968. *Words, Meanings, and Messages*. New York: Academic Press.

Saporta, S. 1965. Review of Psychology, a Study of a Science -- Study II, edited by S. Koch. *Language* 41.95-100.

Schane, S., B. Tranel, and H. Lane. 1974. On the Psychological Reality of a Natural Rule of Syllable Structure. *Cognition* 3.351-58.

Shibatani, M. 1973. The Role of Surface Phonetic Constraints in Generative Phonology. *Language* 49.87-106.

Steinberg, D. and R. Krohn. 1975. The Psychological Reality of Chomsky and Halle's Vowel Shift Rule. E. Koerner (ed.), *The Transformational Paradigm and Modern Linguistic Theory*. Amsterdam: John Benjamins B. V.

Stockwell, R. 1972. Problems in the Interpretation of the Great English Vowel Shift. M. Smith (ed.), *Studies in Linguistics in Honor of George L. Trager*. The Hague: Mouton.

Wang, M. 1974. *Productivity in Generative Phonology*. Unpublished Doctoral Dissertation, University of Texas at Austin.

Wang, W. S-Y. 1968. Vowel Features, Paired Variables, and the English Vowel Shift. *Language* 44.695-708.

Warren, R. 1961. Illusory Changes in Repeated Words: Differences Between Young Adults and the Aged. *American Journal of Psychology* 74.536-546.

Warren, R. and R. Gregory. 1958. An Auditory Analogue of the Visual Reversible Figure. *American Journal of Psychology* 71.612-13.

Warren, R. and R. Warren. 1970. Auditory Illusions and Confusions. *Scientific American* 233.30-36.