THE UNIVERSITY OF ALBERTA

RULE GOVERNED BEHAVIOUR IN ENGLISH INFLECTIONAL MORPHOLOGY©

BY

DIANE J. DENNIS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN PSYCHOLINGUISTICS

DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA

SPRING, 1988

ISBN 0-315-42807-4

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Diane J. Dennis

TITLE OF THESIS: Rule Governed Behaviour in English Inflectional Morphology©

DEGREE: M. Sc.

YEAR THIS DEGREE GRANTED: 1988

(SIGNED) Diane J. Dennis

PERMANENT ADDRESS

11119 - 72 Avenue

Edmonton, Alberta T6G 0B3

Date: 20 April 1988

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled RULE GOVERNED BEHAVIOUR IN ENGLISH INFLECTIONAL MORPHOLOGY©, submitted by Diane J. Dennis in partial fulfillment of the requirements for the degree of Master of Science in Psycholinguistics.

Supervisor

Date: 20 April 1988

To Della and Glenda for answering the phone.

ABSTRACT

A miniature artificial language (MAL) was constructed with inflections analogous in phonological shape to the English past tense and plural suffixes so that the phonological rules that are ostensibly employed in the production of these inflections could be investigated in a novel, controlled context. Monosyllabic nonce words, represented by pictures of imaginary creatures, and two inflections, which were used to indicate developmental stages of these creatures, made up the corpus of this language. Subjects (Ss) were taught the inflections with a set of training stimuli and then asked to provide inflections for a novel set of stimulus items. The stem-final consonants of the words were distributed such that each variant of the inflections could be elicited. The analog for the plural {/z, s, əz/} was the palatal sibilant set {/ž, š, əž/}[a] which indicated that the creature was old enough to walk. The analog for the past tense {/d, t, əd/} was a velar stop set {/g, k, əg/}, which indicated the creature was very old.

In the first experiment Ss were divided randomly into two groups. In Group 1, Ss were taught the inflections with only those stems requiring the [+voiced] variant of the inflections (i.e., /ž/ and /g/). In Group 2, they were taught the inflections with only those stems requiring the [-voiced] variant of the inflections (i.e., /š/ and /k/).

Results indicated that Ss did best on the items that they were trained on. Spontaneous voicing assimilation occurred, but the vocalic variant was rarely produced.

In the second experiment, there were three training groups. In Group 1, Ss were taught only the [+voiced] variant of the inflections /ž/ and /g/. In Group 2, Ss were taught the [-voiced] variants /š/ and /k/. In Group 3, Ss were taught all of the variants for each inflection.

Ss were able to produce the [-voiced] variants spontaneously in all training groups. However, Group 2 Ss generally did not produce the [+voiced] variants as was expected. Spontaneous production of the vocalic variants was rare, except in Group 3, where Ss had been taught these forms. The results of this experiment suggest that the basic or underlying form of the regular plural and past tense inflections of English are the [+voiced] variants, /z/ and /d/. From these variants, we can predict the use of the [-voiced] variants, /s/ and /t/, by a general rule of devoicing. The use of the vocalic variants, /əz/ and /əd/, is apparently governed by a morpheme specific rule, which must be learned.

ACKNOWLEDGMENTS

There are many individuals whose support and interest during the preparation of this thesis made the work an enjoyable exercise.

First, I would like to thank Dr. Bruce Derwing, my supervisor, for introducing the topic to me and for his valuable suggestions on the various drafts which helped to shape the thesis. I would also like to thank Dr. Terry Nearey for his patience while I waded through the statistics and for his interest in the implications of the results, which led to valuable discussions about the thesis. Thank you to Dr. Tom Priestly for serving on my committee and for questions and comments regarding the thesis.

Thanks also to the LING 583 class of 1986 (Grace Wiebe, Ming Ming Pu, Satomi Komai and James Ko) for their input into the data analysis of the results of the first experiment.

Sandy Bellan, Paul Bergen and Grace Wiebe deserve thanks for helping me run the second experiment. Thanks again go to Grace for her continued interest and valuable suggestions in regard to the thesis.

Of course, these experiments would not have been possible without the subjects. I would like to thank the Grades 1, 2, and 3 teachers and students from Delton, Delwood and York schools and the students from the LING 303 class of the 1986 Spring Session for their participation in the first experiment, and to the Grades 3 and 4 teachers and students from Crawford Plains, Ekota and Hillview schools for their participation in the second experiment.

I cannot fail to mention Karen Walker, who introduced me to the "MacIntosh" and then gave me a free rein with hers. Without it, I don't know how I would ever have written my first draft.

Finally, I would like to thank my family for their encouragement and for knowing that I would make it through.

This study was supported by a Province of Alberta scholarship and numerous research and teaching assistantships provided by the Department of Linguistics at the University of Alberta.

TABLE OF CONTENTS

TABLES

FIGURES

CHAPTER ONE

INTRODUCTION

In linguistic theory, rules are posited to account for the systematic variation that occurs in language. This is a perfectly acceptable approach if the theory is restricted to describing the output of a linguistic system, that is, if it is a taxonomic theory. Serious questions do arise, however, if claims about psychological linguistic processes are based solely on taxonomic descriptions of language. For although such descriptions might accurately describe and predict the substance of a language, as Kiparsky points out, they do "not necessarily represent the form of this knowledge in the sense of actually corresponding to the system of rules which is internalized by the speaker" (1982, p.13).

In psycholinguistic theory, a rule should describe some aspect of the knowledge or psychological processes involved in the use of language. Psycholinguistic experiments are therefore designed to demonstrate that a proposed rule has some evidential claim to psychological validity. The goal, in short, is to go beyond the imposition of psychological inferences from the data to achieve their demonstration through data (cf. Baker & Derwing 1982, pp. 194-197). The limitations of the design must therefore be attended to in formulating any interpretation of the results of an experiment.

In the area of English inflectional morphology, many studies have focused on understanding the processes involved in the acquisition and productive use of the regular forms of the past tense and plural inflections. Of particular interest is the question of what strategies are used and how these strategies are generalized into rules in the learning process.

In 1958, Jean Berko conducted a pioneering study on the productive use of English inflections. Prior to her work, most data on these inflections were based on the taxonomic analysis of static language forms. Her study showed that the use of inflections was productive, i.e., the production of the inflected forms of words involved the use of some type of generalized "rules," and that control of these rules seemed to follow a developmental pattern.

What was most innovative about her study was her use of nonce forms to test for the productive use of inflections. By using such forms she was able to eliminate the possibility that children had merely learned the inflected forms of particular words by rote. This method also allows observation of acquisition processes on an intense scale. The diary method of observation, though it provides valuable information, takes considerable time and only a few subjects can be observed in a given study. It is also difficult in naturalistic studies to determine when new inflected forms are actually being encountered and used, since the observer cannot normally spend all of his time with the subjects. Berko's method allows the researcher to provide a situation where novel forms must be produced; where multiple novel forms can be given to a subject in a short period of time; and, as a result, many subjects can be observed. This allows the researcher to make generalizations about the observed performance and gives greater support to some theoretical claims than studies involving only one or two children. Berko's study itself was limited, and the evidence from her data only allowed her to conclude that English inflectional morphology was indeed productive, but not in what way it was. However, the significance of Berko's method was that it opened the door to examining just what processes are involved in the development of the productive use of the English inflections.

One area of English inflectional morphology that has been intensively studied using the "Berko method" is the past tense and plural inflections and their allomorphs. As background, Derwing (1980) considered seven possible theoretical analyses of the distribution of the regular plural allomorphs. These provide possible interpretations for the types of rules involved in the distribution of the plural allomorphs, once a speaker's use of them is fully productive.

The first analysis accounts for the distribution as simply a "List." Plural forms of words are memorized as they are learned, without analysis into stems and affixes. The second ("Whole Stem") analysis states that while speakers may identify the plural suffixes as distinct from the stems to which they are attached, the former are assigned to the latter on an item-by-item basis without any further generalization. Berko's study provided evidence that neither of these analyses was sufficient to account for the processes involved in pluralization. Though some plural forms may, in fact, be memorized, speakers are also able to provide plurals for nonce forms, indicating that further generalization has occurred.

The third analysis that Derwing proposes is that inflections are generalized on the basis of a rhyme analogy. In this analysis the appropriate allomorph is selected by matching the final VC_0 sequence of a nonce word with a rhyming word whose inflection is known. Berko's study could not show the inadequacy of this type of analysis, because the set of nonce forms that she used was limited to those that all had real-word rhymes.

The fourth ("Segment") analysis proposes that the morpheme variants are conditioned only by the final segment of the stem. In formal notation, it appears as:

$$(\text{Plural}) \rightarrow \begin{Bmatrix} \text{-əz} \,/\, \{\text{s,z,š,ž,č,ǰ}\}\_\_\_ \\ \text{-s} \,/\, \{\text{p,t,k,f,θ}\}\_\_\_ \\ \text{-z} \,/\, \{\text{b,d,g,v,d,m,n,ŋ,r,l,y,w,i,e,u,o,a,æ,ə,ɔ,r}\}\_\_\_ \end{Bmatrix}$$

Analysis five, called the "Feature" analysis, states that the variants of the plural morpheme are conditioned only by certain phonetic features of the final segment of the stem. Using the ad hoc feature [+sibilant] to represent the class {s,z,š,ž,č,ǰ}, this rule appears as;

$$(\text{Plural}) \rightarrow \begin{Bmatrix} -əz\ /\ [+\text{sibilant}]\ \_\_\_ \\ -s\ /\ \begin{bmatrix} -\text{sibilant} \\ -\text{voiced} \end{bmatrix} \_\_\_ \\ -z\ /\ \begin{bmatrix} -\text{sibilant} \\ +\text{voiced} \end{bmatrix} \_\_\_ \end{Bmatrix}$$

Just as before, there is nothing in Berko's data to indicate a choice between either of these theories, nor even between them and a sixth ("Phonotactic") analysis. In this sixth analysis, the plural suffix is assigned a "basic" or "underlying" representation /z/, and is supplemented by two phonological rules that correspond to phonotactic constraints of English. The first predicts vowel insertion between two stem-final sibilants, and the second predicts voicing assimilation for stem-final obstruent clusters. The formal representation would be:

(Plural) = /–z/

Rule 1. $ø \rightarrow ə\ /\ [+\text{sibilant}]\ \_\_\_[+\text{sibilant}]\#\#$

Rule 2. $\begin{bmatrix} +\text{obstruent} \\ +\text{sibilant} \end{bmatrix} \rightarrow [\alpha\text{voiced}]\ /\ \begin{bmatrix} +\text{obstruent} \\ +\text{sibilant} \\ \alpha\text{voiced} \end{bmatrix} \_\_\_$

Finally, the seventh ("Neutral Suffix") analysis posits /əz/ as the basic or underlying form, with accompanying vowel-deletion and voicing assimilation rules. This analysis, though it also predicts correct usage, can be judged inadequate because it is not supported by the empirical facts of language acquistion. Specifically, if /əz/ were basic, one would expect it to be the first form to be used and generalized to. Acquisition data shows this not to be the case, and, in fact, it is the /z/ allomorph that exhibits these qualities, while the /əz/ allomorph is the last of the three to be acquired.

Amended and expanded versions of Berko's 1958 study have led to many interesting experimental results which have increased our knowledge of the processes involved in the acquisition and production of the past tense and plural inflections. The analyses outlined by Derwing above will be considered as possible interpretations of the results of the Berko-type experiments that are reviewed in the following section.

Many studies have been conducted using nonce forms to elicit the past tense and plural inflections. The nature of these studies is diverse, some concentrating on developmental trends; i.e., when are rules fully productive? (Innes 1974, Derwing & Baker 1976 & 1979b, Gray & Cameron 1980); what methods or strategies are involved in the productive use of inflections? (Baker & Derwing 1982, Anisfeld & Tucker 1968); is the use of these inflections based on full segments or features (Anisfeld & Gordon 1968, Derwing & Nearey 1986)? All of these studies have provided interesting results and led to further inquiry into the nature of the phonological rules involved in the use of the past tense and plural inflections. The research conducted for this thesis is concerned with extending this prior research by addressing the following three questions: why are /z/ and /d/ the first forms to be used correctly?; why is voicing assimilation second?; and finally, why is the vocalic form of the morpheme the last and seemingly most difficult to acquire?

Before these questions are addressed, a review of the studies that motivate these questions is in order.

CHAPTER TWO

PRIOR RESEARCH

All of the studies reviewed in this section have used the "Berko method" as a starting point for their investigation. These are all cross- sectional studies, done with varying age groups, but usually involving children. Longitudinal studies have also been conducted which examine the acquisition of inflectional morphology (cf. Brown 1973, Cazden 1968, and Ervin 1964), but these will not be referred to here, since the focus of the research here is not the order of the acquistion of the morpheme variants itself, but what motivates this ordering.

Developmental Studies

Berko, 1958

In 1958, Jean Berko conducted a study to determine the nature of the learning of English morphology, asking the question, "Do children possess morphological rules?" (Note that this is Berko's term. It is understood in this study that these "morphological rules" are more likely phonological rules applied during the concatenation of morphemes.) In her study, Berko covered the plural and possessive forms of nouns, the third person singular, the progressive and past tense forms of the verb, and the comparative and superlative forms of the adjective. I will only review the results for the past tense and plural inflections here.

The subjects (Ss) for the experiment were 112 pre- and primary school children, ages 4 to 7. Twelve adults were also included whose responses were considered "correct" and were used as criteria for judging the responses of the children.

The stimuli consisted of nonce words represented by pictures of imaginary creatures or actions. The use of nonce forms ensured that the responses were not previously learned or memorized.

For the plural inflection, 9 nonce forms were used, 4 with the correct response for the inflected form being stem + /z/, 4 with the correct response being stem + /əz/ , and one, "heaf" /hif/, which was considered correct with either a /hivz/ or /hifs/ response. (As noted above, responses were considered "correct" if they corresponded to the adult standard.) One real English word, "glass," was also used to test the ability of preschoolers to use the /əz/ allomorph.

Six nonce forms were used to test for productive use of the past tense: Three stems required the /d/ suffix for a correct response, two required the /əd/ suffix, and one required the /t/ suffix. Two real words, "ring" and "melt," were also used to test for children's use of irregular forms like "rang" and the vocalic allomorph, /əd/, of the regular past tense.

The results of the study showed that there were significant differences between Grade One Ss and preschool Ss on slightly less than half of the inflectional items. For plurals the difference with the highest significance ($p < .01$) was for the real word plural "glasses." The score for preschoolers was 75% correct and for first graders, 99% correct. However, there was no significant difference in scores for the nonce forms that required vowel insertion, all scores being below 40% correct. Berko concluded that though the children have the /əz/ allomorph in their vocabulary, they are not yet able to extend its use to new words. For the other allomorphs, /s/ and /z/, even though there was a significant improvement in scores for some of them, Ss generally had little difficulty extending their use to novel forms.

With the past tense, the difference in scores with the highest significance was for the real word past tense "rang." Here preschoolers scored 0% correct, while first-graders scored 25% correct. The other significant differences ($p < .05$) in scores between preschoolers and first-graders were for the past tense forms "binged" and "bodded." There is no apparent explanation for these differences, since other similar words showed no such significant difference in scores. "Glinged" is not much different from "binged," but there was no significant difference between preschoolers and first-graders responses on this item. Similarly, "motted" is not much different than "bodded;" but there was no significant improvement in scores for "motted" between age groups. What is of interest here is that there was a significant difference between the lowest score for the novel items requiring the /d/ and /t/ allomorphs and the highest score of the nonsense items requiring the /əd/ allomorph.

Berko's study raises many interesting questions; however, its scope is too limited to answer them, as already indicated in Chapter 1. What can be concluded from her results is that there is some extension of morphological or phonological processes occurring in the inflection of the nonsense stems and that a developmental trend could be involved. More detailed studies would have to be conducted to examine the nature of these processes and the nature of the developmental trend.

Innes, 1974

For her master's thesis, Susan Innes (1974) did an extended version of Berko's 1958 study. She dealt only with acquisition of the plural inflection, but the stimuli included all possible single stem-final consonants and a representative vowel, and the age range of Ss was wider than in previous studies. This study replicated Berko's results for corresponding items, which validates it as a tool for evaluating Berko's general hypotheses and conclusions. (In a later study, Rollins 1980 validated the Berko study itself by extending it to a naturalistic play situation.)

Innes tested 120 Ss, ranging in age from 2 to 8 years old. The stimuli consisted of 24 monosyllabic nonce words and 8 real words. The nonce forms were represented by pictures of imaginary creatures. Ss were tested individually and given a "Berko" type test to elicit the plural forms.

The results showed age to be a relatively poor indicator of performance, so Innes regrouped her subjects on the basis of overall performance on nonsense forms. This division led to six performance groups with approximately the same number of Ss. Group I included those Ss with 0-7 total correct responses, Group II, 8-14 correct, Group III, 15-16 correct, Group IV, 17-18 correct, Group V, 19 -21 correct and Group VI, 22-24 correct. On this basis a hierarchical pattern emerged, i.e:, "once a given item is mastered by one group, it continues to be mastered by all successive groups" (p.34). Secondly, those items which are added by the higher performance groups tended to fit into the traditional phonological groupings of non-fricatives, non-sibilant fricatives, and sibilants. Given the nature of the results, Innes posited six stages to describe pluralization behaviour during the acquisition process.

Stage 1: No rules, no plural formation with novel forms. (Group I)

Stage 2: Add /-s/ to stem-final voiceless non-fricative consonants; /-z/ to stem-final vowels, voiced non-fricative consonants, and /v/; and use no overt plural marker with any other unfamiliar words. (Group II)

Stage 3: Add /-s/ to stem-final voiceless non-fricative consonants and /f/; /-z/ to stem-final vowels, voiced non-fricative consonants and /v/; and use no overt plural marker with any other unfamiliar words. (Group III)

Stage 4: Add /-s/ to stem-final voiceless non-sibilant consonants; /-z/ to stem-final vowels and voiced non-sibilant consonants; and use no overt plural marker with any other unfamiliar words. (Group IV)

Stage 5: Add /-s/ to stem-final voiceless non-sibilant consonants; /-z/ to stem-final vowels and voiced non-sibilant consonants; /-əz/ to all stem-final sibilant consonants, except /z/, which takes no overt plural marker. (Group V)

Stage 6: Add /-s/ to stem-final voiceless non-sibilant consonants; /-z/ to stem-final vowels and voiced non-sibilant consonants; and /-əz/ to all stem-final sibilant consonants. (Group VI)

These stages give evidence that the learning of English pluralization rules involves some type of feature analysis; however, some questions arise in regard to the division of the sub-classes that are relevant for each stage. Innes suggests that, at Stage 2, the quality of being fricative identifies the class of stem-final consonants that do not take an overt plural marker. Why then is /v/ an exception? A factor analysis done by Innes supports this division but cannot explain it. Similarly, why, in Stage 3, are both /f/ and /v/ treated unlike the other fricatives? Another question that arises from these results is, "Why don't words with stem-final /z/ take the plural?" These questions do not detract from the value of Innes' study, but do call for further investigation. Innes' results give motivation for more intensive investigation into the nature of the type of feature analysis that occurs with the use of English plural inflection.

Derwing & Baker, (1976, 1979b)

The focus of the study conducted by Derwing & Baker (1976) was to get a better understanding of how the language learner's rules change in the process of acquisition and which factors are most important for the acquisition of linguistic rules in general. They investigated the acquisition of five inflectional morphemes in English: the plural and possessive forms of nouns; the progressive, the past tense and the third person singular, present tense forms of verbs. They used Berko's technique for testing subjects but, like Innes, extended the study across a broader age range and included a greater number of test items, including stem-final consonant clusters (not used by Innes), to get a better representative sample of nonsense words for English.

112 Ss were used in the study, ranging in age from 3 to 9 years old. Stimulus items consisted of the nonce words to elicit regular forms of the inflections, real English words which represented a particular irregular class of plural forms (for nouns) or past tense (for verbs) and a few regular English nouns and verbs. The results of this experiment were consistent with the

results achieved in both Berko's and Innes' studies. I will review only the results for the past tense and plural inflections here.

Derwing and Baker set up tables representing the modal response for a given item by a given age group in order to interpret the results of their study. Items were scored either correct, incorrect or irrelevant (for those items where alterations of the stem-final vowel nucleus, stem-final consonant or consonant clusters occurred). Inflected forms of the nonce words were considered correct if they corresponded to the pattern of inflections for real, regular words, so that the form of the suffix used was phonologically conditioned. (pp.39-40) For both the past tense and plural they found that the "no suffix" response was common for the youngest age group, but after that there was consistent progress towards the predicted adult forms. The response tables show that from the age of 3 onward, nonce stems requiring the [+voiced] suffixes (/-d/ or /-z/) were inflected accordingly. For the [-voiced] suffixes (/-t/ and /-s/) the results were similar, although, with the plural inflection, Ss had some difficulty with inflection of stems with complex consonants and consonant clusters in final position, including a tendency of subjects in the 4 to 6 age range to overextend their use of the /-əz/ suffix for words ending in sibilant-plus- stop clusters. Still, overall responses showed a consistent developmental trend towards the standard response. For stems requiring the vocalic suffix (/-əd/ and /-əz/), the response table shows that either a null response or the vocalic suffix is used (with a few exceptions due to articulation difficulties), with a consistent developmental trend towards predicted usage.

Concentrating on the results for the plural inflection, Derwing & Baker (1979b) discuss the implications these results have in positing rules that are used in the production of English inflectional morphology. Because their study covered a larger and more representative set of nonce stems, they were able to carry their conclusions further than previous studies. In line with Derwing's (1980) proposal of possible analyses, they concluded that Berko's finding with respect to both the "List" and "Whole Stem" analyses was supported. Further, the "Rhyme" hypothesis was proven to be inadequate, because approximately half of their nonce words had no real-word rhymes, yet Ss had no difficulty providing inflections for these forms. Their data indicated (as was also shown in Innes' study) that mastery for one stem type "is highly correlated with mastery of all other stem-types whose final segments are highly similar phonetically" (1979b, p.260). This provides evidence that a "Segment" hypothesis (i.e., "that the stem-classes involved are learned on an item-by-item basis, as a function of their particular final phoneme, or consonant cluster" (p.260).) is probably also false, and supports a hypothesis in which features are used as the conditioning factor for the choice of the morpheme variants. These results are consistent with the "phonotactic" analysis, where vowel insertion and voicing assimilation are predictable by rule. However, they propose a different, "Hybrid," analysis which is also consistent with their data.

> This analysis is one which treats the /əz/ and /z/ allomorphs of the English plural as distinct in the lexicon, but describes the /s/ allomorph as a "derived" or "automatic" variant which results from the application of a general rule of voicing assi- milation to the /z/ allomorph. (pp.260-261)

According to Derwing and Baker this analysis is attractive because it "yields a psychological theory which accounts for the apparent fact that naive speakers of English can readily distinguish the /əz/ allomorph from either the /z/ or /s/, but find it rather difficult to distinguish the /z/ and /s/ allomorphs from one another." (p.261)

Baker & Derwing (1982)

In this study Derwing & Baker undertook to develop "analytical techniques directed toward the demonstration of the presence or absence of subject-determined patterns of responses." (p.9) Innes' (1974) data were used for this study. The results of this analysis were used as a basis for suggesting what strategies were operative in the rules which characterized the performance groups and identifying possible stages of development.

Using distance matrices, subjects were first divided into strategy groups, and then, for each group, treatment of stimulus items was measured. (A distance metric, ranging between zero and 1.0 was used, zero indicating that all subjects placed a given pair of stimulus items in the same cluster, and 1.0 that none of them do so.) This clustering technique allowed them to identify four quite distinct developmental groups among the Ss and gave them an empirical basis for interpreting the rules used within each group. There was also a fifth group with near perfect performance, who, because their performance was predictable from adult patterns, provide no information for evaluation of stages.

In Group 1, Ss do not appear to have any rules; Ss often gave a null response, while the /z/ and /s/ suffixes were used only sporadically. In Group II, some rule for the use of /z/ begins to emerge, and though the use of /s/ is not clearly present, Derwing and Baker suggest that it may be confounded by articulatory problems the children had with the stems and detection problems on the part of the scorer. Thus, the responses for the two beginning groups indicate that the /z/ and /s/ forms appear first, with a possible preference for /z/, but there is no evidence for the use of /əz/.

In Group III, the data suggest that a definite rule has emerged, namely, that /z/, as a target suffix, is attached to all stems and spontaneously devoiced whenever it follows a final consonant that is voiceless. In this group there is inconsistent use of the /əz/ allomorph. It appears that the child finds exceptions to the above rule and begins to experiment with /əz/, sometimes adding it as a duplicate marker, following the /s/ and /z/ allomorphs, or overgeneralizing it to the non-sibilant fricatives.

In Group IV, the single rule above is modified so that sibilants are separated from the other fricatives, and non-sibilants are correctly inflected. Overgeneralization and duplication are eliminated at this stage. Where /əz/

appears, it is in the context of a sibilant; however, the data indicate that performance is not consistent enough for a rule to be implied. The age differences between the subjects in Group IV suggest that this strategy can be retained for quite some time.

From these results Derwing and Baker suggest that rather than simply describing the "consequences" of pluralization, rules should reflect the processes involved in pluralization. They point out that a formal description such as:

$$(\text{Pl}) \rightarrow \begin{Bmatrix} /s/ \; / \; \begin{matrix} \text{C} \\ [\text{-voiced}] \end{matrix} \;\_\_ \\ \\ /z/ \; / \; \begin{matrix} \text{V} \\ \text{C} \\ \begin{bmatrix} \text{+voiced} \\ \text{-sibilant} \end{bmatrix} \end{matrix} \;\_\_ \\ \\ /əz/ \; / \; \begin{matrix} \text{C} \\ [\text{+sibilant}] \end{matrix} \;\_\_ \end{Bmatrix}$$

is only a description of the use of the plural allomorphs. Once again they suggest that an analysis, like the one proposed in their 1976 study, more adequately describes the processes involved in pluralization. See below:

$$(\text{Pl}) \rightarrow \begin{Bmatrix} /əz/ \; / \; \begin{matrix} \text{C} \\ [\text{+sibilant}] \end{matrix} \;\_\_ \\ \\ /z/ \; / \text{ elsewhere} \end{Bmatrix}$$

This rule is accompanied by a general devoicing rule:

$$/z/ \rightarrow [\text{-voiced}] \; / \; [\text{-voiced}]\_\_\_\_$$

This formulation accounts well for the distribution of strategies among performance groups, in that Group I and II can be said to have acquired the most general rule, i.e., the "elsewhere" condition. Group III begins to discover there are exceptions to the general /z/ usage, and Group IV has delimited the target set for the exceptions requiring /əz/. The final group shows mastery of the whole rule.

The results of this analysis coincide with Innes' results ~~but~~ give better empirical support to the division of performance groups as well as to the strategies employed by each group.

Gray & Cameron (1980)

Gray and Cameron conducted a longitudinal study to investigate the acquistion of the morphological rules for the past tense and plural inflections in English. Their subjects were English–speaking children enrolled in French immersion or traditional English school programs. There were 147 Ss in French immersion and 125 in the English curriculum. Ss were tested at the end of Grades One through Four using the 18 plural and past tense items from Berko's (1958) study (see above). Responses were scored correct or incorrect. For items requiring regular forms of past tense and plural a "correct response was determined in accordance with the grammatically correct English model." For those items with an irregular model, both irregular and regular endings were accepted as correct, (e.g., "heafs" and "heaves" were considered correct plurals for singular "heaf," as were "binged" or "bang" accepted as correct past tense forms of "bing.")

A mixed–design analysis of variance was computed on the total number of correct responses. The factors were curriculum groups (2), sex (2), and age (4). No significant differences were found between the French immersion group and the English group. There was a significant difference in the scores for boys and girls ($p < .01$), with the girls performing better than the boys. This is not consistent with the results from other studies but Gray and Cameron did not find any other factor that contributed to this effect. The differences between age groups was significant, with the mean number of correct responses increasing with age ($p < .001$). At each age scores were significantly different from the next. Gray and Cameron also found that "once a child answered an item correctly there was a high probability that the child would continue to respond correctly in succeeding years." (p.177) This corresponds with the cross-sectional studies which found that once an item was correctly inflected by a particular performance group, all higher groups also used the correct inflection.

Looking at the results for the inflections separately, they found that for plurals there was a significantly higher percentage of correct responses for items requiring the /z/ allomorph than for those requiring the /əz/ allomorph ($p < .001$), with an increase in percent correct across the four ages ($p < .001$). Gray and Cameron also found that the difference in difficulty between the two types of items was significant for all age groups, though it is most evident at ages 7 and 8 and decreases at ages 9 and 10. For the past tense, the results were similar. Items requiring /əd/ had significantly fewer correct responses than those requiring /t/ or /d/, with the probability of responding correctly increasing with age. The difference in scores for /əd/ versus /t/ or /d/ remained relatively constant for the first three years, with a decrease at age 10, though the difference was still significant at this age.

For both the inflections they noted that more Ss could supply the correct inflection for the real words requiring the vocalic form (/əz/ or /əd/) of either morpheme than could correctly inflect the nonce forms requiring these allomorphs. Gray and Cameron point out that this "confirms Berko's observation that the presence of an appropriate model in the lexicon does not guarantee that a rule will be applied to nonsense items." (p.178)

This study is valuable in providing longitudinal support for the hypotheses based on cross-sectional studies. The main weakness here, as with

Berko's original study, is that the stimuli are not representative of all possible final segments.

Conclusion

The studies reviewed above indicate that the learning of English inflectional morphology follows a developmental trend and is productive in that inflections are readily supplied for nonce forms. Though productivity in inflectional morphology is affirmed, it must be noted that rote formation (cf. MacWhinney, 1978), as an initial process or strategy, is also supported in both Berko's and Gray and Cameron's studies, where younger subjects could supply the correct form of the plural "glasses" and the past tense "melted" but could not use the vocalic allomorph on nonce forms. Nevertheless, the studies above give some indication as to the nature of the rules that are learned for the productive use of the inflections.

The results of these studies bear on the adequacy of the possible analyses of the distribution of the plural allomorphs considered by Derwing which were mentioned above. As was previously noted, Berko's original experiment showed that the "List" and "Whole Stem" analyses (1 and 2) were inadequate, since they did not account for the spontaneous inflection of novel forms. Innes' results indicated that the stem-final segments fell into traditional phonological groupings in the various stages of acquistion, which at least supports a stem-final "Segment" analysis (4) and suggests that a final segment "Feature" analysis (5) should be considered. These conclusions are strengthened by Derwing & Baker's (1976, 1979b) extensions, which were also able to show the "Rhyme" analysis (3) to be inadequate by using nonce forms with no real-word rhymes. Innes' and Derwing & Baker's studies further suggest that some type of "sub-segmental" feature analysis is used by speakers in producing inflected forms of words, indicating a preference for the "Feature" analysis (5) over the "Segment" analysis (6). The studies reviewed in the following section focus on the strategies employed by learners in the acquisition and production of the past tense and plural inflections and give additional empirical support for analyses which are based on the features rather than full segments.

Strategy Studies

Anisfeld & Tucker (1968)

Anisfeld and Tucker conducted a study "to determine the extent of the child's acquisition of the standard rules of adult speakers of English and to uncover any peculiarities in children's representation of the singular-plural relation." (p.212) Expanding on Berko's study, they tested for both productive and receptive control of the plural rules, that is, they not only tested the subject's ability to pluralize, but also his ability to reconstruct the singular from a plural form. They conducted three experiments: 1) The Role of Numbers in Pluralization, 2) A "Pluralization by Addition" Rule, and 3) The Child's Mastery of the Standard English Pluralization Rules.

The results of the first study showed that children (Ss were Kindergarten students with a mean age of 6 years, 1 month) will use numbers, as an alternative, to indicate plural when they cannot produce the correct adult form. This suggests that young children are at least aware of plurality as a concept and that the plural of nouns is marked in some form.

The results of the second study further support this notion. Two experiments were conducted, using six Kindergarten students with a mean age of 5 years, 11 months as subjects. In the first experiment the S was shown a singular picture and given a nonce name (e.g., "waf"), then S was shown a plural picture (2 or more "wafs") and asked to choose the better of two names for the second picture, the choice being between the original singular name and a longer name from the original root word (e.g., "wafk" or "wafkren"). The results indicated that Ss had a strong preference for the longer forms as plurals.

The second experiment was conducted to test that "pluralization by addition" was the rule and that one did not simply choose a different word for the plural. In this experiment Ss were given a form such as "wafk" as the singular form and asked to choose between forms like "waf" or "wafkren" for the best representation of the plural. Results were nearly unanimous in favour of the longer form.

A third experiment was conducted, using 20 Ss, ranging in age from 4 years, 6 months to 7 years, 2 months, with a mean age of 5 years, 11 months. In this experiment Ss were presented with ten CVC syllables (e.g., "bip") as a singular form, and the choice for the plural form was between a form in which a vowel change had occurred (e.g., "bop") or a singular root with a suffix added (e.g., "bipum"). The longer form was preferred over vowel change in 72% of the trials. This percentage was not as high as in the other experiments but it was significant at $p < .001$. From these results, Anisfeld & Tucker concluded that "even before the child has fully mastered the specific plural suffixes of English, he possesses a general rule to mark the plural by adding onto the singular code." (p.217)

The third study was aimed at investigating the differences in production and recognition of the plural allomorphs. Thirty-six Kindergarten pupils with a mean age of 5 years, 11 months served as subjects. Stimulus materials were brightly coloured pictures of cartoon animals used to represent nonce names. The nonce stems contained the following final phonemes: /f, p, t, k, v, b, d, g, s, z, č, ǰ, š, ž/.

The experiment consisted of three production tasks and three recognition tasks. Subjects were divided so that they either performed the production tasks or the recognition tasks. In all tasks Ss were prompted by the experimenter (E) so that if S gave a correct answer, E said "good," but if the answer was incorrect, E said "no" and encouraged the S to try again. If the S failed, E provided S with the correct answer.

The production tasks required Ss to produce plurals from singular forms or singulars from plural forms. In one task, given the name and picture of a singular stimulus item, Ss were asked to provide the correct name for a plural picture; or, given a plural name and picture, they were asked to provide the singular name. Another task was similar to the one above except that no pictures were used; rather, Ss were asked to pretend that the creatures were present. The third task used pictures and required the Ss to "tell a story" about the creatures (e.g., shown one "gop," S was then shown a few others doing an action and asked to tell E what they were doing.). The first recognition task required the S to provide the correct name for pictures (singular or plural), either by choosing one of two names (e.g., "maj" or "majes"). The second task was to choose the appropriate picture for a given nonce form; and in the third task the S was given two names and two pictures and asked to assign the appropriate names to the pictures. These tasks were directed at showing Ss' ability to recognize the relationship between the inflection and its meaning. Twelve responses were elicited in each task.

The results from this experiment showed that for the Production tasks there was a significant difference among allomorphs and among tasks, although the interaction was not significant. There were no significant differences between responses requiring the /s/ and /z/ allomorphs; however, significantly more errors occurred with the /əz/ allomorph than either /s/ or /z/. The total number of Plural errors (i.e., forming a plural from a singular stem) was significantly higher than Singular errors (i.e., forming a singular from a plural) for all three tasks combined.

The results for the Recognition tasks showed that, for the first two tasks, there was a significant difference in number of errors between /z/ and /s/ and between /z/ and /əz/; fewer errors occurring with the /z/ allomorph. There were no significant differences between the number of Singular and Plural errors.

In terms of productive control of the pluralization rules of English, Anisfeld & Tucker have also found that the /əz/ allomorph is not as readily employed as /s/ and /z/ (cf. studies above). They also found that children performed better in providing a singular form from a plural than when the task was reversed, which indicated to them that Ss were, "to some extent, analyz[ing] a plural word into a stem and a plural marker" (p. 222).

The results of the Recognition tasks provide a different view of the strategies used in acquisition of the plural inflection. The easiest task (only 3 errors in 216 trials) was the final recognition task, where S simply had to match names and corresponding pictures. Anisfeld & Tucker suggest that the strategy used here is simply matching the shorter code with the single object and the longer code with the many objects, corresponding to the results from study 2. This strategy could have also been used in the first recognition task, with the subject simply choosing the longer name for plural and shorter for singular; however, the differences among allomorphs suggests that this strategy, if used, was not used exclusively. The differences in scores between allomorphs has to do with recognition of form. Anisfeld & Tucker explain that this may be the result of the children's experience with the English vocabulary. There are only a few singular words in English that end in consonant + /z/, but words ending in consonant + /s/ are common, so that words ending in consonant + /s/ are not as reliably plurals as are words ending in consonant + /z/. In a similar vein, singular words ending in sibilants, requiring the /əz/ allomorph, already sound like plurals, so the Ss may have regarded them as plurals.

Although the results of this study are suggestive, a major problem with it is that all responses were reinforced. Bryant and Anisfeld (1969) showed that there is a significant learning effect with this type of reinforcement.

Anisfeld & Gordon (1968)

Anisfeld and Gordon conducted a study to "gain... insight into the psychophonological composition of morphological rules." (p.975) They investigated the past tense and plural inflections in English to determine whether or not features are considered as markers of these inflections and, if so, which features serve this purpose. To do this they presented, orally, pairs of nonce words that differed only in their final consonant (e.g., narp vs. narv) and asked subjects to choose one as an artificial past tense or plural form.

Two experiments were conducted, one which dealt with the plural inflection and another which dealt with the past tense inflection. The following sounds were used as final consonants: /b, p, d, t, g, k, ǰ, č, v, f, θ, z, s, š, m, n/. Stem-final /s/ and /z/ were eliminated from the set for the plural study, and /t/ and /d/ were eliminated for the past tense study.

In the plural study, subjects were 64 first graders, 66 fourth graders and 96 adults. The stimulus items were CVC nonce words. The 1st and 4th graders were shown pictures of single cartoon characters and a name was assigned, e.g., "This animal is called a dar." Then the S was shown a picture of 2 or more of the same animal and asked which of the two forms, constructed by appending different consonants to the singular stem (e.g., darm vs. darch), would be a better name for the second picture. Ss had to pronounce both "plurals" correctly before making a choice.

The same procedure was used for adults except they weren't shown pictures. The adults were instructed to choose a form based on their "intuitive impressions" of which form was most like the regular English plural.

In the past tense study, 62 fifth-graders, one fourth-grader and one sixth-grader were used as subjects. The experiment was similar to the plural study, but nonce forms were bisyllabic because, in English, past [illegible] inflections of monosyllables [illegible] to be irregular (p. 976). Ss were shown first of two action drawings and a nonce verb was assigned (E.g."This is a picture of a girl mickering."). The S was then shown the second picture and told that this was the same actor performing the same action except that the event occurred last year. The S was then asked which of a pair of words (e.g., mickerp or mikerm) he/she preferred to describe what happened.

The analysis of the data involved counting the number of times each sound was chosen over all the other sounds and the number of times it was rejected in the comparisons. Chi-square values were computed, comparing the preferences to the rejections.

In the plural study, 1st and 4th graders showed significant preference values for three sounds as representative of plural. These were, in order of preference, /s/, /ǰ/ and /č/, all of which belong to the sub-class of [-grave, +strident] (using Halle's features) which also includes /s/, /z/ and /ž/, sounds not included in the study. This is the class of "sibilant" phonemes in English.

The adults showed either significant preference or rejection for all of the sounds presented in the study. The sounds adults preferred as representative of plural were, in order of preference, /š/, /v/, /ǰ/, /θ/, /č/, /m/, /n/, /f/. Anisfeld and Gordon explain that the preference for /m/ and /n/ is likely due to subjects' being influenced by irregular English plurals and/or foreign plurals (e.g., children and oxen (English); cherubim (Hebrew)). Except for these two phonemes, all of the other sounds preferred were either [+strident], i.e., /č/ and /ǰ/, [+continuant], i.e., /θ/, or both, i.e., /š/, /v/ and /f/. These sounds, along with /s/ and /z/, /ð/ and /ž/ which were not included in this study, all belong to the class of English fricatives.

In terms of significant rejections, Anisfeld and Gordon found only one for the first-graders, /n/. For the fourth-graders, rejection of /t/, /d/, /b/ and /k/ was significant. In the case of adults, all the rejected sounds had negative feature values from the ones they preferred, i.e., they were all [strident, -continuant]. /t/ and /d/ were two of the most frequently rejected sounds.

That these two sounds are used to mark the past tense in English could explain why adults as well as fourth-graders showed a strong dislike for them.

In the pluralization study, preference was shown for fricatives or a subset of them. The past tense study was conducted to determine that these preferences were based on the morphological task and not just on sound preferences. The results from this study show that subjects preferred the sounds /č/ and /ǰ/. These affricates have a fricative component and were preferred in the pluralization study; however, /š/, which was most highly preferred for pluralization was not preferred for past tense. Anisfeld and Gordon point out that the articulation of /č/ and /ǰ/ involves articulation of /t/ and /d/ respectively, as is exemplified by the fact that they are sometimes transcribed as /tš/ and /dž/. For the two reasons noted above, it is more likely that subjects preferred /č/ and /ǰ/ because they include the English past tense markers in their articulation rather than because they have a fricative component.

The differences in the results for the pluralization study and the past tense study indicate that subjects were responding according to the specific morphological task they were asked to perform. Furthermore, they indicate that it was particular features, related to the actual plural and past tense forms, that they attended to in choosing the nonsense forms. Notably, it was not the full feature specification for the plural inflection that guided subjects' choices of artificial plurals, so that, for example, sounds which share the feature [+voiced] with the plural marker /z/, were not chosen over voiceless sounds. "Of the six features characterizing /z/, only +strident and +continuant did significantly influence... Ss' plural choices." (p. 979) What is significant about these two features is that they are more important in differentiating /z/ from other sounds than the other features and they are the only ones that distinguish the plural suffix /z/ from the past tense suffix /d/. Anisfeld and Gordon conclude that subjects based their choices on whether or not the suffix had "those features most specifically relevant to the differentiation of the plural suffix" (p. 979). According to them this provides empirical support for the contention "that linguistic entities are defined by their role in the system and in relation to each other; their absolute physical properties do not directly identify them" (p. 979).

The results of this study bear on the possible analyses that have been put forward for the productive use of inflections in English. They give further evidence that a feature-based analysis, rather than segmental analysis, is appropriate. They also suggest that it is not only the stem but also the inflection which are analyzed in this way.

Conclusion

The two studies reviewed above provide further evidence for determining the adequacy of the possible analyses considered by Derwing (1980). Anisfeld & Gordon's results indicate that both the "Segment" analysis (5) and the "Feature" analysis (6) are inadequate for describing the processes involved in plural formation. First, subjects' choices of "best" plurals was based on whether or not the artificial plural had the features of +continuant and/or +strident, which mark the real plural inflection. This indicates that learners attend to features and not full segments when using inflections. Secondly, subjects consider these features as markers of the inflection, which suggests that it is not only the features in the the final segment of a word that are attended to for the use of inflections, but also the features that mark the inflection. Only the "Phonotactic" analysis (6) and the "Hybrid" analysis are able to account for these strategies.

CHAPTER THREE

THE EXPERIMENTS

The studies reviewed above provide evidence that the productive use of past tense and plural inflections in English is a developmental process involving various stages. They also provide evidence for the strategies employed at the various stages and suggest that a sub-segmental feature analysis is used in the later stages. In these studies, except Anisfeld & Gordon (1968) the investigations deal only with the generalizations that relate to classes of stem final segments and do not consider the feature composition of the inflections themselves. Nearey points out that "[t]he usual generative phonological analysis... supposes that generalization has taken place with respect to the ending as well" (1981, p.367). Anisfeld & Gordon's study did show that features are attended to in the use of inflections, and though they did not investigate the nature of the morphophonological variations of the regular past tense and plural, their study does give motivation for further investigation into what sort of generalizations do take place with respect to features.

Nearey (1981) described an experimental design that could determine whether the phonological representation of the plural involves generalizations with respect to the ending as well as the stem. The method consisted of teaching subjects nonce words (creature names), represented by pictures. Once the names are learned, Ss are given an expanded set, including some of the creatures in their pre-adult state. (Adults are white and young creatures are blue.) The names of the young creatures are the same as adults, except that a diminutive suffix, /Ž/, is added. In this set only creatures whose names end in resonants are used. Once these are learned, the set is again expanded to include names that end in voiceless consonants and sibilants.

In a pilot study, using six adult subjects, spontaneous devoicing occurred after voiceless stem-final consonants, but in the case of sibilant stems, subjects had difficulty and none spontaneously produced the predicted /əž/ suffix. The results here are only suggestive, but the experimental model seems promising for offering insight into the rules governing the productive use of the past tense and plural inflections.

The present study is an expansion of that design, using analogs for both the past tense and plural inflections. The aim of the research here is to further investigate the nature of "sub-segmental" feature analysis as well as the apparent differences between the rules for voicing assimilation and vowel insertion. It is the intention of this study to provide empirical evidence regarding the nature of the phonological rules and representations of the past tense and plural inflections of English.

The use of the experimental design described by Nearey (1981) seems appropriate because it allows for investigation of these phonological rules without reference to the inflections they are used for. Much as experiments which use nonsense words are able to test for the acquisition and productivity of the inflectional morphemes, the use of nonsense words combined with nonsense inflections allows for testing of the rules that are operative in the concatenation of these forms.

Two experiments were conducted in this study. The first experiment involved both children and adults as subjects in the hope that comparisons of responses would provide insight in regard to the use of these rules. The question was whether adults, whose control of inflectional morphology has been in place for some time, would respond differently than children, who have only recently achieved control. Confounding variables in the experimental design prevented any clear interpretation of the results, so a second experiment, using the same basic design as the first, was also conducted. The design of the original experiment was modified for the second experiment to eliminate some of the confounding variables that were present. Adults were not used in the second experiment, for reasons outlined in the following section.

EXPERIMENT 1

METHOD

Subjects

Two groups of subjects (Ss) were involved in this experiment. One consisted of university students enrolled in an introductory linguistics class and the second of Grade 1, 2, and 3 students from Edmonton area public schools. Ss were required to be monolingual speakers of English and to satisfy the conditions of a screening test. Thirty-two children, ages 6 to 9, were in the first group and 16 adults made up the second group. It was assumed that the sex of the Ss would not be a factor in this experiment, since participation in the experiment was dependent on the S's ability to perform the screening task. Grades 1, 2 and 3 were chosen because E was targeting the age group of 6 to 9 years, since it is around this age that speakers begin to show productive mastery of the past and plural inflections (cf. Innes 1974, Derwing & Baker 1976, Gray & Cameron 1980). Specific age or grade was not considered a factor since a screening test was used to determine if Ss had productive control of these inflections. The validity of this assumption is assessed below.

Screening Tests

Screening tests were conducted to select Ss for the experiment. These were designed to establish that Ss had control of past tense and plural inflections and would also be able to articulate the consonant clusters that would appear with inflected stimulus items in the experiment.

In the first task, Ss were given a Berko-type test to ensure that they had productive control of the regular plural and past tense allomorphs. An example of this task is given below:

A. Plural:

E. "This is a ____" (E shows S a picture of an imaginary creature.)

E. "Now there are two of them." (E holds up second picture.)

E. "There are two..." ____ (S gives response.)

B. Past Tense:

E. "Here's Joe. He is _____ing. He ____s every day."

(E shows S drawing of a boy doing strange acrobatics.)

E. "If he ____s every day, then yesterday he..." _____

(S gives response.)

Three monosyllabic nonce stems were used for each task. The stem-final consonants of these forms were selected on the basis of prior studies, which showed them to be the most difficult for speakers to master in use of past tense and plural inflections. It was assumed that if Ss could inflect these forms correctly they would have achieved full mastery of the inflections involved. The forms used were: /oð, nɔθ, twɛž/, for testing the plural inflections and /boð sweθ, sprod/, for testing the past tense inflections. The required plural forms were thus, /oðz/, /nɔθs/ and /twɛžəz/ and the required past tense forms, /boðd/, /sweθt/ and /sprodəd/.

The second task was an imitation task, which was given to determine the S's ability to articulate the clusters that would appear in the experiment. A randomized list of 26 nonce words, given without reference to meaning, was used. Ss were simply asked to repeat the words after E. This list included examples of obstruent clusters assimilated for voicing and sequences of velar stops and sibilants separated by the reduced vowel, schwa. (See Appendix 1 for the list used.)

Each S was screened individually in a quiet room. Ss who performed perfectly on all three tasks were selected to participate in the experiment. Once Ss were selected, they were then taught the uninflected forms of the stimulus items to be used in the experiment. Each S was given a xerox copy, in pink, of the pictures of the creatures, with the uninflected form of the name of the creature printed below in standard orthography. They were asked to take these home and to learn the names.

Stimuli

A miniature artificial language (MAL) was constructed with an inflectional system that was analogous in features to the English past tense and plural inflections. A MAL was used rather than English so that the results of the experiment would reflect phonological generalizations that were independent of morphological structure.

The past tense inflection in English has as its regular allomorphs the dental stop set {/-d, -t, -əd/}. The analog used in the experiment consisted of a velar stop morpheme with the expected allomorphs {/-g, -k, -əg/}. The plural inflection in English has as its allomorphs the dental sibilant fricative set {/-z, -s, -əz/}. The experimental analog for this set was a palatal sibilant fricative morpheme with the expected allomorphs {/-ž, -š, -əž/}.

The corpus of the language included 16 monosyllabic nonce words. The words had the following breakdown of stem-final consonant types: 4 [+sibilant], 4 [+velar], 4 [-voice, -sibilant, -velar], 4 [+voice, -sibilant, -velar]. The nonce words were represented by pictures of imaginary creatures which, Ss were told, came from another planet. The inflections were used to denote developmental stages of these creatures. (Table 1. provides the set of nonce words, their expected inflected forms, and their orthographic representation. Pictures of creatures are in Appendix 4.)

Table 1. Experiment 1, Stimulus Items

Item	PINK Uninflected	BLUE [+sibilant]	GREEN [+velar]
ZABE	/zeb/	/zebž/	/zebg/
KOID	/koyd/	/koydž/	/koydg/
EV	/ɛv/	/ɛvž/	/ɛvg/
NOTHE	/noð/	/noðž/	/noðg/
IPE	/ayp/	/aypš/	/aypk/
SMEET	/smit/	/smitš/	/smitk/
PLURF	/plr̩f/	/plr̩fš/	/plr̩fk/
STRUTH	/strʌθ/	/strʌθš/	/strʌθk/
YURSE	/yr̩s/	/yr̩səž/	/yr̩sk/
GUTCH	/gʌč/	/gʌčəž/	/gʌčk/
OIZ	/oyz/	/oyzəž/	/oyzg/
SWURGE	/swr̩ǰ/	/swr̩ǰəž/	/swr̩ǰg/
BLAKE	/blek/	/blekš/	/blekəg/
GLURK	/glr̩k/	/glr̩kš/	/glr̩kəg/
DROIG	/droyg/	/droygž/	/droygəg/
OOG	/ug/	/ugž/	/ugəg/

A set of three brightly coloured flash cards was made for each imaginary creature. Different colours were used to indicate the particular stage of development of the creature, in order to prompt the use of a particular inflection. Creatures that were babies were coloured pink and represented by the uninflected form of the nonce word. Blue indicated that the creatures had begun to walk and this was represented by the use of the [+sibilant] inflection. Green indicated that the creatures were old and the [+stop] inflection represented this stage. The uninflected form of the name of the creature was typed in standard orthography at the bottom of each card to provide Ss with a prompt, if necessary.

Training and Testing

Training and testing was done within a week of selecting Ss. E was given a quiet room in the school to conduct the experiment and Ss were trained and tested individually.

Ss were first asked to review the names of all of the creatures. Once E was assured that the S knew all of the names, a subset of these items was used for teaching the inflected forms. Ss were randomly assigned to two groups for testing. Group 1 was taught with the items that were [+voiced, -sibilant, -velar] (viz., zabe, koid, ev, nothe), and Group 2 was taught with the items that were [-voiced, -sibilant, -velar] (viz., ayp, smeet, plurf, struth).

Training consisted of showing the S the different coloured creatures and telling a story.

> E. "This is a [noð]. (S repeats name.) When it is born it is pink, but when it starts to walk, it turns blue and its name changes and it is called a [noðž]. (S repeats name.) Then when it gets old, its name changes again and it is called a [noðg]. (S repeats name.)

This procedure was used with all four training items, and then E presented the cards in the order pink > blue > green, having Ss supply the correct form until E was satisfied that the S knew the correct inflections. If Ss had difficulty with an inflection, E would take the set of cards which required that inflection (e.g., greens for the [+velar] inflection) and review all of the inflected names until S was comfortable with the set. E would then repeat the procedure with the other set of inflected items (i.e., blues for the [+sibilant] set).

Once E was satisfied that S knew all of the inflections, the training set was shuffled in with the rest of the pink coloured cards (uninflected forms) and the S was asked to provide names for all of the items. This was done until the S could provide all of the names correctly. At this point actual testing began.

In the test itself, Ss were presented with the full set of stimulus items and asked to provide the correct name for each item. The training set was used as the test stimuli for the inflectional variant that Ss were trained on, with novel items for the other two variants. These were presented in one of two orders, forward and reverse of a single random ordering that was established prior to conducting the experiment.

Responses were taped on a Panasonic portable cassette tape recorder (Model No. RQ-2103) and transcribed by E as they were given. The whole process of training and testing usually took from 20 to 30 minutes for each S.

Ss enjoyed the task for the most part and in general appeared to be comfortable with it. Some of the younger Ss had difficulty paying attention to the task for the whole presentation. If that happened, E would allow them to talk about other things for a few minutes and then redirect them to the task. Sometimes Ss had difficulty remembering which inflection went with which colour. If this happened, E reminded them that pink creatures were babies, blue ones walked and green ones were old, without referring to the phonological shape of the inflection. This information usually was sufficient to get Ss back on track.

Contrary to the assumption of the screening task, articulation seemed to pose the greatest problem. If a S had difficulty articulating an inflected stem, E said the *uninflected* form of the word, had S repeat it and then asked again for the inflected form. This did not always result in a "correct" response, but it usually served to alleviate any anxiety the S might have had.

There was only one S who had to be dropped from the experiment. She had no difficulty with the screening tasks, nor even in training, where inflected forms were articulated for her before she had to produce them herself; however, in the production task she was unable to spontaneously produce the sibilant inflection, consistently replacing /š/ and /ž/ with /s/ and /z/.

RESULTS AND DISCUSSION

Summary of Results

Responses were scored 1 for correct and 0 for incorrect, based on the rule-predicted responses for each inflected stem. (See Table 1.) Since four stimulus items were given to elicit each inflectional variant, a score of 4 indicated 100% correct response for any given variant.

A five factor repeated measures analysis of variance was performed on the data. There were three between-group factors of Age (adults vs. children) x Training ([+voiced] vs. [-voiced]) x Order (forward or reverse). There were two within group factors: Suffix Type and Variant. Suffix Type had 2 levels, velar and sibilant. Variant had 3 levels, voiced, voiceless and vocalic.

When the traditional .05 significance level was adopted, the results showed a Variant X Training X Age X Order interaction ($p \leq .032$) and a Training X Age X Order interaction ($p < .023$). However, no clear interpretation could be found for these interactions involving Order, so it was decided to adopt the more conservative .01 significance level for all analyses in this study.

With $\alpha = .01$, the results of the analysis showed a Variant X Training interaction ($p < .001$) and a Variant main effect ($p < .001$). A Tukey HSD multiple comparisons procedure was performed on the means of the Variant X Training interaction and a graph was plotted (see Table 2. and Figure 1.). The results of the comparisons showed that there were significant differences between each of the inflectional variants within each training group, except for the [+voiced] and [+vocalic] variants in the [-voiced] training group, T2. It also showed that Ss performed better on the items on which they were trained (e.g., [+voiced] or [-voiced]) than on those they were not.

Table 2. Tukey Test — Variant X Training

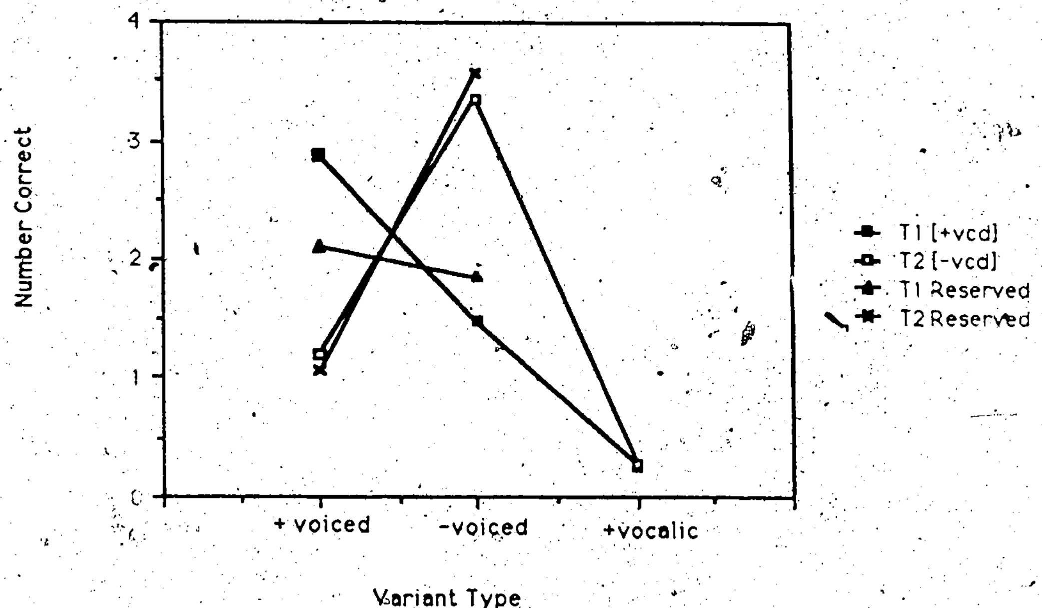
Critical Value ($\alpha = .01$) = 1.019

	T2 [-voiced]	T1 [+voiced]	T1 [-voiced]	T2 [+voiced]	T2 [+vocalic]	T1 [+vocalic]
T2[-vcd]		0.4688	1.8594*	2.1407*	3.0625*	3.0782*
T1[+vcd]			1.3901*	0.6719*	2.5937*	2.6095*
T1[-vcd]				0.2813	1.2031*	1.2188*
T2[+vcd]					0.9210	0.9375
T2[+voc]						0.0157
T1[+voc]						

Scores

T2 [-voiced]	3.3438
T1 [+voiced]	2.8750
T1 [-voiced]	1.4844
T2 [+voiced]	1.2031
T2 [+vocalic]	.2813
T1 [+vocalic]	.2656

Figure 1. Variant X Training Interaction



Reserved Data

It is not surprising that Ss did well on the training items, since, as noted above, they could well have memorized them. In order to get some idea if Ss were able to transfer training to new words, responses for items not included in the ANOVA were scored and compared with the Variant X Training responses. These reserved data are the words with [+velar, +stop] stem-final consonants (viz., /blek, ug, droyg, glɾk/) which were inflected with the [+sibilant] ending, and the words with [+sibilant] stem-final consonants (viz., /yɾs, gʌč, swɾǰ, oyz/) which were inflected with the [+velar, +stop] ending. The scores for these items are shown in Table 3.

Table 3. Tukey Test—Reserved Data — Variant X Training

Critical Value (α = .01) = .9393

	T2 [-voiced]	T1 [+voiced]	T1 [-voiced]	T2 [+voiced]
T2 [-vcd]		1.452*	1.716*	2.500*
T1 [+vcd]			0.264	1.048*
T1 [-vcd]				0.784
T2 [+vcd]				

Scores

T2 [-voiced]	3.560
T1 [+voiced]	2.108
T1 [-voiced]	1.844
T2 [+voiced]	1.060

(These reserved items were not used in the analysis because they would have created an imbalance in the data set, i.e., there would have been 6 responses for the voiced and voiceless variants of the inflections, but only 4 for the vocalic variant.)

The response scores for the [-voiced] training group for this set of data fell into the same pattern as the analyzed data set. What is interesting is that response scores of the reserved data for the [+voiced] training group were not parallel to the analyzed data. Here, the scores for [+voiced] and [-voiced] variants were about the same for the reserved data, while there was a significant difference between these scores in the analyzed data (see Figure 1.).

Discussion

The Variant X Training interaction is not surprising. Since Ss were trained on the the same items they were tested on, it was expected that they would perform best on those items. The lower scores for the items where voicing assimilation was expected to take place are not so easily explained. It was expected that Ss would have little difficulty with them, but in both training groups Ss provided correct responses less than half of the time. Many of these errors were the result of Ss' trying to maintain the phonetic shape of the suffix (e.g., if they were taught the [-voiced] variants, when presented with a green "koid" would articulate /koydk/, or when taught the [+voiced] variant, when presented with a green /ayp/ would articulate either /aypgə/ or /aybg/). The number of correct responses for the vocalic variant were very low, suggesting that the vowel insertion rule is not as readily available as voicing assimilation, which did at least occur spontaneously in some cases.

The response scores for the reserved data are the most interesting. For these items Ss in the different training groups appeared to be using different strategies for transferring training. The [-voiced] group (T2) did as well on non-training items as on training items requiring the voiceless variant and still performed poorly on those items requiring the voiced variant, which was parallel to their performance on the items originally analyzed. (The majority of the errors for the items requiring the voiced variant were due to using the voiceless variant with stem-final voiced consonants.) However, in the [+voiced] training group (T1), the reserved data scores were not parallel to the analyzed data. In this case, the scores for items which required the voiced variants were about the same as the scores for the stems requiring voiceless variants. What seems to be happening here is that T2 Ss simply maintain the phonetic shape of the suffix for all stems, but that T1 Ss were trying to develop a strategy to deal with all stems effectively. This gives some indication that a voicing assimilation rule for English morphology may go in only one direction, that is, it is a devoicing rule.

No firm conclusions can be made here about these differences, but they did give motivation for conducting a second experiment which would focus on the nature of the voicing assimilation rule to see if it is a general rule, applying both for voicing and devoicing, or if it is simply a devoicing rule. The second experiment would also look at the differences between the rules of Voicing Assimilation and Vowel Insertion.

Confounding Variables

It seemed best to treat this experiment as a pilot study and to perform a second experiment which would be modified to eliminate some of the factors involved in the design. The first consideration was that list order should be eliminated as a factor, in order to eliminate the possibility of a significant interaction.

There were a number of factors, aside from order, that seemed to operate to confound the results in this experiment. One of these factors was the difficulty of articulation of the nonce stems. The repetition task in the screening test was given to ensure that Ss would be capable of articulating all of the forms of the inflected stems, but this did not happen. Many Ss had no difficulty with the repetition task, but often had difficulty creating the inflected forms of the stimulus items. Obviously it is easier to pronounce a word in a simple imitation task than it is to produce a novel form spontaneously.

Another problem with articulation occurred with the words having /θ/ and /ð/as stem-final consonants. They were often confused with each other, so that an item ending with one would be articulated with the other, e.g. /strʌð/ for /strʌθ/, or /noθ]/for /noð/. The responses for these forms were scored as incorrect.

/t/ and /d/ stem-final consonants also posed a problem with the [+sibilant] inflection. When these were combined the articulatory results were the affricates /č/ and /ǰ/. Once these shapes appeared, a few Ss began to use these sounds for the [+sibilant] inflection, replacing /ž/ and /š/. These responses were also scored as incorrect.

In the absence of any clear way to "factor out" such errors from the data in this experiment, it was decided to design and run a second experiment that would reduce the likelihood of their occurrence.

EXPERIMENT 2

MODIFICATIONS IN DESIGN

The design of the second experiment was modified to focus on the following questions. Is there a devoicing rule, but not a voicing rule (i.e., what is the nature of voicing assimilation in English inflectional morphology)? What sort of rule is operating for stems that are expected to be inflected with the vocalic variant? Is Vowel Insertion a different type of rule than Voicing Assimilation? The issue that is raised here is the same as that raised by Derwing & Nearey (1986), which is whether these rules represent broad phonological generalizations or whether they are limited to specific morpheme classes.

One of the first considerations in modifying the design of the experiment was to eliminate the unnecessary between group factors. The age factor was eliminated by not including adults in the second experiment. This is not a developmental study, so the differences in responses between children and adults were of no particular interest. I chose to concentrate on children's data for a number of reasons. In studies of second language acquisition, it has been found that there is a possible "critical period" for learning another language (cf. Lenneberg 1967). This theory indicates that children who learn a second language before the onset of adolescence are better able to achieve fluency in that language. Though, in fact, I expect the children to transfer their English "accent" to the new language, so in that sense they are not learning a second phonological system, I expect that the ability of young children to assimilate this second "language" will be better than that of adolescents or adults. These differences are interesting in themselves, but it is not within the scope of this study to examine this issue. The reason why the original data were not simply re-analyzed, leaving out the adults' responses, was that other factors appeared to be confounding the results. The amount of confounding in the first experiment made it seem better to focus on the rules used in one age group first and develop a methodology that could be used later for adult/child comparisons.

Finally, children seemed much more interested in performing the experimental tasks. They usually wanted to spend more time with the stimulus materials after the testing was completed, while adults were generally glad to have it over with. This interest factor was important because Ss had to spend more time with the stimulus materials and in training and testing in the second experiment.

List order was also eliminated as a between group factor. The only motivation for minimizing order variations in the original design was to make transcription easier for the examiner. In the second experiment stimulus items were randomized separately for each S in each trial.

Other confounding variables are not as easy to detect and eliminate, but I found a few elements in the original design that could have made some differences in responses. Learning the inflectional system of a language and the corresponding phonological rules is a process that takes considerable time. In my original design, all training and testing was carried out within one twenty-minute period for each S. With such a time constraint, it is not surprising that there were many errors in the responses. Given more time with the stimulus materials, Ss might be expected to do better in terms of correct responses, in that they would have time to make some analogy, consciously or unconsciously, between the inflections of the MAL and those of English and their corresponding phonological rules. In the original experiment Ss were also expected to produce "correct" inflected forms on the first trial. Given that the Ss had never encountered the inflected forms of these words before, it is not surprising that they often had difficulty articulating them as expected.

For the second experiment I therefore changed the training and testing schedule. Ss were taught a set of inflected training stimuli and the set of uninflected test stimuli on one day and on a second day they reviewed training and were asked to provide inflections for the test stimuli. Ss were also required to go through the stimulus items twice in a session, instead of just once. On a third day Ss repeated the same test given on the second day, which gave four trials for each S. These changes were hoped to make it easier for E to infer what strategies were being employed by the Ss, as well as allow testing for a possible learning effect. Training and testing occurred on consecutive days to help ensure that Ss remembered training items and maintained their interest.

In the preceding section it was mentioned that words with the stem-final consonants /t/, /d/, /θ/, and /ð/ caused articulation errors unrelated to the phonological rules in question. To eliminate this problem, these consonants were not used in stem-final position in the second experiment, where they were replaced by /p/, /b/, /f/ and /v/, respectively.

Another change that was made to deal with articulatory problems was to simplify all nonce stems to CVC's. In the original experiment the nonce words had no systematic phonological shape. There were VC's, CCVC'S and CVC's. Diphthongs and vocalic r's were also used in the V position. This may not have had any effect on Ss performance, but to ensure that the stems were as easy to articulate as possible, there were no word initial consonant clusters, diphthongs or vocalic r's used in the second experiment.

Another confounding variable which made itself evident in the Variant x Training interaction was that the stimulus items in the training set were also treated as test items in the analysis. It is not at all surprising that Ss did significantly better on the items that they were trained on, as they could well have recalled them by rote. Comparisons of the responses for the reserved data, requiring the same endings as the training stimuli, with responses for the training stimuli showed differences in the [+voiced] training group (see Figure 1 and discussion above). In order to control for this in the second experiment, a subset of training stimuli was introduced; these were similar in form to the test stimuli, but the responses for these items were not included in the analysis (see Table 4. for the training sets used for each training group.). In this way, differences in responses based on transference for both the same and different types of stem endings could be evaluated without being confounded by the fact that Ss may have memorized the inflected forms of the training items.

One further change made to the design of the second experiment was to add a third training group as a type of control. This group was trained with all three of the phonological variants of the two inflections. This permitted a comparison of the results of the other two training groups with a group that did not require any extrapolation or rule learning to a novel set of inflections.

Finally, Grade 3 and 4 students were used instead of students from Grades 1-3. Once again, since developmental trends were not the focus of this research, it seemed more convenient to use a more homogeneous age group, and since the younger children had the most difficulty in performing the task, I chose to eliminate them from the second study.

METHOD

Subjects

Subjects for the second experiment were third and fourth grade students from three Edmonton area public schools. There were three training groups, which were divided by school such that all Ss from one school received the same training. This was done to prevent contamination of training groups through contacts during recess or after school. There were 15 Ss from each school for a total of 45 Ss. The mean age of the Ss was 8 years, 10 months.

Screening Tests

The same screening test that was used in the first experiment was employed for selection of Ss in this experiment, the only difference being in the clusters presented for the articulation test. (See Appendix 1 for forms used.)

Stimuli

The stimulus items for this experiment were similar to those used in the first one, except as noted in the preceding section. Group 1 training items all ended in either /b/ or /v/ and Ss were taught the [+voiced] inflectional variants, /g, ž/. Group 2 training items ended in either /p/ or /f/ and Ss were taught the [-voiced] inflectional variants, /k, š/. Group 3 training items ended in /k/, /g/, /č/ and /z/, so that all three variants of the two inflections (i.e., /ž, š, əž/ and /g, k, əg/) could be taught. Testing stimuli included the following breakdown of stem-final consonants: 4 [+sibilant], 4 [+velar], 4 [-voiced, -sibilant, -velar], and 4 [+voiced, -sibilant, -velar] (See Table 4.).

Training

If a S showed productive use of the English plural and past tense and had no trouble with the repetition task, he/she was trained for the experiment. Ss were trained individually, and training immediately followed the screening test.

In training, Ss were given all of the uninflected forms of the stimulus items to be used in the experiment. They were shown the pink coloured picture of the creature which represented the name and E told the Ss a story about them.

> E. "These are creatures from another planet. They are all babies. We can tell that they are babies because they are pink. They look funny, and they have funny names. Let's look at the pictures and I will tell you what their names are, then you can say them after me."

Table 4. Experiment 2. Stimulus Items

Nonce Word	Uninflected	[+sibilant]	[+velar]
ZABE	/zeb/	/zebž/	/zebg/
TIB	/tIb/	/tIbž/	/tIbg/
LEV	/lɛv/	/lɛvž/	/lɛvg/
SEAVE	/siv/	/sivž/	/sivg/
LEP	/lɛp/	/lɛpš/	/lɛpk/
KIP	/kIp/	/kIpš/	/kIpk/
PAFF	/pæf/	/pæfš/	/pæfk/
GIFF	/gIf/	/gIfš/	/gIfk/
SADGE	/sæǰ/	/sæǰəž/	/sæǰg/
MUZZ	/mʌz/	/mʌzəž/	/mʌzg/
GUTCH	/gʌč/	/gʌčəž/	/gʌčk/
BISS	/bIs/	/bIsəž/	/bIsk/
BOOG	/bug/	/bugž/	/bugəg/
DEG	/dɛg/	/dɛgž/	/dɛgəg/
GACK	/gæk/	/gækš/	/gækəg/
BECK	/bɛk/	/bɛkš/	/bɛkəg/

(Training Items [+voiced])

GIB	/gIb/	/gIbž/	/gIbg/
FEB	/fɛb/	/fɛbž/	/fɛbg/
SAV	/sæv/	/sævž/	/sævg/
TEV	/tɛv/	/tɛvž/	/tɛvg/

Nonce Word	Uninflected	[+sibilant]	[+velar]
(Training Items [-voiced])			
GIP	/gIp/	/gIpš/	/gIpk/
FEP	/fɛp/	/fɛpš/	/fɛpk/
SAFF	/sæf/	/sæfš/	/sæfk/
TEFF	/tɛf/	/tɛfš/	/tɛfk/
(Training Items — All Variants)			
GIZZ	/gIz/	/gIzəž/	/gIzg/
FEG	/fɛg/	/fɛgž/	/fɛgəg/
SATCH	/sæč/	/sæčəž/	/sæčk/
TECK	/tɛk/	/tɛkš/	/tɛkəg/

E showed S all of the pictures and reviewed the uninflected forms of the stimulus items with her/him. E then continued the story to introduce the inflected forms with the training stimuli.

> E. "We can tell that these creatures are babies because they are pink, because on the planet where they come from they change colours when they grow up. Another thing that happens is that their names change a little bit.

E shows Ss pictures of the creatures coloured blue.

> E. "When these creatures start to walk they turn blue. Here we have a baby /fɛb/ (E. shows pink coloured creature.) When he starts to walk he turns blue and now he is called a /fɛbž/. Can you say that?

E reviewed all four training stimuli with the S in the same manner. Once the S was comfortable with those, E taught the second inflection in the same manner, explaining that when the creature gets old, it turns green, and its name changes again (i.e., "It is called a /fɛbg/.").

Ss usually had little difficulty with the training and often could supply the correct inflections for the other training stimuli after being taught the first one. (For example, when a S was taught /fɛbž/, on being shown a blue /tɛv/, would say, "/tɛvž/!".)

E then reviewed all the training stimuli with the S, first in order of development, then in a random order. If S had any difficulty in the review, the training would be repeated until S could produce all of the training items with no difficulty. Once E was assured that S was familiar with the inflections, the training stimuli were shuffled in with the pink coloured cards (representing uninflected forms of the testing stimuli) and S was asked to provide names for all of the creatures. Finally, Ss were given a xerox copy of the pictures of all of the creatures on pink paper and told that E would be coming back the next day to show them some more of the creatures. Ss were asked if they would like to come again and their response was generally quite enthusiastic.

Group 1 was taught the [+voiced] variant of the inflections with the training stimuli /fɛb, gIb, tɛv, sæv/. Group 2 was taught the [-voiced] variant of the inflections with the training stimuli /fɛp, gIp, tɛf, sæf/. Group 3 was taught all of the variants of the inflections with the training stimuli /fɛg, tɛk, gIz, sæč/.

Testing

The day following training E returned to perform the experiment. Ss were tested individually in a quiet room in the school. All responses were transcribed by E and taped on a portable cassette tape recorder. Two tape recorders were used during the experiment, one was a Panasonic (Model No. RQ-2103) and the other was a Craig (Model No. J101). Training was reviewed with each S to ensure that he/she was familiar with the inflections. If S had any difficulty, E repeated training by showing S the pictures of the creatures representing the training stimuli and reviewing their names. First the uninflected forms were presented, then the inflected forms were presented in sets, i.e., all the blue ones ([+sibilant] inflection) together and then all the green ones ([+velar, +stop] inflection) together.

The next step was to shuffle all the training stimuli together with all of the pink cards (uninflected forms) and have S provide the names for the pictures. Once Ss could go through the cards with no errors they were asked to do it once again while E taped them. This was done to help Ss become less self-conscious about being taped..

Finally Ss were told that E had some more pictures of the creatures and that these were all the blue and green ones that matched the pink ones they had already seen. E said that she had taught them the names of the creatures that she knew, but she wasn't just sure what to call the other ones and she needed their help to find out what might be the right way. E then asked them to tell her what they thought the names of the creatures might be. At this point Ss were presented with all of the cards, including training stimuli, and their responses were recorded.

After S provided responses for the stimuli once, E reshuffled the cards and asked S to go through them a second time. Responses were again transcribed and taped. After the second trial Ss were asked if they would come back the next day to repeat the task and Ss generally responded favourably.

The next day the second set of trials was conducted. However, on this day, because of time constraints, a second E also performed testing and transcriptions. There were three other Es, all graduate students in linguistics, who helped with the testing and transcriptions. A different second experimenter was used for each training group. Each one sat in on a couple of sessions with E in order to familiarize him/herself with the experimental procedures, and then performed testing on their own. Both E's transcribed responses when they were testing together as one measure of reliability of transcriptions (see Appendix 2. for comparisons). Otherwise, procedures were the same as for the first testing session. Training was reviewed and then stimuli were presented. Each S performed two trials in each session for a total of four trials per S over a period of two days.

In general Ss were comfortable with the task, although in Training Group 3 Ss seemed to be more concerned with getting the "right" answer than Ss in the other training groups. This may have been because they had to learn three variants for each inflection and therefore had more to remember. As one S noted, each time he was presented with a new item, he would try to match it to one he already knew and find the ending that would fit. Assuring the Ss that there was no "right" answer, but that E wanted to find out what they thought would be good usually helped to alleviate tension.

Two Ss had to be dropped from the experiment during the testing. One S, from Training Group 2, missed school because of illness so wasn't available for the second testing session. A second S became so anxious about providing the "correct" response that she couldn't perform the task.

CHAPTER FOUR

RESULTS AND DISCUSSION

DATA ANALYSIS

Scoring system

An initial analysis was done to determine the percentage of correct responses and error types. Responses were coded 1-6 according to the following system:

1 [+voiced] – /ž/ or /g/ used to inflect the stimulus item (e.g., /bugž/, /lɛvg/, /kIpž/ or /pæfg/).

2 [-voiced] – /š/ or /k/ used to inflect the stimulus item (e.g., /lepš/, /gIfk/, /dɛgš/ or/tIbk/).

3 [+vocalic] – /əž/ or /əg/ used to inflect the stimulus item (e.g., /sæčəž/, /bɛkəg/, /lɛvəž/ or /kIpəg/).

4 zero ending (ø) – no suffix added (e.g., /bug/ rather than /bugəg/).

5 V or Asp – reduced vowel or heavy aspiration added to stem (e.g., /bɛkh/, /bugə/, or /bIsšə/).

6 Other – stem final consonant altered or omitted or suffix altered (e.g., /bIš/ rather than /bIsəž/, or /gIfč/ rather than /gIfš/).

Percentages were tabulated for types of responses for each item in each trial.

Analysis of Variance

For the analysis of variance, responses were scored as either correct or incorrect. A four factor repeated measures analysis of variance was performed on the data. The between group factor was Training, with three levels: [+voiced] (T1), [-voiced] (T2) and All Variants (T3). The within group factors were Trials, Suffix Type and Expected Variant. Trials had four levels, each subject having performed four trials. Suffix Type had two levels, velar and sibilant and Expected Variant had three levels, voiced, voiceless and vowel insertion.

The conservative .01 level of significance was adopted for the analyses in this experiment, as was done in the first experiment.

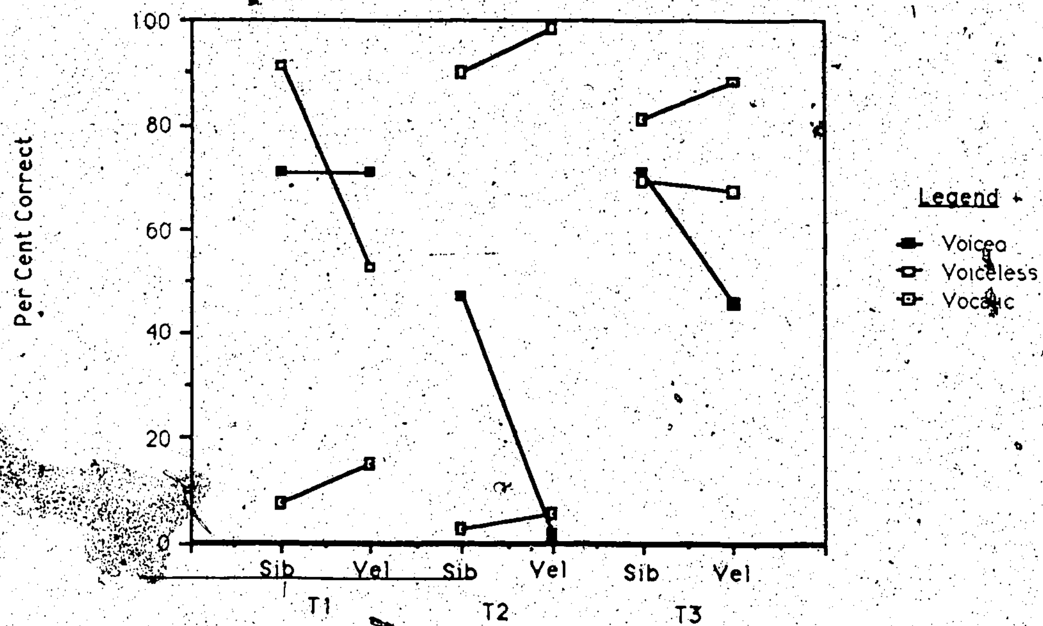
There was a significant 3rd order interaction of Training x Suffix x Variant ($p < .001$). The Training x Variant interaction was significant ($p \leq .001$), as well as the Suffix x Variant interaction ($p < .001$). There were also main effects of Suffix ($p < .001$) and Variant ($p < .001$). The Training x Suffix interaction was not significant. Differences in scores between trials were shown to be non-significant at any level ($p < .01$), so were not considered in the interpretation of the results.

A graph was plotted for the Training X Suffix X Variant interaction (see Figure 2.) and Tukey tests were performed on the means.

Reliability of Transcription

To determine the reliability of the transcriptions, and to make sure that different transcriber's biases did not affect the results of the analysis, comparisons were made of the transcriptions done when both Es tested together. Overall agreement between the alternate transcribers and the author was .85. (See Appendix 2. for the details on each trial for which comparisons were made.) A second analysis of variance was also done on a subset of the data, including only the data from Ss which I had tested in all four trials. The results of this ANOVA showed the same interactions, and similar means for each training group, as were found in the larger set of data. Therefore the results from the larger set of data are reported. (A graph of the means is found in Appendix 3.)

Figure 2. Training X Suffix X Variant Interaction



Training X Suffix X Variant Interaction

Looking at Suffix within Training and Variant, in the [+voiced] training group (T1), the mean scores for the sibilant and velar inflections were the same for the voiced variant; however, for the voiceless variant the scores for the velar suffix were significantly lower than those for the sibilant suffix. The scores for the vocalic variant were higher for velars than for sibilants, but the difference was not significant.

In the [-voiced] training group (T2), the scores tended to be lower for sibilants than for velars for the voiceless and vocalic variants, but not significantly. For the voiced variants, the scores for velars are significantly lower than the sibilant scores.

In the All Variants training group (T3), little difference was shown between the sibilants and velars for the voiceless variant. For the voiced variant the velar scores were significantly lower than the sibilant scores, but with vowel insertion the velar scores are slightly better (non-significantly) than the sibilant scores.

Suffix X Variant within Training

On the graph (Fig. 2), a pattern emerges which suggests that velar and sibilant suffixes are treated differently in all training groups. With the [+voiced] variant, in T2 and T3, sibilant scores are significantly higher than velar scores. This pattern has a possible articulatory explanation. It may simply be easier to carry voicing over two segments when the second one is a fricative. When two stops occur together in English, there must be a total occlusion for the first stop before the second one can be articulated. If the second segment is a fricative, the division between the two segments is not as salient. In T1, the velar and sibilant scores for the [+voiced] variant are equal. This is possibly the case because Ss were taught this variant. However, when we look at the strategies used for the velar inflection in a following section, we find that T1 Ss also had difficulty with the voicing of the velar stop inflection.

The results for the [-voiced] variant are not so clear. In T1, the [-voiced] sibilant scores are significantly better than the [-voiced] velar scores. What seemed to be happening during the experiment was that the sibilants underwent a natural stem-final devoicing, particularly after [-voiced] stem-final consonants (this also occurred with [+voiced] stem-final consonants in some cases). With the velars the training seemed to carry more weight, and Ss made efforts to maintain voicing of the inflection. The main strategy was to add a vowel after the inflection to produce a form like /lɛpgə/ instead of /lɛpk/. The other strategy used added voicing to the stem-final consonant to produce a form like /lɛbg/.

In T2 and T3 there were no significant differences between velar and sibilant scores for the [-voiced] variant. For the vocalic variant, velar scores were always higher than sibilant scores, but this difference was not significant.

ERROR PATTERNS

The ANOVA gave a general view of what was going on with the data, but in order to explore the data further, a descriptive analysis of the raw scores was undertaken for the sibilant and velar inflections separately. The suffix types were separated because the significant differences in response scores for the voiceless variant in T1 and the voiced variant in T2 and T3 suggested that Ss may have used different strategies for the different suffix types. Tables and graphs were constructed to show the relationship of training to use of the expected variant and the types of errors that Ss made. The percent correct for each variant type in each training group was calculated and common error-type patterns were also identified, with percentages calculated for those as well. The types of errors considered were based on the scoring system set up for the initial analysis of percentages of correct responses and error types. In order to get some indication of the effect of training, percentages for correct responses and errors for the training items were also calculated (see Tables 5-10 for all of these results). Since no simple statistical techniques exist for the analysis of repeated measures categorical data, the following investigation is descriptive only and no firm quantitative conclusions can be drawn from it.

Sibilants

T1 ([+voiced] Training)

For both training and testing items with voiced stem-final consonants, 71% of the responses were correct. 25% of the errors on the test items and 23% on training items were the result of devoicing the inflection following the voiced stem-final consonant (e.g., /bugš/ rather than /bugž/). Scores for the voiceless variant were very high, 92%, with only 3% of the responses involving voicing of the inflection (e.g., /lɛpž/ rather than /lɛpš/). The distribution of errors for these two variants show a strong tendency towards devoicing, which supports a phonotactic explanation for the use of the voiceless variant in the production of inflected forms.

Table 5. Sibilants—Percentages of response types — T1 [+voiced] Training

Variant	Response Type [+voiced]	[-voiced]	[+vocalic]	null	V or Asp	Other
[+voiced]	71	25				4
[-voiced]	3	92	1			4
[+vocalic]	4	21	8	19	2	46
Training Items	71	23	2	1		3

By contrast, only 8% of the responses were correct for the vocalic variant. In 21% of the responses the voiceless variant was used (e.g., /bIsš/), and 19% were a null response (e.g., /bIs/). The highest percentage of errors, 46%, consisted of giving an irrelevant, "Other," response (Ss often deleted the stem-final consonant of the word and replaced it with the inflection, e.g., /bIž/ or /bIš/.). These responses suggest that the use of vowel insertion is not motivated by the same type of general phonotactic constraint that is used for the voiceless variant.

Table 6. Sibilants — Percentages of response types — T2 [-voiced] Training

Variant	Response Type [+voiced]	[-voiced]	[+vocalic]	null	V or Asp	Other
[+voiced]	**47**	46	7			
[-voiced]	2	**90**	1	1		6
[+vocalic]	1	45	**3**	19	1	32
Training Items	2	**92**				6

T2 ([-voiced] Training)

In this training group, 47% of the responses for the voiced variant of the sibilant inflection were correct, with 46% of the responses accounted for by devoicing of the inflection (e.g., /zebš/ instead of /zebž/). "Other" responses account for 7% of the errors. As with T1, correct responses for the voiceless variant were very high, at 90%. 2% of the responses involved voicing of the inflection, 1% the result of using the vocalic variant, 1% null responses, and 6% "Other." On training items the distribution was similar with 92% correct responses, 2% of the responses using the voiced variant and the remaining 6% "Other." Like T1, T2 Ss show a preference for devoicing; although almost half of the responses for the voiced variant were correct, the other half simply maintained the devoiced quality of the inflection. This gives further evidence that the phonotactic constraint on voicing assimilation of stem-final consonant clusters goes in the direction of devoicing rather than voicing, i.e., that the voiceless variant is the unmarked form of the inflection.

Only 3% of the responses for the vocalic variant were correct, with 45% of the responses being the voiceless variant, /š/. Nineteen percent of the responses were null responses while 32% were "Other." Here again it is shown that vowel insertion is far less productive than voicing assimilation in inflectional morphology.

T3 (All Variants Training)

Here responses for the voiced variant were 71% correct. The vocalic variant accounted for 24% of the errors (e.g., /tɛvəž/ rather than /tɛvž/). Responses for the training items were 92% correct, 3% voiceless (e.g., /tɛvš/) and 3% vocalic (e.g., /tɛvəž/). The voiceless variant responses were similar, with 69% correct and 29% involving erroneous use of the vocalic variant. The training item responses were, 93% correct and 5% vocalic variant.

Table 7. Sibilants—Percentages of response types-T3 All Variants Training

Variant	Response Type [+voiced]	[-voiced]	[+vocalic]	null	V or Asp	Other
[+voiced]	71	3	24			2
[-voiced]	1	69	29			1
[+vocalic]	1	2	81	2		14
Training Items						
[+voiced]	92	3	3			2
[-voiced]		93	5			2
[+vocalic]			93			7

In T3 response errors for the voiced and voiceless variants followed a different pattern than in the first two training groups. The high percentage of errors due to using the vocalic variant rather than the voiced (24% vocalic) and voiceless (29% vocalic) variants suggest that Ss employed a common strategy of choosing one form that could be used for all instances of new forms. That the vocalic variant was the most commonly picked is not surprising in view of the fact that T3 Ss received twice as much training on this variant as on the other two. The differences in scores between training and test items here, with the high percentage of errors involving use of the vocalic variant suggest that some Ss are using a "List" analysis for those training items requiring the voiced and voiceless variants (i.e., simply memorizing them), and choosing the vocalic variant as the productive form of the inflection. (Note that the number of correct responses for the training items still remained high for all three types.)

The percent correct for the vocalic variant were the highest in this group with 81% correct, and 14% "Other" responses. Responses for the training items were 93% correct and 7% "Other." The relatively high number of "Other" responses and the fact that Ss use the vocalic variant on forms where it is not required suggest that Ss who are trained with this variant are at least able to exercise productive control over it. The high percentage of correct responses for the vocalic variant in this group, compared with the very low scores in the other training groups, gives further indication that the use of the vocalic variant may be governed by a different process than voicing assimilation. That is, the vocalic variant must be taught before it is used productively, whereas voicing assimilation (i.e., devoicing) appears to be relatively "automatic."

Velars

T1 ([+voiced] Training)

In the [+voiced] training group, the responses for the voiced variant of the velar inflection were: 71% correct and erroneous use of the voiceless variant accounted for 27% of the responses (e.g., /zebk/ rather than /zebg/). Training item responses were similar with 71% correct and 26% voiceless. For the voiceless variant 53% of the responses were correct, while 43% involved use of the voiced variant (e.g., /lɛpgə/ rather than /lɛpk/). These responses still show a strong tendency towards devoicing, although there is a high percentage of voiced responses where the voiceless variant is expected. Further examination of the raw data provided an explanation for this phenomenon which is discussed later on in this chapter.

Responses for stems requiring the vocalic variant in T1 were: 16% correct, 22% voiced (e.g., /gækgə/) and 5% voiceless (e.g., /gækk/). The most popular responses involved either the addition of a vowel following the stem (e.g., /bugə/), 26%, or the null response, 24%. The remaining 7% were "Other" responses. Note that, as for the sibilant suffix, the number of *correct* responses for this variant was quite low and was not used with any more frequency or consistency than any other strategy.

Table 8. Velars — Percentages of response types -T1 [+voiced] Training

Variant	Response Type [+voiced]	[-voiced]	[+vocalic]	null	V or Asp	Other
[+voiced]	**71**	27	1			1
[-voiced]	43	**53**	1	2		1
[+vocalic]	22	5	**16**	24	26	7
Training Items	71	26		1		2

T2 ([-voiced] Training)

The responses in T2 show quite a different pattern than T1 responses, indicating a marked preference for the voiceless variant. In this group only 2% of the responses for stems requiring the voiced variant were correct, with 97% involving the erroneous voiceless variant (e.g., /zebk/ rather than /zebg/). This erroneous use of the [-voiced] variant with stems ending in a [+voiced] segment gives an indication that voicing assimilation for English inflections is a progressive devoicing rule. For the stems requiring the voiceless variant, 98% of the responses were correct.

Table 9. Velars — Percentages of response types — T2 [-voiced] Training

Variant	Response Type [+voiced]	[-voiced]	[+vocalic]	null	V or Asp	Other
[+voiced]	2	97	1			
[-voiced]		98	1			1
[+vocalic]		40	6	27	17	10
Training Items		99				1

The responses for velar stop stems requiring the vocalic variant were: 6% correct, 41% voiceless (e.g., /bɛkk/), 26% null, 17% heavy aspiration (e.g., /gækh/), and 10% irrelevant. Though the percentage correct is higher here than for the stems requiring the voiced variant, the distribution of errors indicate that production of the vocalic variant is not an easily learned or predictable strategy.

T3 (All Variants Training)

In the "all variants" training group responses for the stems requiring the voiced variant were: 46% correct, 32% voiceless (e.g., /tɛvk/) and 22% vocalic (e.g., /zebəg/).

Responses for stems requiring the voiceless variant were: 67% correct, 5% voiced and 27% vocalic. As with the sibilants, the high number of errors involving the vocalic variant suggest that Ss are using a strategy of picking a preferred all-purpose variant for generalized use on new forms, while memorizing the inflected forms of the training items. The trend towards devoicing is also supported here when we consider the high number (32%) of voiceless responses for the items with voiced stem-final consonants.

Table 10. Velars — Percentages of response types -T3 All Variants Training

Variant	Response Type [+voiced]	[-voiced]	[+vocalic]	null	V or Asp	Other
[+voiced]	46	32	22			
[-voiced]	5	67	27			1
[+vocalic]	1	2	88	6		3
Training Items						
[+voiced]	88	10	2			
[-voiced]		95	2	2		1
[+vocalic]			100			

Voicing Strategies for Velar Stops

In the analysis above it was noted that in T1 there was a significantly higher percentage of voiced responses for velar stems requiring the voiceless variant than for the sibilant stems, viz., 43% for velars as opposed to 3% for sibilants. To determine what might have motivated this difference, the raw data was re-examined and it was found that for many of the [+velar, +voiced] variants, and for some of the [+velar, -voiced] variants, Ss added a reduced vowel to the variant, resulting in forms like: /zebgə/, /pæfgə/ and /lepkə/. This appears to be a strategy for marking the inflection, and Ss who used the /gV/ form to inflect stems requiring the voiced variant also used it to inflect stems requiring the voiceless variant. Such a strategy allowed Ss to maintain the voiced quality of the inflection without violating the phonotactic constraint which prevents a [+voiced] obstruent from following a [-voiced] obstruent at the end of a word in English. Kenstowicz & Kisseberth (1977) point out that this phenomenon is a natural means for maintaining the voicing of a stem-final segment.

> [S]peakers of [a] language... (where word-final obstruents are devoiced by a morphophonemic rule) make the following "mistake" when they attempt to pronounce a foreign language admitting word-final voiced obstruents: they add schwa to the end of those words. . . that have a word-final voiced obstruent. Behaviour of this sort could be taken as evidence that a word may not end phonetically in a voiced obstruent. (p. 151)

The motivation for the addition of a vowel following the [-voiced] variant /k/ is not as clear, except that it makes the inflected form easier to articulate. It was also not used by as many Ss as the /gV/ form, although it still had the value of marking the inflection (i.e., In T1, 12 out of 15 Ss used /gV/ at least some of the time, while in T2 only 3 out of 15 Ss used /kV/).

In the tables discussed above these CV forms were marked as correct when they followed the appropriate type of stem-final consonant (e.g., /zebgə/ or /lɛpkə/ were scored as correct, while /zebkə/ or /lɛpgə/ were incorrect), but they may actually constitute an alternative error strategy in much the same way that the [+vocalic] responses were considered errors for the stems requiring the voiced and voiceless variants of the inflections. Therefore another set of percentage scores were calculated for the velar inflection, adding CV's as a possible error type (see Tables 11-13).

T1 — CV's incorrect

The use of the CV strategy was most prevalent in this training group. As a result, under the revised analysis, the correct responses for the stems requiring the voiced variant dropped from 71% to 34%, with 39% of the responses being the CV type (e.g., /zebgə/ or with a CV involving /k/, e.g., /tɛvkə/). Correct responses for training items also dropped from 71% to 46%, with 28% of the responses being CV's. The percent correct for the voiceless inflection was also lower, going from 53% to 44%. This was because 9% of the CV responses involved /k/, e.g., /lepkə/, which would have been scored correct in the prior analysis. 41% of the responses were CVs which involved /g/, e.g., /gɪfgə/. This lowered the percentage of incorrect voiced responses to only 2%, and therefore gives further evidence that the intrasyllabic constraint which prevents the voicing of an obstruent following a voiceless obstruent is quite powerful. For the stems requiring the vocalic variant, the number of correct responses remained the same, 16%, but now 21% of the errors are shown to be of the CV type, and only 1% voiced. This gives a better indication that Ss were trying to develop some strategy for dealing with the velar inflection. (It ~~was noted~~ in reviewing the data that most of the Ss who used the CV strategy used it either in combination with the correct use of the vocalic variant, i.e., inflecting some of the test items requiring the vocalic variant with CV and some with the vocalic variant, or went on to use the vocalic variant at a later trial.)

Table 11. Velar CV Strategy—Percentages of response types-T1 [+voiced] Training

Variant	Response Type [+voiced]	[-voiced]	[+vocalic]	null	CV	V or Asp	Other
[+voiced]	**34**	25	1		39		1
[-voiced]	2	**44**		2	50		2
[+vocalic]	1	5	**16**	24	21	25	8
Training Items	**46**	22		1	28		3

T2 CV's incorrect

In this training group responses were also affected by the CV strategy, although not to the same extent as T1. (In T1, twelve out of fifteen Ss used this strategy, at least some of the time, while only three out of fifteen Ss used it in T2.) The shape of the CV responses was always /kV/ in this group. As shown in Table 12, responses for the stems requiring the voiced variant were still predominantly voiceless, 85%, with the CV response accounting for 12% of the responses (e.g., /zebkə/). Correct responses for the stems requiring the voiceless variant went down to 87% correct (still quite high), with 12% of the responses being of the CV type. In the training set correct responses dropped to 90%, with 9% attributed tothe CV response. For the [+velar, +stop] stems, requiring the vocalic variant, 10% were CV responses (e.g., /bugkə/), 30% [-voiced] (e.g., /gækk/), 27% null, and 17% involved heavy aspiration (e.g., /bɛkh/), 10% were "Other," and correct responses remained at 6%. Here, too, Ss seemed to use the CV-strategy for dealing with the stem-final consonant cluster.

T3 – CV's incorrect

The use of the CV strategy also occurred in T3. In this group both /gV/ and /kV/ responses were used, but it is not clear what motivated the choice for either of these forms, since use of both types was distributed between stem types. Table 13, shows that for those stems requiring the voiced variant, the CV strategy accounted for 7% of the responses. The training item responses were not affected by the CV strategy. For the stems requiring the voiceless variant, 8% of the test responses were of the CV type, and 6% of the training item responses were CVs. The responses for the [+velar, +stop] stems, requiring the vocalic variant, were not affected by this strategy, with only 1 out of 240 responses being of this type. It is obvious that even though this strategy is employed to some degree by Ss in this group, it is not as prevalent as in the other groups, presumably because they have been taught the appropriate inflectional variants for each stem type and so have a readily available VC pattern they can employ when they get into phonotactic difficulty.

Table 12. Velar CV Strategy-Percentages of response types-T2 [-voiced] Training

Variant	Response Type [+voiced]	[-voiced]	[+vocalic]	null	CV	V or Asp	Other
[+voiced]	2	85	1		12		
[-voiced]		87	1		11		1
[+vocalic]		30	6	27	10	17	10
Training Items		90			9		1

Table 13. Velar CV Strategy-Percentages of response types T3 All Variants

Variant	Response Type [+voiced]	[-voiced]	[+vocalic]	null	CV	V or Asp	Other
[+voiced]	45	26	22		7		
[-voiced]	3	61	27		8		
[+vocalic]		3	88	6			3
Trainingl tems							
[+voiced]	88	10	2				
[-voiced]		88	2	2	6		2
[+vocalic]			100				

GENERAL DISCUSSION

The results of this experiment give further support for a phonotactic explanation for the use of the voiceless variant in English inflections. Spontaneous devoicing occurred in all of the training groups and most notably in the [+voiced] training group, where scores for the voiced variant and the voiceless variant were equal overall. This gives a further motivation for establishing a base form that is [+voiced]. This would mark the inflection, along with the obvious features of [+dental, + sibilant] for the plural and [+dental, +stop] for the past tense, the voiceless variant being predictable from a general rule of progressive stem-final devoicing, rather than a more general rule of voicing assimilation.

What is suggested by the data is that there is a strong constraint which prevents a voiced obstruent from following a voiceless obstruent within the same syllable (or at least at the end of a word). This constraint is preserved by Ss in a number of ways, depending on training. In the [+voiced] training group, both the CV strategy, and devoicing are used by Ss for inflecting words with stem-final [-voiced] consonants. The CV strategy involves adding a vowel to create a second syllable, so that voicing can be maintained (cf. Kenstowicz & Kisseberth 1977, p.151). As Harms points out, "once voicing ceases following the nucleus (vowel) of any syllable, voicing can no longer resume in the same syllable." (1973, p.11) This supports the interpretation that Ss are adding the vowel to maintain this constraint. In the All Variants training group, this constraint is resolved by using the vocalic variant, i.e., adding a vowel before the inflection to make a second syllable.

In the [-voiced] training group (T2) the difference between velar and sibilant scores for those items requiring the [+voiced] variant suggests that voicing is not as easily maintained in a consonant cluster with a stem-final stop as it is when the stem-final consonant is a fricative. In T2 there were a fairly high number of correct responses for the stems requiring the [+voiced] variant of the sibilant suffix, but this was not the case for the velar suffix, where 97% of the responses involved the erroneous [-voiced] variant. This suggests that there is a greater tendency for voicing assimilation with the sibilant suffix than with the velar stop suffix.

Harms argues that devoicing of the second segment of geminate voiced stops is quite natural:

> To maint[illegible]quires a continued differential of air pressure on bo[illegible]f the vocal cords, vibration of the vocal cords resu[illegible]e air passing between them. Over the duration of [illegible]ate voiced stops, the upper tract pressure approaches [illegible]aryngeal pressure to the point that phonation ceases (1973, p. 8).

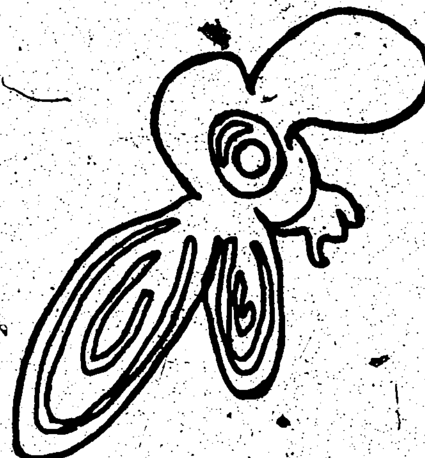
This phenomenon perhaps helps to explain why voicing assimilation scores were so low for the [+voiced] variant of the velar suffix in T2. That is, there appears to be no phonetic motivation for changing the shape of the velar (stop) suffix even if it follows a [+voiced] segment.

The limited use of the vocalic suffixes, except in cases where it was taught to Ss, suggests that these variants must be learned before they can be used productively. The question that arises here is whether the whole variant is learned, or if a vowel insertion rule is learned. What motivates this question is that there were some Ss who did use this variant spontaneously. Overall, the scores for the vowel insertion variant were slightly higher for velars than for sibilants and it occurred more in the [+voiced] training group than in the [-voiced] training group (although neither of these differences were significant). Harms (1973) provides some indication that this phenomenon may be phonetically motivated, although he does not support it as a general phonotactic constraint. In discussing a strategy to overcome the natural devoicing that occurs with geminate stops he states; "It is no doubt possible to block the devoicing phenomenon [in English] by additional gestures geared for that purpose." One of these gestures is the "release of the first stop--thus

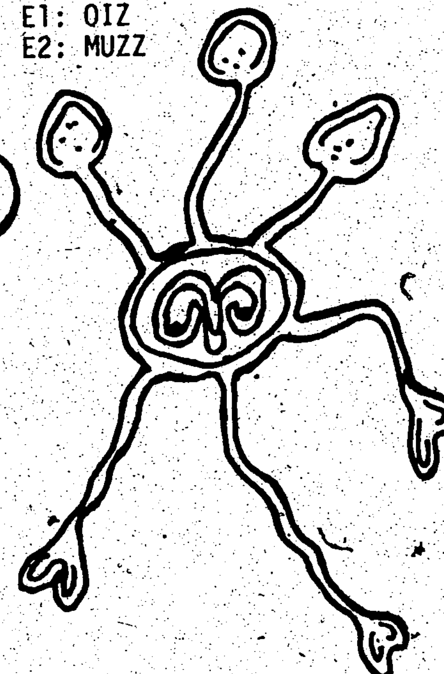
E1: YURSE
E2: BISS

E1: GUTCH
E2: GUTCH

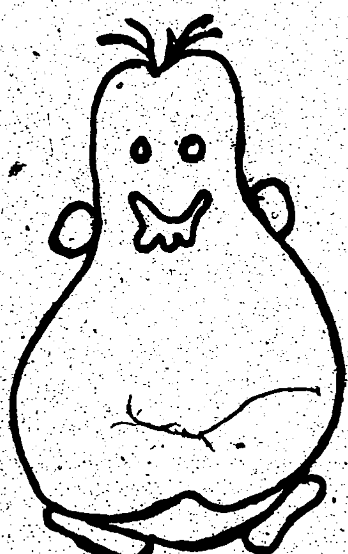
E1: SWURGE
E2: SADGE



E1: OIZ
E2: MUZZ

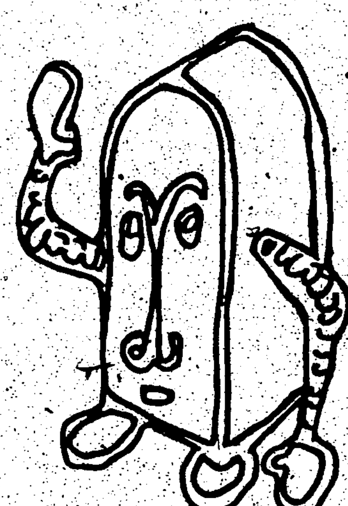


E1: GLURK
E2: BECK



E1: DROIG
E2: DEG

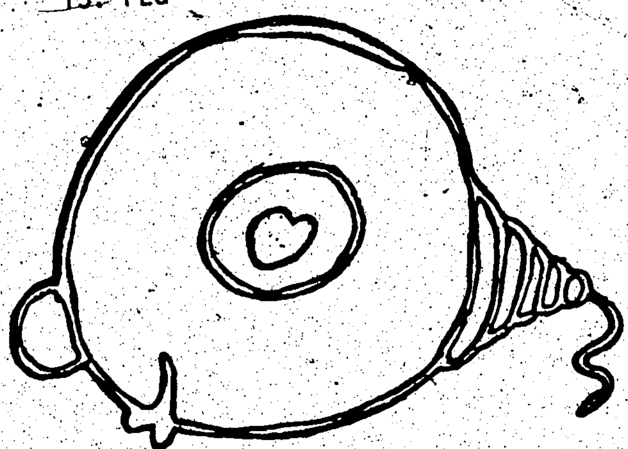
E1: BLAKE
E2: GACK



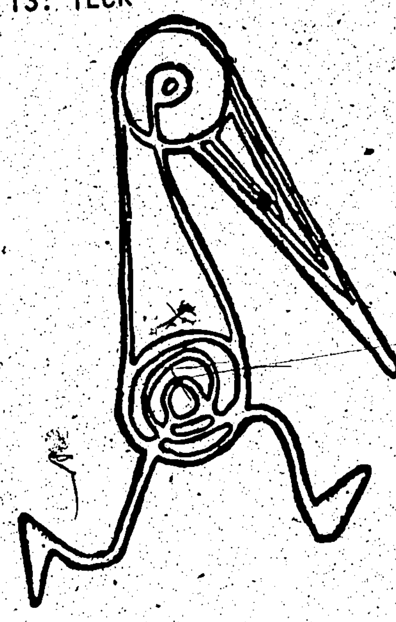
E1: OOG
E2: BOOG

Training Items — Experiment 2

T1: FEB
T2: FEP
T3: FEG



T1: TEV
T2: TEFF
T3: TECK



T1: GIB
T2: GIP
T3: GIZZ

T1: SAV
T2: SAFF
T3: SATCH