# ON TESTING THE PSYCHOLOGICAL REALITY OF PHONOLOGICAL RULES

by

HEATHER JEAN REID

B.A., Simon Fraser University, 1974

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE DIVISION OF AUDIOLOGY AND SPEECH SCIENCES

in

THE DEPARTMENT OF PEDIATRICS

We accept this thesis as conforming
to the required standard

André-Pierre Benguerel

John Gilbert

THE UNIVERSITY OF BRITISH COLUMBIA

August, 1976

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Pediatrics, Div. of Audiology and Speech Sciences

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date 27 August 1976.

ABSTRACT

This investigation was motivated by the growing dissatisfaction with the inconsistent use of empirical methodology in transformational generative phonology (TGP) and by the resulting limited value which TGP has for other scientific fields of study. The investigation is concerned with judging a particular experimental paradigm for its validity as a confirmation/disconfirmation procedure with the intention of consequently confirming or disconfirming the psychological reality of certain phonological rules.

As revealed in the review of the relevant literature, one could justify the violation of TGP's ideal speaker-listener framework--which would result from testing some of TGP's hypotheses--by using real speaker-listeners. Previous testing for the use of certain of Chomsky and Halle's (1968) phonological rules has raised doubts about the validity of claims concerning these rules' psychological reality.

The method used in the present study consisted of requiring subjects to derive and pronounce novel words (without the use of pencil and paper) from existing English stem-words and suffixes aurally presented to them. One group of subjects was exposed to existing English derivations which exemplified sound patterns accounted for by Chomsky and Halle through the rules under investigation. This group was also exposed to example derivations which showed no phonetic change. A second group of subjects were exposed only to examples showing no phonetic change.

Analysis of the results show, first (with respect to the present experiment's design), that the distribution of the number of predicted responses (i.e. the responses predicted by the rules under investigation) in each group of subjects is very similar. It is concluded that each group showed a similar ability in performing the novel derivation task and that the subjects were representative samples of the population under study. The greater occurrence of predicted sound patterns in the responses of the first group of subjects is attributed to that group's exposure to example derivations showing predicted phonetic changes. The overall production of predicted sound patterns in each group cannot be attributed to just a few subjects. A trend appears in which stem-suffix sets which were most often involved in given predicted phonetic changes were the same in both groups' responses.

Conclusions are also drawn with respect to the validity of the experimental paradigm as a valid procedure for confirming or disconfirming the phonological rules in question. First, the logical argument which uses the positive consequences of an hypothesis, known as "the fallacy of affirming the consequent," is invalid. Thus, none of the various possible strategies of sound pattern production which were considered (in order to account for subjects' responses) could be affirmed. Secondly, a valid argument of the type modus tollens can be used when the consequences of an hypothesis are negative. The valid conclusion permitted by this argument is the disconfirmation of the hypothesis. Some of the problems encountered with

this argument are discussed: (a) it is impossible to determine the exact number of times that an hypothesis is disconfirmed in a set of data in which some of the data consist of positive consequences; (b) the argument must be carefully quantified in order to permit valid conclusions to be drawn from data that is derived from real (i.e. non-idealized) conditions of the world; and (c) there exists no criterion frequency of (non-)use for the (dis)confirmation of the 'psychological reality of a phonological rule'. If it were possible to explicitly specify the extension of a rule's use, such a criterion frequency of a rule's (non-)use might be well motivated. In its absence, the psychological reality of phonological rules could not even be disconfirmed in this study.

TABLE OF CONTENTS

Page

# LIST OF TABLES

# LIST OF FIGURES

ACKNOWLEDGEMENTS

I wish to thank Prof. A.-P. Benguerel for his constant guidance, encouragement and humour during the course of this project. In addition, I am appreciative of his work with a digital computer in programming the various statistical measures and in determining the distribution of tetrachoric correlation coefficient values. I would also like to thank Prof. J.H.V. Gilbert for his examination of the thesis. I wish to thank Prof. J.B. Delack who amiably offered advice and suggested readings to aid in the investigation. I am grateful also to Amy Fleming for her help in retranscribing several of the subject interviews and to Marion Jacques for her competent typing of the thesis. And of course I extend my thanks, last but not least, to the students who participated as subjects in the study.

One further acknowledgment: I have greatly appreciated the opportunity during the last two years to study alongside students from different fields.

# CHAPTER I

## INTRODUCTION

Tranformational generative grammar has as its goal the description of a system of rules which specifies the sound-meaning correspondences in language (Chomsky and Halle, 1968, p. 3). The sound-meaning correspondences refer to the relation between an ideal phonetic form and an associated intrinsic semantic interpretation of each sentence within the set of possible sentences for a language. The rule-system specifying a language's sound-meaning relationships is assumed to be a native speaker's internalized "knowledge" of the language. This "knowledge" is also called the "competence" of the native speaker. Whereas a speaker's "performance" refers to the actual use of the language in concrete situations (Chomsky, 1965, p. 3), the knowledge or competence of the speaker is found in an idealized speaker-listener relationship. This relationship involves an idealized speaker-listener who (a) is part of a completely homogenous speech community, (b) knows the language perfectly and (c) is unaffected by grammatically irrevelant conditions such as memory limits, distraction, inattention and non-linguistic knowledge and beliefs in applying his "knowledge" of the language in actual performance.

Chomsky believes that it is only within the framework of the ideal speaker-listener relationship that it is possible for a speaker's performance to directly reflect his competence (Chomsky, 1965, p. 4). The distinction between competence and performance is believed to permit the study of uncontaminated

linguistic knowledge by abstracting away other non-linguistic factors involved in the use of language. As Chomsky recently said, "The scientist wanting to study rules [of grammar] must clean away all of the interactions..." or non-linguistic phenomena from the internalized knowledge of the speaker.

The distinction between competence and performance derives motivation from rationalist philosophy which holds that the mind (or reason) is the sole source of human knowledge (Lyons, 1970, p. 96). In this view it is hypothesized that much of the human's linguistic knowledge is innate. The innate linguistic knowledge is called universal grammar and is thought to be comprised of principles stating what kinds of rules for languages may exist. Exposure to a given language permits the child with his innate or universal knowledge to determine language specific rules or competence.

The ultimate goal of transformational generative grammar seems to be to study universal grammar which hopefully can be inferred from the study of the particular grammars of many languages. In this rationalist tradition it is thought that one may approach the age-old epistemological problem of gaining knowledge of what one knows. That is to say, by stripping away the phenomena of language performance and considering only the idealized speaker-listener's competence, it is thought possible to infer the principles of universal grammar.

Both universal grammar and the grammar for a particular language are mentalistic in nature. Botha (1971, p. 117) quotes Chomsky as saying:

> Thus, at several levels the linguist is involved in a construction of explanatory theories, and at each level there is a clear psychological interpretation for his theoretical and descriptive work. At the level of a particular grammar, he is attempting to characterize the knowledge of a language, a certain cognitive system that has been developed--unconsciously, of course--by the normal speaker-hearer. At the level of universal grammar he is trying to establish certain general properties of human intelligence. Linguistics, so characterized, is simply the subfield of psychology that deals with these aspects of mind. (Chomsky, 1964, p. 24.)

The metaphysical distinction between competence and performance dictates that the rule systems of a language will be considered from the point of view of the ideal speaker-listener relationship. This means that the empirical study of grammar will exclude verification of its theories by evidence gained from experimentation done on real speaker-listeners. Recall that this is because Chomsky is not interested in studying phenomena interacting with and "contaminating" competence in a speaker's performance.

The study of the phonological sub-component of transformational generative grammer is called transformational generative phonology (and is henceforth referred to as TGP). Many researchers in phonetics and phonology are opposed to TGP's metaphysical constraints denying them access to experimental paradigms outside the realm of the ideal speaker-listener. Two general ideas from traditional empiricism are represented in the objections to the prohibition of verification procedures external to the ideal speaker-listener frame. One idea is that all theories must be tested experimentally or else risk maldevelopment. Linell (1974, p. vii) claims that

> A theory runs the risk of degenerating into a mere systematization of data, unless one tries to find empirical interpretations for the various theoretical entities that are proposed.

Ohala likewise proposes that "...all hypotheses require experimental verification" (1973, p.1). In cases of TGP where it has been empirically concluded without experimentation that phonological rules and rule constraints are part of the real speaker-listener's knowledge, Ohala (1974, p. 19) claims that the result seems to be a "...theoretical edifice that is enormous, elaborate and very, very fragile."

A second idea representing objections to the inaccessibility of TGP to experimentation on non-ideal speakers is that a theory must play a role in explaining it as well. The route chosen by Chomsky to explain the data is one where there is first developed a rule system or structural description for a language which is in accordance with the principles of universal grammar (Greene 1972, p. 34). Universal grammar may be derived by studying different languages with the plan of identifying principles common to them all which could plausibly be considered to be part of the child's innate linguistic knowledge. Chomsky also believes that the principles of universal grammar may be discovered through the use of the "evaluation measure" which assigns a value to a grammar or to a sequence of rules. The function of the evaluation measure is to select the most highly valued of competing alternative grammars of a language on the assumption that this grammar will be the one which children would develop in learning the language. Greene (p. 30) describes Chomsky's belief that children must have some innate

linguistic ability permitting them to choose one type of grammar which is appropriate for analyzing language in general. Chomsky holds, for instance, that this one type of grammatical analysis that children are programmed to develop must be universal to all languages. "This universal grammatical theory would give an account of the grammatical forms and relations that are common to all languages..." (Greene, 1972, p. 31). Universal grammar therefore accounts for or 'explains' (in what Linell (1974, pp. 147-149) calls a weak sense of "'explanation'") the grammatical forms and relations of a language specific grammar which in turn accounts for or explains the observable phenomena. According to Linell, generative theory presupposes a sense of explanation in which observable phenomena are subsumed under a "'covering'" theoretical principle. The theoretical principle must correctly predict the observable phenomena.

Linell believes that explanations such as these are weak since correct predictions can be produced by false theory just as well as by true theory and by theories which are intended to represent reality in varying degrees. Explanations of phenomena (by means of a theoretical principle which correctly predicts the phenomena) can be strengthened if independent reasons are found which lead one to believe that the theoretical principle is true. Some linguists prefer to 'explain' the data first in this stronger sense and then proceed to develop universal principles for language. For instance, Lindblom doubts the explanatory power of Chomsky and Halle's

(1968) phonological theory wherein linguistic form has primacy over the "variables of language use and its substantive bases" (1971, p.5). Lindblom (1971, p.8) disagrees with Chomsky and Halle that one should try to develop an abstract theoretical apparatus to account for phenomena without relating the postulated mental structures and processes to the physiological mechanism. Concerned with explaining the physical phenomena of language, Lindblom (pp.7-8) suggests re-evaluating the notion of 'linguistically relevant facts' (and linguistically relevant phonetic facts) while keeping in mind the possibility of assigning phonetics a less peripheral role in linguistic inquiry. His alternative to the abstract theory of Chomsky and Halle would be a theory which uses phonetics to predict phonological phenomena by beginning with hypotheses for physically based preconditions of speech communication and development.

Ohala (1974, pp. 1-3) states that the first task in phonology is to discover the sound patterns and that the second task is to give a causal explanation of some aspect of the patterns. Language and speech, he reasons (Ohala 1974, pp. 18-19), are physically and psychologically real systems and therefore are limited systems. He believes that "The problem, then, is to constrain the range of hypotheses we entertain in the same way that the real world is constrained." Since he believes there are stricter or more numerous constraints on what can be explained by physical phonetic factors, Ohala recommends the use of phonetics (over psychology, history or other possible

sub-disciplines of phonology) as a tool for explaining aspects of phonological rules.

An implication of the above objections to TGP's immunity to experimental verification on non-ideal speakers is the limitation of the number and form of independent research strategies available to those who would study phonology within the idealized speaker-listener frame. The absence of research strategies of a non-formal nature is a possible risk for the generality of any theory which, from the point of view of traditional empiricism, must be derived from greater numbers of "'independent motivation, argumentation or reasoning'". Empirical science in general requires, in Botha's words (1971, pp. 229-230 and fn. 29), that "...postulation and use of...concepts must be motivated by showing that they are required for diverse reasons, no two of which are interdependent."

The main objectives of this project are first to carry out an experiment on real speakers to test the psychological reality for that group of speakers of some of Chomsky and Halle's (1968) general phonological rules. Another objective is to draw conclusions on the validity of the experimental paradigm itself as a confirmation of disconfirmation procedure. The experiment will consist of giving a word-derivation task--that of suffixation--to adult, native English speakers to see if the phonetic form of their responses reflects (or does not reflect) the use of certain general phonological rules.

It will be helpful to further clarify the motivation for and the goals of research of those groups of linguists who are

opposed in their views toward verification procedures in TGP. The experiment of the present study would not be included by Chomsky and Halle as a method of validation within the framework of TGP. Therefore, following the recommendation of Botha (1968, p. 104), attempts will be made (a) to explicitly state how experiments on non-ideal speaker-listeners can (or cannot, as the case may be) apply to theories on ideal speaker-listeners; (b) to give explicit criteria for confirming or disconfirming the mentalistic claims of certain phonological rules; and (c) to give an explanation (as well as the method for arriving at that explanation) of the occurrence of both positive and negative evidence for the phonological rules within the responses of single speakers and groups of speakers.

# CHAPTER II

# REVIEW OF THE LITERATURE

## 2.1 The meaning of the phrase 'psychological reality of phonological rules'

### 2.1.1 Introduction

In a recent paper entitled 'Perspectives in Phonology', Fischer-Jørgensen (1975, p.221) characterizes the present state of phonological theory as one in which most of the basic assumptions are the object of serious criticism and in which many points are being revised by professed adherents of the theory. Some of the points on which linguists disagree are involved in "the very basis of generative phonology, the claim that the description has psychological reality" (Fischer-Jørgensen, 1975, p. 219). Before describing the attempts of this and other projects to find evidence for or against the psychological reality of phonological description, comment must be made on the various interpretations of the phrase 'the psychological reality of phonological rules.'

### 2.1.2 The psychological nature of the description

It was mentioned in the introduction that both the universal grammar and the grammar for a particular language are mentalistic in nature. Botha (1971, p. 116) points this out in saying that the object of study of a particular grammar and that of the general linguistic theory are mental faculties. The mental entity which a "language-particular" (that is, language-specific) grammar represents is linguistic competence, the internalized rule-system by means of which the ideal speaker-listener relates the sound signals of an indefinite

number of sentences to their respective semantic interpretations and vice versa. Universal grammar is meant to represent the 'faculté de langage' of an ideal speaker-listener which is taken to be a language independent mental capacity enabling the speaker-listener to acquire linguistic competence.

Chomsky does not make clear which aspects--form, substance, structure, or anything else--of the 'mental faculties' are characterized by the universal and language-specific grammars. Other linguists have been more explicit concerning the relation of grammars and mental entities. "Katz regards the relationship between the structure of the linguistic theory and that of the mental mechanism as one of isomorphism" (Botha, 1971, p. 119). Kiparsky, on the other hand, believes that a grammar should correctly represent both the substance and the form of the fluent native speaker's language-specific knowledge. Botha concludes that the isomorphic relation between theory and mental faculty found in the approaches of Katz and Kiparsky means that linguistic theory for them is a conceptual analogue to mental structures.

In his metaphysical assumptions for the linguistic theory's representation of psychological reality, Linell proposes that the grammar's internal structure "<u>be isomorphic to</u> the speaker's <u>underlying psychological reality</u>" with respect to forms (such as the different forms or lexical units which are there) and structures (such as properties of the forms, relations between and generalizations over the forms, etc.) (Linell, 1974, p.1).

### 2.1.3 Rules in the description

Another fundamental concept in TGP which elicits disagreement amongst linguists is the notion of 'rule'. Linell (1974, pp. 30-32) reports that 'rule' is a vague notion in science and in ordinary language.

> In general, a 'rule' is a proposition, the formulation of which includes indications of a class of human actions, a class of persons who perform these actions and a modality for the actions involved (required, forbidden, permitted). (Linell, 1974, p. 30.)

Linell refers to Black who states that the uses of the term range from "'explicit norms'" to the "'degenerate'" sense in which 'rule' is almost synonymous with "'generalization'" and "'general assertions about matter of fact.'"

Linell mentions some performance evidence cited by Black for the 'psychological reality' of rules. The evidence consists of speakers'

> readiness to correct themselves and others, their willingness to believe that there is a rule even if they do not know what it is, and their endorsement of rules after they have been formulated by an onlooker. (Linell, 1971, p.31.)

To this Linell adds that second language learners realize that it is not enough merely to be understood, as there are rules for what counts as correct which go beyond that. Linell mentions further possible support for the existence of rules which is supplied by Searle. Searle proposes that phonological rules are not just regularities which may be described by an observer; he feels that they are also "'constitutive'" of language since "'we recognize departures as mispronunciations'" and since the rule projects to cover new cases (Linell, 1974, p.31).

A question which is fundamental to the use of 'rule' in linguistic theory is whether language behaviour is rule-guided or whether, on the other hand, it is merely rule-conforming (Linell, 1974, p.31). In the latter case the rules would only represent regularities that an observer could extract from language behaviour. Chomsky believes that linguistic behaviour occupies an intermediate position between conscious rule guidance and mere conformity to rules. This belief might be paraphrased by saying that the speaker is directed by rules but is not consciously aware of the rules.

The sense of rule which Linell prefers is the weak sense in which rules are only regularities in the observed data. He reaches this conclusion after deciding that there are probably many types of awareness covered by the verb 'know' when linguists say that a speaker 'knows' the 'rules' of his language. That is to say, Linell does not appear to be prepared to consider 'rule' as being constitutive of human language until there is clarification of the term 'to know' (a rule).

Linell qualifies his notion of the "weak" or "regulative" sense of rule by rejecting Katz's idea that rules are isomorphic to causally efficient neurophysiological processes. He states that "rules" and "causal processes" are on different levels of explanation in the same way as are "reasons" and "causes." His final notion of 'rule' is that rules

> govern linguistic behaviour in the sense that the speaker chooses (consciously or unconsciously, deliberately or habitually) to follow them. Linguistic competence is not a causally efficient force in speech production; rather it defines the linguistic conditions which must be met...(pp.31-32).

Linell's interpretation of rules governing linguistic behaviour is also ambiguous in the sense that it is not clear whether the speaker chooses either "consciously" and "deliberately" or "unconsciously" and "habitually"; or whether it would be possible for the speaker to choose "consciously" and "habitually" or "unconsciously" and "deliberately"; and whether the choosing is consistently of one nature, whatever that might be.

2.1.4. Knowledge of rules

Rules are the things which a native speaker is assumed to know, or in other words, to be aware of. In order to discuss further the sense of rule it is appropriate to discuss linguists' use of the following frequently interchanged terms: "knowledge," "to know," "awareness" and "to be aware of". Once again, as in the case of "rule," the words are assigned different meanings which are not always explicitly stated by the authors.

Zimmer (1969, p.309) refers to the phrase "to know a linguistic rule" in the sense of knowing a mathematical formula. He does not say whether knowing a formula involves the sense of 'knowing how' to use the formula, the sense of 'knowing what' the formula is or the sense of merely 'knowing that' the formula exists.

Hockett (1968) interprets Chomsky's (1965) sense of "knowledge" to be an epistemological sense, whatever that might be. In Language and Mind (1972) Chomsky refers to "knowledge" as an underlying system of beliefs. An epistemological sense of

"knowledge" is referredto by Carterette and Friedman (1974, p.7). They state that the Cartesian (that is, rationalist) doctrine of perception and belief holds that knowledge is a psychophysical judgement that is self-warranted.

Chomsky (1965) states that when a speaker "knows" the grammar of his language he has "knowledge" of that grammar which cannot explicitly be described by the speaker. He believes that this knowledge, considered as part of the mind's properties and content, may not even be accessible to the speaker's introspection (Linell, 1974, p.13).

Linell (1974, p.162, fn.23) describes another later discussion by Chomsky of his use of the word "'knowledge.'" Chomsky purports to meaning neither "'knowledge that'" nor "'knowledge how.'" Rather, he intends "'something in between'" such as "'tacit or unconscious knowledge.'" Just as in defining the sense of 'rule,' Chomsky here situates his notion in between two polar senses and then restricts the indefinite meaning by referring to the tacit nature of the object of definition (such as 'rule' or 'knowledge').

Some authors' interpretations of 'knowledge' reflect the view which sees language as "something which the speaker uses for communication with other people and with himself, and... [which] is shaped by its functions in communication" (Linell, 1974, p.27). They define and use linguistic 'knowledge' from a "behavioural perspective." For instance, Hockett (1968, p. 63) states his preference as the "know how to" sense of the word

'know.' Rather than use the ambiguous word 'knowledge' Hockett substitutes the word "habit." "The term 'habit' is little more than a paraphrase of the know how to...sense of 'to know'." Hockett finds the phrase "to have the habit of one's grammar" as less misleading, from an empirical point of view, than Chomsky's "have knowledge of" sense. This is because although a habit may provide the basis for and the predictions from a given generalization, it may not, as part of empirical evidence, endure forever and so is subject to modification. In other words, a habit is not a "Law of Nature" as Hockett feels Chomsky's sense of self-warranted "knowledge" is apt to imply.

Linell (1974, p.27) contends that linguistic knowledge should include what he calls "grammatical competence" and "general linguistic competence." Grammatical competence refers to the 'know that' sense of knowledge. The speaker 'knows' a system of obligatory conditions which utterances must meet in order to be recognized as grammatical. Linell qualifies his sense of 'knowledge that' by stating that this knowledge is not always explicit. "Sometimes the speaker knows explicit rules, but typically he is not able to formulate rules explicitly" (p. 162, fn.23). General linguistic competence, on the other hand, concerns the speaker's skill or ability "to manipulate the language actively, with precision and variation within the grammatical rules." This sense of knowledge is of course the 'know how to' sense and is indicative of a 'behavioural perspective' on language.

Wang (1968, p.707) believes that a speaker has different degrees of "awareness" (or 'knowledge') of linguistic rules. He suggests that the disparities in awareness may be dependent on several factors: the formal character of the rule, the complexity of the rule, the number of items in the lexicon for which the rule is relevant, a combination of the preceding factors or possibly some other factor. Unfortunately Wang does not say which sense of knowledge he is referring to--the 'know how to' sense, the 'know that' sense or some amalgamation of the two.

Finally, Ladefoged (1970) expresses a concern with the term "knowledge" which is founded on his belief that a speaker's "knowledge" may have different sources: a preliterate source, a spelling-influenced source and a grammatically (that is to say, linguistically) influenced source. Ladefoged complains that the vaguesness of the notion of "knowledge" prevents the grammarian from knowing which knowledge of which different sources should be included in or excluded from phonological models.

2.1.5 The reality of a theory

It has been seen that many linguists do not agree on their interpretations of "psychological" theory, "rule" and "knowledge." Some interpretations are vaguely defined. In particular, Chomsky's interpretations, which are most relevant to the aims of this study, are indefinite. He makes no statement regarding which aspect of the mental faculty the psychological linguistic descriptions should refer to. His description of

"rule" as an unconscious belief that implicitly guides the speaker is somewhat more clear. But the interpretation of "knowledge" as being "in between 'knowledge that' and 'knowledge how'" (Linell, 1974, p.162, fn.23) leaves his readers with the task of imagining the nature of this intermediary concept of knowledge.

The vague nature of the phrase under discussion is yet increased by the permissive definition of the 'reality' of a theory accepted by philosophers of science. Linell (1974, p.12) refers to a statement made by Harré that a "'realistic'" or "'representational'" concept of a theory can be supported without claiming that all of the components of the theory are real. Yet representational theory contends that "the theoretical entities and processes refer to real (though non-observable or inaccessible) entities and processes which are assumed to stand in a causal relation to the observable phenomena" (Linell, 1974, p.155, fn.1). Thus the theory is permitted to depict or represent only certain causal relations in an inaccessible reality.

Linell states that "Chomsky, Halle, Postal and other generativists conceive of generative phonology as a representational theory of psychological reality" (1974, p.12). He adds that Chomsky and Halle prefer to think of the relation between competence and performance as rather abstract and so they might describe their theory as being one of "moderate" realism. They therefore wish to take advantage of the flexibility of representational theory to represent only certain

causal relations in reality.

On the other hand, a claim is sometimes made for all of the components of generative grammar to have mentally real referents. Linell cites a statement by Katz who says that "'every aspect of the mentalistic theory involves psychological reality'" (Linell, 1974, p.13). He also mentions Whitaker's belief that "'hypotheses about the phonological, syntactic and semantic structure of language in fact represent an underlying psychological reality'" (Linell, 1974, p.13).

In conclusion, the terms used in the phrase 'the psychological reality of phonological rules' are often assigned different senses by different linguists. The whole phrase therefore takes different senses depending on the linguist who uses it or on the reader who must fill in for himself some of the indefinite points of the phrase. The terms and the statement when used by Chomsky are vague on certain points mentioned earlier. This is due in part to the statement and its terms having been derived from a body of theory which is riddled with theoretical components which are inexplicitly described. One naturally asks, first, why such an inexplicitly defined theory is maintained and second, how the theory is maintained in an empirical inductive approach to problem solving by an orthodox school of linguists.

## 2.2 Whether experiments on real speakers can apply to theories on ideal speaker-listeners

### 2.2.1 Motivation for adhering to the inexplicit theory based on the ideal speaker-listener frame

Linguists who differ in their opinions on the use of the

ideal speaker-listener frame for verification of generative grammar's theories none the less would probably share Chomsky's view that

> man is endowed with a number of special faculties which play a crucial role in our acquisition of knowledge and enable us to act as free agents, undetermined (though not necessarily unaffected) by external stimuli in the environment (Lyons, 1970, p.98).

They do not, however, agree with Chomsky's categorical rejection of Skinnerian behaviourist psychology and his consequent dismissal of experiential phenomena such as instrumental and impressionistic data gained from observing the language performance of a real speaker. Chomsky's opponents feel that it is important to consider 'experience' for numerous reasons. One reason is that the use of language, in communication with others and with self, probably has a role in shaping language. Chomsky might answer that the use of language may shape language-specific grammars but would not influence universal grammar, the discovery of which is the goal of his theory. From a synchronic point of view, language may not affect the human "faculté de langage." However, speaking from a broad diachronic view, that is to say, an evolutionary view, language use would be expected to play a role in shaping the innately endowed grammar in man. To resolve this potential area of disagreement linguists should decide whether the cognitive system of man, in which the "faculté de langage" is presumed to play some role, is to be considered on an evolutionary or a static basis.

Another reason why experiential language data is not

ignored by some of Chomsky's opponents is the belief, mentioned before, that the range of hypotheses on language must be constrained in the same real way that language systems are delimited in humans. Since language is physically observable, the data used for hypothesis formation should also deal with the physically observable phenomena of speech and language.

Another motive for considering experiential data is their reproval of Chomsky's repudiation of behaviourist theory as a means of explaining human behaviour. Chomsky rejects the behaviourist approach since it does not deal with (nor try to describe) the principles of innate knowledge. However, just because behaviourist theory is "primitive" in the foregoing sense does not require it to be rejected by a type of "how-else argument." According to Botha, the "how-else" argument of Chomsky and Halle would state that the abstract idealized listener-speaker frame is the correct theoretical approach since it is the only approach which can describe the principles of innate knowledge. Botha refutes the "how-else" argument by pointing out that

> The fact that, in a given field of inquiry, there is at a specific moment only one theory that works and that no alternative can be proposed is no more than a coincidence (1971, pp.125-126).

The different approaches to the study of TGP give importance either to abstract mental principles and entities, on the one hand, or to the sounds actually pronounced and perceived (Linell, 1974, p.150). Transformational generative theory, with its obscure and indefinite aspects, is maintained

by certain linguists because it successfully propagates abstract mental principles and entities which are rationalistic constructs. Chomsky (1965) regards his general linguistic theory as "a specific hypothesis, of an essentially rationalist case, as to the nature of mental structures and processes" (p.53).

Chomsky and Halle (1968) feel that since the empirical verification of abstract mental entities will require indirect and subtle means of verification, their abstract theory should be maintained. They say (p. 332) that "for the foreseeable future, the study of language and mental processes will have to be carried out at such a level of abstraction if it is to make significant progress."

### 2.2.2 Whether Chomsky and Halle's mentalistic and empirical theory can be maintained

#### Steinberg doubts the validity of the theory's mentalistic nature

Criticism of transformational generative grammar in general and of TGP in particular as valid mentalistic and empirical theories is growing. Steinberg (1975) claims that Chomsky has invalidated his position as a mentalist by including some "formalisms" in his intended mentalistic theory. A "mentalistic" theory is held by Steinberg to be one in which all aspects of the theory are held to be psychologically significant; that is to say, all aspects of the theory relate to ideas of the speaker. A "formalistic" theory, on the other hand, requires only certain aspects of the theory to be psychologically significant. It was stated earlier that Chomsky and Halle's

concept of the 'reality' of a theory was a 'representational' concept in which, indeed, not all components of the theory are claimed to be real.

Steinberg proposes that a certain functioning of Chomsky's theoretical grammar--the competence process of constructing derivations--does not relate to the psychological function of the speaker:

> while a theoretical grammar is said to construct derivations according to a process which begins with the symbol S and proceeds to the Base and from there to the Transformational, Semantic, and Phonological (sub)components, a grammar internalized by a speaker is not hypothesized to have this process. Derivations are said to be constructed by speakers in some other way. Chomsky indicates that a speaker would require a set of heuristics or [else] Use rules, in addition to the grammar, in order to produce derivations (1975, p.246).

The fact that psychological significance is not given to the process by which the grammar provides derivations results in inconsistency for a theory which is supposed to be concerned with mentalism. This formalism in Chomsky's mentalistic theory also results in arbitrariness as no justification is given for treating the process underlying the grammar as a psychologically insignificant component. Steinberg believes that even if justification were given for incorporating formalisms into a mentalistic theory, (for instance Chomsky might claim that his theory was meant to be a representational one of reality), the positing of those formalisms could still result in the psychological invalidity of several components of the theory. For instance, given that the underlying process of competence is denied psychological significance, the nature of the process

affects the character and composition of structures in a derivation. The resulting derivations would be psychologically invalid. In addition, he states that it would be absurd to think that a psychologically invalid process could provide a determining basis for the discovery of psychologically valid rules. Steinberg also believes that the organization of the grammar--that is, the input and output relations among the syntactic, semantic and phonological components--is rendered psychologically invalid. This follows first from the fact that the grammatical organization is in accord with the mentally invalid generative process which underlies the grammar; second, from the fact that the grammar's derivations are a reflection of the generative process and grammatical organization. Finally, Steinberg proposes that linguistic generativity cannot be psychologically meaningful in linguistics when the derivations which a grammatical theory generates are psychologically invalid.

Steinberg offers two difficult remedies for "correcting the inherently contradictory theory which Chomsky has developed" (p.250). One solution is for Chomsky to "renunciate mentalism" and declare that the grammar and its output derivations are non-psychological formalisms. Another solution involves a radical revision of the theory, which Steinberg only briefly outlines, so that it conforms with how speakers might reasonably be claimed to construct derivations.

Botha questions TGP's status as a mentalistic and as an empirical theory

TGP's status both as a mentalistic and as an empirical

theory has been censured by Botha in numerous works (1968, 1970 and 1971). In general, his criticism is based on a line of reasoning which points out that the patterns of argumentation used to justify linguistic hypotheses are invalid.

In The Methodological Status of Grammatical Argumentation (1970) Botha derives his criteria for a valid argumentation pattern from Toulmin. (Although Toulmin's arguments are mainly concerned with jurisprudence, Botha claims that his conclusions on the logical structure of an argument are valid for argumentation in empirical science (1970, p.19).)

The components of Toulmin's well-formed argument are:

(1) a claim or conclusion,
(2) data,
(3) a warrant,
(4) a qualifier,
(5) conditions of rebuttal, and
(6) a backing.

The claim or conclusion is an hypothesis of which the correctness or incorrectness is affirmed or denied. The data are the facts provided to support or contradict the claim. The warrant is a "rule" or "inference-licence" given to show why the data can be taken as support or contradiction of the claim. Botha describes the warrant as a "bridge-like hypothetical statement that has the form: if X, then Y" (p.20). Different warrants can be seen to confer different degrees of force on the claim which they attempt to authorize. The qualifier is therfore built into the argument to specify explicitly the degree of force which the data confer on the claim through the use of the warrant. In the case where a particular qualifier is used, the conditions of rebuttal state the reasons for using

it. That is, the conditions of rebuttal specify the circumstances under which the warrant has no authority. Finally, to know why a warrant is generally valid, the grounds from which it derives its validity is specified explicitly, as a "categorical assertion of fact," by the backing. See Figure 1 for Botha's exemplification of the components of this argument.

A well-formed argument requires that all of the components of the argument be explicitly presented and related in a proper way. To ensure that the conclusion of an argument is correct, the argument must both be well-formed and have correct statements presenting the data, backing and conditions of rebuttal.

Botha finds that the invalidity of many linguistic arguments stems in part from the lack of proof of the correctness of the backings of grammatical warrants. He concludes that grammatical argumentation fails to qualify as a confirmation procedure also because "the correctness of many claims about hypothetical data cannot be established beyond doubt" (1970, p.62).

To discuss whether TGP is valid as a mentalistic theory Botha investigates in what ways the mentalistic hypotheses of the theory can be verified. First he outlines the basic types of norms used in empirical science for the verification of theories. Then he enumerates the means by which mentalist linguists attempt to verify their theories. Finally, he discusses whether or not the verification procedures are acceptable from a logical and epistemological point of view.

The norms of truth used in validating empirical theories

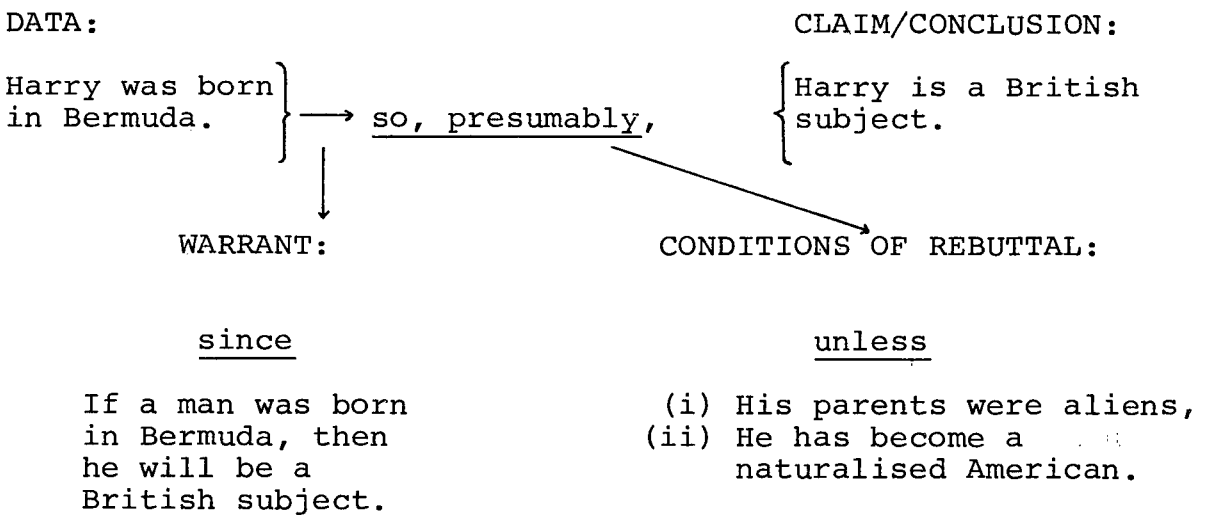



Figure 1. An example of an argument. Figure from Botha, 1970.

are norms of correspondence, norms of coherence and pragmatic norms. Correspondence norms are based on a concept of truth in which "the truth of a statement depends on whether the state of affairs referred to by this statement corresponds to what it asserts about the state of affairs" (Botha, 1971, p.122). Therefore a statement will be true if it "fits the facts." Correspondence norms are thought to be "more fundamental" than pragmatic norms or coherence norms since science is generally governed by a reality principle. The truth of correspondence norms is called empirical truth since the correspondence between what a theory asserts and the state of affairs is determined by experience.

Norms of coherence are based on the epistemological thesis that the truth of a theory depends on whether it fits into a body of theory or a system of previously-established scientific statements. With these norms of truth, a statement is considered to be true if it "coheres with others, and if these are known to be true" (Botha, 1971, p.123). Botha mentions two problems arising with the use of coherence norms. One is that an explicit criterion is needed by which to determine whether two theories "cohere." Another problem is that of determining which of two or more opposing bodies of internally consistent theory is the "valid" theoretical system.

Finally, pragmatic norms specify a concept of truth wherein the truth of a theory is reflected by the way in which the theory guides and stimulates the continuation of scientific inquiry. Botha states, however, that although scientists do

appreciate theories for their "fruitfulness," they do not equate fruitfulness with truth. Consequently pragmatic norms are not advocated by many philosophers of science as conceptions of truth.

Botha goes on to describe the modes of argument used by mentalist linguists to verify their theories. (His detailed discussion and schematization of these arguments may be found in Chapter 4 of Methodological Status of Tranformational Generative Phonology (1971).) One mode of argument is the "how-else argument" in which linguists argue that their theory is valid because there is presently no other theory which accounts for the data. Botha rejects this argument on a number of counts: it contains no backing; the norm of truth used is the "primitive" pragmatist norm; and there is a possibility for other appropriate theories to be later developed by other scholars.

A common mode of verification in linguistic study is testing the predictions of grammars. Norms of correspondence are used in this verification procedure. The warrant for the argument in this verification process can be stated in two ways. In one case, the warrant might state that a theory is correct if the data it predicts are correct. Botha, however, states that a theory which correctly predicts some events does not necessarily correctly describe or represent the structure of the mechanism whose operation results in the predicted events. He cites examples from the history of science and from transformational generative grammar to support his rebuttal

(1970, p.130). These examples are instances of competing hypotheses which all predicted correct data. He also cites instances where a given hypothesis has been preferred over rival hypotheses even though the former's predictions were incorrect.

On the other hand, the warrant could be interpreted as the statement that "unobservable events cause measureable effects." This is a statement of "causal efficacy" used in ontological arguments--that is, arguments which attempt to prove the existence of some unobservable entity. Psycholinguistic experiments often use ontological arguments. The warrant, as a statement of causal efficacy, is logically invalid and is known as "the fallacy of affirming the consequent"(Botha, 1971, p.134). The warrant is generally of the form "if an empirical generalization 'g' follows from hypothesis 'h', and if 'g' is true, then 'h' is true." It is not valid, from a logical point of view, to simply assume that 'h' implies 'g' and 'g' implies 'h' without careful quantification of the logical statement.

Other types of arguments which are based on norms of coherence are carefully outlined by Botha. These arguments obtain their evidence from numerous sources such as history, psychology, metatheory, neurophysiology, and other fields of study. Each argument discussed is refuted as a valid verification technique for mentalistic hypotheses. In general it might be said that the criticisms of these arguments stem from the arguments' failure to meet conditions of "well-formedness" mentioned earlier. (The reader is referred once again to

Chapter 4 (Botha, 1971) for further details.)

After demonstrating the conditions disqualifying grammatical arguments as valid verification procedures, Botha amplifies the serious nature of the situation in two ways. First, he proposes that even if the arguments using coherence and correspondence norms were valid, the supporting evidence of psycholinguistic testing is "scant" and often "mutually contradictory" (Botha, 1971, p.167). Secondly, Botha's discussions of grammatical argumentation assume that the theory's formal devices "categorically" characterize linguistic competence. A categorical characterization of competence rejects the notion that the speaker's acqusition of language is instantaneous. Rather, the categorical statement of linguistic competence assumes that language acquisition takes place over a period of time and requires that mentalistic linguistic theories take into account the order in which primary linguistic data are used by the child.

Chomsky and Halle specify that the mentalistic assertions of their hypotheses are not categorical statements of linguistic competence but are instead "conditional assertions" (1968, p.331) of competence. The conditional assertion of competence relies on the a priori assumption that a speaker acquires language instantaneously. Chomsky and Halle state that it would be more difficult to study the categorical statement of a speaker's competence and the mentalistic nature of hypotheses resulting from that statement since the order in which primary linguistic data are used by the child would have to be taken

into account.

Botha then argues that if there are thus far no valid procedures of verification for those hypotheses (which he discussed) involving the more complex categorical statement of linguistic mentalism, there could not presently exist any procedure to test Chomsky and Halle's conditional statement of mental reality for TGP.

Botha claims to reject the thesis of mentalism in transformational generative grammar. He does claim, however, that transformational generative grammar as a mentalistic linguistic theory is untestable at present. He suggests that the theory's mentalistic claims might become testable if linguists would identify and correct the weak methodology of argumentation. In other words, the arguments must meet the conditions of well-formedness and the statements of the arguments' warrants and backings must be correct.

Consequently Botha goes on to laboriously consider the empirical status of TGP as a non-mentalistic theory. Chomsky and Halle (1968) state explicitly that they wish their theory of TGP to be an empirical one. They say that even with the assumption that the grammar shall consider deviations from grammaticalness as "correct," the problem to be solved is one of an empirical nature: a set of formal devices and an evaluation measure must be selected which

> jointly meet the empirical condition that the highest valued grammar of the appropriate form is, in fact, the one selected by the child on the basis of primary linguistic data (Chomsky and Halle, 1968, p.331).

They continue to stress the empirical nature of the theory in saying that such a theory which

> specifies formal devices and an evaluation procedure can be proven false...by confronting it with empirical evidence relating to the grammar that actually underlies the speaker's performance. There is such a grammar, and it is an empirical problem to discover it...However difficult it may be to find relevant evidence for or against a proposed theory, there can be no doubt whatsoever about the empirical nature of the problem.

Before considering the empirical status of TGP as a non-mentalistic theory Botha describes the aim of a theory which makes no claims for mentalism. Such a theory aims to construct formal devices which can correctly "account for" the data which constitute linguistic reality. The devices do not need to correctly describe in some way the mental mechanism underlying linguistic data.

Botha approaches the issue of TGP's merits as an empirical theory by outlining the factors which determine and bear an influence on its empirical status. One factor which determines that a theory is empirical is the fact that it is testable. Botha's requirements for a testable theory are borrowed from Hempel (Botha, 1971, p.174). Hempel believes that a theory has "testability" if explicit test implications can be derived from that theory and if the experimenter can state what would constitute favourable and unfavourable evidence for those test implications. Testability would be confirmed by confronting test implications with experimental evidence.

Another variable which influences the empirical status of a theory are certain factors which can adversely affect the theory's testability. Such a factor would arise in a theory

which is so inexplicitly stated that no test implications can be derived from it. Another factor is the possibility that it could be impossible to specify just what evidence--both relevant and reliable--would be able to confirm or disconfirm a theory. In addition, it is possible for a theory to be constructed so as to enable it to be upheld even in the face of evidence which contradicts it. In this situation ad hoc hypotheses are incorporated into the theory for the single purpose of protecting the theory against contradictory evidence.

In Chapter 5 Botha (1971) investigates whether the evidence available for some of TGP's formal devices is relevant and reliable. Also, he discusses the inclusion of ad hoc hypotheses in certain principles of the theory. His main conclusions on the availability of evidence are, first, that the evidence is indeed not available for the hypothesized form and function of the evaluation measure. (Briefly, the evaluation measure is a simplicity metric used for evaluating rules. "The 'value' of a sequence of rules is the reciprocal of the number of symbols in its minimal representation " (Chomsky and Halle, 1968, p.335).) TGP often depends on the evaluation measure to play a key role in linguistic verification procedures as the warrant of the grammatical argument.

Secondly, he finds that Chomsky and Halle (1968) do not specify what evidence would confirm phonetic representations if their composite phonetic features were viewed as being either mental instructions to the articulatory apparatus or as representing aspects of vocal tract behaviour. Further, the

empirical status of Chomsky and Halle's phonetic representations is found by Botha to be questionable when the representations are considered to be a description of perceptual reality. This follows from what he considers to be irrelevant physical evidence, as well as irrelevant and unreliable perceptual evidence, which is offered by Chomsky and Halle to confirm or disconfirm the representations as referring to perceptual reality. Botha also points out several difficulties inherent in developing an indirect mode of validating phonetic representations on the basis of the notion "systematic theory." This validation procedure would involve a limited number of data about the placement of primary stress being offered as evidence for the correctness of complex stress contours. The pattern of argumentation based on the notion of "systematic theory" is not proposed and defended by Chomsky and Halle; its existence is only suggested by them and thereafter explored by Botha.

Botha discusses four theories in TGP from the point of view that they contain ad hoc hypotheses which block adverse evidence from applying to them. (These theories are the principle of the transformational cycle, linguistic universals, the theory of exceptions, and devices accounting for stress phenomena. For the details of his discussion the reader is referred to Botha, 1971, Chapter 5.) The ad hoc hypotheses or "blocking devices" involved in the four cases he describes serve to block the process of testing the theories which incorporate them by neutralizing contradictory evidence. Therefore the theories' testability is denied and Botha disqualifies

them as empirical theories.

Chomsky and Halle feel that it is quite usual in empirical science to employ blocking devices to protect hypotheses of great generality. Botha refutes their attempt to justify the use of blocks by saying that the mere use of a principle in empirical science does not imply that it is logically or epistemologically sound.

It is important to mention that Botha feels that his conclusions, if correct, do not "show that TGP is a completely unempirical approach to the study of phonological properties of natural language" (1971, p.247). He states clearly that his discussion dealt with only a small number of theoretical devices used in phonological theory. "If some of the hypotheses constituting a theory have a questionable empirical status, it does not automatically follow that the remainder are also suspect with respect to their testability" (p. 247).

### 2.2.3 How experiments on real speakers cannot apply to theories about ideal speaker-listeners

Given that doubt has been cast on the validity of TGP as a mentalistic theory and on the empirical status of some components of TGP, certain ideas behind the requests for experimentation outside the ideal speaker-listener frame are reinforced. Those ideas are, first, that all theories should be tested, since a theory does not have empirical status unless it is testable and since an untested theory risks maldevelopment. The second idea is that all theories should be tested and verified in as many independent ways as possible in order to increase the theories' generality.

The individual studying TGP is therefore greatly tempted to expose various hypotheses to experimentation on the true source of all linguistic data, the real speaker. Yet the bothersome obstacle to the verification of transformational generative grammar's theories through language performance data remains: in principle, a theory system which is based on an idealized speaker-listener is immune to evidence gained from the performance of a real speaker. A real speaker, recall, does not meet the conditions of coming from a "'homogeneous'" speech community; it is quite possible that he does not know the language "'perfectly'"; it is possible that he will not remain uninfluenced by memory limits, distractions, inattention and non-linguistic knowledge and beliefs. Short of hoping that Chomsky and his followers will have a sudden turn-about and either renounce the mentalistic nature of transformational generative grammar or discard the ideal speaker-listener framework, one is inclined to sit back and ponder Botha's systematic dismantling of the theory system and hope for a miracle.

### 2.2.4 Possible justification of experimentation external to the ideal speaker-listener frame

However, it may be possible to reach some compromising solution to the situation which, given the truth of Botha's conclusions on the non-empirical status of some of TGP's hypotheses, will continue to propagate untestable theories. Would the possibility of rendering testable some of TGP's hypotheses by experimentation on a real speaker be worth the risk of violating the theory's stipulation for an idealized speaker framework? The questions to be answered are what can

be lost and what can be gained in such a step towards transgressing TGP's metaphysical assumptions and which is of greater value (or for that matter, which is of less damage) to the theory.

That which would be immediately lost by submitting theories of TGP to verification procedures external to the idealized framework is consistency in dealing with the metatheoretical assumptions of the theory. But this consistency may have already dissipated. Steinberg (1975) proposed that Chomsky, without justification, maintained formalisms within his supposedly mentalistic framework. Botha meticulously demonstrated the varying abilities of hypotheses to be testable and hence empirical in a theory system which purports to be concerned with the "doubtless empirical problem" of writing grammars. In addition, the rules and principles of TGP hypothesized by linguists, the hypothetical grammatical models with apparent descriptive and explanatory adequacy and eventually the metaphysical assumptions of TGP would be submitted to the scrutiny of various independent verification procedures and so of course could be modified or abandoned altogether.

That which could be gained, on the other hand, would be an increase in the explicitness of descriptions of all aspects of the theory. This would be achieved by means of independent sources of evidence reflecting back on the theories and forcing their explicit modification. The explicitness would also be brought about by means of the need for continuing experimentation to develop on the basis of test implications which could

only be derived from theories that were more fully specified. The gain would likely be a self-propagating cycle of the development of more explicit theory. Testing the hypotheses and modifying them specifically would produce the opportunity to derive further test implications which might, in turn, result in further modification and specification of the theory.

Ohala points out Chomsky's admission that it will be

> necessary to discover conditions on theory constructions, coming presumably from experimental psychology or from neurology, which will resolve the alternatives [that is, alternative hypotheses] that can be arrived at by the kind of speculative theory constructions linguists can do on the basis of the data available to them (J. Ohala, 1970, p.10).

His and Halle's most consistent position, however, is that

> for the foreseeable future, the study of language and mental processes will have to be carried out at such a level of abstraction [the abstraction of an ideal speaker-listener who instantaneously acquires language] if it is to make significant progress (Chomsky and Halle, 1968, p.332).

Chomsky and Halle, however, fail to specify the point that TGP must reach before speculative theory constructions can be submitted to independent means of verification. Here is another reason to begin testing theory constructions, if only to decrease the risk that the mentalistic theory will develop erroneously.

2.3 Independent verification procedures for phonological rules

Ways of approaching the current challenge of verifying the psychological reality of generative phonology are reviewed by Fischer-Jørgensen (1975). To her, the notion of the reality of mentalistic claims presupposes that the speaker somehow "has command of [some equivalent of] the units and rules set up by

the linguist" (p.224). She calls this reality "functional psychological reality" and surveys the types of linguistic data from which inferences can be drawn for the problems of the theory's mental realism. The sources of data which she cites are from various types of linguistic change such as diachronic linguistic data, accomodation of loanwords and acquisition of first and later languages; speech errors, such as slips of the tongue and aphasic disturbances; metrics and rhyme, phonetic puns and linguistic games; orthography; and direct experimentation. Following are some cases where inferences were drawn from various types of linguistic behaviour on the psychological reality of certain aspects of TGP.

### 2.3.1 Direct experiments as a means of verification of phonological rules

This means of verifying the psychological reality of phonological rules defies the ideal speaker-listener frame of TGP since the experiments are done on non-ideal speaker-listeners with the performance data being considered as directly relevant to the speakers' competence.

An experiment involving the use of nonsense words is reported by Zimmer (1969). Zimmer states that speakers may not have an awareness of the principle underlying some phonological regularity seen in their language if the phonological process accounting for the regularity is not synchronically productive. Further, he suggests that it may sometimes be impossible to determine the relative productivity of aspects of linguistic structure.

He sets out to test subjects' awareness of three Turkish

morpheme structure conditions (MSC's) involved in Turkish vowel harmony. The palatality MSC holds that vowels in bases of words agree in palatality (for instance, /iti/ dog and /eti/ meat ). The labial consonant MSC states that after /a/, a high vowel agrees in labiality with an intervocalic labial consonant (for example, /karpuz/ watermelon and /avuč/ palm of hand). Several (62) morphemes were found conforming to this MSC and many of them were frequently used in everyday speech. In addition, 13 uncommon morphemes were found to be counter-examples of the labial consonant MSC. These morphemes had the form /...a...$C_1$...i.../ where $C_1$ is a labial consonant. They can be accounted for by the general labiality MSC. This rule states that a non-first vowel in the base agrees with the labiality of the first vowel in the base if both are high vowels--except if the preceding vowel is /a/ and a labial consonant intervenes between the two vowels.

The first test consisted of pairs of single morpheme Turkish nonsense words. Each pair was designed to test one of the three MSC's mentioned above. The list of words was mailed as a questionnaire to Turkish students at Berkeley with the instructions that they were to pronounce the word pairs aloud and check which word sounded like a possible Turkish word. They were permitted to check both words if both sounded like possible Turkish words. Twenty-three students replied and their responses were considered.

Five of the 16 word pairs were relevant to the labial consonant MSC. Only 23.5% of the responses for these 5 pairs were

in accord with this rule's prediction, while 50.4% of the responses were not those which would have been predicted by the rule and 26.1% of the responses showed no preference for either word in the pairs. Two pairs of words were not relevant to any of the MSC's under consideration. Nine pairs were relevant either for the general labiality MSC or the palatality MSC. Zimmer considered the results of the responses to word pairs which were relevant to these two MSC's together, he said, since their responses were what one might expect on the basis of regularities in the harmonic portion of the lexicon. In these 9 word pairs, 88.4% of the responses were those predicted by the MSC's, 4.4% of the responses were in disagreement with the rules' predictions and 7.2% were indicative of no preference for either word in the pairs.

The second test consisted of 18 pairs of nonsense words recorded twice by a native Turkish speaker. The time allowed subjects for producing responses was limited as the presentation of word pairs occurred with intervals of less than four seconds between the pairs. A second presentation of the list was given. It involved the same word pairs but the order of the words in each pair was reversed. There was a fifteen minute interval between the two test presentations during which the examiner and the subject talked of matters unrelated to the test.

Sixteen native speakers of Turkish were asked to indicate a unique preference by checking an "(a)" (for the first word) or "(b)" (for the second word) on a sheet of paper after listening to each pair of words.

In order to test subjects' awareness of an MSC not related

to vowel harmony, Zimmer included two pairs of words where one member had a consonant cluster occurring neither in Turkish nor in loan words accepted by Turkish. The answers were in favour of the member with the reduced consonant cluster; only 5 responses favoured the words with unreduced clusters.

The results of the second test showed 84% of responses relevant to the palatality MSC to be in accord with the rule; 65.6% of responses agreed with the general labial consonant rule; and 48.1% of responses which were relevant to the medial labial consonant MSC were in accord with that rule.

Zimmer draws two conclusions. First, despite the great number of non-harmonic loan words in Turkish, subjects still showed a "considerable preference" for nonsense words conforming to the palatality and general labiality MSC's. Zimmer explains this preference by proposing that subjects need these phonological rules to use their language productively. Second, analysis of individual subject's responses indicated a negative correlation between the preference of words accounted for by the labial could be divided into two groups. In each group the subjects tended to ignore one MSC and more often evaluate the word forms on the basis of the other MSC. Furthermore, the two MSC's conflicted with each other. Zimmer believes that subjects preferred to use only one of two overgeneralized--rather than both of two complementary--MSC's concerning the aspect of labiality of vowels in vowel harmony. The degree of interference between conflicting MSC's was measured in terms of the percentage of responses by a given subject where the first

rather than the second MSC was applied. For instance, the interference of the generalized labiality MSC with the labial consonant MSC ranged from 40% to 100%. To confirm the hypothesis of conflicting MSC's Zimmer suggests testing a large number of subjects repeatedly over an extended period of time to see if the amount of interference of an MSC for a given subject appears stable.

Since the results showed that one group of subjects more frequently used an overgeneralized labial consonant MSC while another group preferred to use an overgeneralized labiality MSC (that is, the general labiality MSC with the "except if" clause dropped), Zimmer concludes that his subjects had different phonological models. In addition, their phonological models were incorrect as the overgeneralized use of either rule cannot account for all of the word forms in Turkish. Zimmer attributes the erroneous overgeneralizations to the possibility that unproductive phonological rules are involved in the MSC's under discussion. In this case he feels that the subjects probably learn lexical items in their fully specified form. The MSC's in question do not fill in values for incompletely specified segments.

Zimmer finally states that

> A precise statement of the relevant MSC's, which we might want to incorporate in a phonology of Turkish, apparently does not provide a realistic analog of the knowledge internalized by the native speaker, and of course fails completely to reflect the differences which seem to exist between speakers (p.320).

An experiment done by M. Ohala (1974) involved unusual

derivations of existing words. Ohala was not convinced that purely structural evidence was sufficient proof that grammars are psychologically real constructs. She attempts to provide experimental evidence for the psychological reality of a certain segment in underlying forms which does not always appear in surface structures.

Ohala states that for a group of forms in Hindi it is reasonable to posit an underlying form containing a schwa. Before a medial consonant cluster the morpheme could drop its schwa by a ə-deletion rule of the following form: ə → ø/VC__CV. This system could account for words such as [pəkəṛ]/[pəkṛa] (catch/caught), [pʰisəl]/[pʰislə] (slip/slipped), etc. Certain suffixes such as -iya (which forms adjectives from nouns) block the application of the -deletion rule and produce such forms as [kə:səṛ]/ [kə:səṛiya].

Ohala explains that a possible alternative way to account for the forms mentioned above is to use a[ə]-epenthesis rule. Then the underlying forms would be /pəkr/, /kəs:r/, etcetera and the [ə] would be inserted in the following way: $\emptyset \rightarrow ə / VC\_\_\_C\left\{\begin{matrix}\#\\ +C\\ +iya\end{matrix}\right\}$. Unfortunately this rule would incorrectly insert schwas in all Hindi words with final consonant clusters. It would be too extravagant to mark the Hindi words showing final culsters as not being receptive to the [ə]-epenthesis rule.

Ohala proposes that evidence for the presence or absence of schwa in underlying forms could be tested by observing the derivations of nonsense words. For instance, the nonsense

words [məsək] and [məsk] could be derived by using the morphemes -õ and -iya. If the [ə]-epenthesis rule was used, the addition of -iya to both stems would produce [məsəkiya:] for the underlying forms of both stems would be /məsk/. If speakers used a ə-deletion rule, the underlying forms would be /məsək/ and /məsk/ and the responses could be expected to be [məsəkiya:] (since the ə-deletion rule is blocked by -iya) and [məskiya:] respectively.

After finding that one adult native speaker of Hindi gave [məskõ:] as the response when the oblique plural morpheme -õ was added to both [məsək] and [məsk] and that the addition of -iya yielded [məsəkiya:] and [məskiya:], Ohala felt that the ə-deletion rule was likely used. She sets out to test whether a schwa should be posited in underlying forms of certain morphemes which have clusters at the phonetic level but which have no alternating word forms containing [ə].

The forms used were nouns in common use in everyday speech: for example, [$g^{h}$õ:sla:] (nest), [kẽ:kra:] (crab). The task was to add -iya to the nouns, producing non-existent but semantically reasonable forms. In a preliminary test seven college-educated native speakers of Standard Hindi added -iya to five nouns. They were verbally given two warm-up examples; of the two, only one stem and its derivation contained schwas. Five of the seven subjects consistently gave responses containing schwas indicating that the underlying schwa was psychologically real. The other two gave forms with consonant clusters only.

A second test was performed in which 27 native speakers of

Standard Hindi added -iya to 30 common Hindi nouns. (Of the 30 words 9 were of no interest to the test but were inserted to prevent subjects from answering automatically in a certain pattern.) Subjects were told that adding -iya might result in words that don't exist in Hindi but that nevertheless make some sense. "Occasionally" Ohala gave the meaning of a new derivation to prevent the subject from adding -iya automatically to nonsense words.

Assuming that subjects would forget their responses after an intervening activity, Ohala assigned the subjects a brief task after completing the test. Then, claiming to have misrecorded some items, she asked the subjects once again to derive certain of the words. Two new words were added which have existing forms in Hindi containing -iya. One of the existing Hindi derivations with -iya did and one did not contain a schwa. The existing derivations were given at the beginning of the second list to see whether the speakers would be influenced by their form. The responses were reocrded throughout both test presentations by noting a ø if subjects did not add or retain a schwa.

The results showed that most subjects preferred to keep the bases unchanged. To investigate whether subjects were influenced by the two cue words, the change in distribution of ø and schwa responses was considered for the words following the cue words. The data showed a small change in the incidence of the two types of responses; for the majority of the words, the change was in the direction of conformity to the cue words'

patterns. Chi-square tests of the significance of the difference between the observed distribution of responses and the expected distribution if subjects had responded identically both times showed that subjects' responses were not significantly affected by cue words.

With the intention of making some interesting points, Ohala tentatively assumes the sample of the responses in the test to be representative of derivations which might be made in everyday speech. Three subjects' responses predominantly contained a schwa. She groups these subjects into a "ə-dialect". Another three subjects' responses were predominatly lacking schwas and were grouped into a "ø-dialect". The remaining 21 subjects showed considerable variation in their responses.

Ohala observes that for those subjects in the ə-dialect, the assumption that -iya blocks the ə-deletion rule is wrong. She hypothesizes that the environment of the ə-deletion rule is still changing from the context mentioned earlier to a more general environment. The variation in subjects' responses might indicate, she feels, that the change has not been completed.

For the 21 subjects giving mixed responses, -iya sometimes does and sometimes does not block the ə-deletion rule. There is no clear-cut evidence for the absence or presence of an underlying schwa in the test's words. Ohala suggests that the words do perhaps have an underlying schwa since the sporadic appearance of the schwa in surface forms could be caused by the fact that the -iya suffix is being removed as a block to

the ə-deletion rule.

Like Zimmer, Ohala observes that "speakers who produce forms which are identical phonetically may still have quite different grammars" (1974, p.223).

J. Ohala (1973) reports testing to see if speakers use certain sound patterns productively. Proof of the productivity of sound patterns would be, according to Ohala, the demonstrated ability of speakers to extend the pattern to new words or to existing words in a new situation. If the sound pattern was shown to be productive, Ohala feels that the next step in the study would be to determine how it was productive. He considers two possible means of productivity: independent phonological rules such as are hypothesized by Chomsky and Halle (1968) and analogical phonological rules which require explicit reference to existing words which manifest a given sound pattern. For example, if a speaker is asked to derive a new word using the stem word slave and the suffix -ity he might produce [slævIti]. Ohala believes that the speaker may have produced [slævIti] through an analogy with the word pair sane/sanity. On the other hand the speaker may have used an independent phonological rule such as V → [-tense]/__$C\breve{V}CV$[1] which Chomsky and Halle (1968) propose.

Ohala tests the productivity of six phonological rules proposed by Chomsky and Halle (1968) by giving subjects a word derivation task involving suffixation. There were 63 stem-words.

---

[1] $\breve{V}$ was used by Ohala, and shall be used in the discussion of the present study, to represent $\begin{matrix} V \\ [-\text{stress}] \end{matrix}$.

Forty-one of these were test words, that is, words which Ohala thought would likely undergo a phonetic change in the stem when subjects combined them with a suffix. Twenty-two were "filler" words, or words which were not expected by Ohala to undergo phonetic change when a suffix was added to them. The filler words were included in the list to prevent subjects from thinking that a phonetic change in the stem of the derived word should be necessary. Subects' responses to the filler words were not included in the analysis.

A total of ten suffixes were used. Some of the suffixes (such as -ment, -hood and -dom) were not expected to trigger a phonetic change in the stem of the new word. Altogether ten stem-words were to be combined with these three filler suffixes and the responses to their combinations were also not included in the analysis. Therefore a total of 32 out of 62 responses in the suffixation task in Ohala's experiment were not counted in his analysis.

After the introduction of a new suffix which was to be added to a following set of stem-words, two examples of the use of that suffix in English were given to the subject. Most of the examples of suffixation exhibited no phonetic change in the stem (for example, odd/oddity); two example derivations demonstrated a change in vowel quality in the stem wherein tense vowels became lax (for instance, detain/detention, explain/explanatory); some examples showed two changes wherein a previously unstressed schwa was realized as another lax vowel which carried primary stress (solid/solidify).

The subjects were 26 Berkeley undergraduate or graduate majors in art or architecture. They were informed that their help was needed in constructing an extrapolated or extended dictonary of English and were reminded of the formation of new English words through suffixation and given an example of it. The subjects were directed to pronounce the word resulting from the addition of the suffix and word supplied to them. In order to distract subjects' attention from the pronounciation task, subjects were also directed to say what the new word would mean and to tell whether they would be likely to use the word.

In administering the test, the subjects were interviewed individually. The examiner introduced a particular suffix by pronouncing it and spelling it aloud. Then two examples of that suffix's use in existing English words were given. Finally several stem-words (to which that suffix had not been added before in English) were presented, one at a time, permitting the subject to perform his tasks after the presentation of each stem-word. Ohala does not mention in his report that the time allowed for responding was limited. In fact (personal communication), there was no time limit in which subjects had to respond.

The entire test was presented orally to individual subjects. The subjects' responses were tape-recorded and later phonetically transcribed. In the case where a subject gave two responses and a preference between them, the preferred form was counted. If no preference was indicated the last form given

was counted.

In general, the results showed the independent rules under investigation to be unproductive. In addition, as M. Ohala found in investigating new derivations in Hindi, there was an "overwhelming" tendency for J. Ohala's subjects to keep the stem-words phonetically unchanged.

The majority of responses which were relevant to the rule represented by V → [-tense]/__CV̆CV retained a tense vowel in the stem. To discuss whether subjects used their first responses as a model for a response that followed, Ohala considered the adjacent words methane + -ity and sustain + -ity as well as define + -ity and iodine + -ity. The tetrachoric correlation coefficient measure used with responses for these words showed that the correlations of either both lax or both tense vowels in the stems of each of the above mentioned pairs of derivations were highly significant. Therefore it appeared to Ohala that subjects were basing their responses for sustain + -ity and iodine + -ity on their responses for methane + -ity and define + -ity respectively. Ohala concludes that subjects were likely using analogical phonological rules by using their first response as a model for the second in each of the pairs of derivations.

The sound pattern of one type of vowel tensing, V → [+tense]/__CiV̆ (for example, mammal/mammalian), was shown to be very unproductive with less than 4% of responses exhibiting the predicted tense vowel in the stem of relevant derivations.

Another case of vowel tensing, V → [+ tense]/__V (for example, algebra/algebraic) showed little evidence of productivity with only 13% of responses having the predicted tense vowel. The responses which occurred more frequently were the deletion of the stem-final vowel and the insertion of a glottal stop between the stem-final vowel and the initial vowel of the suffix. Ohaha's impression of these results is that the ə/ø alternation might be "the dominant strategy in the English 'conspiracy' to avoid vowel clusters" (p.43) as there are a number of existing English words showing this pattern (such as cholera/choleric, parabola/parabolic).

The 's-voicing' sound pattern, wherein

$$/s/ \rightarrow [+\text{ voice}] / \begin{matrix} V \\ [+\text{ tense}] \\ [-\text{ high}] \end{matrix} \_\_\_ V$$

(for example, gymnast/gymnasium), was very unproductive with such eligible stems as space and fleece where the preceding vowel was produced as a tense vowel.

Only 29% of responses that could have exhibited velar softening, the pattern wherein $/k^{d}/$[2] → /s/ /__I (for example, critic/criticism), did so. Ohala describes the percentage for velar softening as indicative of "marginal" productivity for the rule, given the conditions of the experiment.

Ohala also sought to discover whether the underlying vowel in the second syllables of obtain and pertain are marked for eligibility for vowel shift when the vowels are laxed. Ordinarily, underlying tense vowels such as those in divine and

---

[2]Ohala (1973, p.44) uses $/k^{d}/$ to refer to what Chomsky and Halle call "derivable" /k/ which is in opposition to ordinary /k/'s (such as that in keep which does not soften).

serene undergo vowel shift only if they remain tense. Occasionally, however, as in words such as detain/detention and retain/retention, Chomsky and Halle (1968) find it convenient to hypothesize that underlying tense vowels are eligible for vowel shift even after being laxed (by the rule that laxes vowels appearing before consonant clusters).

Ohala presented the stem-words pertain and obtain on two occasions. On each occasion, a different example of an existing derivation was given. The two derived examples (explain/explanatory, detain/detention) had different vowels in the second syllables of their stems. Subjects did not show consistent behaviour in deriving each of pertain and obtain with the suffixes -ion and -atory as they would likely have done had there been a specified underlying vowel in each of these stem words. Instead, subjects appeared to make their derivations according to the pattern provided them in the example dervations. In all, 9 of 26 subjects produced [ɛ] and [æ] as the second vowel in the new word in accordance with the example supplied them. Ohala concludes that "It is clear that the treatment of the vowel in these words depends not on its underlying specification but rather on what existing, surface alternation the derivation is patterned after" (p.44).

Perhaps a better way to verify the influence of example derivations showing appropriate surface alternations would be to give two tests to two different groups of subjects. One group would be given examples of derivations relevant to the derivations they were required to make. The other group would

not be exposed to these leading examples, Then a check could be made to find the total number of all subjects' responses wherein a phonetic change was made according to the examples provided. This number could be compared with the number of responses with identical phonetic changes made by subjects who were not supplied with the examples. Then some idea of the influence of subjects' exposure to leading examples of derivations could be gained. From this influence one might draw inferences on the possible use of analogical rules in the derivations tasks.

For instance, if the subjects who were exposed to examples of derivations illustrating given sound patterns happened to produce more responses showing those sound patterns than the alternate group of subjects, one might conclude that the presence or absence of leading examples influenced the phonetic form of subjects' utterances. The next task would be to analyze the data to see if the influence of those examples was caused by the examples' illustration of surface phonetic alternations or by their illustration of abstract phonological rules. This might be done by checking to see if subjects produce vowels of similar quality in the stems of derivations originally composed of identical stems but different suffixes. The different suffixes would, of course, serve as an environment in which the same independent phonological rule was hypothesized to apply. For example, Ohala checked to see whether the second vowel in the stem-words <u>obtain</u> and <u>pertain</u>, (which were each administered twice to be combined once with

the suffix -ion and once with -atory), underwent the same phonetic change in each derivation using a different suffix. He found that the responses of individual subjects for either pertain or obtain did not agree in the quality of the second vowel. Rather, as was mentioned earlier, the vowels were in accordance with the vowel alternation demonstrated in the example derivations. Ohala concludes that subjects were not aware of an identical underlying vowel in either obtain or pertain. This conclusion is based on the assumption that the same rule would be applied to the specified underlying vowel in the derivations involving, for example, pertain, producing an identical surface vowel in those derivations. Ohala thinks that the treatment of the vowel in each stem-word was not determined by an underlying vowel and a phonological rule. He believes that subjects instead directed their attention to the surface alternation of the vowel in the example derivations and to the surface vowel of the stem-word. Subjects might then have produced the new derivation by analogical rule.

Ohala suggests a possible means of operation for analogical rules:

> The essential part of the procedure is finding a form in the lexicon of existing words which can serve as a model for the derivation of the input stem. The model stem and input stem must be divided into dissimilar parts and similar parts....Then the dissimilar part of the model stem is subtracted from the model derivation and replaced with the dissimilar part of the input stem, i.e., [ɛkspl-] is subtracted from [ɛksplǽnətɔri] leaving [-ænətɔri] and then replaced by the dissimilar part of the input stem [ʌbt-] which yields [ʌbtǽnətɔri] (1973, p.46).

Another means of operation for analogical rules which also involves the surface phonetic forms of appropriate words is the proportional equation mentioned by Esper (1973, p.40). The proportional equation requires three words, each of which must either have a "formal" or "material" correspondence with one other of the words in the equation. The form of the equation is as follows: explain : explanatory :: obtain : _____, and may be read as explain is related to explanatory in the same way as obtain is related to _____.

Whereas Ohala considers his hypothetical operation of analogical rules to have mentalistic significance, Esper believes that the proportional equation has only a formalistic status. The latter states that "logicians and linguists may diagram analogies as proportional equations, but it is absurd to suppose that such operations are implied by analogical changes in language" (1973, p.153).

Returning to the discussion of Ohala's test results, it should be noted that one rule, V → [1 stress]/ __ C + affix, (a part of the stress rules posited by Chomsky and Halle (1968)), was shown to be "highly" productive for certain forms. Ohala feels that the productivity of this rule in words derived with the suffix -ian may be accounted for by the examples of this suffix which exhibited the stress pattern. Yet, the rule was also productive in words derived with -ity, whose example derivations did not show the stress pattern. In addition, two words showing "marginal" productivity of the rule (thermos, Thomas) contain reduced final vowels and have no

alternating word forms with reduced vowels. One word (human) with a reduced vowel in the final syllable of the stem showed great productivity for the stress rule. It, however, does have an alternating form in which this "same " vowel is unreduced and carries primary stress (that is, humanity). According to Ohala, "This suggests that if a particular vowel in the stem of a word is only realized as an unstressed completely reduced vowel, then it is difficult for speakers to put primary stress on it in derivations" (p.43).

Ohala concludes that the productivity of a given rule varies depending on the particular rule and on the examples of the sound pattern which are given to the subjects to illustrate the rule. He feels that in this study analogical rules were stronger than independent rules in forming new derivations. As was mentioned above, however, this hypothesis should be tested when the results of a test such as Ohala's can be compared with the responses of a similar group of subjects who have not been supplied with examples on which to possibly model their derivations. Administering such a test as Ohala's and omitting the leading examples might also provide the opportunity to check for subjects using analogical rules by referring not to the examiner's examples but to examples which they may have drawn from their lexicon.

Ohala expects that patterns not showing high productivity in this test could be made to show greater productivity if appropriate examples were supplied to subjects. Finally, Ohala concludes that the presumed underlying forms in certain

words (such as obtain and pertain) appeared to not be what certain linguists would have thought.

Steinberg and Krohn's (1975) study attempts to find evidence for or against the productivity of the Vowel Shift Rule (VSR) proposed by Chomsky and Halle (1968). The general nature of the rule can be illustrated by considering the words extreme and extremity, whose second vowels are [$I^y$] and [ė] respectively. The abstract representation of both vowels is /$\bar{e}$/. Steinberg and Krohn say that "in the case of extreme, the underlying /$\bar{e}$/ undergoes Chomsky and Halle's Dipthongization ($\bar{e} \rightarrow \bar{e}^y$) and then their Vowel Shift Rule ($\bar{e}^y \rightarrow I^y$). In the case of extremity, the underlying /$\bar{e}$/ undergoes a laxing rule ($\bar{e} \rightarrow e$)" (Steinberg and Krohn, 1975, p.234). The VSR is a general rule. This means that it applies to any underlying phonological representation of a lexical item having the proper structural description, except when the item is specially marked to not undergo the rule. Steinberg and Krohn believe that since underlying phonological representations are posited to accommodate the application of the VSR, an extensive revision of Chomsky and Halle's underlying forms would be required if the VSR's generality was proven invalid.

The authors mention three experimental studies which have been done to verify the VSR and which produced evidence contrary to the theory that speakers internalize this rule. (One of these was the unpublished study done by J. Ohala (1973) which was discussed above.)

Like the experiments of M. Ohala (1974) and J. Ohala (1973), that of Steinberg and Krohn involved novel derivations. In

general, the experiment's subjects were required to select, within the context of a meaningful sentence, one appropriate suffix of two given them, attach it to a base word and pronounce the novel derived form. The task of selecting the appropriate suffix was intended to distract the subjects' attention from pronouncing the new word.

Five vowels which were predicted by Chomsky and Halle to undergo change (when the base word in which the vowel occurs is combined with a suffix) were used in the test. The vowels and their alternations were: [$\bar{a}^{y}$] - [i] (divine/divinity), [$\bar{i}^{y}$] - [$\bar{e}$] (extreme/extremity)[3], [$\bar{e}^{y}$] - [æ] (sane/sanity), [$\bar{o}^{w}$] - [a] (verbose/verbosity) and [$a^{w}$] - [ʌ] (pronounce/pronunciation). Five suffixes were selected: -ic, -ical, -ify, -ity, -ish. All but -ish are predicted by Chomsky and Halle to cause vowel alternations in derivations of the base words. The suffix -ish was used to see if it would trigger the alternation seen in Spain/Spanish.

The base forms were 26 in number and were composed of ordinary English words. If the subjects chose the appropriate suffix, each suffix (except -ian) would be combined with five base forms. An additional base form was used with the suffix -ian.

---

[3]Steinberg and Krohn appear to use [$I^{y}$] and [$\bar{i}^{y}$] interchangeably for the vowel in the second syllable of extreme and [e] and [$\bar{e}$] for the vowel in the second syllable of extremity.

The test directions and stimuli were tape-recorded. The 26 items were arranged in random order. Subjects were further distracted from the pronunciation task by being told that the test's purpose was to investigate suffix preference. Each base word and the two suffixes were introduced after a brief paragraph in which the last sentence was deleted. The subject was required to repeat the last sentence aloud and "Fill the blank with the word (for example) maze plus either -ic or -ity". The time the subjects had in which to respond was limited as the experiment was pre-recorded. This time limit was not specified by the authors. One might conservatively estimate that the time limit was probably within four seconds as the subjects, after hearing the stem and the suffixes, heard the final sentence of the paragraph, a pause, the question "Ready?" and a click before they were to begin their response. Even at that point the pronunciation of the new word was delayed by the subject having to repeat the final sentence before filling in the blank at the end of the sentence with the pronunciation of the new derivation.

Before the actual test began, subjects were required to repeat aloud all of the base forms and suffixes for the examiner to ensure that they received the intended stimuli and to determine their pronunciation of the forms.

The subjects were native English speakers randomly chosen from introductory psychology courses. They were tested individually and their responses were tape-recorded. Scoring of the responses was done by two people who transcribed the subjects'

pronunciation of the base words, the suffixes and the derived words.

In Experiment I, 12 male and 12 female subjects participated. The experiment was conducted entirely auditorily. Experiment II, in addition to the auditory stimuli discussed above, involved stimuli of an orthographic nature. The 8 males and 8 females participating in Experiment II were divided into two groups of 4 males and 4 females each. These groups were submitted to different test conditions and were called Condition 1 and 2 respectively. The subjects of Condition 1 received, for each paragraph, one card on which the base word and the two suffixes were written. The subjects of Condition 2 received two cards for each paragraph. One card was identical to that in Condition 1. On the second card were written the two derived forms. The subjects of Experiment II were instructed to turn over and (silently) read the card(s) when a new suffix was introduced at the end of the paragraph.

Responses were discarded if they contained odd stress, a deleted or an added syllable, "disjuncture" between stem and suffix, a fabricated suffix or if no response was given. Steinberg and Krohn do not say why the above types of responses were discarded. Granted their criteria for disqualifying certain responses leaves the data with a much simpler nature. It can then be analyzed into three categories: responses with phonetic changes in the stem predicted by Chomsky and Halle, responses with phonetic changes not predicted by Chomsky and Halle, and responses with no phonetic change. The discarded

responses may have been considered as 'accidents' since, of the 50 discards out of 504 responses in Experiment I, 29 were attributed to 1/6 (that is, 4) of the subjects. In Experiment II, only 3 of the total 168 responses were discarded. (The total numbers of responses mentioned above do not include responses for the suffix -ish. They were not included in the final analysis because only 1 of 120 responses in Experiment I and 1 of 80 responses in Experiment II showed a phonetic change in the stem.) It is possible that many of these responses are not 'mistakes' but are the best answers the subjects could have produced. In this sense, they should have been considered as valid data along with the responses showing predicted and unpredicted changes and no change in the stem.

The results of both experiments will be considered together since the authors report that the difference between the distribution of responses for the two Experiments was statistically significant. Of the total valid responses for the two Experiments, 3.5% of the responses showed vowel changes in agreement with the VSR. An additional 6.6% of the responses showed phonetic changes not in accord with the VSR. A trend of no change in the vowels of the derived words was most predominant -- 89.9% of responses fit in this category.

There were 27 responses predicted by the VSR and 20 of these involved the base vowel [$\bar{a}^{y}$]. Of the 51 unpredicted vowel

changes, 29 were made in response to base items containing the vowel [ī$^{y}$]. Steinberg and Krohn report that the frequencies of the predicted responses to [ā$^{y}$] and of the unpredicted responses to [ī$^{y}$] are significantly greater than the frequencies of predicted and unpredicted responses to the other base vowels. The 27 predicted vowel changes occurred more frequently with the suffixes -ic (13 times) and -ical (5 times). The difference in frequency between -ic and -ical is not significant; however, the difference between -ic and the other suffixes (which were involved 3 times each with predicted vowel changes) is statistically significant. The authors suggest that the significant suffix differences are not due to an effect of the -ic suffix alone as most of the -ic responses occurred in conjunction with the base vowel [ā$^{y}$]. Since all of the -ic responses with [ā$^{y}$] base vowels were responses to a single word, sapphire, they say that "the possibility remains that the observed differences are due...to the effect of some idiosyncratic feature of that particular word" (p.250).

Steinberg and Krohn conclude that the findings of their study indicate that vowel alternation as predicted by the VSR is "largely non-productive." Therefore, if one still assumes that speakers have rules to account for vowel alternations, the "VSR accounts only for exceptions, i.e., to the creative pattern, which we state here; there is no (productive) vowel change in derived forms" (p.252). They justify this conclusion in saying that the notion of the productivity of a rule "is essential for distinguishing between generative phonology

(the creative generativity of living language) and etymology." Therefore they think it should be assigned an important role in judging the regularity of a rule.

They also conclude that since vowel alternation was seldom produced and since the VSR's prerequisite rule of vowel laxing seldom appeared to apply (in the predicted trisyllabic environment and in base forms taking the -ic suffix), the validity of claims for the generality and the psychological reality of the VSR is highly dubious.

Steinberg and Krohn go on to question whether alternations of sound patterns existing in the language are to be accounted for at all by rules. They suggest that it is possible instead that representations of both the base and the derived form are listed in the lexicon and that no rules are involved in vowel alternations of existing related word forms.

Finally, they discuss and dismiss some potential objections to the experiment and their conclusions. One objection might be that phonological rules do not operate for novel derived forms as they are meaningless. Steinberg and Krohn refute this criticism by pointing out what they consider to be evidence for the meaningful nature of the derived forms. They point out, first, that subjects generally selected the suffix which was appropriate to the context of the sentence. Secondly, subjects generally assigned correct stress to the novel derivations. The latter argument is perhaps weak. Speakers might easily be shown to assign English stress patterns in attempting to pronounce foreign phrases which were meaningless to them.

One might also object that a non-formative word boundary appeared between the base form and the suffix when the VSR was supposed to operate, blocking the application of the rule. However, the authors point out that if this were the case, the Main Stress Rule would also be blocked. The Experiments' results show that in general, stress was correctly assigned to the new derivations.

2.3.2 Evidence from other sources which is relevant to the findings of direct experimentation

Zimmer (1969) and M. Ohala (1974), in their attempts to directly test speaker's awareness of phonological rules, both found that speakers did not all appear to use the same phonological rules. Sherzer (1970) was writing a structural description of a word game played by the Cuna Indians of Panama. Like Zimmer and M. Ohala, he also concluded that speakers of a same dialect had different phonological models. He observed that some models might have been, from a linguistic point of view, "better" or "more correct" than the others. Zimmer also found that his speakers' rules for the aspect of labiality in vowel harmony were not "correct."

In any experiment which attempts to test subjects' awareness of rules, the examiner should not limit his expectations to finding, as Chomsky and Halle would lead one to believe, that speakers have identical sets of phonological rules and that these rules are as efficient as those of the linguist in accounting for phonological data. Evidence which may be useful in explaining speakers' different and "incorrect" phonological

models is reported by Wang, Chen and Wang, and Hsieh.

Wang (1969) surveys diachronic linguistic data for phonological change in many languages and concludes that there are "cross-currents of sound change vying for the phonological future of every word in the lexicon" of a language (p.23). If the sound changes are not considered over a long enough time, he doubts whether one could tell if the tendencies toward change are sporadic, persistent, old, new or receding. In addition Wang hypothesizes that rules for phonological change, while being phonetically abrupt, would be lexically diffuse. In other words, the change would occur abruptly for a given word but might not simultaneously apply to all of the appropriate words in the lexicon. Another inference drawn from the data he considers is that phonological rules may often be dependent on morphological and syntactic factors. He suggests finally that understanding the "complex dynamic situation" of competing cross-currents of sound change will require the "careful analysis each of the various interwoven factors--the physiological, the structural, the societal, and yet others" (p.24).

Chen and Wang (1975) consider diachronic data from Chinese, English and Swedish as well as evidence from language acquisition to describe how a sound change implements itself. Their claim is that this implementation occurs through lexical diffusion of the sound change; the change propagates gradually from morpheme to morpheme. They also comment on the "actuation" of a sound change, or why it assumes a particular form and follows a schedule. Chen and Wang's source of evidence for this problem

is consonantal attrition in Chinese dialects and Indo-European languages. They claim that the actuation of a phonological process is to be sought in the "inherent constraints of the physiological and perceptual apparatus of the language user" (p.255). They provide experimental evidence (both perceptual and physiological) in support of their theory.

Chen and Wang reveal that the aims of their study were not only to prove a certain point but to exemplify the philosophy of experimental research in linguistic study. They believe that their approach proved "more rewarding than the pursuit of 'explanatory adequacy' in terms of a-prioristic and sterile notions of simplicity and economy" (p.279). The "pursuit of 'explanatory adequacy'" which they mention refers of course to the research strategy of Chomsky and Halle (1968). The experimental research performed by Chen and Wang is the type of alternative approach linguists such as Ohala (1974) and Lindblom (1971) have in mind when they criticize the poor methodology and explanatory power of Chomskyan theory.

In addition to historical data as a source of evidence from linguistic change is the data available in studying the phonological development in children acquiring language. In order to support Wang's concept that sound change is lexically gradual Hsieh (1972) cites child language acquisition as a data source with two advantages. He believes that children's phonological development provides cases of sound change that are both completed within a period of a few years (and so are practicably observable) and that are immune to borrowing from other diaects.

## 2.4 Conclusions

1. The experimental verification of some of TGP's hypotheses, which requires violation of the ideal speaker-listener frame, can be justified.
2. Certain phonological rules formalized by Chomsky and Halle (1968), at least in given experimental conditions, are not productive (J. Ohala, 1973; Steinberg and Krohn, 1975). The validity of claims for those rules' psychological reality and generality is therefore questionable.
3. Different speakers have different phonological models (according to evidence from direct experimentation by Zimmer (1969) and M. Ohala (1974) and from observations of a linguistic game by Sherzer (1970)). These phonological models, furthermore, are not always "correct" from the point of view of an economical linguistic description.
4. Phonological change diffuses slowly across the lexicon (according to evidence presented by Wang (1969), Chen and Wang (1975) and Hsieh (1972)).

# Chapter III

## METHOD

### 3.1 Aims of the investigation

It was mentioned in the introduction that one of the objectives of this study is to verify the psychological reality for a given group of speakers of some of Chomsky and Halle's (1968) phonological rules. This aim might be fulfilled by providing subjects with the opportunity, under different test conditions, to produce novel utterances and by checking to see whether the phonetic forms of those utterances reflect the use of certain phonological rules.

It was pointed out that the phrase "psychological reality of phonological rules" is assigned different senses by different linguists. In the experimental studies mentioned earlier, a rule (or sometimes an underlying phonological representation of a word) was considered to be part of the speaker's mental mechanisms underlying his speech production and perception if the rule was used in the speaker's generation of responses in the test situation. The 'use' of a given rule in past studies has been based on the assumption that the appearance of a phonetic form (which is predicted by the application of that rule to some underlying representation) implies that the speaker employed the rule. The studies previously discussed referred to the psychological reality of a rule for individual speakers and also groups of speakers. The extent of the rule's use in the test situation, (or in other words, the percentage of times that the rule was used with appropriate stimuli), is called

the rule's productivity for the test. However, a minimum amount of productivity which would qualify the rule as being psychologically real either for one or for many speakers, was never specified. It is important to decide upon a specific amount of productivity to be demonstrated by a rule in order for it to qualify as being psychologically real. This minimum productivity is ultimately needed to judge the validity of claims for including the rule in a mentalistic grammar for a specific language. Some of the rules whose psychological reality will be tested in this study are considered to be general rules by Chomsky and Halle (1968) and so would be expected to demonstrate a high degree of productivity before their inclusion in a grammar is justified. Other rules which will be considered in this investigation are not described by Chomsky and Halle as general rules. A specific criterion of productivity, justifying the inclusion of a rule in a grammar, might depend on such considerations as the number of speakers, for example, an individual, a small group within a dialect, all the members of a dialect, etc., for which the grammar was meant to be representative.

In the present study, discussion will center on whether the use of an "independent phonological rule" (such as those hypothesized by Chomsky and Halle) can be assumed if the speaker produces a response predicted by the hypothetical application of that rule to some underlying representation of the word. Also to be discussed is whether the use of the independent phonological rule can be refuted (as was done by J. Ohala, 1973)

if it can be shown statistically that the forms of the speakers' responses have high correlations with the forms of adjacent responses or of recent derivations. In addition, the possibility for the application of a rule (other than that hypothesized by Chomsky and Halle) to result in the response predicted by them will also be considered. Recall the notion discussed by Botha (1970): the fact that a theory is able to correctly predict data is not sufficient evidence for the proof of that theory. He believes that it is no more than a coincidence that at a given moment there is only one theory that "works" and that no alternative can be proposed.

Steinberg and Krohn (1975) appear to distinguish between the "psychological reality" and the "validity" of the Vowel Shift Rule for their group of subjects. The rule was shown to be "largely non-productive" in the responses of their subjects. Steinberg and Krohn concede that the rule may be psychologically real for their speakers. They deny, however, that the rule has "generality." They state that "C&H's [Chomsky and Halle's] claim that the VSR is a psychologicall real and general rule is...highly dubious" (p.252). Judging from their concession that speakers may have such rules as the VSR to account for vowel alternations which are exceptions to the more general creative pattern of no vowel change in derived forms, the authors probably mean that Chomsky and Halle's claim that the VSR is psychologically real as a general rule is "highly dubious."

At this time it is impossible to non-arbitrarily state in

advance, for the purpose of this study, a minimum amount of productivity of a rule which would qualify that rule as being psychologically real for a group of speakers and as being validly assigned the role of a general or 'less-than-general' rule in the grammar for a group of speakers.

3.2 Experimental plan

In general, subjects were required to combine, without the use of pencil and paper, English stem-words with suffixes in order to create a new word and to thereafter pronounce it aloud. The subjects' spoken responses provided data on the frequency of production of certain sound patterns in novel utterances. The data should permit one to infer whether certain types of rules were used in the production of the new words.

The purpose of the pilot experiments was to check several aspects of the methodology of the proposed main experiment. The problems encountered with certain methods and the conclusions drawn from the results of the pilots will eventually be discussed.

In each of the two pilot experiments (called Pilot I and Pilot II, respectively) one group of subjects (Group a[1]) were provided with two example derivations for each suffix. The example derivations demonstrated the use of the suffix in English. Some of the examples exhibited phonetic changes in

---

[1]Group a shall refer to Group a' and Group a'' of Pilot I and Pilot II, respectively. Group b shall likewise correspond respectively to Group b' and Group b'' of Pilot I and Pilot II.

the stem and/or in the suffix of the derivation and some did not. Another group of subjects (Group b) in each of the pilot experiments was provided with no derivations exemplifying the use of the suffixes. The arrangement of the experiments into groups and the contents of each group's experiment is outlined in Table I.

In the main experiment, one group of subjects (Group A) were provided, for each suffix, with one example derivation that showed no phonetic change, as well as one example demonstrating a possible phonetic change, in their stems and/or suffixes. Often the phonetic changes illustrated by the example derivations were those predicted by Chomsky and Halle to occur when an hypothesized phonological rule was applied to an hypothesized underlying representation.

Ohala (1973) hypothesized that rules showing low productivity in his test could be made to show greater productivity if appropriate example derivations were supplied to the subjects. Therefore some of the example derivations used by Ohala were substituted with more appropriate examples illustrating the phonetic changes predicted by Chomsky and Halle to occur with the use of their hypothesized rules. These more appropriate examples were included in the list of examples supplied to the subjects of Group A in the main experiment.

3.3 Subjects

In all, 41 subjects participated in the studies. As can be seen in Table I, a total of 9 subjects participated in the

| Experiment: | Pilot I | | Pilot II | | Main | |
|---|---|---|---|---|---|---|
| Group | a' | b' | a'' | b'' | A | B |
| **Contents:** | | | | | | |
| Practice task | yes | yes | yes | yes | yes | yes |
| example derivations with phonetic change | yes | yes | yes | yes | yes | yes |
| example derivations without phonetic change | yes | yes | yes | yes | yes | yes |
| Experimental task | | | | | | |
| example derivations with phonetic change | yes | no | yes | no | yes | no |
| example derivations without phonetic change | yes | no | yes | no | yes | yes |
| **Subjects:** | | | | | | |
| Number of female | 0 | 3 | 1 | 1 | 13 | 10 |
| Number of male | 3 | 0 | 1 | 0 | 4 | 5 |
| Total number | 3 | 3 | 2 | 1 | 17 | 15 |

Table I. Arrangement of the experiments into groups, the contents of each experiment and the number of subjects participating in each experiment.

pilot experiments and 32 subjects took part in the main experiment.

Twenty-eight of the total number of subjects used in the experiments were females and 13 were males. Table I also shows the numbers of female and male subjects participating in each experiment.

It was intended that the subjects participating in the experiments should meet the following qualifications: (a) they should have been born and/or have lived most of their lives in British Columbia; (b) English should be their native language; (c) they should have taken no linguistics courses; (d) they should be working towards or have completed a Bachelor of Arts or Bachelor of Education degree at one of three universities in the lower mainland of British Columbia.

Some of the subjects did not fulfill all of the above qualifications. First of all, with respect to qualification (b), three subjects reported that they first began to learn another language simultaneously with English. Before the subjects began primary school, however, English was being used more frequently than the alternate language. The subjects reported at the time of the experiments that they possessed only minimal facility in the listening and speaking skills of the other language. Secondly, ten subjects were not students within Arts or Education Faculties. Five subjects were studying in a Faculty of Science, two were studying in a Faculty of Commerce, one was studying in a Faculty of Physical Education and Recreation and two were studying in a Faculty of Fine Arts.

Thirty-nine of the total 41 subjects were within the age bracket 18-29 years. The remaining 2 subjects were in their mid-forties.

All 41 subjects had been exposed to the study of a language other than English. Altogether, 20 subjects had studied a second language in secondary school only and 21 had studied another language at university as well. In the main experiment 7 subjects in Group A and 8 subjects in Group B had studied a second language in secondary school only. Ten subjects in Group A and 7 in Group B had studied a second language in secondary school as well as at university. Finally, in addition to the 3 subjects mentioned earlier who had begun to learn English simultaneously with another language, 2 subjects in Group A of the main experiment had been exposed to a language other than English at home some time during their lives.

3.4 Composition of the word lists

The list of word forms used in all of the experiments consisted of English words (henceforth called stem-words to differentiate from derivations), suffixes and derivations (composed of stem-words and suffixes). The total list of these forms comprised two successive sub-lists: a section of stem-words, suffixes and derivations which served as stimuli for a practice task and a list comprising the experimental task stimuli.

3.4.1 Practice task stimuli

The practice task was included in both the pilot and main experiments in order to give subjects practice at deriving new

new words using existing stem-words and suffixes. Ohala's format, wherein a suffix and two example derivations precede the presentation of a group of stem-words, was used. The responses for the practice task were not to be included in the analysis of the responses from the experimental task.

The list of word forms used as stimuli in the practice task consisted of 9 stem-words, 3 suffixes and 6 existing couples of stem-words and derivations. Only 3 of the 9 stem-words were greatly expected to undergo a phonetic change in the stem when they were combined with a suffix. These were located towards the end of the practice task in an attempt to proceed from easier to more difficult derivations. Not all of the 6 example derivations showed a phonetic change in the stem. This was arranged to prevent subjects from expecting all derivations to either undergo or not undergo a phonetic change. Only two of the examples exhibited any phonetic change. The two examples industry/industrious and angel/angelic both showed a change in the location of stress and also a change in the quality of one and two vowels, respectively.

Two example derivations were used with each suffix. The particular stem-words to be combined with the suffix -ic and -ous were expected to possibly undergo a phonetic change. The inclusion of stem-words likely to undergo changes when derived with suffixes in the practice task was arranged to introduce subjects to the notion that a suffix being illustrated by an example showing a phonetic change in the stem did not

require the derivation task to necessarily involve a phonetic change. The list of word forms comprising the practice task is given in Appendix I.

3.4.2 Experimental task stimuli

The list of stimuli in the two pilots was identical to that used by J. Ohala (1973) and is given in the Appendix. Although there were 60 different stem-words in the list, three of these (obtain, pertain and trade) were stated twice, making a total of 63 stem-words. These stem-words were arranged into twelve groups. There was one suffix available to be combined with the stem-words in each group. (Although there were only 10 different suffixes in Ohala's list, two of the suffixes (-ic and -ity) were listed twice, making a total of 12 suffixes available for stimuli to accommodate the 12 groups of stem-words.) In addition, the list contained 20 different examples of existing stem-words and derivations. Two of these (odd/oddity and Darwin/Darwinian) were each given twice, bringing the total number of example derivations to 22. This meant that 10 groups of stem-words were exemplified with two example derivations while two groups were illustrated with only one example.

Twenty-two of the stem-words and 3 of the suffixes (which were combined with an additional 10 stem-words) in Ohala's list were described by him as "filler" words and "filler" suffixes, respectively. Ohala did not expect derivations involving these word forms to show any phonetic change and did

not include the responses for these stimuli in his analysis. The purpose of including these stimuli was to prevent subjects from believing that the addition of a suffix necessarily required a phonetic change in the stem. In addition, 8 of the 20 different example derivations showed a phonetic change in the stem. One example (explain/explanatory) showed only a vowel change. One showed only a consonant change (fuse/fusion). Another showed change in a vowel and in a consonant (detain/detention). Five examples demonstrated a change from an unstressed schwa to another lax vowel which was stressed (stupid/stupidity, Darwin/Darwinian, solid/solidify, artist/artistic, realist / realistic). One example derivation demonstrated a change from an unstressed schwa to another vowel with primary stress as well as a change in a consonant (music/musician).

Although the list of stimuli used in the main experiment included the stem-words and suffixes mentioned above, the list of example derivations was not identical to Ohala's list. In order to exemplify each suffix with one derivation showing a phonetic change in the stem and one derivation showing no change, 9 examples in Ohala's list were replaced by different example derivations. Four of the new examples demonstrated a change in vowel quality. The examples sane/sanity and rate/ratify were adopted to illustrate vowel laxing, while mammal/mammalian and algebra/algebraic were intended to exemplify a vowel tensing rule. One example, critic/criticism, demonstrated the velar softening rule. Four of the new examples, Kant / Kantian, false/falsify, rebel/rebellion and graph/graphic were included

to ensure that each suffix was exemplified by one example derivation exhibiting no phonetic change. All of the stimuli mentioned above are listed in Appendix I.

It should be noted that the combination of two stem-words and their suffixes (between + -ity and probe + -ity) in the experimental task could result in words that exist in English. The word betweenity is rarely used. According to Oxford English Dictionary, it is pronounced with a tense vowel in the second syllable position. The word probity is cited by Chomsky and Halle (1968, p.181) as an exception to the words to which the vowel laxing rule V → [-tense]/ __ C$\breve{V}$CV applies. These stem-words were retained in the set of forms to be combined with the suffix -ity in order to see how subjects' pronunciations of their derivations compared with the existing pronunciations.

### 3.4.3 Verification of the appropriateness of combining the stem-words and suffixes

Studies by Nicholson (1916) and Marchand (1960) demonstrate that existing English derivations show trends for stem-words and suffixes to have been restricted in their combinations according to such factors as the language of their origin and their grammatical function. For instance, Marchand states that the suffix -ity forms abstract substantives from its combination with adjectives (most of which are of Latin origin) and with nouns.

The grammatical function and the language origin of each of the stem-words and the suffixes in both the practice task and the experimental task were examined in preparation for their

being retained in the lists. This information was obtained from the works of Nicholson and Marchand mentioned above, from The English Word Speculum (Dolby and Resnikoff, 1967), from the Oxford Dictionary of English Etymology (1966) and from Webster's English Dictionary. It was then possible to verify whether the stem-words' characteristics were appropriate for the individual derivational tendencies of the suffixes with which they were to be combined. Most of the stem-words seemed to match the qualities required by their suffixes. Some of the stem-words, however, had seemingly inappropriate characteristics which could have influenced the form of their derivations with the suffixes specifically assigned to them in the experiments.

Some stem-words' grammatical functions were inappropriate for the specific suffix. The verbs define and sustain were combined with the suffix -ity which forms substantives from nouns and adjectives. Less than 10 derivations comprised of a verb stem and -ity were found in Dolby and Resnikoff. The verb trample was combined with the suffix -ify which forms verbs from its combination with nouns and adjectives. Only two derivations formed with -ify were found (1 in Marchand and 1 in Dolby and Resnikoff) to contain a verb as a stem. The noun sapphire was combined with the suffix -atory. This suffix actually has the form -ory (according to Marchand and the Oxford English Dictionary of Etymology) and has been combined with verbs of Latin origin ending in -ate.

In some instances the inappropriate characteristic of the stem-word arose from its language origin. The noun Buddha, of

Sanskrit origin, may not be an appropriate stem for the suffix -ic since all the words in Dolby and Resnikoff ending in V + -ic were of Greek or Latin origin. The noun trade, of Germanic origin, was combined with the Latin suffix -ian(us). The suffix -ian is most often combined with proper nouns (Marchand) or, as was observed in Dolby and Resnikoff, with nouns only of Greek or Latin background.

The semantic features of some of the stem-words could possibly serve as a problematic source for their derivation with certain suffixes. Apple is an inanimate noun. It was combined with the suffix -hood which was observed in Dolby and Resnikoff to be predominantly combined with animate nouns. Only one derivation with -hood, whose stem was inanimate, was listed in Dolby and Resnikoff. The abstract noun joy was combined with the suffix -dom. The suffix -dom appeared in Dolby and Resnikoff to be derived only with noun stems which had the semantic quality 'animate.'

All of the stem-words mentioned above were retained in the experimental task's word list. The responses to these stem-words and their suffixes were observed while keeping in mind the possibly inappropriate characteristics discussed.

## 3.5 Administration and recording of the experiments

### 3.5.1 The stimuli

The type of stimuli given to each group of subjects is outlined in Table I. The subjects of the two pilots were given the word forms of the practice task and those in J. Ohala's list of stimuli. In each pilot, however, there were two groups of

subjects who were given different test conditions. One group of subjects (Group a) in each pilot were given all of the example derivations in Ohala's list of test stimuli. The other group of subjects (Group b) in each pilot heard no example derivations other than those in the practice task. For them, Ohala's example derivations were omitted.

As was mentioned earlier, the subjects of the main experiment were also members of one of two groups (either Group A or Group B). Each group was administered the entire practice task. The experimental task's word-stems and suffixes were taken from Ohala's list. When it came to example derivations, however, Group A was given two--one example demonstrating phonetic change and one showing no phonetic change--to illustrate each suffix. Group B, on the other hand, received only one example derivation, showing no phonetic change, for each suffix in the experimental task.

It should be noted that J. Ohala repeated only one of a suffix's example derivations upon presenting that suffix to subjects for the second time. In the first condition of each of the experiments described here (that is, in Groups a', a", and A) where two example derivations were supplied with each suffix, both of the suffix's examples were repeated to the subjects when the suffix was presented on a second occasion. This was done in an effort to keep the conditions of the experimental task the same for each set of stem-words with which the repeated suffixes were to be combined.

3.5.2 Directions to the subjects

The directions to each group of subjects may be found in Appendix I. Except for a few alterations, they are the directions cited by J. Ohala (1973) which were used in his test. In brief, the directions for all of the studies had the following points in common. Subjects were told that the examiner needed their help in forming an extended or extrapolated dictionary of English. They were reminded of a process called "suffixation" by which new words in English are often formed and were given an example of a word derived by suffixation. They were then told that the examiner would read aloud a suffix, some examples of English words containing that suffix and some English words which had not been used with the suffix before. Finally, they were asked to create and then pronounce the new word formed by the combination of the suffix and the English word provided them, to define the word and to say whether they would use the word.

The differences in the directions given to the subjects of the different experiments mainly concerns the different words used to convey the process of suffixation. For instance, both groups of subjects in Pilot I were told that new words can be formed by the "addition" of suffixes or endings to the end of existing words. They were also then asked to pronounce the word resulting from the "addition" of the suffix to the stem-word. The responses of the subjects of Group b' indicated that subjects were likely following the directions explicitly and adding or tacking on the suffix to the stem-word without making any phonetic changes regardless of any tendencies they felt to

making phonetic changes. For instance, there was a predominant trend of no phonetic change in the responses of these subjects. In addition, two of the subjects during the test asked the examiner in rhetorical fashion whether the subject was to "just add on the ending" to the stem-word. Therefore in Pilot II and in the main experiment, the word "addition" was changed to "combination" in reminding subjects of the process of suffixation. The subjects were also asked to take the word and the ending and "put them together somehow" in order to create the new word. It was hoped that the implications of no or of minimal phonetic change involved in the notion of "adding" the suffix to the stem would thus be eliminated. The idea of "combining" or "somehow putting together" the word-stem and the suffix was intended to broaden subjects' expectations to include the possibility for phonetic change to be involved in the word creation task.

Another difference in the directions consisted of reminding subjects in the main experiment that the study in which they were participating was similar to a "survey" in that all of their responses were valid and would not be considered as "right" or "wrong." The subjects had been told this once before when arrangements were made for their participation in the main experiment. The examiner explicitly referred to the experiment as a "survey" for two reasons. One reason was based on a subjective impression that the subjects of both pilots were concerned with the validity of their responses. For instance, some subjects in the pilots repeatedly apologized for their

"poor knowledge of English" and often responded during the interview in a hesitant manner with anxious expressions on their faces.

The directions for the two groups of subjects in each of the two pilots differed slightly in a way which accounted for the absence of derivation examples in the experimental task. These differences may be seen in the directions which are listed in Appendix I. The cutback in the number of examples given to Group B in the main experiment was not mentioned in the directions to the subjects as each suffix in the experimental task was exemplified once.

3.5.3 Order of presentation in the experiments

The contents of each experiment (except Pilot II b" ) were presented in the following order. First the directions were read to the subject. Then the examiner presented the word forms of the practice task and then those of the experimental task. There was no break between the practice task and the experimental task. Subjects were not informed that their responses for the first nine stem-words were only practice derivations that would not be counted in the results.

The order of presentation within both the practice and experimental tasks was the following. First the suffix and then one or two example derivations were presented. Then a word-stem was given. Finally the subject made his/her responses for the combination of the word-stem and suffix. That is, he/she pronounced the new word, attempted to define it and then said whether he/she might use it. The next stem-word was

presented, the subject made his responses, and so on. This order of presentation of the suffix, the example derivation(s) and the stem-words is identical to that used by Ohala in his test.

In Pilot II b", the examiner's instructions, and the subject's opportunity, for defining the new words and for saying whether the subject would use the new words were delayed. When the subject had pronounced all of the words created from the stimuli in both the practice and experimental tasks' lists, she was informed that the examiner would give the suffix, the example derivation, the stem-word and then the new words the subject had suggested. (The new words were pronounced by the examiner according to the phonetic transcription in which the words had been recorded.) Finally the subject was directed to tell what the new word would mean and whether or not she would use the new word.

The delay in defining and commenting on the use of the new words was interposed in an attempt to see if the predominant trend of no phonetic change in the responses of Pilot I b' 's subjects had been influenced by the possibly difficult job of having to define and comment on the subject's use of the new word immediately after pronouncing the word. Phonetic change did occur more frequently in the responses of the single subject in Group b" than in the responses of Group b' 's subjects. However, this separation of the pronunciation task from the other two required responses almost doubled the time needed to complete the interview. Therefore the method was not

adopted for use in the main experiment.

3.5.4 Method of presentation

With one exception, the studies were presented orally by the examiner. The directions and the word forms were read aloud by the examiner from a sheet on which the subject's responses were to be recorded. In the pilot experiments the suffixes, in addition to being pronounced aloud, were spelt aloud. The suffixes were presented once at the beginning of the list of stem-words which were to be combined with that suffix.

The one exception to the oral presentation occurred in the main experiment. In this case, the suffixes were not pronounced and spelt aloud. Rather, each suffix was presented in printed form on a 3 inch by 5 inch card. The subject retained the card in front of him throughout the time needed to respond to the combination of the suffix with each of the stem-words. The card was turned face-down before the card with the next suffix was presented. The suffixes were presented in printed form in the main experiment because it was found that subjects in the pilot experiments had trouble remembering the suffix when it was spoken and spelt only once at the beginning of a series of stem-words. In addition, the fact that a given suffix is pronounced differently in different derivations made it difficult to decide on how the examiner should pronounce the suffix in isloation without biasing the subjects' pronunciation of the suffix in different new words.

All the subjects were interviewed individually. As was done by J. Ohala, in the case where a subject produced more than one

pronunciation of a new word without indicating his preference, the examiner asked the subject which pronunciation he preferred.

3.5.5 Recording the data

The subjects responses were recorded in two ways: by a tape recording and by phonetic transcription. Tape recordings were made of the examiner's presentation of and the subjects' responses to the practice task and to the experimental task stimuli. The definitions of the new words and the comments on the use of these words were also tape recorded to distract subjects' attention from the pronunciation task. The recording equipment used was an Ampex Micro 10 tape recorder powered either by batteries or by a Philips CP 9140 A power source. The Ampex Micro 10 microphone was clipped onto the subject's clothing just below the neck. Most of the subjects (in all the experiments except Pilot II b'') required about 40 minutes to complete the experimental session.

Subjects' pronunciations of each new word were recorded in broad phonetic transcription on a response form during the interview. The symbols of the International Phonetic Association (1972) were used in this transcription. The transcription was broad in the sense that in addition to the segmental symbols, only the signs for the extended length of a sound [:], for the syllabicity of a contoid [,] and for primary stress falling on a syllable ['syllable] were used.

To further distract subjects from the pronunciation task, comments made by the subject in defining the new words and their decisions about possibly using the new words were randomly noted on the response form.

## Chapter IV

## ANALYSIS AND RESULTS OF THE DATA

### 4.1 Checking the reliability of the transcriptions

Since the phonetic transcriptions of responses were done by the experimenter during the interview, an attempt was made to informally check the reliability of the transcriptions of the main experiment. The tape recorded responses of 3 subjects each from Group A and from Group B were played on a high quality reproduction system and were transcribed by a second person. This person was directed to transcribe, as the experimenter had done during the interview, the differently pronounced responses for the given stem-suffix sets and to note the pronunciation preferred by the subject.

Only the responses to the experimental task whose stimulus was relevant to one of the rules, and whose form was preferred by the subject, were considered. Disagreements between the two transcriptions of responses were counted. The disagreements involved: (a) a difference in the location of primary stress; (b) a difference in the number of syllables in the response; (c) a difference in a consonant and/or a vowel which was predicted (by Chomsky and Halle, 1968) to undergo phonetic change; (d) a difference in which response was preferred by the subject. The two transcriptions showed disagreement in an average of 3-4 responses for the total of 39 items which were considered in each of the six interviews.

Two interviews (one from each group) were arbitrarily chosen

to be transcribed a second time by the experimenter and the second transcriber in order to check the reliability of each person's transcriptions. There were three disagreements between the two second transcriptions of one interview and four disagreements between the two second transcriptions of the other interview. These disagreements represented an increase of one disagreement for each interview over those found after the first trial of transcriptions. The differences between the second transcriptions did not involve many of the same responses which produced disagreement in the first trial. This meant that each transcriber was, on occasion, inconsistent in transcribing some responses over the two trials.

It was not anticipated that the transcriptions by each person would become consistent after further trials of listening to and transcribing the tape recorded responses. In the case of the two interviews which were transcribed a total of four times, disagreements were resolved when one phonetic form for a given stem-suffix set occurred once and the other form occurred three times. In the two interviews, a total of 5 disagreements in responses were resolved in this way. The original transcriptions of the examiner were retained for the remaining disagreements.

4.2 Estimations of the lengths of response latencies

Word derivation experiments conducted by Zimmer (1969) and Steinberg and Krohn (1975) involved time limits for subjects to make their responses. Since the present study did not impose a limit on the time which subjects had to pronounce the

new word, an estimate of the range in the length of response latencies and an estimate of median and upper quartile response latencies were made in two interviews (one lasting the greatest and one lasting the least amount of time) from each of Group A and Group B.

In each interview the time latency between the end of the examiner's presentation of the stem-word and the beginning of the subject's preferred response was measured for each item in the practice and in the experimental task. Considering the shortest and the longest interviews does not of course guarantee that the range of response latencies or the median or the upper quartile response latencies found in those interviews will be respectively least and greatest for a group of subjects. It should, however, give at least a fair idea of the range to be encountered. The results can be seen in Table II

| | Group A | | Group B | |
|---|---|---|---|---|
| overall duration of interview | 30 min. (shortest) | 42 min. (longest) | 28 min. (shortest) | 45 min. (longest) |
| response latencies (in sec.) | | | | |
| range | .7-6.2 | 1.0-48.8 | .2-24.5 | 1.2-33.7 |
| median | 1.3 | 3.8 | .8 | 3.3 |
| upper quartile | 1.8 | 8.5 | 1.3 | 6.3 |

Table II. Estimated range, median and upper quartile values of response latencies.

The results show a considerable range of response latencies for the subjects within each group. The median response latency values indicate that half of the preferred responses for the subjects of both groups were probably made within a period of about 4 seconds. This period of time is similar to the time limits (mentioned earlier) which were afforded the subjects in the experiments of Zimmer (1969) and Steinberg and Krohn (1975). The upper quartile values, however, vary considerably between the shortest and longest interviews of each group.

## 4.3 The "productivity" of the phonological rules

### 4.3.1 Introduction

Authors such as Zimmer (1969), Krohn (1972), J. Ohala (1973) and Steinberg and Krohn (1975) refer to the "productivity" of phonological processes accounting for the phonological regularities in language. These authors tend first to refer to the "productivity" of a phonological regularity, of a sound pattern or of a vowel alternation. Only afterwards do they discuss the "productivity" of a formal rule which has been hypothesized to describe the regularities in sound patterns observed at the surface structure level. They do this because, as was mentioned in Chapter 3, they equate the "productivity" of a rule with the frequency of speakers' presumed 'use' of that rule, which they in turn infer from the "productivity" or occurrence of a given sound pattern. In the present study also, the frequency of occurrence of given sound patterns will be discussed before trying to infer the subjects' 'use' of certain phonological rules.

Krohn states: "By productivity, we are referring to the ability of a rule to apply to new items that have been added to a grammar" (1972, p.17). It seems peculiar to use the words "productivity" as well as "ability" with reference to non-animate nouns such as "phonological regularity," "sound pattern" and "rule." Therefore the word "productivity" will only be used in the present discussion with reference to the ability of individual subjects or groups of subjects to produce a sound pattern or to apply a rule to new items. Furthermore, it will be specified whether the ability of subjects involved in this "productivity" refers to the ability only to produce a sound pattern by unspecified means or to the ability to produce a sound pattern by using a certain rule. Otherwise, reference to a sound pattern will concern the notion of that pattern's frequency of occurrence. Reference to a rule will concern the sound patterns' implications for the corresponding frequency of use of that rule.

### 4.3.2 The occurrence of sound patterns in each group of subjects

One purpose for considering the different sound patterns and the number of times that they occurred in the responses of each group was to determine if the different experimental conditions experienced by each group influenced the responses. The stem-suffix sets whose responses were considered for the occurrence of different sound patterns were those considered by Ohala (1973). In addition, some sets which had not been considered by Ohala happened in this study to elicit responses exhibiting the predicted sound pattern and were therefore

included in the analysis. All of the stem-suffix sets mentioned above are given in Appendix II and comprise what shall be referred to as the 'long' list.

Most of the stem-suffix sets considered in the analysis were preceded in the conditions of Main Experiment A by an example derivation demonstrating the predicted sound pattern. (In order for an example derivation to qualify as 'preceding' a stem-suffix set, the example had to be presented at the beginning of the group of sets in which the set under consideration occurred.) Some sets (noted in Appendix II) were not preceded by an appropriate example derivation. The purpose of considering the occurrence of sound patterns was to determine if the responses were influenced by subjects' exposure or lack of exposure to appropriate example derivations. Therefore it seemed important to consider as well the occurrence of sound patterns in a shorter list of stem-suffix sets which included only those sets which had been preceded by an appropriate example derivation in Main Experiment A. This shorter list will be called the 'short' list.

It is possible that (a) the variety of sound patterns and (b) the percentage of occurrence of a given sound pattern in the total responses considered will vary under the influence of the number of responses considered in the analysis. Steinberg and Krohn (1975) eliminated several forms of responses (see Chapter 2) from the data which was considered in their analysis. Unlike the subjects participating in Steinberg and Krohn's experiments, the subjects in the present study were not

systematically limited in the time which they had to make their responses, and so probably always produced the response they intended to make. If they did make a slip of the tongue and were conscious of it, they had the time to correct their response. Therefore all of the responses made by the subjects in the present study were analyzed. It was believed that considering all of the responses would more fairly reflect subjects' abilities to produce predicted (and unpredicted) sound patterns and therefore to use predicted (and unpredicted) rules in novel utterances.

Another variable which can influence the variety of sound patterns and their frequency of occurrence is the choice and the number of categories to which the different responses are assigned. The categories were based upon the different types of responses produced by the subjects for the stem-suffix sets relevant to one rule. The responses differed according to the vowel or consonant in the stem of the derivation which had been predicted to undergo phonetic change. For instance, methane + -ity elicited responses containing different vowels such as are seen in [məðænəti ], [məˈθe͡ɪnəti ], [məˈθɪnəti ] and toxic + -ism resulted in the production of different consonants such as can be observed in [ˈtɑksəsɪzm̩] and [ˈtɑksəkɪzm̩]. Some responses (such as sustain + -ity [sʌstəˈnɪnədi ] exhibited an additional segment or segments and were counted in an 'Addition' category. Responses were counted in a 'Substitution' category when the stem of the response differed from that of the stimulus yet appeared to closely resemble another existing

stem. For example, [bu'dɪstɪk] (derived from Buddha + -ic but resembling Buddhist + -ic) and [æd'vətɪzm̩] (derived from adverse + -ism but resembling advert(ise) + -ism) were counted in a 'Substitution' category. A 'Deletion' category was used for responses such as ['sowfɪzm̩] (derived from sofa + -ism) and [baɪ'alədʒɪk] (from biologist + -ic) which appeared to drop one or more segments.

Some responses (e.g. [baɪ'alədʒɪstɪk] (from biologist + -ic) and ['θərməsɪjən] (from thermos + -ian)) showed, according to the primary stress rule[1] tested in this study, an incorrect placement of primary stress. It may seem presumptuous to categorize responses being considered as relevant to certain rules, according to another rule which is itself being investigated in this study. However, those vowels which are predicted to undergo a phonetic change (according to all of the rules predicting a phonetic change in a vowel, with the exception of the rule V → [+tense] / __ V) happen to be those vowels which are located in the syllable immediately preceding the suffix. These vowels are therefore those which, according to the primary stress rule, receive primary stress when derived with certain suffixes[2]. In order to avoid unfairly counting responses in one or another vowel category when the vowel was not predicted to change unless it was stressed, the primary stress rule was given the benefit of the doubt, that is, responses with an unpredicted stress

[1]The formalized primary stress rule considered here is V → [1 stress] / __ C + affix.

[2]Chomsky and Halle list some affixes which are exceptions to this rule, one of which, -ism, is combined with certain stem-words in this study.

pattern were placed in a separate category. The separate category was felt to be warranted by the fact that Groups A and B each produced the predicted stress pattern in about 67% of their responses (to the long list of stimuli) which were hypothetically eligible for a change in the stress location.

The 'Wrong Stress' category was not mutually exclusive from each of the other categories. For instance, a response could contain a predicted (e.g. tense or lax) vowel yet show 'incorrect' stress location: for example, [ bi'dɑ͡ɪən] from bed + -ian and ['sæfɪɹətɔɹi] (from sapphire + -atory). The response [məθə'nɪnəti ] (from methane + -ity) could be counted in either of the 'Addition' or the 'Wrong Stress' categories while [bɑ͡ɪ'ɑlədʒ͡ɪk] (from biologist + -ic) could be counted in either of the 'Deletion' or 'Wrong Stress' categories. Such responses were counted in the 'Addition' and 'Deletion' categories respectively. In the 'Addition' case, one could not tell if the subject considered the stem of the derivation to be [sʌstən-] or [sʌstanɪn-] and therefore could not say whether the stress rule was correctly applied. In the 'Deletion' case, on the other hand, one could not know if the location of stress was assigned before or after the syllable was dropped and therefore one could not fairly judge the final stress assignment.

The responses to the stem-suffix sets which were considered for the primary stress rule were counted in categories with stress falling on different specified syllables. For example, ['θɚməsijən] (from thermos + -ian) was counted in the category where stress fell upon the initial syllable. Two responses with

the form ['tɑmɪjəsən] (derived from Thomas + -ian) necessitated the adoption of the category called 'Metathesis' since one could not tell whether the assignment of stress occurred before or after the transposition of segments in the stem and the suffix. (This category was used for data relevant to the primary stress rule as well as to the rule which tensed vowels preceding CiV̆.) Two responses of the form ['æd'vəsɪzm̩] (from adverse + -ism) required the adoption of a category called 'Equal Stress' in which two syllables contained heavy stress of equal strength.

Each group's responses to the stimuli which were relevant to a given rule were first categorized in matrices which considered the responses to one stem-suffix set. The columns of the matrix contained the different categories of responses to the set while the rows referred to the responses of the subjects in the group. These matrices served as the basis for all later measures of the occurrence of sound patterns.

The cumulative number of predicted responses in each group

Each subject's predicted responses for a given rule were counted for the long list of stimuli. 'Predicted' responses or sound patterns shall henceforth refer to those sound patterns predicted by the phonological rules under discussion which were hypothesized by Chomsky and Halle (1968). For each rule, the subjects in each group were assigned an average rank according to the number of predicted responses they produced. However, since a different number of stem-suffix sets was considered for each rule, finding the unweighted mean rank of the seven (average) ranks injustly gave equal weight to each

(average) rank. Therefore each rule was assigned a weight equal to the number of stem-suffix sets considered relevant to the rule. The number of sets considered for the /s/-voicing rule varied for each subject according to the number of eligible responses (containing [ V + s + V ]) he produced. The /s/-voicing rule was assigned a weight of 3 in both groups even though the average number of eligible responses per subject was 2.5 in Group A and 2.3 in Group B. A weighted mean rank was found for each subject[2].

The unweighted and weighted mean ranks of each subject were then compared. The two ranks never differed by more than 5.0 mean ranks in either group of subjects. Often the two mean ranks of a subject were equal. Another informal check on the two mean ranks was made by grouping subjects according to a combination of the 'statuses' (i.e., low, medium, high) of the average ranks which subjects achieved for each rule. When the unweighted and weighted mean ranks were than considered for these grouped subjects, both ranks corresponded fairly well to the status of the grouping, with only a few cases where,

---

[2]The formula used for the weighted mean rank was:

$$\text{weighted mean rank} = \frac{\sum_{i=1}^{i=7} w_i \cdot r_i}{\sum_{i=1}^{i=7} w_i}$$

where $w_i$ refers to the weight of a given rule and $r_i$ refers to a subject's average rank for that rule.

according to either mean rank, the subject should have been placed in another grouping. Therefore one might conclude that the fact that a varying number of stem-suffix sets were considered for the different rules did not greatly influence subjects' mean ranks in the group's productivity of all of the predicted sound patterns.

The difference in the two groups' overall productivity of predicted sound patterns can be shown by comparing the cumulative histogram of each group's total number of predicted sound patterns. The responses to the short list of stimuli were used in order to take the opportunity to attribute the difference in the two groups' cumulative number of predicted responses to the different experimental conditions. The subjects were ordered according to their mean unweighted ranks. The subjects with the lowest and the highest ranks in Group A were eliminated in order to keep the number of subjects in the two groups equal. The results can be seen in Figure 2. The subjects of each group are plotted along the abscissa with the subject with the 'worst' rank (i.e. with the smallest number of predicted responses) at the far left and the subject with the 'best' rank (i.e. the greatest number of predicted responses) at the far right. The cumulative number of predicted responses is plotted along the ordinate. There clearly is a difference in each group's production of predicted responses. In order to attribute the difference to the difference experimental conditions, one needs to know that the two groups of subjects were equally

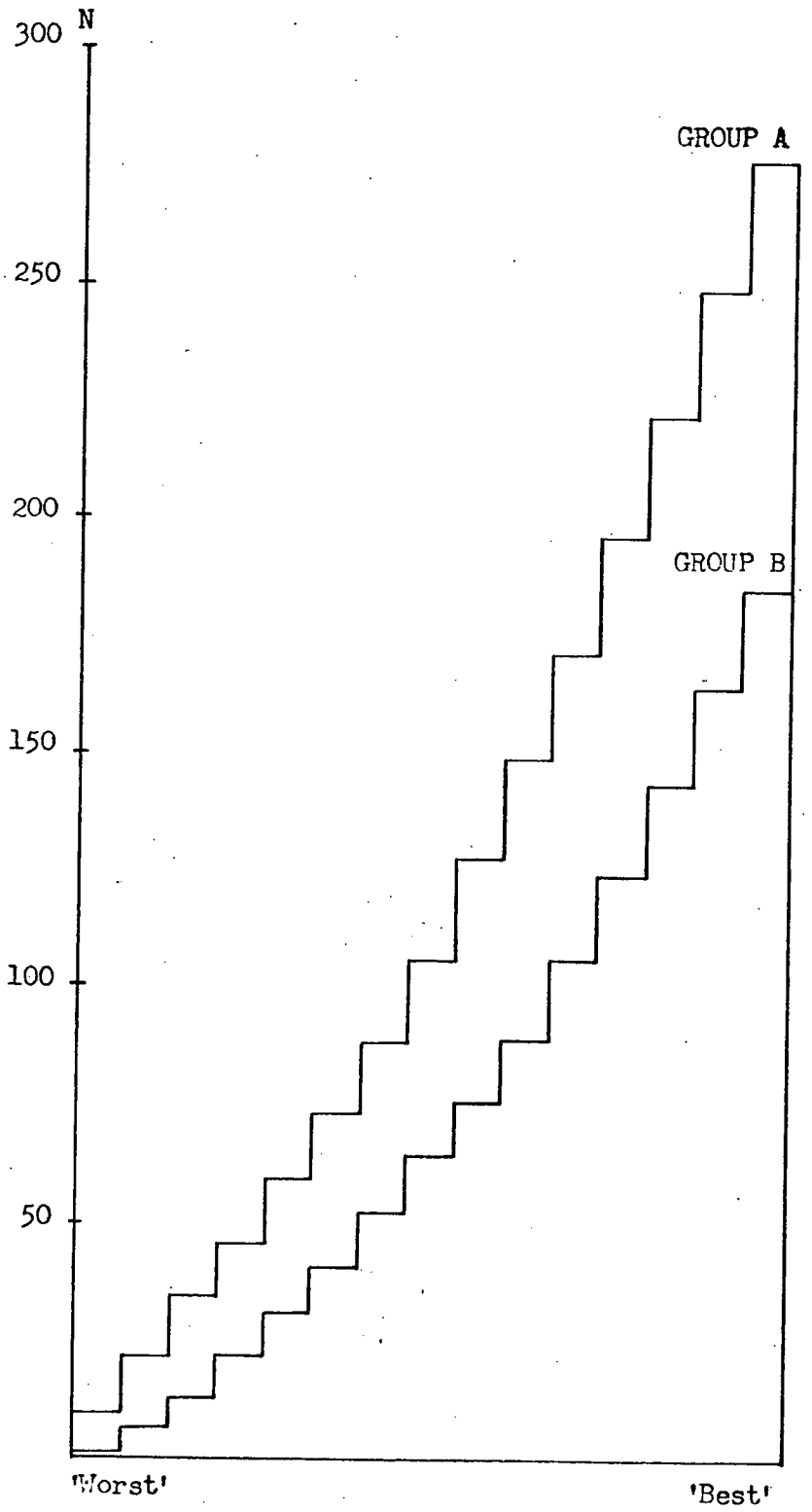



Figure 2. Cumulative number (N) of predicted responses (short list) for each group.

'adept' in the practical matter of producing novel derivations. Such an assumption could be verified statistically if several groups of subjects participated (in different orders) in the two experiments on two different occasions. This type of check was beyond the scope of this study. However, a less formal check on the difference in the word-derivation skill of each group can be made by comparing each group's cumulative histogram of percentage of predicted changes versus subject rank. The data from the short list of stem-suffix sets were used. (The data from the long list of stimuli produced curves which are very similar to those shown for the short list.) The subjects of each group were ranked according to the total number of predicted responses they made for the six rules under consideration in the short list. Once again, the subjects in Group A with the highest and lowest ranks were eliminated. For each group, the cumulative percentage of predicted responses were plotted on the ordinate. The subjects were plotted in rank order on the abscissa (with the worst ranking subject on the far left and the best ranking subject on the far right). The curves are shown in Figure 3.

The curves fit very closely, suggesting that the total number of predicted responses was distributed similarly among the members in each group. This distribution which is seen in the two groups is likely what would be expected if the groups represented a normal sampling of the population--that is, a few subjects produced a large number of predicted responses, a few produced an intermediate number and a few produced a small

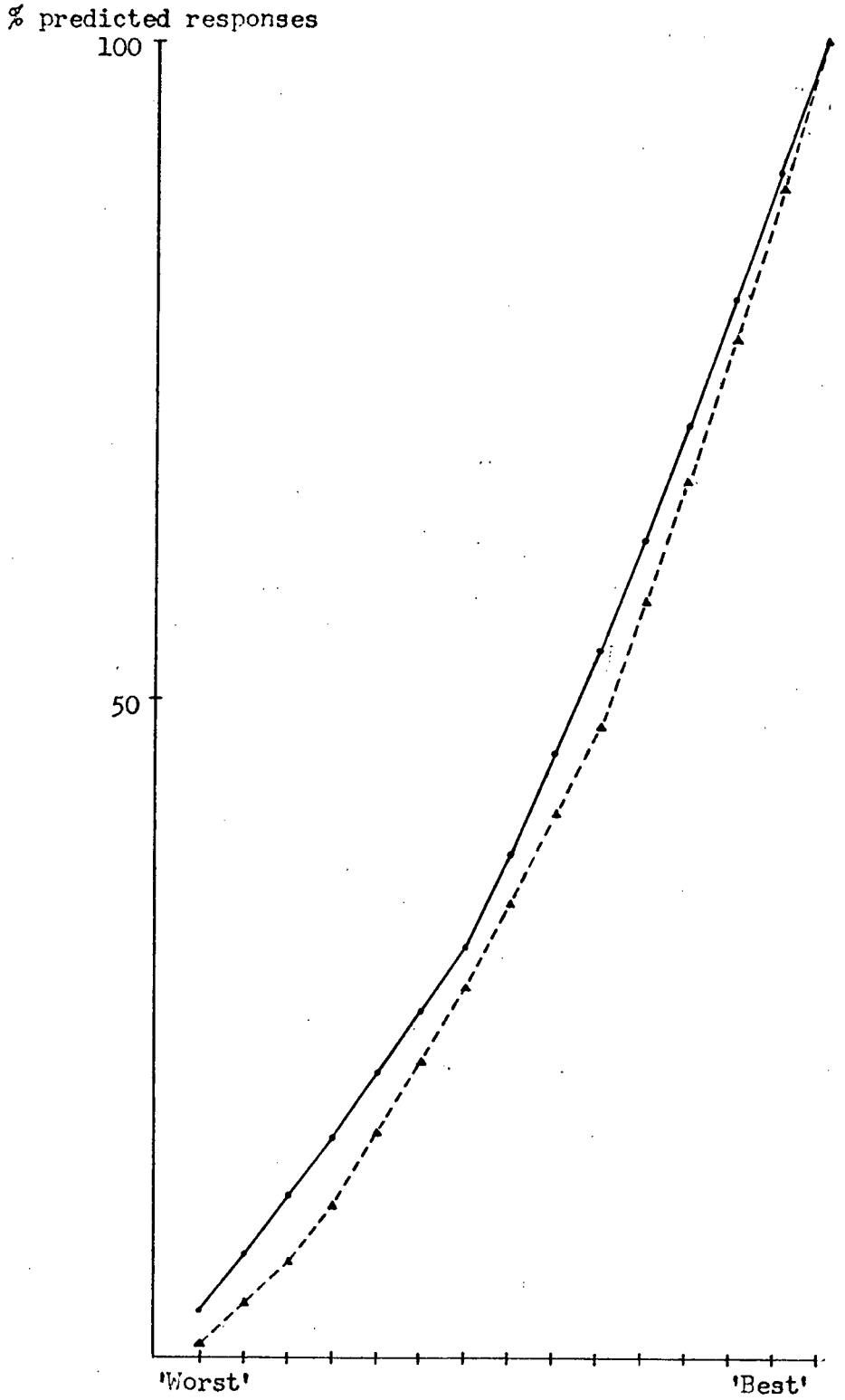



Figure 3. Cumulative percentage of predicted responses (short list) for Group A (——) and Group B (---).

number.

One could likely conclude that the two groups were similar in their ability to perform the experimental task. In addition, each group of subjects was likely representative of the population of young university students from which they were drawn. So, in turn, one could probably conclude that the difference in the cumulative number of predicted responses in the two groups was a result of the different experimental conditions: namely, the presence of absence of example derivations showing relevant sound patterns.

Whereas Steinberg and Krohn (1975) found that a small number of subjects were responsible for the production of predicted responses, in the present study 13 out of 15 subjects in each group were responsible for approximately 80% of their group's predicted responses. The data used to construct Figure 3 showed that 13 subjects in Group A produced 80.4% of that group's total number (short list) of predicted responses while 13 subjects in Group B produced 77.3% of that group's predicted responses.

The various sound patterns and their frequency of occurrence in each group

The responses counted in each category for a given stem-suffix set were totalled and entered in a matrix which contained the categorized responses to all of the stimuli relevent to a given rule. The columns of the matrix referred to the categories of responses and the rows referred to the stem-suffix sets. The percentage of occurrence of each given category of response with respect to the total number of responses

(considered for a rule) was found. The results, seen in Tables III through IX, show for each group of subjects the percentage of occurrence of given categories of responses for each rule. Results for the short list of stimuli are shown on the left side of each column while the results for the long list of stimuli are shown on the right side of each column in the tables. The data which will be discussed in the text below with reference to each rule will be that derived from the short list.

The vowel laxing rule (henceforth $VLR_1$): V → [-tense] CV̆CV

As can be seen in Table III, the sound pattern predicted by $VLR_1$ (e.g. ['spæsəti] from space + -ity) occurred in a greater percentage of responses in Group A (38.1%) than in the responses of Group B (20.4%). The 'tense vowel' sound pattern (e.g. ['bɹa͡ɪbəti] from bribe + -ity and [sup'ɹimə͡fa͡ɪ] from supreme + -ify) occurred most often in the total responses of each of the two groups. However, this pattern occurred more often in the responses of Group B (74.9%) than in those of Group A (57.2%). Both groups produced a few responses categorized as 'Additions' (e.g. [ʃli'sɪdəti] from fleece + -ity). The segments which were added in the new derivations were [-ɪn-], [-ɪd-], [-e͡ɪ-] and [-æs-]. They always appeared between the stem and the suffix -ity. The occurrence of the responses containing [-ɪn-], [-ɪd-] and [-e͡ɪ-] might in part be accounted for by the number of existing English derivations containing one of these syllables and the suffix -ity (e.g. validity, timidity;

| | TOTAL RESPONSES | | LAX VOWEL | | TENSE VOWEL | | ADDITION | | DELETION | | SUBSTITUTION | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | GROUP A | | | | | | |
| no. | 283 | 300 | 108 | 115 | 162 | 170 | 6 | 6 | 5 | 5 | 2 | 4 |
| % | 100 | 100 | 38.1 | 38.3 | 57.2 | 56.6 | 2.1 | 2 | 1.8 | 1.7 | 0.7 | 1.3 |
| | | | | | | GROUP B | | | | | | |
| no. | 255 | 270 | 52 | 57 | 191 | 201 | 6 | 6 | 3 | 3 | 2 | 2 |
| % | 100 | 100 | 20.4 | 21.1 | 74.9 | 74.4 | 2.3 | 2.2 | 1.2 | 1.1 | 0.7 | 0.7 |

Table III. V → [-tense]/__C$\breve{V}$CV. The number of responses in each category and the percentage of occurrence of each category for the lists of stimuli relevant to the rule mentioned. The data for the short list of stimuli are on the left side of each column while the data for the long list are on the right side.

vicinity, virginity, femininity; omneity, spontaneity, deity and homogeneity). Subjects may have analogized their responses after the surface structure of words such as these. Or perhaps they used an independent phonological rule, by which those existing words were partly generated, to produce such responses.

The insertion of the four syllables initially listed above might be due to a strategy which called for segment or syllable reduplication. This follows from observing, first, that [-In-] was inserted after a sequence of [-V+n-] in [məθə'nɪnəti ], [sʌstə'nɪnəti ] and [ bɪtwi'nɪnəti ]. Secondly, [-ej-] was inserted after a syllable containing a similarly tense diphthong in [tɹaɪ'teɪəti ]. Thirdly, [-æ+s-] occurred after the sequence of [-ej+s-] in [spei'sæsəti ].

Both groups also produced a few responses in the 'Deletion' category (e.g. ['sæfətɔɹi ] from sapphire + -atory). The deletion of a syllable might be due to subjects' conscious or unconscious desire to make one or more phonetic changes in the second syllable of the stem but being unable to make a change which would be 'satisfactory' to them. They might therefore drop the 'awkward' syllable. On the other hand, they may have had a 'distaste' for words containing more than four syllables.

Group A produced responses in the 'Substitution' category. In this case, two subjects responded with [ bi'twɪksədi] for between + -ity. This response may have occurred because subjects wished to lax the vowel preceding the suffix. On the other hand, they may have produced [- ks-] in the new derivation by considering the surface structure of betwixt, a word that is

close in sound structure and meaning to the stem between. In addition, the two stems occur together in the phrase, "betwixt and between," and so might have some collocating association.

One subject in Group B placed stress on a wrong syllable in just one response.

The primary stress rule (henceforth PSR):
V → [1 stress]/__C + affix

The sound pattern predicted by the PSR (e.g. [aɪˈoʊdɪnəti] from iodine + -ity)occurred in an almost equal percentage of the responses of both groups: 64% in Group A and 61.4% in Group B (cf. Table IV). Furthermore, this sound pattern was the most frequent one in both groups' responses. The response category which ranked second in the number of responses was the 'Deletion' category (e.g. [baɪˈalədʒɪk] from biologist + -ic). The percentages of occurrence of 'Deletion' responses were also quite close for the two groups, being 22.8% in Group A and 19.2% in Group B. One might mention and try to explain the occurrence of other sound patterns produced for the stimuli under consideration for the PSR. However, such explanation, as in the cases of the explanations for the unpredicted categories of responses for $VLR_1$, is not exhaustive and would only be speculative.

The vowel tensing rule (henceforth $VTR_1$): V → [+ tense]/___V

Table V shows that the 'tense vowel' sound pattern (e.g. [ziˈbɹeɪək] from zebra + -ic) predicted by this rule occurred most often (41.1%) in the responses of Group A while the 'Deletion' category of response (e.g. [ˈzibɹɪk]) occurred most often in Group B (43.3%). A clue to the common occurrence of

| | TOTAL RESPONSES | | ŚSSS+SS | | SŚSS+SS | | SSŚS+SS | | SSSŚ+SS | | SSSS+ŚS | | DELETION | | ADDITION | | SUBSTITUTION | | META-THESIS | | EQUAL STRESS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | GROUP A | | | | | | | | | | | |
| no. | 153 | 221 | 4 | 13 | 1 | 1 | 0 | 0 | 98 | 149 | 5 | 5 | 35 | 40 | 2 | 3 | 8 | 9 | 0 | 0 | 0 | 1 |
| % | 100 | 100 | 2.6 | 6.3 | 0.7 | 0.4 | 0 | 0 | 64 | 67.4 | 3.3 | 2.2 | 22.8 | 18 | 1.3 | 1.3 | 5.2 | 4.1 | 0 | 0 | 0 | 0.7 |
| | | | | | | | | | | | GROUP B | | | | | | | | | | | |
| no. | 135 | 195 | 10 | 18 | 1 | 1 | 1 | 1 | 83 | 132 | 8 | 8 | 26 | 30 | 1 | 1 | 4 | 4 | 1 | 1 | 0 | 1 |
| % | 100 | 100 | 7.4 | 9.2 | 0.7 | 0.5 | 0.7 | 0.5 | 61.4 | 67.6 | 5.9 | 4.1 | 19.2 | 15.4 | 0.7 | 0.5 | 2.9 | 2 | 0.7 | 0.5 | 0 | 0.5 |

TABLE IV. V → [1 stress]/__C + affix. (See Table III's legend.) The 'S' in SSSS+SS in the categories above stands for 'syllable.'

| | TOTAL RESPONSES | | TENSE VOWEL | | LAX VOWEL | | ADDITION | | DELETION | | SUBSTI-TUTION | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | GROUP A | | | | | | |
| no. | 34 | 51 | 14 | 16 | 0 | 0 | 1 | 6 | 12 | 22 | 7 | 7 |
| % | 100 | 100 | 41.1 | 31.1 | 0 | 0 | 2.9 | 11.8 | 35.2 | 43.1 | 20.6 | 13.7 |
| | | | | | | GROUP B | | | | | | |
| no. | 30 | 45 | 12 | 13 | 1 | 2 | 1 | 9 | 13 | 18 | 3 | 3 |
| % | 100 | 100 | 40 | 28.9 | 3.3 | 4.4 | 3.3 | 20 | 43.3 | 40 | 10 | 6.6 |

TABLE V. V → [+ tense]/__V. (See Table III's legend.)

this sound pattern in both groups might be the fact that (as was mentioned by Ohala) there are several words existing in English that demonstrate this sound pattern, e.g. saliva/salivic, guerrilla/guerrillic, cholera/choleric, oedema/oedemic and vulva/vulvic. The percentage of occurrence of the 'tense vowel' response in Group B (40%) was, however, almost as great as its percentage in Group A. In addition, the predicted 'tense vowel' responses in Group B occurred almost as often as the most common response (the 'Deletion' category of response, 43.3%) in that group. The 'Substitution' responses (20.6% in Group A and 10% in Group B) all had the form [bu'dɪstək] (from Buddha + -ic). One possible explanation of this sound pattern is that subjects were considering the stem-word to be Buddhist. This may not be surprising, as all three English derivations of Buddha (e.g. Buddhist, Buddhistic, Buddhistical) contain the stem Buddhist. The 'Addition' responses involved the insertion of [ʔ] or [kt] between the stem-word Buddha and the suffix -ic. This strategy was probably used to separate the sequence of the two vowels resulting from the combination of Buddha + -ic. One might also note that there are a number of English derivations ending in [-æktɪk] (e.g. didactic, prophylactic, climactic) whose sound patterns the subjects may have recalled and extended by some means.

The vowel tensing rule (henceforth $VTR_2$):

$$V \rightarrow [+\text{ tense}] / \_\_ C_i\check{V}$$
$$[-\text{high}]$$

Although the predicted tense vowel sound pattern (e.g. [θɚ'mowsɪjən] from thermos + -ian) occurred in a higher

percentage of the responses of Group A (23.5%) than in Group B (8.3%), it was not, by far, the most common category of response in either group (cf. Table VI). The 'tense vowel' response had the second highest percentage of occurrence in Group A and only the third highest percentage in Group B. A sound pattern containing a lax vowel (e.g. [θɚ'məsijən] from thermos + -ian and [tə'mɪsijən] from Thomas + -ian) occurred most often in each groups' responses, in fact three to four times as often as the second ranking pattern. The 'lax vowel' category of response occurred almost as often in each group (69.9% in Group A and 70.8% in Group B).

A speculative explanation of the 'Deletion' responses (e.g. ['krɪsmijən] from Christmas + -ian) might be that subjects were dropping what they felt were 'awkward' syllables. In doing so, they could eliminate the 'predicted' need (a) for a change in the location of primary stress in all of the two-syllable stem-words and (b) for a change to a tense vowel in Christmas + -ian and thermos + -ian (due to the double consonant context).

In Group B, responses counted in the 'Wrong Stress' category (15%) numbered almost twice as many as those in the predicted 'tense vowel' category. Six of eight 'Wrong Stress' responses made by Group A and six of eighteen made by Group B involved primary stress being located on the first vowel of the suffix -ian (e.g. [təmə'sijən], [fɹə'dijən] and [ɹowbə'tijən]). One speculative explanation is that subjects may have been extending, by some means, the sound pattern seen in English derivations

| | TOTAL RESPONSES | TENSE VOWEL | LAX VOWEL | ADDITION | DELETION | META-THESIS | WRONG STRESS |
|---|---|---|---|---|---|---|---|
| | | | | GROUP A | | | |
| no. | 136 | 32 | 87 | 1 | 8 | 0 | 8 |
| % | 100 | 23.5 | 63.9 | 0.7 | 5.8 | 0 | 5.8 |
| | | | | GROUP B | | | |
| no. | 120 | 10 | 85 | 0 | 5 | 2 | 18 |
| % | 100 | 8.3 | 70.8 | 0 | 4.1 | 1.6 | 15 |

Table VI. V → [+ tense] / __ CiV̆. The number of [-high] responses in each category and the percentage of occurrence of each category for the single list of stimuli relevant to the rule mentioned.

such as plebian, European and Jacobean, whose stems end in a consonant. This sound pattern also occurs in many existing derivations whose stems end in a vowel (such as Chaldean and Galilean).

The velar softening rule (henceforth VSOR): /$k^d$/ → s/__ I

The predicted 'velar softening' sound pattern (e.g. [dow'mɛstəsɪzm] from domestic + -ism) showed the highest percentage of occurrence in Group A's responses (52.9%) but only the second highest percentage of occurrence in Group B's responses (42.2%) (cf. Table VII). Conversely, the category of response which retained the [k] of the stem (e.g. [dow'mɛstəkɪzm] from domestic + -ism) ranked second in the percentage of occurrence (33.3%) in Group A's responses but first in percentage of occurrence (48.8%) in the responses of Group B.

The underlying vowel in the second syllable of obtain and pertain

Chomsky and Halle state that ordinary underlying vowels undergo the vowel shift rule (VSR) only if the vowel is [+ tense]. They propose that some vowels (such as those in the underlying forms of retain and detain) remain eligible for the application of the VSR even if the vowel has been laxed in the environment directly preceding a double consonant. Ohala believes that if

> the underlying vowel in the second syllable of obtain and pertain are marked in eligibility for vowel shift on laxed vowels, then one would expect obtain + ion to possibly be rendered [ʌbtɛ́nšən], but obtain + atory should then likewise be rendered [ʌbtɛ́nətɔɹi] (1973, p.44).

| | TOTAL RESPONSES | | /s/ | | [k] | | ADDITION | | DELETION | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | GROUP A | | | | | |
| no. | 51 | 68 | 27 | 30 | 17 | 20 | 1 | 2 | 6 | 16 |
| % | 100 | 100 | 52.9 | 44.1 | 33.3 | 29.4 | 1.9 | 2.9 | 11.7 | 23.5 |
| | | | | | GROUP B | | | | | |
| no. | 45 | 60 | 19 | 20 | 22 | 28 | 1 | 1 | 3 | 11 |
| % | 100 | 100 | 42.2 | 33.3 | 48.8 | 47 | 2.2 | 1.7 | 6.6 | 18.3 |

TABLE VII. /$k^d$/ → s/__I. (See Table III's legend.)

If the underlying vowel is not marked to undergo the VSR after being laxed, then [æ] might be expected to appear in the second syllable of the surface structure of obtain + -ion and obtain + -atory. The suffix -atory is predicted to cause the vowel preceding it to become lax according to $VLR_1$. The suffix -ion is predicted to lax the vowel preceding it after the epenthesis of a dental consonant (between a stem such as retain, ending in a dental nasal, and a suffix) which creates a consonant cluster. The rule which predicts the lax vowel is V → [-tense]/__CC, henceforth called $VLR_2$. If the subjects were unable to, or did not wish to, extend the sound pattern in which the vowel is laxed in the environments preceding CV̆.CV or CC, then presumably the tense vowel [e͡ɪ] in the stem-word would be retained in the subjects' responses. The appearance of [e͡ɪ] in Group B's responses to stimuli combining with the suffix -ion could lead to two conclusions. First, the subjects may not have laxed the vowel because they could not or did not extend the sound pattern wherein a vowel is laxed before a consonant cluster. On the other hand, the subjects may not have extended a consonant cluster sound pattern in the derivation since they were given no leading example derivations and since the written form of the suffix which they were given did not include a "t".

For the present, discussion of the responses to pertain and obtain will be limited to the occurrence of lax or tense vowels which would reflect on the predictions for a 'lax vowel' sound pattern as well as the opportunity to insert a [t] or

a [ʃ] before the suffix -ion. It will be assumed that subjects can extend the double consonant sound pattern. In other words, responses without the double consonant pattern will not be excluded from the analysis for tense or lax vowels. Comments on the underlying vowel in question in these stem-words will be discussed in a later section.

Table VIII shows that the predicted 'lax vowel' sound pattern (e.g. [pə'tɛnətɔɹi] from pertain + -atory and [ʌb'tɛnʃən] from obtain + -ion) occurred most often (70.5%) in the responses of Group A. One 'lax vowel' response in Group A, [pə'tɛnjən], occurred in the environment preceding CiV̆ which is predicted by Chomsky and Halle to elicit a tense vowel. Twenty percent of the responses of Group B were counted in the 'lax vowel' category. All of the lax vowel responses made by Group B to the stimuli combining with the suffix -ion occurred directly preceding a consonant cluster environment [__nʃən] which is hypothesized to elicit a lax vowel.

The 'tense vowel' response (e.g. [ʌb'teɪ̯njən] and [pə'teɪ̯nətɔɹi]) occurred most often (76.6%) in the responses of Group B. Some of Group A's responses (27.9%) contained tense vowels. Both groups of subjects also produced a few 'Deletion' responses such as ['pətətɔɹi].

The /s/-voicing rule (henceforth /s/VR):

$$/s/ \rightarrow [+\text{voice}] \,/\, \underset{[+\text{tense}]}{V} \_\_ V$$

The sound pattern predicted by this rule cannot strictly be included in the analysis which considers the effects of the different experimental conditions. None of the stimuli

| | TOTAL RESPONSES | LAX VOWEL | TENSE VOWEL | DELETION |
|---|---|---|---|---|
| | | GROUP A | | |
| no. | 68 | 48 | 19 | 1 |
| % | 100 | 70.5 | 27.9 | 1.5 |
| | | GROUP B | | |
| no. | 60 | 12 | 46 | 2 |
| % | 100 | 20 | 76.6 | 3.3 |

Table VIII. $V \rightarrow [-\text{ tense}] / \_\_ \left\{ \begin{matrix} CVCV \\ CC \end{matrix} \right\}$. Ability to extend the double consonant sound pattern in the context __ -ion was assumed. (See Table VI's legend.)

considered for this rule were preceded in Main Experiment A by example derivations demonstrating the 'voiced /s/' sound pattern.

Some idea of the subjects' ability to extend the sound pattern to novel utterances might be gained from considering the numbers in Table IX. In new derivations in which [s] was preceded by a tense vowel, Group A produced the predicted sound pattern (e.g. [θɚ'mowsijən] from thermos + -ian) in 7% of their responses while Group B failed to produce the sound pattern. One might speculate that the reason that even a few of Group A's eligible responses exhibited the 'voiced-/s/' sound pattern was that Group A had been exposed to example derivations showing phonetic change and were, on the whole, producing more responses involving phonetic changes. Perhaps in a limited sense, therefore, the responses eligible to be analyzed for the occurrence of the voiced /s/ can be included in the comparison of each group's responses.

Further evidence for the sound pattern occurred in responses to prose + -ify wherein a few subjects in each group laxed the vowel in the stem of the derivation. One of these four 'lax vowel' responses made by Group A and one in three produced by Group B devoiced the [z] and produced [pɹɑsəfɑɪ].

Summary

The sound patterns predicted by each of the rules occurred more often in the responses of Group A whose subjects were exposed to example derivations which demonstrated all but one

| | TOTAL ELIGIBLE RESPONSES | [-VOICE] | [+VOICE] |
|---|---|---|---|
| | | GROUP A | |
| no. | 42 | 39 | 3 |
| % | 100 | 93 | 7 |
| | | GROUP B | |
| no. | 31 | 31 | 0 |
| % | 100 | 100 | 0 |

Table IX. s → [+voice]/ V ___ V.
[+tense]
(See Table VI's legend.)

of the predicted sound patterns. Since the two groups were equally adept in handling the experimental task, the greater percentage of occurrence of predicted sound patterns in the responses of Group A can likely be attributed to that Group's exposure to leading example derivations. This hypothesis will be further investigated later in relation to the means by which subjects extended given sound patterns.

If the results of the /s/VR are a reliable indication for each group's tendency to make predicted phonetic changes, the trend towards or away from making phonetic changes (predicted by some rules) might also bear an influence on the occurrence of predicted sound patterns for another rule, independent from the others. One might also consider the relation of the total number of phonetic changes (whether predicted or unpredicted) to the number of predicted changes made. It could be that as the number of responses involving a phonetic change increases or decreases, so does the number of predicted phonetic changes. Finding the total number of responses involving a phonetic change in the stem of the derivation would be very difficult. It would be desirable to categorize responses according to the number of phonetic changes made. For example, should not [mɛθəˈnɪnəti] (from methane + -ity) count as having been submitted to more phonetic changes than [məˈθéɪnətɪ]? It is difficult to tell which phonetic changes are discrete from one another and so assign to them a unit value. With stimuli such as sustain + -ity and pertain + -ion it would also be impossible to tell whether a subject left the primary stress on the

hypothetically appropriate syllable because he was oblivious to the stress pattern or because he felt that the stress should remain in the same location in the new derivation.

Despite the apparent trend for predicted sound patterns to occur more often when subjects are exposed to example derivations illustrating the sound patterns, one cannot specify whether the sound patterns are produced through the use of analogical phonological rules, the independent phonological rules under investigation or through some other means.

## 4.4 Attempts to determine the means by which a sound pattern was produced

### 4.4.1 The influence of example derivations on 'similar' stem-suffix sets

A further check to see if the more frequent occurrence of predicted responses in Group A was due to the presence of relevant example derivations was attempted by considering each group's chi-square value for the number of predicted and unpredicted responses for certain stem-suffix sets. The stem-word of each set was similar to the stem of the preceding example (with the relevant sound pattern) in that it contained the same strategic vowel or consonant as that in the example. It was expected that the difference in the number of predicted responses might also have some implications for the subjects of Group A having used a given example derivation as a surface structure form on which to model their responses to a given stem-suffix set. That is, if the occurrence of the predicted sound pattern was significantly greater in Group A's responses to a given stem-suffix set, there would be a possibility that the

difference was caused by subjects' exposure to an example derivation illustrating the predicted sound pattern with the vowel or consonant which was common to the example and the stimulus.

The chi-square measure was taken on several stem-suffix sets by entering the number of predicted responses and the number of remaining responses made by one group to a given stem-suffix set in one row of a 2 x 2 matrix. The matrix is exemplified in Table X using the data from the two groups' responses to slave + -ify.

| | No. Predicted Responses | No. Remaining Responses | Tot. |
|---|---|---|---|
| Group A | 8 | 9 | 17 |
| Group B | 4 | 11 | 15 |
| Tot. | 12 | 20 | 32 |

Table X. The matrix used for the chi-square measure of both groups' responses to slave + -ify.

Due to the small number of each group's responses to a stem-suffix set many of the matrices contained entries with a small number. The questions as to how small expected frequencies could be before using certain chi-square correction formulae and before abandoning the use of the chi-square measure altogether were answered differently in different sources. Therefore two correction formulae were used, in addition to the formula for the tetrachoric correlation coefficient (see fn. 3, p.126), to

determine whether the values obtained for a given set consistently remained statistically significant with respect to confidence level. The sets whose values did remain consistent are mentioned below. The level of significance mentioned for each set is the most conservative one obtained for it.

There were statistically significant differences in the number of predicted responses for only a few sets. Only one of those preceded by mammal/mammalian (Thomas + -ian ($p < .05$)) showed a significant difference. A possible explanation for obtain + -ion ($p << .001$), pertain + -ion ($p < .05$), obtain + -atory ($p < .05$) and pertain + -atory ($p < .005$) all having reached significant levels is, first, that the stem-words of the stimuli and those of the examples ended in [-ein]; secondly, they were composed of stems and prefixes. Likewise, a possible explanation for methane + -ity and sustain + -ity, which were preceded by sane/sanity, having reached levels of significance (both at $p < .005$) is that their stems contained the same V + C sequence as the example's stem-word.

Predicted responses occurred significantly more often in Group A's responses to the stem-suffix sets which were similar to their preceding example derivations. This conclusion might at first glance appear as evidence that the subjects of Group A were using the surface structure of the example derivations (particularly of those whose stems contained [-ein]) as models on which to base analogous responses. This idea is reinforced by the fact that many of the stem-suffix sets whose numbers of

predicted responses did not reach significant levels had only a vowel in common with their preceding example derivation's surface structures. However, the possibility still remains that the example derivations served to illustrate the independent phonological rule formalized by Chomsky and Halle. For instance, sane/sanity could have illustrated $VLR_1$ more effectively for pairs such as methane + -ity and sustain + -ity than for pairs whose stem-words did not contain [ẽɪ].

4.4.2 The influence of a previous response on a later response

J. Ohala sought to check whether analogical phonological rules were the means by which a sound pattern was productive by considering the correlation of responses to successive or nearby stem-suffix sets. The statistical measure that he used and that was used in this study is the tetrachoric correlation coefficient[3]. In the present study, the responses considered were those that were counted in two sound patterns that (a) were common to both groups' responses to the two stem-suffix sets and (b) most frequently occurred in both groups' responses to the two sets. For each group, the data for a pair of stem-suffix sets were arranged in a matrix such as the following (which uses data from methane + -ity and sustain + -ity):

---

[3] The tetrachoric correlation coefficient measures the relatedness of events tallied in a 2 x 2 matrix, using the formula:

$$r_t = \sin \frac{\pi}{2} \cdot \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}}$$

The distribution of $r_t$ was obtained by generating (2500 times) random 2 x 2 matrices, for each value of N (sum of the entries) from 10 to 17 and from 30 to 32. The levels of confidence were inferred from these distributions.

Group A

| | (methane + -ity) tense vowel | (methane + -ity) lax vowel |
|---|---|---|
| tense vowel (sustain + -ity) | a = 1 | b = 1 |
| lax vowel | c = 3 | d = 10 |

Table XI. Example of a matrix arrangement for an individual group's responses to a pair of stem-suffix sets.

The pairs of stem-suffix sets whose responses showed a significant positive correlation involve significantly more pairs of responses which are either both sound pattern 'x' or both sound pattern 'y'. Those pairs whose responses showed a significantly negative correlation involved significantly more pairs of responses which were comprised of two different sound patterns. When the correlation coefficient was non-significant, the responses to the pairs were usually comprised of more equal numbers of 'both the same' responses and 'different' responses.

It is possible, in the cases of the stem-suffix sets (particularly successive sets) whose responses showed a significant positive correlation, that the form of the first response influenced the form of the following response. This could support both the hypothesis of an independent phonological rule or that of a model whose surface structure was the object of analogy. Alternatively, the influence of the earlier response

could have occurred in still some other way. However, a high statistical correlation does not imply a causal relation between the two events considered. Therefore, checking the correlation of similar responses in pairs of stem-suffix sets cannot even prove a cause and effect relationship between the sets, let alone infer the means by which the relationship took place.

The pairs of stem-suffix sets, their positions (relative to each other) in the experimental task, the sound pattern categories (i.e., parameters) considered and the results are listed in Table XII. The results indicate whether the values for the tetrachoric correlation coefficient and for the chi-square measure both remained at a confidence level better than $p = .05$. (The chi-square measure's values were taken into consideration as well since the tetrachoric correlation coefficient reaches its maximum or minimum values of 1.0 or -1.0 when one entry in a matrix is 0.) No conclusions can be drawn in these data concerning the relation of the levels of significance of the pairs of stem-suffix sets to their location (relative to each other) and to their common characteristics.

### 4.4.3 The influence of different example derivations on responses involving one stem-word

J. Ohala (1973) exposed his subjects to two example derivations whose sound patterns differed in such a way as to suggest that the VSR had been used in the derivation of one example (detain/detention) but not in the other (explain/explanatory). Several of his subjects' responses to different

| | PAIRS OF STEM-SUFFIX SETS | PARAMETERS | GROUP | BOTH $\chi^2$ AND $r_t$ VALUES REACH .05 LEVEL OF SIGNIFICANCE |
|---|---|---|---|---|
| SUCCESSIVE SETS | (identical vowel) | | | |
| | define + -ity/iodine + -ity | tense vowel vs lax vowel | A | no |
| | " | " | B | no |
| | methane + -ity/sustain + -ity | " | A | no |
| | " | " | B | no |
| | trade + -ity/space + -ity | " | A | no |
| | " | " | B | yes |
| | sublime + -ify/pride + -ify | " | A | no |
| | " | " | B | no |
| | (different vowel) | | | |
| | iodine + -ity/methane + -ity | " | A | no |
| | " | " | B | no |
| NEARBY SETS SETS IN DIFFERENT STIMULI GROUPS | (identical vowel or consonant) | | | |
| | bed + -ian/Fred + -ian | " | A | yes |
| | " | " | B | no |
| | Christmas + -ian/thermos + -ian | " | A | no |
| | " | - | B | - |
| | toxic + -ism/public + -ism | [k] vs [s] | A | yes |
| | " | " | B | no |
| | zebra + -ic/Buddha + -ic | tense vowel vs Deletion | A | no |
| | " | " | B | yes |
| | (different vowel) | | | |
| | define + -ity/methane + -ity | tense vowel vs lax vowel | A | yes |
| | " | " | B | no |
| | (identical vowel) | | | |
| | obtain + -ion/obtain + -atory | " | A | yes |
| | " | " | B | yes |
| | pertain + -ion/pertain + -atory | " | A | no |
| | " | " | B | no |
| | probe + -ity/prose + -ify | " | A | no |
| | " | " | B | no |

Table XII. $\chi^2$ and $r_t$ levels of significance of different pairs of stem-suffix sets.

derivations (containing one stem-word) followed the sound patterns of the conflicting example derivations. Since the subjects' responses provided contradictory evidence for the use of a single identical underlying form for a given stem-word, Ohala rules out the use of the phonological rules hypothesized by Chomsky and Halle. He does not consider the possibility that some other independent phonological rules were used. Instead, he turns to another means of production, that of analogical phonological rules, and concludes that this was the means in use since the subjects' responses supported this hypothesis.

In the present study Group A's responses "followed" the leading example derivations quite frequently. Nine of 17 subjects derived obtain and 10 of 17 subjects derived pertain by producing the lax vowel [ɛ] when the preceding example was detain/detention and [æ] when the example was explain/explanatory. The fact that the vowels in the responses to the pairs of stem-suffix sets were not the same involved two implications for Chomsky and Halle's phonological theory. The responses could indicate either (a) that Chomsky and Halle's hypothesized single underlying form was not used or (b) that the underlying form was used but their hypothesized rules were not applied to it. However, rejection of Chomsky and Halle's theory does not serve, as Ohala would have it, as an unconditional warrant for accepting the theory of analogical rules. Such a "how-else" argument would preclude other possibilities for the production of the sound patterns to have been influenced by the example derivations.

One asks if the subjects of Group A possibly felt obliged to

follow the example derivations, and if so, one wonders what sound patterns they would have produced had they not felt this way. An answer might be found by considering the responses of Group A which were not similar to the examples' sound patterns as well as the responses of Group B, whose subjects were not exposed to the examples.

One finds that the two groups derived each of the two pairs of stem-suffix sets inconsistently (that is, without the same vowel in each member of the pair of derivations) in other ways than that mentioned above. Most often the pair of derivations involving a given stem contained a tense vowel in one response and a lax vowel in the other. This occurred in 7 of the total 34 pairs of stem-suffix sets in Group A and in 8 of the 30 pairs in Group B. In only 3 of the 7 pairs in Group A and in 5 of the 8 pairs in Group B did these inconsistent derivations agree with the sound patterns predicted by Chomsky and Halle. For instance, a tense vowel would appear in the environment preceding CiV̆ (e.g. [pə'tê͡ɪnjən] from pertain + ion) and a lax vowel would appear in the environment before CV̆CV (e.g. [pə'tɛnətɔɹi] from pertain + -atory). On the other hand, 4 of the 7 pairs in Group A and 3 of 8 pairs in Group B involved 'tense' or 'lax vowel' responses (e.g. [pə'tê͡ɪnjən] from pertain + -ion and [pə'tê͡ɪnətɔɹi] from pertain + -atory in hypothetically appropriate environments. Other sets of inconsistent responses included one derivation containing a tense vowel and the other derivation with a deleted syllable in Group A while in Group B, 2 pairs of sets were comprised of one derivation with a lax vowel and the other with a deleted syllable.

A few subjects in each group consistently used the same tense or lax vowel in their responses to a pair of stem-suffix sets. In response to the pair of sets involving obtain, 3

subjects in Group A produced 'tense vowel' sound patterns and 1 produced 'lax vowel' sound patterns (containing the vowel [ɛ]). In response to the pairs of derivations involving pertain, 3 subjects in Group A and 8 in Group B produced 'tense vowel' sound patterns. However, most (25 of 26) of the consistent pairs of responses mentioned above contain one member response which contradicts the sound patterns predicted by Chomsky and Halle (e.g. a tense vowel preceding CV̆CV or a lax vowel preceding CiV̆). While the occurrence of identical vowels might have seemed to support the hypothesis for identical underlying forms, the occurrence of unpredicted sound patterns in one member of 26 consistent pairs of responses showed contradictory evidence for the hypothesized phonological rules ($VLR_1$ and $VTR_2$).

Ohala (1973, p.45) uses the low frequency of occurrence of the predicted sound patterns in his study as evidence that subjects infrequently used the phonological rules in question. He believes, on the basis of the results of the tetrachoric correlation coefficient measure of successive and nearby responses as well as the results of responses to obtain and pertain, that subjects were using analogical phonological rules. Ohala finally concludes that "analogical phonological rules seem to be stronger in determining the form of a new derivation than are independent phonological rules" (1973, p.45). Recall, however, that it was shown in this study that no conclusions could be made on the particular means of sound pattern production by considering the occurrence of different sound patterns and the existence of high statistical correlations. In other words,

Ohala's conclusion precludes the possibility that independent phonological rules (other than those formalized by Chomsky and Halle) or that some other means of sound pattern production, not yet proposed in linguistic theory, were being used by the subjects.

## 4.5 Further observations

### 4.5.1 Subjects' productivity with sound patterns and with phonological rules

Other authors investigating the psychological reality of phonological rules discuss whether the rules under investigation were frequently used by subjects and so were shown to be "productive." In essence, they are saying that a hypothesized phonological rule was used if the sound pattern it predicted occurred in subjects' responses. This form of argument which they use to make positive statements about the rule's use and it's productivity commits a logical fallacy (discussed in Chapter 2) known as "the fallacy of affirming the consequent." As pointed out by Caws (1965), this argument cannot be used to verify hypotheses. Therefore one cannot discuss whether the subjects are productive with the rules under investigation. A positive statement about subjects' "productivity" will have to be restricted to their productivity with particular sound patterns.

For example, one could discuss whether any of the predicted sound patterns have a 'special status' in so far as they were produced, and could in similar situations be expected to be produced, more often than other sound patterns. Then if any

did have a 'special status,' it would be interesting in further study to check for the frequent extension of these sound patterns in other situations to see if subjects more generally have a special "awareness," whatever that might be, of these sound patterns. If, furthermore, subjects did extend these sound patterns in other situations, then it would be interesting to try to discover why they did.

In order to assign a 'special status' to a sound pattern produced by subjects, it is necessary to decide on a minimum percentage of occurrence for the sound pattern. This criterion frequency of occurrence should depend on the probable distribution of the possible sound patterns for a given stimulus. However, it is not possible at this time to know their probability of occurrence since the number of sound patterns and their frequencies of occurrence depend on such factors as the presence or absence of example derivations, on the various stem-suffix sets, their number, as well as the time permitted for responding, the duration of the experiment, memory factors, etcetera.

A more informal way to approach the problem would be to consider how many subjects produced the sound pattern predicted by a given rule in at least 2/3 of his/her responses. Two-thirds is a severe criterion that would likely insure that the special status of a sound pattern was not overestimated. Two-thirds was chosen as the criterion since one-half would be just at chance level in the case where only two sound

patterns were possible. Using this criterion, the responses of Group B were considered as this group seemed equally 'adept' at the production of novel derivations as Group A. In addition, they were not exposed to leading example derivations and so their responses may have been closer to the responses which would be produced by speakers in a natural situation when there is a need for a novel derivation. Group B's responses to the long list of stimuli were considered in this case since it is preferable to know their production of sound patterns in response to the greatest number of stimuli possible. Data for $VLR_1$ and $VTR_2$ were derived from eligible responses to obtain and pertain as well as from each rule's separate list of stem-suffix sets. In this case, the relevant data from both lists of stimuli were combined for each rule.

Nine subjects reached the criterion of 67% in producing the sound pattern predicted by the PSR. Five subjects produced the 'tense vowel' sound pattern predicted by $VTR_1$ in over 2/3 of their responses. Four subjects reached the criterion percentage in producing the velar softening sound pattern predicted by VSOR. No subjects produced the 'lax vowel' sound pattern predicted by $VLR_1$, the 'tense vowel' pattern predicted by $VTR_2$ nor the voiced-/s/ pattern predicted by /s/VR in 2/3 or more of his/her responses.

To be safe, one could conclude that only that sound pattern which a majority of the subjects extended (that is, the sound pattern predicted by PSR), by unknown means, in 2/3 or more of their responses could be given 'special status'

for the subjects of Group B. In addition, however, four subjects produced 6 responses to sets involving obtain and pertain that were relevant to $VLR_2$. All 6 of the responses were the predicted 'lax vowel' sound pattern. It is possible, then, that this sound pattern may also be eligible for special 'status' for the subjects of Group B.

A valid argument may be used in the case where the consequences are not what the hypothesis says they ought to be. In Caw's words, "while a hypothesis cannot be verified (that is, shown to be true) by reference to its consequences, it may be falsified conclusively if the consequences fail to occur" (1965, p.112). So examining the responses in which the consequence (i.e., the predicted sound pattern) failed to occur will indicate how often the hypothesized rule was not used. Therefore one may say how often subjects were 'unproductive' with a given rule. The percentage of responses (from the long list) which showed unpredicted sound patterns (i.e., any sound pattern other than the predicted one) is listed for each rule in Table XIII. Only the responses of Group B were considered as this group was not submitted to the influence of leading example derivations. The data for each of $VLR_1$ and $VTR_2$, from eligible responses to pertain and obtain and from each rule's separate list of stimuli, were once again combined for each of those rules.

The results show that, as a group, the subjects of Group B did not use the following rules in a clear majority of the responses considered: $VLR_1$, $VTR_2$, VSOR and /s/VR.

| RULE | TOTAL NO. ELIGIBLE RESPONSES | UNPREDICTED RESPONSES | % UNPREDICTED RESPONSES |
|---|---|---|---|
| $VLR_1$ | 300 | 237 | 79 |
| PSR | 195 | 63 | 32.3 |
| $VTR_1$ | 45 | 32 | 71.1 |
| $VTR_2$ | 144 | 110 | 76.4 |
| VSOR | 60 | 40 | 66.7 |
| $VLR_2$ | 6 | 0 | 0 |
| /s/VR | 31 | 31 | 100 |

Table XIII. The percentages of occurrence of unpredicted sound patterns in the responses (long lists) of Group B which were relevant to each rule.

The occurrence of unpredicted sound patterns was also considered for each subject in Group B to determine if some subjects failed to use a given rule more often than other subjects. For Group B, the range in the percentages of occurrence of unpredicted sound patterns and the mean percentage are listed for each rule in Table XIV. In this case also, the data from the responses to obtain and pertain and that from each rule's separate list of stimuli were combined for each of $VLR_1$ and $VTR_2$. The ranges are quite broad for all rules except /s/VR and $VLR_2$. This suggests, first, that the subjects did vary in the number of responses in which they failed to use a given rule. Secondly, since the subjects' lack of productivity with each rule varies, it is possible that their strategies for novel word derivations and hence their phonological models are not exactly the same.

### 4.5.2 Stem-suffix sets most and least often involved in predicted sound patterns

It was observed in this study that the stem-suffix set (within the group of those sets considered for each rule) which most often underwent the predicted phonetic change tended to be the same for each group of subjects. The stem-suffix sets which most often underwent the predicted phonetic changes (in each group) and the percentages of times that the sets underwent the changes are listed in Table XV.

One can probably conclude from the results listed in the table that since the stem-suffix set (which most often showed the sound pattern predicted by a given rule) frequently remained the same in both groups' responses, the different

| RULE | % OCCURRENCE OF UNPREDICTED RESPONSES | | | |
|---|---|---|---|---|
| | RANGE | | MEAN | AVERAGE DEVIATION |
| | highest | lowest | | |
| $VLR_1$ | 100 | 35 | 79 | 13.6 |
| PSR | 61.6 | 7.7 | 33.3 | 15.9 |
| $VTR_1$ | 100 | 33.3 | 71.1 | 27.2 |
| $VTR_2$ | 100 | 60 | 76.1 | 8.5 |
| VSOR | 100 | 25 | 66.7 | 28.9 |
| $VLR_2$ | 0 | 0 | 0 | 0 |
| /s/VR | 100 | 100 | 100 | 0 |

Table XIV. Ranges in the percentages of unpredicted responses, the mean percentages and the average deviations from the means for Group B (long list).

| RULE | GROUP A | | | | GROUP B | | | |
|---|---|---|---|---|---|---|---|---|
| | Short List | | Long List | | Short List | | Long List | |
| | stem-suffix set | % | stem-suffix set | % | stem-suffix set | % | stem-suffix set | % |
| $VLR_1$ | define + -ity | 76.4 | → | 76.4 | → | 60 | → | 60 |
| PSR | robot + -ian | 88.2 | methane + -ity<br>iodine + -ity<br>robot + -ian | 100<br>94.1<br>88.2 | robot + -ian | 100 | methane + -ity<br>iodine + -ity<br>robot + -ian | 100<br>100<br>100 |
| $VTR_1$ | zebra + -ic | 64.7 | → | 64.7 | → | 46.6 | → | 46.6 |
| $VTR_2$ | cop + -ian | 47 | → | 47 | → | 20 | → | 20 |
| VSOR | domestic + -ism<br>public + -ism | 64.7<br>64.7 | →<br>→ | 64.7<br>64.7 | →<br>→ | 46.6<br>40 | →<br>→ | 46.6<br>40 |
| $VLR_1$, $VLR_2$ | obtain + -ion | 82 | → | 82 | pertain + -ion | 26 | → | 26 |

Table XV. Stem-suffix sets whose responses most often showed the predicted sound pattern. The % is based on the number of times (out of the total number of responses to that set) that the set elicited a predicted response. Note that /s/VR is not listed since Group B did not produce any response predicted by it.

experimental conditions did not play a strong role in determining these sets. Rather, the given stem-suffix sets must have been determined as a result of the sound pattern for which they were considered, the stem-word or the suffix comprising the set or a combination of these factors.

For example, the set define + -ity may have demonstrated the predicted 'lax vowel' sound pattern so often as a result of the sound pattern being demonstrated in the existing English derivation divinity. This derivation is comprised of a stem, which is in minimal phonemic contrast with define, and of the same suffix. So the stem and the suffix may have played a role in inducing the lax vowel sound pattern. Another possibility is that many of the existing English derivations in which the stem-word define is involved contain a lax vowel in the second syllable of the stem: e.g. definition, definite, definitely, definiteness, definitive. In this case the stem itself may have played an important role in eliciting the 'lax vowel' sound pattern since that stem contains a lax vowel when derived with several different suffixes.

In the data from the short and the long lists which were considered for each group of subjects, the stem-suffix set which least often underwent the predicted phonetic change remained the same only in the case of two rules. In the sets considered for $VLR_1$, the stem-suffix set bribe + -ity showed the predicted change least often in both groups' data from both the short and long lists of stimuli. In the sets considered for the PSR, the stem-suffix set in all four lists of stimuli was

Buddha + -ic. This set also showed the least occurrence of phonetic change in the short lists' data comprised from Group A and B's responses for $VTR_1$.

As in the case of the sets showing the most frequent occurrence of predicted phonetic change, the sets consistently showing the predicted changes the least often may be determined from a combination of such factors as the characteristics of a given set's stem-word and suffix. For instance, the low frequency of Buddha + -ic undergoing the predicted change in primary stress and the predicted tense vowel change might be explained by the fact that the stem-word is of Sanskrit origin while the suffix combines with proper nouns usually only of Greek or Latin origin. Another possibility is that none of the English derivations involving Buddha (e.g. Buddhist, Buddhistic, Buddhistical) retain and alter the final vowel of the stem-word.

### 4.5.3 Suffixes and vowels most often involved in predicted sound patterns

Steinberg and Krohn (1975) specified the suffixes in their study which were most often involved in derivations showing the predicted sound pattern. It was possible for them to do this since in their experiments an equal number of stem-words was derived an equal number of times with each suffix. All of the derivations were considered for the occurrence of a given sound pattern predicted by the VSR. The conditions of the present study prevent the suffixes most often involved in predicted sound patterns from being specified. Each suffix was used a different number of times with the stem-words

considered relevant to each rule. None of the stem-words was paired an equal number of times with each suffix.

Steinberg and Krohn's experimental design also permitted them to identify which vowel of the stem-words most often underwent a predicted phonetic change. In their design, the given vowels were all (with one exception) represented an equal number of times in the corpus of stem-words and were all intended to be paired with each suffix an equal number of times. In the present study, however, all of the given vowels occurring in the list of stem-words did not occur an equal number of times in the experimental task list nor in those stem-words considered for each rule.

## CHAPTER V

## DISCUSSION AND CONCLUSIONS

Three secondary objectives for this study, mentioned at the end of Chapter I, were fulfilled by the literature review and by the results discussed above in an unforeseen manner.

First, in addition to stating "how," it was stated "why" experiments on non-ideal speaker-listeners could apply to theories on ideal speaker-listeners. "How" such experiments could apply to theories on ideal speaker-listeners was limited to a discussion of the experiment undertaken in the present study and shall be further discussed below. The reason "why" such experiments should be conducted in relation to theories on ideal speaker-listeners was said to be essentially the lack of scientific methodology in generative linguistic theory and the consequently limited value of generative linguistics to other scientific fields of study. One might still feel uneasy about testing Chomsky and Halle's phonological theory outside the idealized speaker frame for the basic reason that this procedure draws on some of the very notions (that is, "behavioural" notions) which Chomsky set out to disprove. Any residual uneasiness might be dispelled by considering Kuhn's (1962) belief that arguments for particular scientific

paradigms[1] are necessarily circular:

> When paradigms enter, as they must, into a debate about paradigm choice, their role is necessarily circular. Each group uses it's own paradigm to argue in that paradigm's defense. The resulting circularity does not make arguments wrong or ineffectual. The paradigms' arguments will show what scientific practice will be like for those adopting the paradigm (p.93).

If one scientific paradigm has reached the "crisis" stage (wherein the explicit and fundamental generalizations of the paradigm are questioned and other conditions, mentioned by Kuhn, occur), then a new paradigm--whether an entirely new one or a compromise of former ones--should be adopted. The testing of Chomsky and Halle's phonological theory in the present investigation is based on the presumption that, as in crises of other scientific paradigms, the present crisis situation in verifying generative phonological theories (by structural evidence) calls for at least some compromise of the opposing paradigms. Kuhn suggests that "paradigm choice can never be unequivocally settled by logic and experiment alone" (1962, p.93). According to him, in the end the question of paradigm choice is answered in terms of criteria outside of normal science by the assent of the relevant community who dictates which problems are more significant to have solved.

---

[1] Kuhn uses "paradigm" in the sense of study patterns used by scientists in a particular field of research. A paradigm is characterized by the unprecedented achievement of scientists employing the study patterns which "attracts an enduring group of adherents away from competing modes of scientific activity" and by the open-ended quality of the study patterns which leaves all sorts of problems to be resolved (Kuhn, 1962, p.10).

Secondly, the specification of explicit criteria which could serve to confirm or disconfirm the mentalistic claims of the phonological rules was discussed. While evidence was specified which could serve to disconfirm the use of phonological rules, none could be given to confirm the use of the rules. This was due to the nature of the valid logical argument, modus tollens, used in scientific investigation, which is only able to disconfirm and thus eliminate incorrect hypotheses. It was shown that confirmation and disconfirmation of the use of phonological rules were not conclusions which could stand in a complementary relationship. In other words, in the case where the hypothesized use of a rule was not disconfirmed, it's use was not consequently confirmed.

Thirdly, the logical impossibility of specifying evidence which would confirm the use of independent phonological rules (or any other means of sound pattern production) prevented the recognition, and hence the explanation, of the means of production which were used. The few attempts to explain the occurrence of certain sound patterns were, indeed, speculative and inexhaustive.

Thus one of the main objectives of this study, which was to draw conclusions on the validity of the experimental paradigm as a confirmation or disconfirmation procedure, was fulfilled. It was pointed out that the form of the logical argument used in the investigation could at best be used to disconfirm given hypotheses. The other main objective, which was to decide

whether certain phonological rules were psychologically real for a given group of speakers, was more difficult to fulfill. Three problems currently stand in the way of confirming or disconfirming the psychological reality of the rules under investigation.

One problem concerns the unknown maximum frequency with which a phonological rule can fail to be used in the responses of a group of speakers (or in those of an individual speaker) and still be described as psychologically real. This problem might be resolved if generative phonologists could specify the extent to which hypothesized rules are used. (Chomsky and Halle (1968) merely state that their rules are applied to all prerequisite underlying structures and then give what is presumably an inexhaustive list of exceptions to the rules' application.) Before phonologists specify the exact extent to which phonological rules are used, the adoption of a maximum frequency for the non-use of a rule can only be ad hoc and, due to the problem to be discussed next, could cause one to draw false conclusions.

The second problem is that even if the rule's extent of use was clearly specified, a conclusion on the psychological reality of the rule could not always be made. Indeed, the exact number of times that a given rule was not used in a group of responses remains unknown when the predicted sound pattern occurs at least once in those responses. Therefore a situation could occur in which the rule's use was disconfirmed in a number of responses x% short of the criterion percentage

and the predicted sound pattern (and the unknown means of production) occurred in at least x% of the responses. For example, if the criterion percentage of 'non-use' at which a rule was disconfirmed as psychologically real was 80%, 77% of the responses could possibly contain unpredicted sound patterns while 23% could possibly show the predicted sound pattern. In the case, on the other hand, wherein the stimuli considered relevant to a rule elicited unpredicted sound patterns in 100% of the responses, one could state the exact number of times that the rule was not used by the group. In this study this occurred in the case of the responses which were considered for the /s/VR. However, there was some potential evidence for the extension of the predicted sound pattern when one subject laxed a tense vowel preceding [z] and devoiced the [z] as well. This evidence could possibly weaken the certainty that the rule failed to be used at all by the group.

A third problem involves the fact that the percentages of unpredicted sound patterns for the rules were derived from different numbers of responses. It would not be fair to consider the psychological reality of the rules when the decisions for or against the psychological reality were based on a different number of responses relevant to the rules and consequently on percentages (of unpredicted sound patterns) that did not vary in a similarly discrete way.

Still another difficulty which might be considered in the attempt to determine rules' psychological reality is that the logical argument <u>modus tollens</u> accepted for this investigation

has a potential flaw which, if Chomsky and Halle's idealized framework were not the frame under consideration, could allow the wrong conclusion to be drawn. Consider Caw's (1965) exemplification of the modus tollens argument. Someone hypothesizes that if a man takes arsenic, he will die. "Suppose the man...does not die: we may then conclude with certainty that he did not take arsenic, or at least not enough" (Caw, 1965, p.112). At least two conclusions are possible: either the man did not take arsenic or else he did not take enough arsenic to kill him.

Similarly, more than one conclusion could be drawn from the negative consequence of hypothesized phonological rules if the rules were removed from Chomsky and Halle's idealized framework. For instance, suppose that the following points in Chomsky and Halle's theory were abandoned: (a) that all speakers of an hypothesized homogenous speech community possess identical phonological models; (b) that the use of a phonological rule implies the use of a specific prerequisite underlying form; and (c) that a given rule is applied to all underlying forms with prerequisite structures. Then in the case where the predicted sound pattern did not occur, the argument modus tollens would be incapable of permitting a single conclusion disconfirming the use of the rule. A possible conclusion could be, as was suggested by Sherzer (1970) and Zimmer (1969), that speakers of the same dialect have different phonological models. Therefore different rules or other strategies of sound pattern production could be used by

different speakers in response to the same stimulus. This possibility might be explained by the idea put forth by Wang (1969), Hsieh (1972) and Chen and Wang (1975) that phonetic changes (and hence phonological rules) diffuse gradually across speakers' lexicons. Therefore speakers might not apply a given phonological rule to all of the word forms with the prerequisite underlying form. Nor might all speakers apply the same phonological rule, or use the same strategy for sound pattern production, in response to the same word form.

In order to derive one conclusion (that is, disconfirmation of the hypothesis) from negative consequences, the statement of the hypothesis should likely be quantified. For example, Caw's example of an hypothesis could be stated: If a man takes enough arsenic, he will die. The hypothesis tested in this investigation might be stated: If and only if a given phonological rule is used, then a particular sound pattern will be produced.

The present investigation might be improved, first, by limiting the time for subjects to respond in order to prevent them from repeating one form of a response to a stimulus several times, thus shortening the duration of the interview. A time limit which might be generous enough to permit subjects to respond without making "slips of the tongue" might be 4 seconds, the estimated period within which half of the responses were made by the two groups in the present study. Secondly, subjects' responses could be transcribed after, not during, the interview so that the transcriber might have a better

opportunity to judge how the phonetic forms of responses differed. Thirdly, those stem-words and suffixes which in Chapter III were mentioned to have characteristics inappropriate for their combination in a derivation might be replaced by stem-suffix sets whose members did have appropriate characteristics. Fourthly, the example derivation Kant/Kantian, showing no phonetic change in its stem, was unfamiliar to most subjects and could be replaced with Marx/Marxian.

One result of this investigation which warrants further study is that both groups produced a predicted sound pattern (in those responses relevant to PSR and also in those relevant to $VTR_1$) almost equally as often. One could check to see if this result remained when different groups of subjects and/or longer lists of stimuli were involved. If the occurrence of predicted sound patterns continues to be very close for the two groups, one might suspect that the presence or absence of leading example derivations is not influencing the groups' numbers of predicted responses. To investigate this hypothesis, the leading example derivations preceding Group A's list of stimuli could be replaced to see if other leading examples make a difference to the group's numbers of predicted responses. In addition, more than two groups could be submitted to the lists of stimuli for each rule while exposing each group to a different number of leading examples. The results might suggest that each group's production of predicted sound patterns is limited in that this production would not increase after a certain number of leading examples had been supplied them.

The trend for the stem-suffix set most often involved in a predicted response to be the same in each group's lists of data could be further investigated to determine if such a trend occurs (a) when more subjects participate in the study and/or (b) when the lists contain more stem-suffix sets. If such a trend did remain in the two conditions mentioned above, a few stem-suffix sets in a list of stimuli of given length could be replaced, one at a time, with other sets. This might serve as a check to see whether the number of times that sets were involved in predicted responses remained fairly constant. Such a check might also have implications for the possibilities that the stem-suffix set most often involved in predicted responses is somehow self-determined or is somehow determined by other sets in the group of stimuli.

On another occasion, the corpus of stem-suffix sets could be controlled for the number of sets which were relevant to each rule, the numbers of different strategic vowels represented in the stem-words and the number of stem-words, as well as the number of stem-words containing given strategic vowels, assigned to each suffix. This should help to obtain equal numbers of responses which are to be considered for the various rules and should facilitate the identification of the suffix and of the strategic vowel which are most often involved in predicted sound patterns.

Another method which could be used to check for the extension of predicted sound patterns would be to present subjects with existing or novel derivations and require them to

retrieve the component stem-word and suffix. This method was used by Hsieh (1975) with the purpose of comparing its results with those of the novel derivation task.

Although the consequences of the novel derivation task cannot be used to confirm any hypothesized means of sound pattern production, the consequences might be valuable as evidence for which sound patterns speakers are able to, or wish to, extend in certain phonetic contexts. Linguists might be interested in knowing which sound patterns are or are not extended for the purpose of assigning different statuses to the phonological regularities of a language which they describe. Such information could be used in a grammar which is at some intermediate position between ones which are purely formalistic and ones which are mentalistic.

It has been pointed out that the validity of the modus tollens argument, employed when the consequences of an hypothesis are negative, limits the conclusions that may be drawn in the ontological experiment particular to this study. The strategies of sound pattern production used by the subjects could not be determined. It was only possible to determine, through negative consequences, when the hypothesized rule was not used. Thus, when some of the consequences were positive, it was impossible to determine the exact number of times that the hypothesis (i.e., the use of a given rule) was disconfirmed.

It has been suggested above that one cannot specify a maximum frequency for the non-use of a phonological rule, (above which that rule will no longer be considered to be

psychologically real), until the hypothesized extension of rules is made clear by generative phonologists. Since a criterion frequency of non-use of a rule could not at this time have been well motivated, the psychological reality of given phonological rules for an individual or for a group of speakers could not be determined.

If the modus tollens argument were to be used with phonological theories which do not employ certain ideals of Chomsky and Halle's (1968) theoretical framework, then the unquantified statement of the argument's hypothesis would be compatible with the existence of more than one possible conclusion.

With respect to the particular experimental design employed in this study, the similar distribution of numbers of predicted responses in each group showed that the two groups were similar in their ability to perform the experimental task, and were therefore representative samples of the population under study. Within each group, the overall production of predicted sound patterns was not attributable to just a few subjects. The overall occurrence of predicted sound patterns was shown to be greater for the group of subjects (Group A) who were exposed to leading example derivations. Due to the two groups' similar abilities in handling the experimental task, the greater occurrence of the predicted responses in the total of Group A's responses was attributed to that group's exposure to leading example derivations.

However, in the group's responses which were considered for just one rule at a time, the occurrence of predicted sound

patterns was not, in the case of each rule, much greater for Group A. The reason for the closeness in the two groups' numbers of sound patterns which were predicted by each of PSR and $VTR_1$ requires further investigation.

The trend for the same stem-suffix set (in both groups' long and short lists of stimuli) to be most often involved in a predicted phonetic change could only be explained speculatively. Much further investigation would be required to determine whether such a trend would always exist and if so, to successively eliminate some of the hypotheses proposed to account for the trend.

It appeared that a small variation in the number of stem-suffix sets which were considered for a rule did not greatly influence subjects' mean ranks for the overall production of predicted responses.

Since the probability of occurrence of sound patterns could not be determined, the procedure used to assign a 'special status' to certain predicted sound patterns (those predicted by PSR and possibly also that predicted by $VLR_2$) was *ad hoc*.

## BIBLIOGRAPHY

Botha, R. 1968. The Function of the Lexicon in Transformational Generative Grammar. (Janua Linguarum, Series Maior, Nr. 38.) The Hague: Mouton & Co.

Botha, R. 1970. The Methodological Status of Grammatical Argumentation. (Janua Linguarum, Series Minor, Nr. 105.) The Hague: Mouton & Co.

Botha, R. 1971. Methodological Aspects of Transformational Generative Phonology. (Janua Linguarum, Series Minor, Nr. 112.) The Hague: Mouton & Co.

Carterette, E. and M. Friedman, eds. 1974. Handbook of Perception, Vol. One. Historical and Philosophical Roots of Perception. New York: Academic Press.

Caws, P. 1965. The Philosophy of Science, a systematic account. Princeton: Van Nostrand.

Chen, M. and W. Wang. 1975. 'Sound change: activation and implementation.' Language 51, 2 (June): 255-281.

Chomsky, N. 1964. Current Issues in Linguistic Theory. (Janua Linguarum, Series Minor, Nr.38.) The Hague: Mouton & Co.

Chomsky, N. 1965. Aspects of the Theory of Syntax. Massachusetts: M.I.T. Press.

Chomsky, N. 1968. Language and Mind. New York: Harcourt, Brace and World.

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

Dolby, J. and H. Resnikoff. 1967. The English Word Speculum. Vol. Three. The Hague: Mouton & Co.

Esper, E. 1973. Analogy and Association in Linguistics and Psychology. Athens: University of Georgia Press.

Fischer-Jørgensen, E. 1975. 'Perspectives in phonology.' Annual Report, Institute of Phonetics. ARIPUC 9: 215-236. University of Copenhagen.

Greene, J. 1972. Psycholinguistics, Chomsky and Psychology. England: Penguin Books, Ltd.

Hockett, C. 1968. The State of the Art. (Janua Linguarum, Series Minor, Nr. 73.) The Hague: Mouton & Co.

Hsieh, H.-I. 1972. 'Lexical diffusion: evidence from child language acquisition.' Glossa 6, 1: 89-103.

Hsieh, H.-I. 1975. 'How generative is phonology?' The Transformational-Generative Paradigm and Modern Linguistic Theory, ed. E. Koerner, et al. Amsterdam: John Benjamins B.V.

Krohn, R. 1972. 'The vowel shift rule and its productivity.' Language Sciences 20: 16-18.

Kuhn, T. 1962. The Structure of Scientific Revolutions. Chicago: University of Chicago Press.

Ladefoged, P. 1970. 'The phonetic framework of generative phonology.' Working Papers in Phonetics-UCLA 14 (March): 25-32.

Lindblom, B. 1971. Numerical Models in the Study of Speech Production and Speech Perception. Some Phonological Implications. Department of Speech Communication, Royal Institute of Technology, 10044 Stockholm 70, Sweden and Institute of Linguistics, Department of Phonetics, University of Stockholm.

Linell, P. 1974. Problems of Psychological Reality in Generative Phonology. A Critical Assessment. Reports from Uppsala University Department of Linguistics (RUUL), Nr. 4.

Lyons, J. 1970. Chomsky. Great Britain: Wm. Collins & Co. Ltd.

Marchand, H. 1960. The Categories and Types of Present-Day English Word-Formation. Wiesbaden: Otto Harrassowitz.

Nicholson, G. 1916. English Words with Native Roots and with Greek, Latin or Romance Suffixes. Chicago: University of Chicago Press.

Ohala, J. 1970. 'Aspects of the control and production of speech.' Working Papers in Phonetics-UCLA 15.

Ohala, J. 1973. 'On the design of phonological experiments II.' Mimeographed version of an unpublished paper read at the Winter 1972 Linguistic Society of America meeting in Atlanta.

Ohala, J. 1974. 'Phonetic explanation in phonology.' Paper delivered at the Natural Phonology Parasession, Chicago Linguistic Society, 18 April, 1974.

Ohala, M. 1974. 'The abstractness controversy: experimental input from Hindi.' Language 50, 2 (June): 225-235.

Onions, C., ed. 1966. Oxford English Dictionary of Etymology. London: Oxford University Press.

The Principles of the International Phonetic Association. 1949 (Reprinted 1972). International Phonetic Association, University College, Gower Street, London WC1E 6BT.

Sherzer, J. 1970. 'Talking backwards in Cuna: the sociological reality of phonological descriptions.' Southwestern Journal of Anthropology 26: 343-353.

Steinberg, D. 1975. 'Chomsky: from formalism to mentalism and psychological invalidity.' Glossa 9, 2: 218-252.

Steinberg, D. and R. Krohn. 1975. 'The psychological validity of Chomsky and Halle's vowel shift rule.' The Transformational-Generative Paradigm and Modern Linguistic Theory, ed. E. Koerner, et al. Amsterdam: John Benjamins B.V.

Wang, W. 1968. 'Vowel features, paired variables, and the English vowel shift.' Language 44, 4 (December): 695-708.

Wang, W. 1969. 'Competing changes as a cause of residue.' Language 45, 1 (March): 9-26.

Zimmer, K. 1969. 'Psychological correlates of some Turkish morpheme structure conditions.' Language 45, 2 (June): 309-321.

# APPENDIX I

## DIRECTIONS AND LIST OF WORD FORMS FOR EACH EXPERIMENT

### Directions

#### Pilot I

I would like your help in a project of mine which is to construct an extended or extrapolated dictionary of English. As you know, English allows the formation of new words by the addition of suffiexes or endings such as -ness, -ity and so on to the end of existing words. For example, odd can have -ity added to it to give oddity.

We'll proceed as follows: I'll give a particular suffix, then give some examples of its use with existing words, and then give a number or words which that suffix has not been added to before. (Later on, I'll give just the suffix and then words to which the suffix has never been added.)[1]

Would you please pronounce the word that would result from the addition of that suffix, tell me what you think this new word might mean, and then tell me whether or not you would be likely to use this new word.

#### Pilot II

I would like your help in a project of mine which is to construct an extended or extrapolated dictionary of English. As you know, English allows the formation of new words by the combination of suffixes or endings with existing words. For example, the word odd combined with the ending -ity gives oddity.

We'll proceed as follows: to begin with, I'll give a particular ending, then give some examples of its use with existing words, and then give a number or words which that ending has never been combined with before. Later on, though, I'll give just the ending and the words which the ending is to be combined with.)[2] So, to help me in my survey, would you please take the word and the ending, put them together somehow and pronounce the new word. (Then would you tell me what you think the new word would mean and lastly tell me whether or not you would be likely to use this new word.)[3]

Now let's return to the new words you proposed. I'll give the ending, the word, and the new word you suggested. This time, to help me in my survey, would you tell me what you think this new word would mean and then tell me whether or not you would be likely to use this new word.

#### Main

May I remind you that this experiment is like a survey in that the responses that you give are neither right nor wrong; rather, they are all valid.

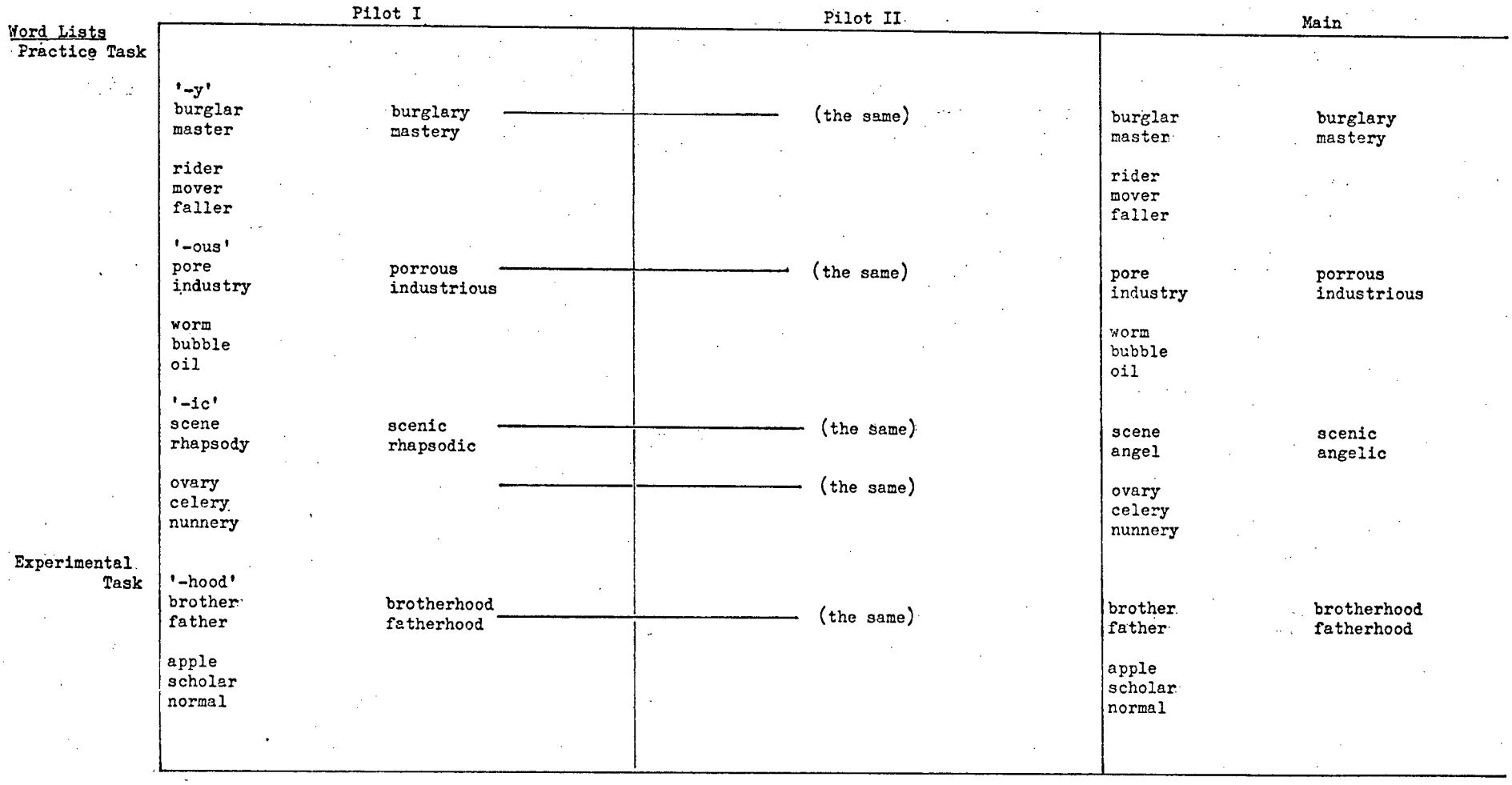
I would like your help in a project of mine which is to construct an extended or extrapolated dictionary of English. As you know, English allows the formation of new words by the combination of suffixes or endings with existing words. For example, the word odd combined with the ending -ity gives oddity.

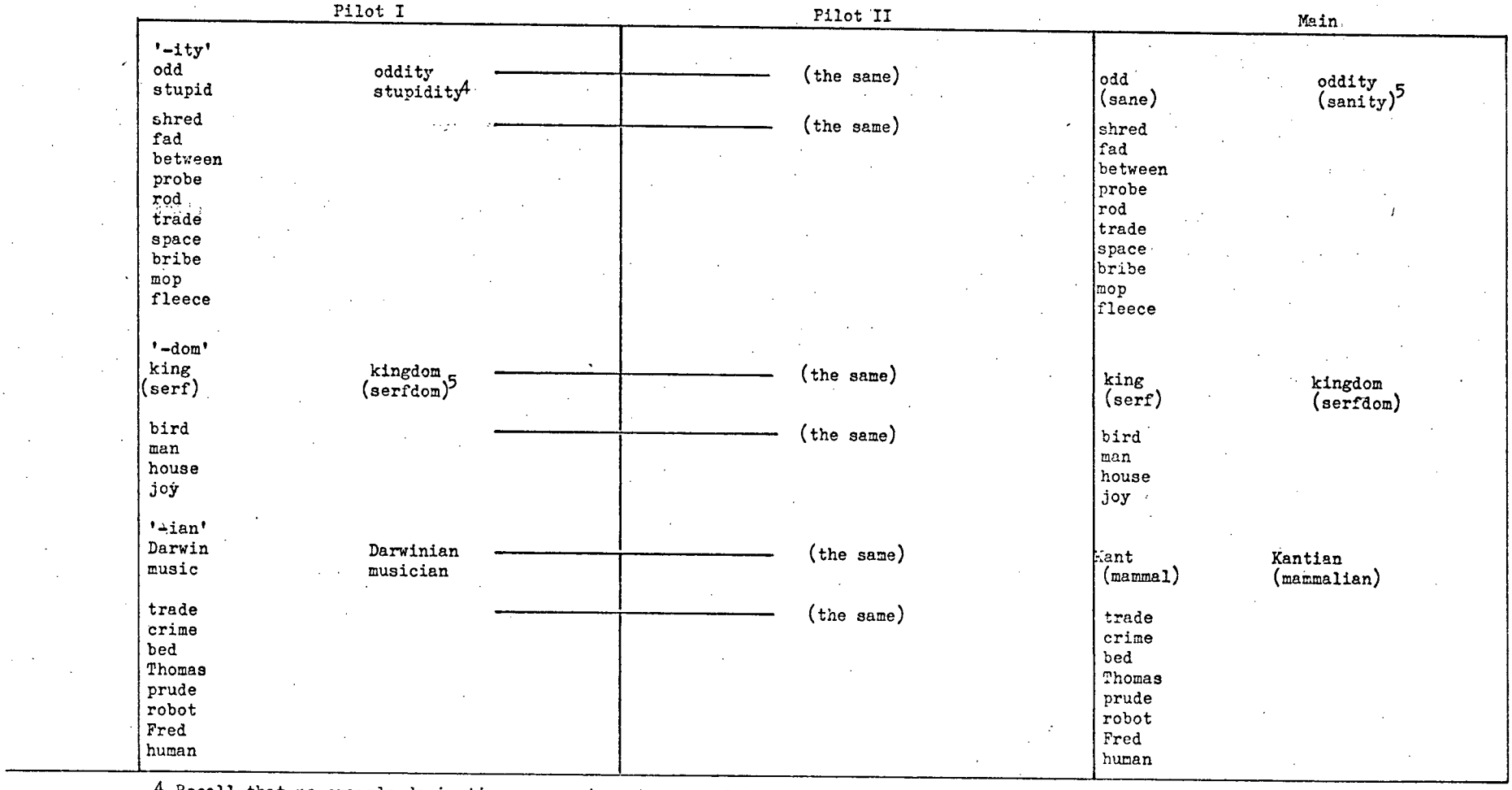
We'll proceed as follows: I'll give a particular ending, then give a number of words which that ending has never been combined with before. To help me in my survey, would you please take the word and the ending, put them together somehow and pronounce the new word. Then would you tell me what you think the new word would mean, and lastly tell me whether or not you would be likely to use this new word.

1 The directions to Groups a' and b' in Pilot I were identical except that the sentence enclosed in parentheses was included in the directions to Group b' only.

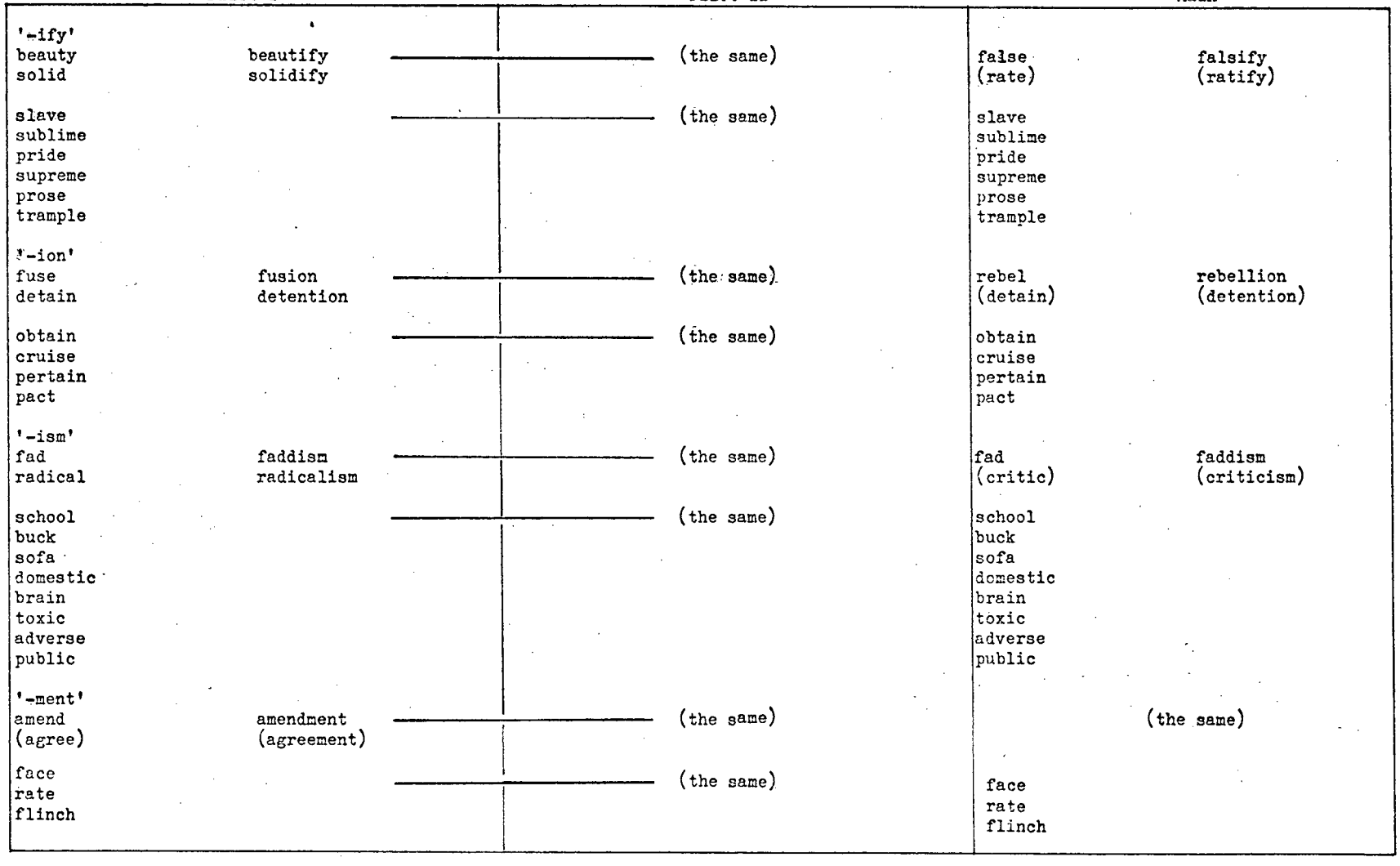
2 The directions to Group b'' in Pilot II included the sentence in parentheses. The directions to Group a'' did not.

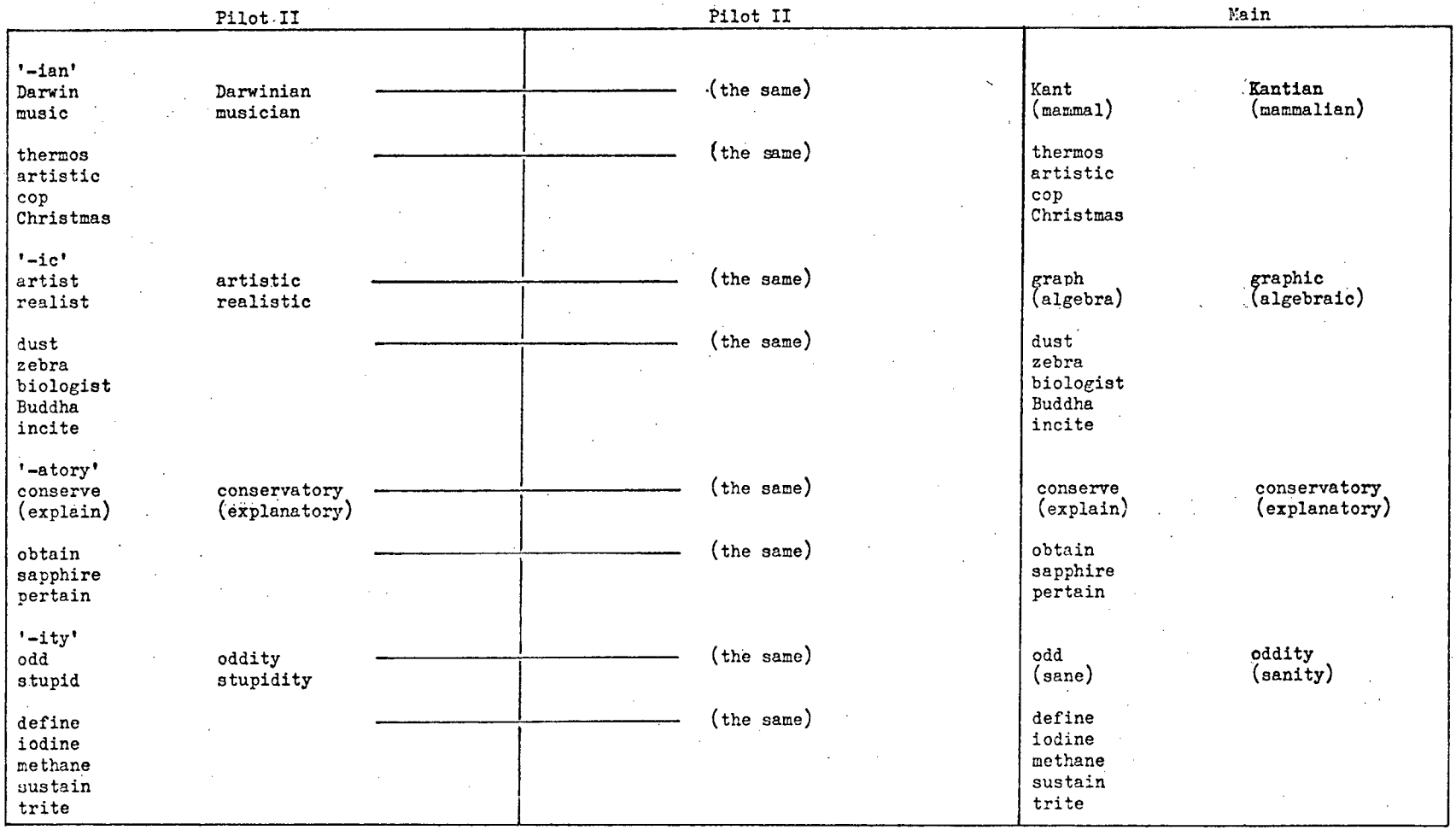
3 The directions to Group b'' in Pilot II ommitted the sentence in parentheses and included the next paragraph.


Word Lists
Practice Task
Experimental Task
Pilot I
Pilot II
Main
'-y'
burglar
master
rider
mover
faller
burglary
mastery
(the same)
burglar
master
rider
mover
faller
burglary
mastery
'-ous'
pore
industry
worm
bubble
oil
porrous
industrious
(the same)
pore
industry
worm
bubble
oil
porrous
industrious
'-ic'
scene
rhapsody
scenic
rhapsodic
(the same)
scene
angel
scenic
angelic
ovary
celery
nunnery
(the same)
ovary
celery
nunnery
'-hood'
brother
father
apple
scholar
normal
brotherhood
fatherhood
(the same)
brother
father
apple
scholar
normal
brotherhood
fatherhood




4 Recall that no example derivations were given to group b in the two pilots.
5 The example derivations enclosed in parentheses were witheld from Group B in the main experiment.


Pilot I
Pilot II
Main
'-ify'
beauty
solid
beautify
solidify
(the same)
false
(rate)
falsify
(ratify)
slave
sublime
pride
supreme
prose
trample
(the same)
slave
sublime
pride
supreme
prose
trample
'-ion'
fuse
detain
fusion
detention
(the same)
rebel
(detain)
rebellion
(detention)
obtain
cruise
pertain
pact
(the same)
obtain
cruise
pertain
pact
'-ism'
fad
radical
faddism
radicalism
(the same)
fad
(critic)
faddism
(criticism)
school
buck
sofa
domestic
brain
toxic
adverse
public
(the same)
school
buck
sofa
domestic
brain
toxic
adverse
public
'-ment'
amend
(agree)
amendment
(agreement)
(the same)
(the same)
face
rate
flinch
(the same)
face
rate
flinch


Pilot II
Pilot II
Main
'-ian'
Darwin
music
Darwinian
musician
(the same)
Kant
(mammal)
Kantian
(mammalian)
thermos
artistic
cop
Christmas
(the same)
thermos
artistic
cop
Christmas
'-ic'
artist
realist
artistic
realistic
(the same)
graph
(algebra)
graphic
(algebraic)
dust
zebra
biologist
Buddha
incite
(the same)
dust
zebra
biologist
Buddha
incite
'-atory'
conserve
(explain)
conservatory
(explanatory)
(the same)
conserve
(explain)
conservatory
(explanatory)
obtain
sapphire
pertain
(the same)
obtain
sapphire
pertain
'-ity'
odd
stupid
oddity
stupidity
(the same)
odd
(sane)
oddity
(sanity)
define
iodine
methane
sustain
trite
(the same)
define
iodine
methane
sustain
trite

# APPENDIX II

## STEM-SUFFIX SETS CONSIDERED FOR EACH RULE

Rule: V → [-tense]/__CV̆CV

between + -ity
trade + -ity
space + -ity
bribe + -ity
fleece + -ity
define + -ity
iodine + -ity
methane + -ity
sustain + -ity
trite + -ity
probe + -ity
*incite + -ic
sapphire + -atory
slave + - ify
sublime + -ify
prose + -ify
pride + -ify

Rule: V → [+tense]/ V

*sofa + -ism
zebra + -ic
Buddha + -ic

Rule: /k$^{d}$/ → s/__I

domestic + -ism
toxic + -ism
public + -ism
*artistic + -ian

Underlying vowels in pertain and obtain

obtain + -ion
pertain + -ion
obtain + -atory
pertain + -atory

Rule: V → [1 Stress]/__C + affix

Thomas + -ian
robot + -ian
human + -ian
thermos + -ian
artistic + -ian
Christmas + -ian
*adverse + -ism
zebra + -ic
biologist + -ic
Buddha + -ic
*sapphire + -atory
*iodine + -ity
*methane + -ity

Rule: V → [+tense]/__CiV̆ (V: [-high])

bed + -ian
Thomas + -ian
robot + -ian
Fred + -ian
human + -ian
cop + -ian
thermos + -ian
Christmas + -ian

Rule: /s/ → [+voice]/ V[+tense] __V

*space + -ity
*fleece + -ity
*Thomas + -ian
*thermos + -ian
*Christmas + -ian
*artistic + -ian

---

*marks those stem-suffix sets which were not preceded in Main Experiment A by an example derivation demonstrating the sound pattern predicted by the rule for which the sets were considered.