EUGÈNE JOHN BRIÈRE

A Psycholinguistic Study of Phonological Interference

MOUTON

A PSYCHOLINGUISTIC STUDY
OF
PHONOLOGICAL INTERFERENCE

JANUA LINGUARUM

STUDIA MEMORIAE
NICOLAI VAN WIJK DEDICATA

edenda curat

C. H. VAN SCHOONEVELD
INDIANA UNIVERSITY

SERIES MINOR

NR. 66

1968
MOUTON
THE HAGUE · PARIS

A PSYCHOLINGUISTIC STUDY OF PHONOLOGICAL INTERFERENCE

by

EUGÈNE JOHN BRIÈRE
UNIVERSITY OF CALIFORNIA

1968
MOUTON
THE HAGUE · PARIS

LIBRARY OF CONGRESS CATALOG CARD NUMBER 68-13339

Printed in The Netherlands by Mouton & Co., Printers, The Hague

ACKNOWLEDGEMENTS

This work was supported in part by the Center for Research in Languages and Linguistics, UCLA, Professor Jaan Puhvel, Director, and in part by the Human Communication Laboratory, Department of Psychology, UCLA, Professor Edward C. Carterette, Director.

To Professors Laurence C. Thompson, University of Washington and Peter Ladefoged, UCLA, I extend my deepest thanks for their constant direction and patient instruction throughout the entire project.

Willie Martin, Senior Electronic Technician, UCLA phonetics laboratory, gave freely of his time to assure me that the temperamental machines were kept to a very tight schedule; James Farley Anthony, Senior Research Fellow, University of Edinburgh, Scotland, gave patient explanations and constructive criticism throughout the project.

Since the suggestions and criticisms offered by my senior colleagues were not always heeded, I must claim sole responsibility for this work and admit that the outcome undoubtedly suffered as a result of my stubborness.

Special thanks must be extended to my mother, Laurette, my wife, Adelle, and my sons, John, Mike, and Dan, for accepting my long absences from home, for being amazingly quiet on the rare days that I was present, and more important, for remembering who I was when the project was finally completed.

PREFACE

Linguistics and psychology have existed as separate disciplines for many years. Although the investigation of language *per se* is clearly the major concern of linguists, many psychologists have investigated various aspects of verbal learning and verbal behavior.

Linguists have been involved primarily with developing productive theories and rigorous descriptive and analytical methods that would account for the linguistic phenomena which they found in the various languages of the world. Because linguists demanded a high degree of rigor within their discipline and because linguists felt that they had little to say (in any interesting way) about certain aspects of language, linguistic studies, until very recently, were limited to very narrowly defined areas of language. Certain issues, e.g., meaning, were termed 'exolinguistic', that is, outside the legitimate scope of investigation by the linguist. The narrow restrictions set by linguists on the areas they were willing to investigate frequently annoyed scholars from other disciplines. A psychologist friend of mine once said to me, "Unfortunately, linguists seem to be annoyed at the fact that *people* use language to say something to each other".

Psychologists, on the other hand, have long been aware of the fact that *people* learn and use languages. Further, psychologists are known to have developed a high level of sophistication in rigorous experimental procedures and methods for measuring behavior. Since the production and perception of language is a form of behavior, it seems obvious that many psychologists would be interested in determining the underlying parameters that would account for this behavior. Some psychologists have been concerned with the very questions linguists were frequently reluctant to investigate, e.g., first and second language learning and interference.

As the linguists broadened their areas of investigation to include such questions as learning and interference, the need for sophistication in experimentation and behavioral measurement became obvious. Meanwhile, some psychologists, primarily concerned with verbal learning and verbal behavior, soon felt the need for sophistication in the analysis of language.

The empirical observations and experimental investigations of verbal learning and verbal behavior which resulted from the merging of skills and information from both disciplines became known as psycholinguistics.

The investigation reported in this book is truly psycholinguistic because (1) an attempt was made to map the linguistic parameters over the psychological parameters and (2) an experimental method was devised to determine, in behavioral terms, a hierarchy of proactive interference between two competing phonological systems. The problem, the experiment, the results and their implications and the questions ensuing therefrom are discussed in detail in the following pages.

TABLE OF CONTENTS

LIST OF TABLES

I

INTRODUCTION

For some time, both linguists and psychologists have noted that interference occurs when a speaker of language X attempts to learn language Y. This interference has been attributed to the fact that between any two languages there are similarities and differences on all levels of analysis. The degree of interference that would ensue from the partial similarities and the complete differences should be a function of the degree of the similarity or difference between the two competing categories, one in the learner's native language (N) and the other in the target language (T). For example, the phonological categories of two different languages can be similar or different with respect to (1) classificatory features; (2) collective or individual membership of the established categories; (3) distributions of the categories throughout the respective systems. Linguists have assumed that by a contrastive analysis of N with T the relevant categories could be defined precisely and, by comparing them, the areas of interference between N and T could be predicted.

The linguist defines interference in terms of articulatory or classificatory features, allophonic membership of the phonemic categories, and the distributions of the phonemes within the two systems N and T. The psychologist defines retroactive or proactive interference, between first and second learning situations, in terms of the degrees of similarities between the original and interpolated learning material and in terms of convergent versus divergent learning structures. Since the linguistic parameters of the competing categories, generally, are not necessarily the same as the psychological parameters, the hierarchy of difficulties predicted by the linguist may be completely different from the hierarchy of difficulties predicted by the psychologist.

To test empirically the amount of interference that would ensue from competing phonological categories, a composite language consisting of Arabic, French, and Vietnamese utterances was written and taught to a given group of subjects (Ss), monolingual American English speakers. The hierarchy of difficulties was first theoretically defined in linguistic terms of partial similarities and complete differences of competing categories and, to the extent possible, in psychological terms of convergent, divergent, and new learning structures.

A composite language had the advantage of permitting the investigation of many more types of learning structures than would actually be found in any one natural language and yet maintained the advantage of using native speakers as models and testers of the S's responses, which would not be possible when using a purely artificial language.

A hierarchy of difficulties was then empirically obtained in terms of the total number of correct test trials, performed by Ss, over a given number of trials for each phonological category tested in T. The ensuing hierarchy of difficulties of learning the T phonological categories ranged from the 'easiest', consisting of the largest number of correct responses, to the hardest, consisting of the fewest number of correct responses.

A statistical analysis of variance was performed, and individual and grouped categories were tested for statistically significant differences between individual categories or groups of categories. The theoretical bases, the experiment, and the results are discussed in detail in the following chapters.

II

INTERFERENCE AS DEFINED BY LINGUISTIC SCIENCE

Linguists have attributed the ease or difficulty of learning phonological categories, experienced by a speaker of language X attempting to learn language Y, to: (1) the competing phonemic categories of the N and T systems, (2) the allophonic membership of the phonemic categories, and (3) the distributions of the categories within their respective systems. It has been noted that the higher the degree of similarity between the N and T phonological categories, the easier it is for the speaker to learn the T phonological categories, and the converse has been held to be true. Degrees of similarity and difference between categories have been defined in linguistic terms in the following manner.

On the phonetic level of analysis, the sound represented by the orthographic *-p-* in American English (AE) *spin* is different from the word initial sound in *pin*. Although both share many similar articulatory features (both are voiceless, bilabial, fortis,[1] stops) the *-p-* of *spin* is unaspirated whereas the *p-* of *pin* is aspirated. Thus, one might say that the two sounds differ in one articulatory feature only, i.e., aspiration. Yet, AE speakers cannot use the two

[1] We believe that previous 'definitions' of the binary features of fortis-lenis have been inadequate as testable definitions since they have failed to specify the precise articulatory or acoustic parameters of the features. *Cf.* the discussion of fortis-lenis by Roman Jakobson, Gunnar Fant, and Morris Halle, *Preliminaries to Speech Analysis* (Cambridge, Massachusetts, The MIT Press, 1963), pp. 36-39 and 57-61. The use of the terms fortis-lenis throughout this paper is made without any attempt at a concise definition, assumes that fortis-lenis, distinctions are somehow associated with degrees of tension in the muscles of the vocal organs, and adheres to the attitude conveyed by Bertil Malmberg *Structural Linguistics and Human Communication* (New York, Academic Press Inc., 1963), when he says [p'] "is *supposed to be* a 'fortis' ..." (p. 78) and [p] is "*probably* a fortis" (p. 79) (Italics mine).

different *p's* to distinguish between words. In this case, the feature of aspiration is called a NONDISTINCTIVE OR REDUNDANT FEATURE because aspiration (1) is not used to distinguish one sound from another in the AE phonological system: (2) is predictable in terms of the context in which it occurs.[2] Further, when the occurrence of one of two (or more) phonetically similar sounds, e.g. [p'] and [p], can be predicted in terms of a definable context, and when these two sounds are mutually exclusive within the defined context, then the two sounds are said to be in COMPLEMENTARY DISTRIBUTION.

The feature of voicing, on the other hand, has been considered a DISTINCTIVE FEATURE in AE since the presence or absence of voicing serves to distinguish one sound from another in the AE phonological system and "is capable of distinguishing one meaning from another"[3] e.g., the *p-* of *pin* is different from the *b-* of *bin* in that the former is voiceless and the latter is voiced and the difference in meaning between the two morphemes *pin* and *bin* has been attributed to the overt difference between *p-* and *b-*, i.e., voicing.[4] Since /p/ and /b/ are perceived as two different sounds by AE speakers, occur in the same environment, and cause a difference in meaning, they are said to CONTRAST and to constitute PHONEMES. Bloch and Trager define a phoneme as a "class of phonetically similar sounds, contrasting and mutually exclusive with all other similar classes in the language".[5] The different sounds (PHONES), e.g. [p'] and [p], to which native speakers react as one sound, *p*,

2 George L. Trager and Henry Lee Smith, Jr., *An Outline of English Structure* (Washington, D.C., American Council of Learned Societies, 1957), p. 33, describe the context in which aspirated, voiceless stops occur as "initial, internal before stressed vowels, final in 'free variation' with ... unreleased stops".

3 Bernard Bloch and George L. Trager, *Outline of Linguistic Analysis* (Baltimore, The Waverly Press, Inc., 1942), p. 24.

4 This description of voicing as the overt difference between the /p/ in *pin* and the /b/ in *bin* is an admitted oversimplification often accepted by linguists in describing hierarchies of difficulties of learning phonological categories but frequently rejected by phoneticians who have been aware, for quite some time, that English "voiced" stops, in initial and final positions, are "often completely devoiced". See Daniel Jones, *An Outline of English Phonetics* (New York and Cambridge, England, E. P. Dutton and Co., Inc., and W. Heffer and Sons, Ltd., 1956) (first published 1918).

5 Bloch and Trager, *ibid.*, p. 40.

are considered members of the same phoneme, /p/, and are called ALLOPHONES of the phoneme class. It is a linguistic convention to write phones and allophones in square brackets [ ], and phonemes between slant lines: / /.

In addition to contrast and complementary distribution there is the concept of FREE VARIATION, in which two different sounds occur in an identical environment but the occurrence of one or the other (1) is not predictable in terms of the environment; (2) does not make a difference in meaning. An example of free variation at the allophonic level is [tip'] and [tip] where the alternating occurrence of aspirated versus non-aspirated (unreleased) allophones of /p/ is not predictable in terms of the context and, further does not make a difference in meaning.

Bloch and Trager's definition of a phoneme implies a SYSTEM of distinctive, contrasting phonemic classes of allophonic members which are distributed according to definable rules. Hockett defined the phonological system of a language as "not so much a set of sounds as it is a network of difference between sounds".[6] It has been assumed by linguists that it is this very existence of a system of distinctive and non-distinctive features which causes interference when the speaker of one language attempts to learn another language in which the phonological system is composed of partially similar and completely different distinctive and non-distinctive features.

It is believed that the speakers of a language learn to attend only to those features which are distinctive and to ignore those (features) which are redundant. Bloomfield, in reference to distinctive features says that "the speaker has been trained to respond only to these features and to ignore the rest of the gross acoustic mass that reaches his ears".[7] Once a speaker has learned to attend to certain features and to ignore others, he approaches other languages through his own "grid" of distinctive versus nondistinctive features.

[6] Charles F. Hockett, *A Course in Modern Linguistics* (New York, The MacMillan Company, 1958), p. 24.

[7] Leonard Bloomfield, *Language* (New York, Henry Holt and Company, 1933), p. 79.

Troubetskoy said: "Le système phonologique d'une langue est semblable à un crible à travers lequel passe tout ce qui est dit. ... Les sons de la langue étrangère reçoivent une interprétation phonologiquement inexacte, puisqu'on les fait passer par le 'crible phonologique' de sa propre langue".[8]

Since the speaker of N will interpret the phonological system of T in terms of the features of N, interference is expected. Hans Wolff offers a short but interesting discussion of a hierarchy of difficulty encountered by speakers of Puerto Rican Spanish learning English, and notes: "For purposes of teaching and learning we distinguish two stages: I new phonemes, II familiar phonemes with new allophones and new distributions: Such a grouping represents stages of difficulty in learning the phonemes of English, with stage II the most difficult".[9] In other words, the assumption is that it is easier for everyone to learn a completely new phoneme, e.g., AE /v/ or /z/ for a Puerto Rican Spanish speaker since these sounds do not exist in Puerto Rican Spanish, than it is to learn a partially

[8] N. S. Troubetskoy (J. Cantineau, translator), *Principes de Phonologie* (Paris, Librairie Klincksieck, 1955), p. 54. In addition to the authors already cited, the following, very limited, selected discussions are listed as a ready reference for the psychologist who may be interested in pursuing in greater depth the "evolution" of the concepts of phoneme and linguistic system: Edward Sapir, "Sound Patterns in Language", *Language*, 1 (1925), pp. 37-51; Edward Sapir, "La Réalité psychologique des phonèmes", *Journal de Psychologie Normale et Pathologique*, 30 (1933), pp. 247-265 (both of the Sapir articles were reprinted in David G. Mandelbaum (ed.), *Selected Writings of Edward Sapir* (Berkeley and Los Angeles, University of California Press, 1949), pp. 33-45 and 46-60; Morris Swadesh, "The Phonemic Principle", *Language*, 10 (1934), 117-129; W. Freeman Twaddell, "On Defining the Phoneme", *Language Monograph*, no. 16 (1935); Bernard Bloch, "Phonemic Overlapping", *American Speech*, 16 (1941), pp. 278-284, Rulon Wells, "De Saussure's System of Linguistics", *Word*, 3 (1947), pp. 1-31; Martin Joos, "Description of Language Design", *Journal of the Acoustical Society of America*, 22 (1950), pp. 701-708 (the latter five articles and the second Sapir article are reprinted in Martin Joos (ed.), *Readings in Linguistics* (Washington, D.C., American Council of Learned Societies, 1957); Kenneth L. Pike, *Phonemics* (Ann Arbor, University of Michigan Press, 1947), pp. 57-67; Noam Chomsky, "The Logical Basis of Linguistic Theory", in *Proceedings of the Ninth International Congress of Linguists* (The Hague, Mouton & Co., 1964).

[9] Hans Wolff, "Partial Comparisons of the Sound Systems of English and Puerto Rican Spanish", *Language Learning*, vol. III, nos. 1 and 2 (1950), p. 38.

similar class in the target language that will involve negative transfer caused by the N system, e.g., AE /p/, /t/, and /k/ with aspirated allophones, as opposed to Puerto Rican /p/, /t/, and /k/, which only exist as non-aspirated stops or AE final /b/, /d/, or /m/ where the very similar Puerto Rican phonemes do not occur or are very restricted in final position.

Weinreich defines interference phenomena as "those instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language" and then adds that "the greater the difference between the systems, i.e., the more numerous the mutually exclusive forms and patterns in each, the greater is the learning problem and the potential areas of interference".[10]

Through a contrastive analysis of the two languages, Romansh and Schwyzertütsch, he carefully compares the phonological categories of both languages from the standpoint of distinctive versus nondistinctive features, allophonic membership of the phoneme classes, and distribution of the classes, as they are defined in each language respectively. He then analyzes the errors made by bilinguals when speaking the second language into four basic types of errors and establishes a hierarchy of importance of the classifications ranging from the least important on the phonetic level to the most important on the phonemic level.[11] He adds the

[10] Uriel Weinreich, *Languages in Contact* (New York, Publications of the Linguistic Circle of New York, 1953), p. 1.

[11] Weinreich defines the four basic types of errors as: "(1) UNDER-DIFFERENTIATION OF PHONEMES occurs when two sounds of the secondary system whose counterparts are not distinguished in the primary system are confused. ... (2) OVER-DIFFERENTIATION OF PHONEMES involves the imposition of phonemic distinctions from the primary system on the sounds of the secondary system, where they are not required. ... (3) REINTERPRETATION OF DISTINCTIONS occurs when the bilingual distinguishes phonemes of the secondary system by features which in that system are merely concomitant or redundant, but which are relevant in his primary system. ... (4) ACTUAL PHONE SUBSTITUTION, in the narrow sense of the term, applies to phonemes that are identically defined in the two languages but whose normal pronunciation differs. ... *It ought to be stressed that the above classification emerges not from the raw data directly, but from their phonemic analysis.* Viewed impressionalistically, the fact cited as (2) and (3) might not warrant being termed interference at all". *Ibid.*, pp. 18-19. (Italics mine).

note of caution that not all cases of phonic interference can be identified by a single one of the four basic types and points out that "the complicating possibility of HYPERCORRECTNESS which may operate both in listening and in speech, and which is subject to experimental testing, must always be allowed for".[12] He cites here as an example of hypercorrection on the phonological level Markwardt's report of Spanish speakers' mishearings of English final /-n/ as [ŋ] which is explained by "the bilingual's excessive caution against underdifferentiating /n/ and /ŋ/, a phonemic distinction which Spanish does not possess". It is clear that cases of hypercorrection are also attributed to conflicting systems.

Lado[13] feels that there is a hierarchy of difficulties in learning the phonological categories of a foreign language but does not specify precisely what this hierarchy is and his discussion of difficulties is essentially a paraphrase of Weinreich. He defines the areas of difficulties in terms of: (1) the distinctive versus the nondistinctive features of the two systems; (2) the allophonic membership of the phonemes; (3) the distribution of the phonemes. At the easiest end of the scale are the sounds that are physically similar to those of the native language, that structure similarly to them, and that are similarly distributed. Presumably these sounds will be learned by simple transfer without any difficulty. On the other hand, sounds that are physically different from the sounds in the native system, that structure differently, and that are distributed differently will be the most difficult for the student.

In Moulton's book, based on a contrastive analysis of English and German, he not only considers interference on the basis of competing systems but also adds a purely phonetic consideration. He says: "Phonemics will underlie much of our description of the sounds of English and German; without it we cannot say how sounds are related to each other or even how many of them there are in each language. At the same time, we must not underrate the importance of phonetics; many of the difficulties which our students

[12] *Ibid.*, p. 19.

[13] Robert Lado, *Linguistics Across Cultures* (Ann Arbor, The University of Michigan Press, 1957), pp. 1-51, *passim.*

have with German pronunication are of a phonetic rather than of a phonemic nature."[14]

In an article devoted specifically to classification of pronunciation errors,[15] Moulton classifies errors made by native speakers of AE when learning German, under four major categories, *viz.* phonemic, phonetic, allophonic, and distributional errors. Subsumed within these rubrics are eighteen types of errors, seven of which are phonemic, one phonetic, four allophonic, and six distributional. It should be noted, however, that in spite of the addition of purely phonetic considerations and the warning that a contrastive analysis should be pragmatic as well as theoretical,[16] the classification of errors (and the implied hierarchy of difficulties ensuing therefrom) is still based primarily on the theoretical concepts of (1) conflicting systems of distinctive and nondistinctive features; (2) allophonic membership of the phoneme classes; and the more empirically testable (3) distribution of the classes within their respective systems.

Stockwell and Bowen[17] have the most explicit and complete hierarchy of difficulties published. For the AE speaker learning Spanish, eight different types of learning structures are posited by comparing both languages in terms of phonological categories that have been defined as OPTIONAL (phonemes), OBLIGATORY (allophones), and ø (complete absence of a sound in one of the languages). The ensuing hierarchy of difficulties is based on

[14] William G. Moulton, *The Sounds of English and German* (Chicago, The University of Chicago Press, 1962), p. 5.

[15] William G. Moulton, "Toward A Classification of Pronunciation Errors", *The Modern Language Journal*, XLVI, no. 3 (March 1962), 101-109.

[16] A pragmatic level of contrastive analysis would simply involve listening to and recording the errors actually made by students in a learning situation. These errors can then be matched with the predicted errors based on a theoretical analysis of both languages. Where the actual errors agree with the predicted errors we can assume that the theoretical analysis was correct. Where the actual errors disagree with the theoretically predicted errors, we can assume that the analysis was inadequate and that empirical tests need be made to try to determine the actual cause of the difficulties.

[17] Robert Stockwell, and J. Donald Bowen, *The Sounds of English and Spanish* (Chicago, The University of Chicago Press, 1965).

additional considerations such as distribution of the classes within the respective systems and FUNCTIONAL LOAD which is defined as the extent to which a given sound is used to distinguish one word from another. Stockwell and Bowen organize their eight types of learning structures into three magnitudes of difficulty. The first magnitude of difficulty, i.e., the hardest for an AE speaker learning Spanish, contains the following types in the following order (1): AE ø / Spanish Ob., e.g., [-β-]; (2) AE ø / Spanish Op., e.g., *erre*, /r/; (3) AE Op. / Spanish Ob., e.g., [d-] ~ [-ð-].

It has been seen here that the majority of linguists, who have talked about interference phenomena between the N and the T, have based their predictions of degrees of difficulty of learning phonological categories primarily on the theoretical constructs of "systems of distinctive versus redundant features" and "phonemic class memberships". Even the concept of "distribution of phoneme classes" is often primarily in terms of a theoretically defined system. This paper makes the following assumptions:

(1) While the theoretical considerations are undoubtedly extremely important, descriptions and predictions of difficulty based solely on theoretical analyses at the phonemic level will be inadequate. A more complete description of phonological categories in terms of their specific articulatory features on the phonetic level is necessary. For example, saying that a 'new' phoneme in the T will be easier (or harder) than the T phonemes that are partially similar to the N phonemes is inadequate as a prediction because of the over simplification inherent in the statement. There can clearly exist a hierarchy of difficulties even among 'new' T phonemes and lumping them all together as 'new' is misleading. Arabic /x/, /ḥ/, and /γ/ are all new phonemes for the AE speaker, yet each may be significantly harder to learn than the others due to the different articulations involved at the phonetic level of production and perception.

(2) Previous attempts to quantify degrees of difference and similarity between phonemes (thereby implying a quantifiable degree of ease or difficulty for both encoder and decoder) solely in terms of the Jakobson, Fant, and Halle distinctive feature matrix

have been inadequate.[18] In the distinctive features classification of phonemes, identifying any articulatory feature as a redundant feature (therefore unnecessary in the classificatory matrix) certainly satisfies the objectives of simplicity and economy sought by the phonemicist; but has the inherent serious danger of omitting important information in determining the degrees of difficulty involved. Frequently it is precisely that articulatory feature which is being ignored as redundant for classification that may be paramount in determining the degree of difficulty when both encoder and decoder are confronted with the phonetic reality of a target language sound. Once again, it is felt that descriptions in terms of articulatory features are superior to that of a description in terms of distinctive features in determining a hierarchy of difficulties in learning phonological categories.

[18] For examples see, Sol Saporta, "Frequency of Consonant Clusters", *Language*, 31 (1955), 25-31 and the answer to Saporta, Geoffrey Bursill-Hall, "Frequency of Consonant Clusters in French", *The Journal of the Canadian Linguistic Association*, 2 (1956), pp. 66-78.

III

INTERFERENCE AS DEFINED BY PSYCHOLOGY OF LEARNING

The COMPETITION OF RESPONSE theory as formulated by McGeoch[1] posits that two response systems acquired in succession by a subject will both remain available at the time of recall. If both responses are associated with similar or identical stimuli, competition occurs and the stronger response is given. Similarity of experimental context also favors competition. This led to the INDEPENDENCE HYPOTHESIS, which asserted that the associative strength of the responses in the first list are not changed by the interpolations of a second list, and to the hypothesis of RESPONSE DOMINANCE, which posits that the correct response in one list will be displaced by a stronger incorrect response from the other list. The independence hypothesis is considered 'untenable' in light of the experimental evidence.[2] But the concept of response dominance has stimulated a great deal of experimentation among psychologists and has held exciting possibilities for psycholinguists concerned with bilingualism, language learning, and verbal behavior.

Prior to 1942, most of the psychologists who had concerned themselves with retention and forgetting, investigated the effect of interpolated learning on original list learning in terms of retroactive inhibition (interference).[3] RETROACTIVE INTERFERENCE is

[1] J. A. McGeoch, *The Psychology of Human Learning; An Introduction* (New York, Longmans, 1942).

[2] Leo Postman, "The Present Status of Interference Theory", in Charles Cofer (ed.), *Verbal Learning and Verbal Behavior* (New York, McGraw-Hill Book Company, Inc., 1961), p. 153.

[3] For example, *cf.* J. A. McGeoch, "The Influence of Degree of Interpolated Learning Upon Retroactive Inhibition", *American Journal of Psychology*, 44 (1932), pp. 695-708; Eleanor Gibson and J. J. Gibson, "Retention and the Interpolated Task", *American Journal of Psychology*, 46 (1934), pp. 606-610;

described by Underwood[4] as the influence the learning of subsequent lists has on the retention of the originally learned list. Whitely and Blankenship[5] were, in a sense, exceptions for their time in that they investigated the effects of PROACTIVE INTERFERENCE, which is defined as the decremental effects on retention of subsequently learned lists resulting from prior learning.

It is worth noting that the concepts of competition of responses and retroactive and proactive inhibition were essentially defined through the experimental procedure of PAIRED ASSOCIATE LEARNING which is, essentially, the learning of one list, say numbers, nonsense syllables, or words, which later becomes a stimulus for a response from a second list which may or may not be similar. In other words, in paired associate learning the first list may consist completely of consonantal trigrams which are to be paired with three digit numbers on the second list. The direction of most of the studies was primarily in terms of the inhibitory effects the second list learning had on the first list.

Different variables were investigated such as: the concept DIFFERENTIATION which, according to Gibson, is the relative degree of learning two responses to the same or similar stimuli or the differentiation of stimuli by attaching different responses to them;[6] RESPONSE GENERALIZATION which involves the concept that the interlist interference (or facilitation) is a function of response similarity (which led to Osgood's prediction that the amount of retroactive interference would vary inversely with the degree of

---

J. A. McGeoch, "Studies in Retroactive Inhibition: VII Retroactive Inhibition as a Function of the Length and Frequency of Presentation of the Interpolated Lists", *Journal of Experimental Psychology*, 19 (1936), pp. 674-693.

[4] B. J. Underwood, "The Effect of Successive Interpolations on Retroactive and Proactive Inhibition", *Psychological Monographs*, 59, no. 3 (1945), p. 1.

[5] P. L. Whitely and A. B. Blankenship, "The Influence of Certain Conditions Prior to Learning Upon Subsequent Recall", *Journal of Experimental Psychology*, 19 (1936), pp. 496-504.

[6] Eleanor J. Gibson, "A Systematic Application of the Concepts of Generalization and Differentiation to Verbal Learning", *Psychological Review*, 47 (1940), 196-229.

response similarity);[7] and RESPONSE MEDIATION, which is the direct mediation of second list responses by first list responses, i.e., the A-B, A-B' paradigm in which responses from the first list are practiced and strengthened during the interpolated learning of the second list. In other words, when response similarity in high, direct mediation not only provides additional practice on the first list during interpolated learning but also maintains the second list responses at high strength since "responses which are highly similar are also likely to be strongly associated".[8]

Since 1957 attention has been shifting to a focus on proactive interference — a logical sequel to Underwood's demonstration that extra-experimental interference is likely to be proactive rather than retroactive.[9] Postman's example of EXTRA-EXPERIMENTAL INTERFERENCE is the A-B, A-C paradigm, "where A is a stimulus term in the experimental list, B is a response associated with A through linguistic usage and C is the response to A prescribed in the experiment.... The evidence... makes it reasonable to assume that the extinguished habit, A-B, will gradually recover as a function of time and compete with A-C at the time of recall".[10] In other words, a subject's linguistic habits will be a strong extra experimental source of proactive interference.

This brief résumé of some of the studies of interference and facilitation is not meant to imply that proactive interference was completely ignored before the Underwood paper. Though the majority studies were in terms of retroactive interference, psychologists examined, to a lesser degree, differentiation, and response generalization on proactive interference and facilitation of second list learning. Further, most of the statements made by linguists, when positing the native language system as being the cause of interference when learning a second language, are essentially statements defined in terms of proactive interference, *viz.*

[7] Charles E. Osgood, "The Similarity Paradox in Human Learning; A Resolution", *Psychological Review*, 56 (1949), pp. 132-143.

[8] Postman, *ibid.*, p. 159.

[9] B. J. Underwood, "Interference and Forgetting", *Psychological Review*, 64 (1957), pp. 49-60.

[10] Postman, *ibid.*, p. 167.

the effect of learning the first 'list' on learning the second 'list'. Some linguists readily accepted the concepts of divergent, convergent, and unrelated learning structures (as reported by Jenkins in 1954),[11] which are stated in terms of the effects the N system has upon learning the T system. The three types are probably most clearly shown in the following chart:

	First Learning	Second Learning	Test
Type 1 Divergent	S1—R1	S1—R2	S1 < R1, R2
Type 2 Convergent	S1—R1	S2—R1	S1, S2 > R1
Type 3 Unrelated	S1—R1	S2—R2	S1—R1 S2—R2

In the first type, DIVERGENT STRUCTURE, two (or more) responses must be made to the same stimulus. Interference is expected in both places and the amount of interference is considered a function of the degree of dissimilarity of the responses — the more antagonistic the responses are to each other, the greater will be the degree of interference. In the second type, CONVERGENT STRUCTURE, one response is made to two (or more) different stimuli and facilitation can be expected. The amount of facilitation is a function of the degree of similarity of the stimuli. In the third case, UNRELATED STRUCTURE, there will neither be undue interference nor facilitation due to the nature of the structure of the learning situation, which is to make two unrelated responses to two unrelated stimuli. Jenkins says: "The first situation broadens the occasion for the use of a single response and, hence, requires less information".[12] This, then, constitutes a hierarchy of degrees of proactive inter-

[11] James J. Jenkins, "The Learning Theory Approach", in Charles Osgood and Thomas Sebeok (eds.), *Psycholinguistics* (Baltimore, Waverly Press, Inc., 1954), pp. 20-34.
[12] *Ibid.*, p. 25.

ference, therefore a hierarchy of degrees of difficulty of learning the second task, ranging from the very most in the divergent structure to the very least in the convergent situation.

Some linguists have been quick to assume that the linguistic parameters of the N and T categories involved in a given learning situation would form a one-to-one correspondence with the underlying phychological parameters. Although linguistic examples of convergent and divergent structures are given below, we believe that (1) Sl in N $\neq$ Sl in T; and that (2) Rl in N $\neq$ Rl in T. To discuss the concepts of divergence, convergence, and unrelated parallelism in going from N to T one must take into account certain assumptions that linguists have made about different languages. It must then be assumed that: (1) any N stimulus can only be similar to, and never truly 'equal' to, any T stimulus; (2) any 'S' in the formula might best be understood as "situational context" which can be defined in distributional terms rather than actually meaning 'stimulus'; (3) any R in N is probably (certainly on the phonetic level) unlike any R in T. In other words, we must qualify the statements N Sl = T Sl and N Rl = T Rl to mean no more than "similarity of context" and "similarity of response" at some, but not all, levels of analysis.

The usual linguistic example of the divergent structure is the sort of learning situation that exists when an AE speaker attempts to learn Hindi (H) voiceless bilabial stops. On one level of analysis AE /p/, when contrasted with H /p/ and /p'/ can be considered a divergent learning situation.

From the standpoint of distribution, the AE speaker must now learn to produce two phones in contrast that had previously been in complementary distribution. Essentially, this is the problem of making a different response to a similar stimulus. If Sl means a syllable-initial context and S2 means a syllable-medial or syllable-final context, and R1 = [p'] and R2 = [p], then the AE speaker's use of the aspirated or unaspirated allophones of /p/ can be described in terms of the equation S1——R1 and S2——R2.

When confronted with H /p'/ and /p/, the AE speaker's behavior now supposedly represents the classical equation used of divergent structures and not just one divergence but two, for his expected behavior in H can now be described as a situation demanding divergent behavior in two different contexts, i.e., $S1 \left\langle \begin{matrix} R1 \\ R2 \end{matrix} \right.$ and $S2 \left\langle \begin{matrix} R1. \\ R2 \end{matrix} \right.$ In other words, the 'split' category, so identified on the -emic status of phones, in AE and H, in contrast is both a problem of unlearning, i.e., inhibiting a previously made response, and learning, in addition, a 'new' response, i.e, replacing the previously made response by a response that is new and different in this particular environment.

A linguistic example of, supposedly, 'pure' convergence would be the situation involved when a Vietnamese (V) speaker attempts to learn French /t/ voiceless dental stop. On the phonemic level of analysis, V /t'/ and /t/, when contrasted with F /t/, represent a case in which the V speaker will now always make the same response to different stimuli. His /t'/ and /t/ are now collapsed into one consistent response F /t/ on all occasions. Linguists feel that facilitation will occur for the V speaker.

Learning a new category (a ø — 1 category linguistically) would represent the structure of the unrealted learning equation in which undue interference or facilitation is expected to occur. Though theoretically harder than the convergent structure described in the last paragraph, it is supposed to be less difficult than the divergent structure. Thus, the AE speaker learning V /ë/ should find this less difficult than learning H /p'/ and /p/.

The theoretical position of the psychologists in establishing the new, but unrelated, structure as being somewhat less difficult than the divergent structure does not seem to take into account whether a new phoneme to be learned in the T is composed of a regrouping of articulatory features already existing in the first system or learning a completely new set of features. An example of regrouping existing articulatory features would be the learning situation existing when AE speakers learn, say, V /ë/, a higher-mid back unrounded vowel. AE has a higher-mind front unrounded vowel, /e/, and a higher-mid back rounded vowel, /o/, but no back unrounded vowels.

Although learning V /ë/ does constitute a new learning task for the AE speaker (ø-1), the problem is to regroup the existing features of 'unrounded' and 'back' and now articulate them together to get the back unrounded /ë/. This is surely different from the learning structure that exists when the AE speaker learns French /R/, a voiced uvular trill. From the standpoint of theoretical abstractions, the situation is still a new unrelated structure, ø-1, but in this case the only feature existing in AE is voice, and the features of uvularity and trilling are completely new features that must be learned. It is clear that numerical abstractions of any given learning situation can be misleading and, in the examples just given, a so called ø-1 learning structure does not always = the ø-1 definition of another learning structure.

A composite T phonological system was compiled to incorporate different kinds of learning structures for the AE speaker. Divergent convergent, and different kinds of ø-1 categories were recorded on tapes and tested for degrees of difficulty of learning. The linguistic description of the T phonological system follows in chapter IV.

IV

THE T SYSTEM PHONEMIC INVENTORY

Since there was actually only one occurrence on the master tape[1] of each phone being investigated in the T, and since there were only fourteen "lexical" items in the entire corpus, any phonemic statement made is clearly arbitrary and other interpretations of what actually constituted '-emic' status in the new system might be made. The following 'phonemic' interpretation, however, if somewhat different from the 'usual' phonemic analysis, seems justifiable in light of establishing operational definitions of the underlying linguistic and learning structures. The phonological system of the T was defined as follows: (1) All sounds in the T system are considered significant sounds though not all sounds were investigated for difficulty of learning; (2) Every sound on which Ss received specific instruction and which had an accompanying written symbol as a referent that was different from the rest is called a phoneme. (3) The [ey]of [eym] and the [ɛ] of [ɛm], both assigned only one written symbol as a referent, ɛ, are considered allophones in free variation. (4) Allophonic membership of the phonemic categories and the articulatory descriptions thereof are based: (a) partially on previous descriptions by scholars working specifically with Arabic, French, and Vietnamese,[2] (b) on the

[1] Any references to experimental data throughout this chapter, e.g., "master tape" and "written referent", are explained in detail in chapter VII, "Experimental Method".

[2] The following works were consulted before choosing specific sounds and later as bases for descriptions of the sounds in T. *Arabic:* Walter Lehn and William Slager, "A Contrastive Study of Egyptian Arabic and American English: the Segmental Phonemes", *Language Learning*, IX, nos. 1 and 2 (1959), pp. 25-34. Nancy Kennedy, *Problems of Americans in Mastering the Pronunciation of Egyptian Arabic* (Washington, D.C., Center for Applied Linguistics, 1960), *passim:* Roman Jakobson, "Mufaxxama-The 'Emphatic'

phonetic reality of the sounds actually uttered on the master tape, (c) on the constraints imposed by the composite T system itself.[3] (5) Although all of the sounds of the T are included in the chart, only those sounds being investigated for degree of difficulty are given 'complete' descriptions and the other are labelled as sounds "not investigated" (NI).

As has been mentioned in chapter II, it is felt here that descriptions of sounds solely in terms of distinctive features is inadequate for determining interference between two systems. The following descriptions, therefore, include information concerning articulatory features that would not normally be included in a phonemic description, since some of the information is often considered unnecessary for classification or redundant on the basis of complementary distribution. The descriptions of the T sounds being investigated will include: (1) information as to phonemic or

---

Phonemes in Arabic", in Ernst Pulgram (ed.), *Studies Presented to Joshua Whatmough* (The Hague, Mouton and Company, 1957), pp. 105-117; Richard S. Harrell, *The Phonology of Coloquial Arabic* (New York, American Council of Learned Societies, 1957), *passim;* French: Lilias Armstrong, *The Phonetics of French* (London, G. Bell and Sons, Ltd., 1932), *passim;* Bertil Malmberg, *Le Système consonantique du français moderne* (Copenhague, C. W. K. Gleerup, Lund, 1943), *passim;* Robert Politzer, *Teaching French: An Introduction to Applied Linguistics* (Boston and New York, Ginn and Company, 1960), pp. 45-72; Albert Valdman, *Applied Linguistics French*, Simon Belasco, General ed. (Boston, D.C. Heath and Company, 1961), pp. 91-111. *Vietnamese:* M. B. Emeneau, *Studies in Vietnamese (Anamese) Grammar* (Berkeley and Los Angeles, University of California Press, 1951), pp. 1-44; R. B. Jones, Jr. and Huynh Sanh Thong, *Introduction to Spoken Vietnamese* (Washington, D.C., American Council of Learned Societies, 1957), pp. 1-7; Laurence C. Thompson, "Saigon Phonemics", *Language*, 35, 3 (1959), pp. 454-476.

[3] This consideration of the T system is important in light of allophonic descriptions of the phonemes since the limited corpus necessarily produced different allophonic memberships than have previously been given by other linguists for a specific language. For example, Harrell lists several allophones of /ḥ/ and /a/ in Arabic which obviously do not apply in this particular T system, since there is only one occurrence of each sound; the distribution is always the same; and the phonemes always consist of only one allophone each. Merely for the simplicity of charting, certain sounds were grouped together as allophonic members of the same phonemes, e.g., the members of /š/ and /a/, but the justification for doing so is a strictly mechanical consideration and the sounds so grouped were not being investigated for any hierarchy of difficulty.

allophonic status; (2) rather complete articulatory information; and (3) distributional information. It was found that the more complete descriptions, as used here, lend themselves to an easier contrastive analysis with the AE system when trying to determine the causes of transfer or interference. No attempt was made to separate 'phonetic' from 'phonemic' information by the use of parentheses enclosing 'phonetic' information. In other words, no hierarchy of status of information is established and all information given in the description of any given T sound being investigated is considered of equal importance.

T System Consonants

	Bi-labial	Dental	Palatal	Velar	Uvular	Pharyn-geal	Laryn-geal
Stops		t'					
		t					
	b	d					
Continu-ants			š		x	ḥ	h
			ž		γ		
Nasals	m	n		ŋ			
Semi-Vowels	w		y				

In the following descriptions an example of the T utterance is frequently given but it must be remembered that the only 'meaning' for each T utterance was, in fact, the written symbol accompanying the spoken stimulus. It will be noted further that the sounds not being investigated for hierarchy of difficulties have no accompanying written symbol as a referent even though they have been called 'phonemes'.

Bi-Labials

/b/ Phoneme with one allophone [b], a voiced stop. NI. T /bɛ̃/.
/m/ Phoneme with one allophone [m], a voiced nasal. NI. T /eym/.

/w/ Phoneme with one allophone [w], a voiced semi-vowel. NI. T /hawa/.

Dentals

/t'/ Phoneme with one allophone [t'], a voiceless aspirated lenis dental stop. Occurs syllable initial and has a written referent *th*. T /t'ao/.

/t/ Phoneme with one allophone [t], a voiceless non-aspirated fortis dental stop. Occurs syllable initial and has a written referent *t*. T /tao/.

/d/ Phoneme with one allophone [d], a voiced stop. NI. T /dəsü/.

/n/ Phoneme with one allophone [n], a voiced nasal. NI. T /ne/.

Palatals

/š/ Phoneme with two allophones: [s], a voiceless grooved fricative occurring before front rounded vowels, /ü/, and [š], a voiceless alveo-palatal grooved fricative occurring elsewhere. Both sounds were NI.

/ž/ Phoneme with one allophone [ž], a voiced alveo-palatal grooved fricative. Occurs syllable initial and has the written referent *ž*. T /žue/.

/y/ Phoneme with two allophones [i], high front unrounded non-syllabic semi-vowel before /-m/ and [I], lower high front unrounded, elsewhere. NI.

Velar

/ŋ/ Phoneme with one allophone [ŋ], a dorsal medio-velar nasal. Occurs syllable initial and has the written referent *ŋ*. T /ŋay/.

Uvulars

/x/ Phoneme with one allophone [x], a voiceless uvular fricative. Occurs syllable initial and has the written referent *x*. T /xawa/.

/γ/ Phoneme with one allophone [γ], a voiced uvular fricative. Occurs syllable initial and has the written referent *γ*. T /γawa/.

Pharyngeal

/ḥ/ Phoneme with one allophone [ḥ], a voiceless pharyngeal fricative. Occurs syllable initial and has the written referent *ḥ*. T /ḥawa/.

Laryngeal

/h/ Phoneme with one allophone [h], a "breathy voiced"[4] laryngeal fricative. Occurs syllable initial and has the written referent *h*. T /hawa/.

T System Vowels

	Front Nasal	Front Unrounded	Front Rounded	Central	Back Unrounded	Back Rounded
High			ü		ï	u
Mid		e				o
Lower mid	ɛ̃	ɛ				
Low				a		

Front Nasal

/ɛ̃/ Phoneme with one allophone [ɛ̃], a lower-mid front unrounded vowel, with nasalization throughout the duration. Occurs

[4] Peter Ladefoged defines "breathy voice" as "vocal cords slightly apart and with the edges loosely vibrating as in intervocalic [h]". "Some Possibilities In Speech Synthesis", *Language and Speech*, VII, part 4 (1964), p. 206. A spectrogram of the Arabic syllable-initial /h/ on the master tape would seem to bear out this analysis since the /h/ shows very clearly distinctive formant structures beginning with F1 through F5 and only slight turbulence throughout the spectrum. This contrasts quite markedly with the spectograms of /x/, which show 'raggedly' arrayed energy around F2 and F4 only and considerable energy above 5kcs. A spectrogram of /ḥ/ has concentrated energy around F3 and F4 only. (Cf. Appendix III).

syllable final, i.e., after consonant and before pause, C—#, and has the written referent ε̃. T /bε̃/.

Front Unroundeds

/e/ Phoneme with one allophone [e^], a high upper mid front unrounded vowel. Occurs syllable final, C—#, and has the written referent *e*. T /ne/.

/ε/ Phoneme with two allophones: [eˇ], a low upper mid front unrounded before —/y/, and [ε], lower mid front unrounded, elsewhere. Both allophones have the same written referent, ε, and are treated as allophones in free variation. Both occur syllable initial, #—C. T [eˇym] and T [εm].

Front Rounded

/ü/ Phoneme with one allophone [ü], a high front rounded vowel. Occurs syllable final, C—#, and has the written referent ü. T /dəsü/.

Central

/ə/ Phoneme with one allophone [ə], a central mid slightly rounded vowel. NI.

/a/ Phoneme with two allophones: [a], lower-low central unrounded after /ŋ/; and [a], upper-low central unrounded elsewhere. NI.

Back

/ɪ̄/ Phoneme with allophone /ɪ̄/, an upper high back unrounded vowel. Occurs syllable final, C—#, and has the written referent ï. T /šɪ̄/.

/u/ Phoneme with one allophone [u], an upper high back rounded vowel. NI.

/o/ Phoneme with one allophone [o], an upper mid back rounded vowel. NI.

Concatenation Rules

The T system utterances may be composed of the following consonant (C) and vowel (V) sequences: (1) #CV# or #VC#, e.g., /ne/ and /ɛm/; (2) #CVC#, e.g., /ŋay/; (3) #CVV# or #VCC#, e.g., /žue/ and /ɛym/; (4) #CVCV#, e.g., /hawa/.

Summary of T System

The T system has fifteen phonemic consonants and nine phonemic vowels. Of this total of twenty four, only fourteen were investigated for underlying hierarchy of difficulties as follows: from Arabic /h/, /ḥ/, /x/, and /γ/; from French /ü/, /ž/, /ɛ̃/, and /e/; from Vietnamese /t'/, /t/, /ĩ/, /ŋ/, [eˇ] and [ɛ]. All were chosen to reflect a specific learning problem for an AE speaker.

V

AE PHONEMIC INVENTORY

A complete description of every member of the phonemic inventory of AE is unnecessary here since there are many competent and complete linguistic descriptions already written and many of the sounds in the AE system clearly do not come under consideration as sources of interference or facilitation due to the peculiar structure of the T system (e.g., AE /θ/ and /ð/ are simply not involved in any of the learning structures). In chapter VI, there are given complete descriptions of the AE sounds which 'match' the T sounds being investigated for hierarchy of difficulties in an attempt to determine as precisely as possible the different variables 'causing' transfer or interference when the AE speaker attempts to learn the T system. Following are phonemic consonant and vowel charts of AE with some comments about certain details as they are pertinent to this discussion.

AE Vowel Chart

	Front	Central	Back
High	iy		uw
Lower High	I		U
Higher Mid	ey		ow
Lower Mid	ε	ə	
Higher Low	æ		ɔ
Low		a	

In addition to the glides /iy/, /ey/, /uw/, and /ow/ the following diphthongs are part of the phonemic vowel system: /ay/, /aw/, and /ɔy/.

AE Consonant Chart

		Bi-Labial	Labio-Dental	Apico-dental	Alveolar	Alveo-palata	Velar	Glottal (laryngeal)
Stops	Vl	p			t	č	k	
	Vd	b			d	j	g	
Continuants	Vl		f	θ	s	š		h
	Vd		v	ð	z	ž		
					l			
Resonants		m			n		ŋ	
		w			r	y		

All resonants in AE, whether lateral, /l/, nasal /m/,/n/,/ŋ/, or median /w/, /r/, /y/, are considered voiced, therefore voicing is not included in a phonemic description. At the phonetic level of description, voiceless resonants do occur, e.g. [płey]-/pley/-*play* has a voiceless resonant, whereas [bleym]-/bleym/-*blame* has a voiced resonant. The presence or absence of voicing, however, can be predicted in terms of complementary distribution.

VI

CONTRASTIVE ANALYSIS OF AE/T SYSTEMS

The rationale for the type of linguistic description used in the following contrastive analysis is given in Chapter II, p. 13. The contrastive analysis of AE and T sounds seemed best approached by grouping the T sounds in to six basic types of learning problems, the rationale for which will be explained throughout the chapter. In terms of T sounds, the six basic groupings are: (1) /ž/ and /ŋ/; (2) [eˇy] and [ɛˇ]; (3) /h/, /ɛ̃/; (4) /ü/ and /ı̄/; (5) /e/, /x/, /γ/, and /ḥ/; (6) /t'/ and /t/.

Group 1 — /ž/ and /ŋ/. Primarily distributional problems.

AE /ž/	=	T /ž/
Voiced	=	Voiced
Alveo-palatal	=	Alveo-palatal
Grooved	=	Grooved
Fricative	=	Fricative
Phoneme	=	Phoneme
Syllable initial	=	Syllable initial

There is a 1-1 correspondence between the AE and T sounds from the standpoint of articulatory features, phonemic status, and privilege of occurrence within the respective systems. It should be noted that previous predictions of difficulties, by Lado and Politzer,[1] for the AE speaker learning French "word initial"

[1] "English speakers will transfer their /ž/ phoneme with its limitations into French and will thus have difficulty with learning the word initial /ž/ in that language." Lado, *ibid.*, p. 17; and "Some students have also difficulty with the French /ž/ as in *général* in initial position, perhaps because the /ž/ phoneme of English as in *measure*, ... does not occur initially." Politzer, *ibid.*, p. 49.

/ž/, are in terms of distributional statements based on the 'word' as a prime. This notion is rejected here, and the syllable will be used as the prime unit for all distributional statements. On the basis of the 1-1 correspondence between AE /ž/ and T /ž/, one would expect positive transfer, rather than interference, and therefore T /ž/ should present very little difficulty for the AE speaker.

AE /ŋ/	=	T /ŋ/
Voiced	=	Voiced
Velar	=	Velar
Nasal	=	Nasal
Phoneme	=	Phoneme
Syllable final	≠	Syllable initial

The only difference between AE and T /ŋ/ is that AE /ŋ/ never occurs syllable initial[2] whereas T /ŋ/ always does. This should cause the AE speaker some difficulty in learning the T category. We can expect the underlying learning structure to present more difficulties than T /ž/ but fewer difficulties than many of the other sounds showing far fewer correspondences with the AE "phonological counterparts" (which can be a partially similar sound or simply ø).

[2] In an attempt to justify this statement in behavioral terms, an experiment was performed in which monolingual AE speakers were presented with a corpus of written monosyllabic and polysyllabic AE words printed on 4 × 6 cards. Ss were instructed to "say the words aloud but, where possible, to break the words into parts", each part to be given in time with a single beat of a metronome ticking at 80 beats per minute. This method permitted the experimenters to analyze the Ss' assignments of pre-selected intervocalic consonants to the syllable immediately preceding the consonant, (1), e.g., /mɛž + ər/, to the syllable immediately following the consonant, (2), e.g., /mɛ + žər/, or to both syllables, (B), e.g., /mɛž + žər/. Out of 100 occasions, no S ever assigned /ŋ/ to syllable 2 after a pause, i.e. there was not one single instance of any S separating *ringing*, for example, into /rI + ŋIŋ/. /ž/, on the other hand, was assigned to syllable 2, after a pause, 107 times out of 125 occasions. This lends support, in behavioral terms, to the hypothesis the /ŋ/ never occurs in syllable initial position in AE but that /ž/ frequently does. The complete experiment cited in this footnote is reported in detail in Brière, Eugène J., Russell N., Campbell, and Soemarmo, "A Need for the Syllable in Contrastive Analyses", *Journal of Verbal Learning and Verbal Behavior*, forthcoming.

Group 2 — [eˇy] and [εˇ]. Phonemically convergent; phonetically requiring a slight change of feature for each.

AE /εy/		T [eˇy]
Front	=	Front
Unrounded	=	Unrounded
Glide	=	Glide
Mid-Mid to High	≠	Higher-Mid to High
Phoneme	≠	Allophone
Syllable Initial	=	Syllable initial

Two basic areas of disagreement between the two sounds are the phonemic versus the allophonic status of each and the tongue height. On the phonemic level of description two AE phonemes /ey/ and /ε/ are now collapsed into one T phoneme /ε/ consisting of the two allophones [eˇy] and [εˇ]. Since the AE sounds are phonemic, however, one can expect overdifferentiation of the T sounds at the levels of production and perception, therefore the theoretical concept of convergence does not seem too important a consideration at the phonetic level. The difference of tongue height, however, could cause difficulty. It is assumed here that the degree of difficulty will depend upon the phonemic tolerance permitted (a) in the encoder's system and (b) in the decoder's (T) system. Though complete positive transfer is not really possible, the learning problem is such that not much difficulty is expected for the AE speaker. This assumption is based not on the notion of convergent structures but on the existence of correspondences between features on the phonetic level.

AE /ε/		T [εˇ]
Front	=	Front
Unrounded	=	Unrounded
Lower-mid	≠	Low Lower-mid
Monophthong	=	Monophthong
Phoneme	≠	Allophone
Syllable initial	=	Syllable initial

As in the immediately preceding descriptions, not much difficulty is expected for the AE speaker, since the tongue height, just slightly different, is the only area of phonetic difference between the two sounds. Convergence is considered neither an aid nor a detriment at the phonetic level.

Group 3 — /h/ and /ɛ̃/. Partially similar AE allophones now being learned as phonemes in new positions. (The following contrast is based on the most similar AE allophone compared with T phoneme.)

AE [h+]		T /h/
Partial breathy voice	≠	Breathy Voice
Laryngeal	=	Laryngeal
Fricative	=	Fricative
Allophone	≠	Phoneme
Only Intervocalic	≠	Syllable initial
AE [ɛ̃]		T /ɛ̃/
Front	=	Front
Unrounded	=	Unrounded
Lower-Mid	=	Lower-Mid
Partially Nasalized	≠	Fully Nasalized
Allophone	≠	Phoneme
Only before nasals	≠	Syllable final

Both of these categories show similar problems for the AE speaker. The AE categories show differences from the T phoneme from the standpoint of phonemic status, privilege of occurrence, and degree of a specific feature, yet, they correspond in other features. In other words, allophones in AE, partially similar to the T phonemes, must be articulated in a completely different position and changed slightly from the standpoint of articulation, *viz.* AE [h+] must be more fully voiced and AE [ɛ̃] must be more fully nasalized throughout the duration of the vowel.

Although these problems are similar, they are not quite the identical. AE [h] in syllable initial is completely voiceless. The speaker must, therefore, inhibit the voiceless allophone in the syllable initial position, substitute his partially breathy voiced allophone [h+], which never occurs in this position, and lengthen the amount of breathy voice. The AE partially nasalized [ɛ̃] occurs only before nasal consonants; so there will be no necessity for the AE speaker to inhibit any incorrect response in syllable final position since he never has a partially nasalized allophone there. On a theoretical basis one might predict that the T /h/ would be harder to learn since the speaker has the additional problem of inhibiting an incorrect response whereas in learning the T /ɛ̃/ he simply has to produce an allophone in a different environment and intensify an already existing feature that forms part of the allophonic description of the most closely corresponding AE sound. (T /t/, discussed below in terms of the popular analysis of 'split' category might well be included in this rubric, the rationale for which is discussed in p. 46, below.)

Group 4 — /ü/ and /ı̄/. Regrouping of existing features into completely new combinations.
For the AE speaker, the 'same' type of learning problem is involved when attempting to learn /ü/ or /ı̄/. AE has high front unrounded and high back rounded vowels, /i/ and /u/ respectively, therefore learning /ü/ is the problem of learning how to round AE /i/ and learning /ı̄/ is the problem of learning how to unround /u/. The problem is essentially that of regrouping existing features into new combinations. Although these should probably be learned in equal amounts of time, the /ı̄/ was included as a check against /ü/ since teaching back unrounded vowels to AE speakers has frequently been claimed by classroom teachers to be more difficult than teaching front rounded vowels. By including /ı̄/ presumably a test of statistical significance would result showing whether /ı̄/ is actually significantly harder than /ü/ or the reported observations were simply due to chance circumstances. The a priori position taken was that /ı̄/ would be somewhat harder than /ü/ although no difference could be determined in the underlying learning parameters for the two vowels.

Group 5 — /e/, /x/, /γ/, and /ḥ/. Sounds non-existent in the AE system either as phonemes or allophones and involving a reduction of features or the addition of at least one feature that is completely new to the AE system.

Although all of these sounds represent a ø-1 learning situation (i.e., there is no similar sound in AE, therefore ø, and a new category in the T, therefore 1), there is clearly a difference in the degree of difficulty involved in these sounds for the AE speaker. T /e/ is a higher-mid front unrounded monophthong phoneme, occurring in syllable final. The usual argument has been that the AE speaker will substitute his 'nearest' phoneme, commonly predicted to be AE /ey/, thereby requiring a reduction of the second part of the glide to produce the monophthong /e/. It was felt here, however that the AE speaker, without the encumbrance of the orthographic interference of the written word, might substitute the closest monophthong /I/ or /ε/. This would then represent a different learning task from that predicted if AE /ey/ were substituted. If /I/ or /ε/ were most frequently substituted, the problem would no longer be one of inhibiting the "second part" of a response but rather one of changing the position of the tongue height: lowering, in the case of an /I/ substitution, or raising, in the case of an /ε/ substitution. The apriori position was that this sound would not be difficult for AE speakers. Clearly, the underlying learning parameters for T /e/ are different from the other members of this group.

T /x/ is a voiceless uvular strongly fricative phoneme, occurring in syllable initial. T /γ/ is a voiced uvular fricative phoneme, occurring in syllable initial. Since AE has no similar phonemes or allophones, distribution can be disregarded as a consideration. The matter of the uvula as an articulator is, however, most important. AE does not make use of the uvula for the articulation of any of its speech sounds; thus, the speaker is confronted with the task of perfecting the articulation of a new sound on the basis of an articulatory feature that is completely different from anything he has in his system. The major distinction between /x/ and /γ/ is that the former is voiceless and the latter voiced. If Nancy Kennedy[3] is right, the voiceless sound should be learned much more

[3] Kennedy, *ibid.*, pp. 12 and 16.

rapidly than its voiced counterpart. Both sounds were included to permit testing of her previously stated observations.

T /ḥ/ is a voiceless pharyngeal fricative phoneme, occurring in syllable initial, with no AE correspondent on the phonemic or allophonic level. This also represents a ø-1 category but it is somehow different from /e/, /x/, and /γ/. Presumably the underlying learning parameters should be similar to /x/ and /γ/ but Kennedy's report cited this sound as one of the most difficult for AE speakers to learn. It was hoped that this test might reveal whether pharyngealization is a more difficult articulatory feature for AE speakers to learn than uvularity.

Group 6 — /t'/ and /t/. Phonemically divergent; from the phonetic point of view, /t'/ might be included with group 5 and /t/ might be included with group 3.

AE [t']		T /t'/
Voiceless	=	Voiceless
Alveolar	≠	Dental
Aspirated	=	Aspirated
Fortis	≠	Lenis
Stop	=	Stop
Allophone	≠	Phoneme
Syllable initial	=	Syllable initial
AE [t]		T /t/
Voiceless	=	Voiceless
Alveolar	≠	Dental
Non-Aspirated	=	Non-Aspirated
Fortis	=	Fortis
Stop	=	Stop
Allophone	≠	Phoneme
Syllable Medial	≠	Syllable Initial

As described in chapter III, these categories are usually considered to represent a divergent (split) learning structure for the AE speaker. The rationale for this analysis is based on considerations at the phonemic level. It is posited that when AE /t/, with allophones [t'] and [t], 'splits' into two phonemic categories in the T, sometimes /t/ and sometimes /t'/ in the same phonological environment, then the AE speaker is confronted with a divergent learning structure. On the phonetic level of analysis, however, this may not be a case of divergent learning, as defined by the psychologist, at all.

Both T /t'/ and /t/ differ from AE [t'] and [t] in essentially three ways. The first two are the same for both pairs: both T sounds are dental, whereas both corresponding AE sounds are alveolar; both T sounds are phonemic whereas the corresponding AE sounds are allophonic. Both the T and AE aspirated apical stops occur in syllable initial position but T /t'/ is lenis while AE [t'] is fortis. In contrast, T /t/ differs from AE [t] in distribution but the sounds in both languages are fortis. Since all AE voiceless stops are supposed to be fortis, positive transfer of that particular feature is possible when the AE speaker is confronted with T /t/ but not possible with T /t'/. In the latter case, the AE speaker must inhibit a feature, fortis, and learn to substitute a feature that is normally concomitant only with voiced stops, i.e., lenis. Since the V speaker seems to pay more attention to the fortis-lenis distinction than to the alveolar-dental distinction,[4] the learning problem involved in T /t'/ seems more difficult than that in T /t/. On the basis of comparisons on the phonetic level, one might well place T /t'/ in group 5 (new sounds for the AE speaker) and T /t/ in group 3 (partially similar allophones in new positions).

These sounds were included in the composite language to try to

[4] During a pilot study, using similar materials and procedures, conducted before this experiment was undertaken, the experimenter practiced with V speakers during which time he would deliberately change, without warning, various aspects of the pronunciation of certain sounds to see what the native speaker reaction would be. A substitution of a non-aspirated for an aspirated sound or the substitution of a fortis for a lenis sound (even if all other features were kept correct) was immediately and consistently rejected. The substitution of an aveolar versus a dental articulation, however, was almost never detected and, if all other features were correct, alveolar apical stops were judged as "near native performances" by the V informants. This leads the experimenter to consider the possibility that there is a hierarchy of importance of features for the decoder which (in the case of the V listener) places fortis-lenis on a higher level of distinctive importance than dental-alveolar. Additional research will soon be undertaken to try to answer this question for various languages.

determine whether they actually represent a divergent learning structure, and, if so, to test whether these two sounds, taken as a class, would be significantly more difficult than the other phonological categories presented as learning problems. The a priori position was that T /t/ would be slightly easier to learn than T /t'/.

The following chapter reports in detail the experimental method used to test the various linguistic and learning hypotheses discussed here and in the preceding chapters.

VII

EXPERIMENTAL METHOD

Subjects. Ss were chosen on the basis of a three-step selection procedure designed to control some of the exolinguistic variables which could have affected the results. Available for the experiment were a number of students at UCLA. Those wishing to participate were given an initial interview.

The first part of this interview elicited each student's age, previous education, previous foreign language training, language spoken at home by parents and grandparents, place of birth, places of residence from the date of birth through the senior year in high school, and place of birth and place of most consistent residence of parents.

The second part made use of lexical items reported by Trager-Smith[1] as having specific dialect variations, such as *land*, *law*, *wash*, *Mary*, *merry*, and *marry*. The students were asked to pronounce the written items and their responses were transcribed phonetically by the experimenter.

The third part consisted of a modified form (completely oral-aural) of the Eunice Pike test for predicting phonetic ability,[2] in which the student is asked to mimic orally each of fourteen items pronounced by the experimenter. Each item contains a specific phonological category, not usually found in American English, e.g., /ač0/ and /aᵽo/ contain voiceless /0/ and voiceless bilabial fricative /ᵽ/, respectively. The highest score possible was 28.

The actual sample population then chosen consisted of 20 graduate and undergraduate students (9 males and 11 females)

[1] George L. Trager and Henry Lee Smith, Jr., *ibid.*, *passim.*

[2] Eunice Pike, "A Test for Predicting Phonetic Ability", *Language Learning*, vol. IX, 1-2 (1959), pp. 35-43.

attending the University of California, Los Angeles, and ranging in age from 18 through 26. The average amount of formal education included one completed year of college training. Ss were monolingual speakers of 'general' American English with no noticeable dialectal variations in their pronunciations of the test items given in the second part of the interview. Parents and grandparents were monolingual American English speakers, mostly from Los Angeles County. The only foreign languages formally studied were Italian, Latin, and Spanish, by one, eight, and eleven Ss, respectively. Ss had had an average of two years of high school plus one year of college training in one of the foreign languages cited. All subjects had scored 20 points or above on the phonetic ability test. Ss were paid $ 2.00 per session and told that they would be given a bonus of $ 4.00 as soon as they had produced 4 consecutive correct test trials.[3]

Informants. Native speakers of each of the three languages used in the experiment (Arabic, French, and Vietnamese) were needed as informants to serve as models, testers, and general consultants throughout the project. Foreign students wishing to participate were interviewed and queried concerning their levels of education, their majors in school, the number of years spent in the United States, the number of years during which Arabic, French, or Vietnamese had served as the basic language of communication, their various places of residence, and the amount of time spent in each.

Prospective informants were then taken to a language laboratory containing individual booths and asked to listen to a tape on which utterances were recorded in each of the three languages. The 'same' utterances were recorded by: (1) a native speaker, (2) a non-native speaker making an attempt to pronounce the utterance as accurately as possible, and (3) a non-native speaker making a deliberate attempt to distort the utterance by using American English sound substitutions that were sometimes similar to, at

[3] Only four Ss reached the criterion performance set for the bonus. There were three males and one female.

other times completely different from those that would have been produced by a native speaker in the sequences involved. The prospective informants were then asked to judge the utterances by placing a written check in one of three columns opposite the written version of each sequence they heard. The columns were labeled '1-Good', '2-Acceptable', and '3-Not Acceptable'. 'Good' was defined as "native or near-native proficiency"; 'acceptable' as "easily understandable but with a non-native 'accent'"; 'not Acceptable' as "difficult to understand or involving a substitution of a completely different sound". The experimenter made a similar judgment chart for all languages. The informants who were finally chosen not only were able consistently to tell a native speaker from a non-native speaker but also showed a high degree of agreement among themselves and with the experimenter on all three kinds of judgments.

Four informants were chosen for each language. One was to serve as the speaking model on the tapes, two were to serve as testers, and one was to serve as a general consultant. Among the Arabic informants were included one speaker each of Egyptian, Moroccan, and two speakers of Palestinian Arabic. The French informants included one speaker each of Belgian, Congolese, Moroccan, and Parisian French. Vietnamese informants included two speakers of the Hanoi dialect and two speakers of the Saigon dialect. All informants had attended universities in their native languages. All French and Vietnamese informants had, at one time, attended universities in France and Saigon, respectively. Male speakers of Palestinian Arabic, Parisian French, and Saigonese Vietnamese were used as the speaking models on the tapes.

TAPES

Master Tape. — Informants recorded in a sound-proofed room using a high quality Sony condenser microphone and a Sony dual-track 777 recorder which is maintained to have a frequency response of ± 2 dbs from 50 to 10,000 cps. The master tape consisted of four Arabic, six Vietnamese, and four French utterances as follows: *Arabic:* /hawa/, /ḥawa/, /xawa/, /ɣawa/; *Vietnamese:* /t'ao/, /tao/, /šï/, /ŋay/, /eym/, /ɛm/; *French:* /dəsü/, /žue/, /bɛ̃/, and /ne/. Exactly five seconds were allowed between each utterance on the tape. From this master tape, twelve learning tapes and twelve testing tapes were made.

Learning tapes. — All of the Arabic, Vietnamese, and French items were numbered respectively as sections 1, 2, and 3. By using two Sony 777 recorders, one as a playback and one as a recorder, the order of the sections (but not the order of the individual items within each section) was rearranged so that there were two learning tapes each of the orders 1-2-3, 1-3-2, 2-1-3, 2-3-1, 3-1-2, and 3-2-1.[4]

Testing Tapes. — Every individual item on the master tape was assigned a number from 1 through 14. A random order chart of fourteen items and twelve different orders was compiled. Once again, by using one Sony 777 as a playback and one as a recorder, the order of individual items on the master tape was changed according to the random order chart. The resulting twelve tapes were labeled "testing tapes". One each of the twelve testing tapes was spliced after each learning tape with a brightly colored lead in between. Three reels of eight tapes each were compiled, four learning tapes alternating with four testing tapes.

Spectographic Analysis of Individual Items. — Two spectograms were made on a Kay Sonograph of each utterance on the master tape, one over a 4.5 kcs range and one over 8 kcs. All spectrograms were analyzed by the experimenter before attempting to teach the individual phonological categories.

Locale and Material. The entire experiment was conducted in a sound-proofed recording chamber. Ss and the experimenter were

[4] The reordering of the sections was done to insure equal presentation of all three sections at the beginning, the middle, and the end of the list. The order of the individual items within the sections was maintained so that members of all minimal pairs would immediately follow each other regardless of the order of the sections on any individual learning tape.

seated at a table on which were placed a high quality Sony condenser microphone, a high quality loud speaker and two sets of remote controls, permitting simultaneous operation of both Sony 777 recorders placed outside the chamber. The recorders could be observed through a window in the chamber. Each S was given a chart with columns numbered 1-4 and rows numbered from 1-14. The experimenter had four sets of 2×4 flash cards on which were printed the symbols that were to be associated with the words heard from the loud speaker. The set of cards that was given to the S matched the order of items on the particular learning tape being played at the moment, e.g., on learning tape 1 the order of the sections was 1-2-3 and the order of items on card set 1 was /h/, /ḥ/, /x/, /γ/, /t'/, /t/, /ï/, /ŋ/, /ɛ/, /ɛ/, /ü/, /ž/, /ɛ̃/, and /e/.[5] In addition, the experimenter had a master chart for each subject, which included spaces for transcribing all sounds made by the S on all trials.

Procedure. — The learning and testing tapes, the utterances of which constituted the stimuli, were played on one of the recorders. The stimuli were heard through the loud speaker in the chamber and were simultaneously recorded on the second recorder. Each S's responses, immediately following the stimuli, were also recorded on the second recorder. This produced a series of tapes for each subject consisting of learning and testing stimuli immediately followed by their corresponding responses for each trial.

Each S accomplished his work in three consecutive sessions, 24 hours apart. Sessions averaged 30 minutes each and were divided into a learning period, devoted to instruction and practice (involving four learning tapes) and a testing period (utilizing 4 testing tapes).

[5] The symbols chosen were essentially a phonemic transcription of each phonological category being tested for difficulty. This 1-1 correspondence between sound and referent was established in an attempt to focus attention on the target sound and to minimize the varying degrees of interference encountered in paired-associate learning when non-sense syllables, numbers, or drawings are used as one or both of the pairs, since all have been shown to have different and constantly changing associative values.

At the beginning of the first session, Ss were instructed: (1) to listen carefully to the word that came from the loud speaker, (2) to respond immediately[6] into the microphone mimicking the sound as closely as possible, (3) to look at the 2×4 cards in the order given and to memorize the symbols thereon, (4) to memorize the words on the loudspeaker as the names of the symbols in the particular order given. They were directed to turn the uppermost cards after each sound. Learning tape 1 was then played. Ss responded orally and turned the cards as they responded. The experimenter transcribed each response phonetically in the appropriate square on each S's master chart. After the learning tape, Ss were instructed that they would now (1) hear the same words in a different order, (2) respond immediately into the microphone, once again mimicking the sound as closely as possible, (3) try to write, in the appropriate column on the chart provided, the symbol that had been associated with each particular word on the learning tape just played.[7] Testing tape 1 was then played. Ss responded orally and tried to write symbols at the same time. The experimenter transcribed phonetically all responses on the master chart.

After the testing tape, instruction was given on all of the sound categories that were being tested. A timer was used throughout the experiment to ensure exact timing of the instruction and practice periods. For the first session, five minutes each were

[6] Fluent performance in a language consists of automatic, rather than carefully intellectualized responses. If this experiment were to be a test of hierarchy of the difficulties that occur in actual language learning, it was thought best to encourage an automatic response at the very beginning stages of learning. A second reason for encouraging an immediate response was the mechanical consideration that the Kay Sonograph can record only sequences lasitng 2.4 seconds or less. By keeping both stimulus and response for any one item within this time limit, they can be more easily compared, since both appear on a single spectrogram for any given trial.

[7] For the purposes of this study, there was no interest in the paired-associate learning that ensued during the experiment. It was felt, nevertheless, that forcing the Ss to concentrate on remembering and writing the symbols would more closely resemble actual language performance in which a speaker usually concentrates on the content (referent) of the message rather than on the specific physiological processes that make up the phonological expression of the content.

given to Arabic and French sounds and $7\frac{1}{2}$ minutes were given to Vietnamese sounds. Instruction consisted of telling the Ss what the target sounds were, showing how the individual sounds compared with similar English sounds (if any) and how the target sounds compared with each other in this particular system, giving a physiological description of each, showing and explaining a sagittal drawing of each sound, giving examples of each sound, and instructing the Ss to practice after each sound was produced by the experimenter.[8] Learning tape 2 followed by testing tape 2 were then played.

Throughout the experiment: (1) a learning tape was always immediately followed by a testing tape and (2) a five minute instruction period always preceded a learning tape except for learning tape 5, which marked the beginning of the second session.[9]

After the first session, all instruction periods focused on only those sounds that were missed on an immediately preceding testing series. Psysiological descriptions and sagittal drawings were reviewed. By manipulating the learning tapes Ss were able to listen to the native speaker on the tape and then repeat after him. Ss were trained in perception of minimal pairs by scrambling the order of the learning tapes (through the use of fast forward and fast reverse remote controls), having Ss identify a particular sound and then produce it. Specific problems were handled by giving specific instructions, e.g., Ss were told to repeat after the experimenter in a series of lengthenings of the medial /ŋ/ in /rɪŋɪŋ/ and after an exceptionally long /rɪŋ: : ɪŋ/, the target utterance /ŋay/.

Testing. — Informants chosen for the experiment were first trained by judging tapes taken from a small pilot study that had been

[8] The experimenter practiced every day with an informant from each of the three languages being used. His pronunciation of all items was judged "near native" by all informants throughout the experiment.

[9] No instruction was given at the beginning of the second session in order to enable the experimenter to compare the scores of each subject on the first and last trials of the first session with the uninstructed-trial scores at the beginning of the second session. By so doing, it was hoped that a measure of forgetting could be achieved. The results, unimportant to the considerations in this paper, will be reported elsewhere.

conducted before the main experiment began. They were given 8 hours of training in listening and judging before taking part in the actual experiment. Informants were trained to listen only to the specific sound being tested, e.g., the response /žuɛ/ to the stimulus /žue/ was counted as correct on the basis of the correct production of /ž/ which was the target sound. This procedure was followed throughout the actual experiment. The Ss individual tapes, containing both stimulus and response, were then evaluated as follows:

(1) Two native informants for each language listened to four individual tapes each day. Tapes averaged 9½ minutes each. The informants sat in individual booths in a language laboratory and the experimenter played the tapes on an Ampex 112. Utterances were heard simultaneously on a high quality speaker at the front of the room and on earphones in the individual booths. Informants were allowed to use whichever source seemed the most comfortable to them. Informants were given individual language charts on which the specific sounds of a particular language were written as heads of rows, while the judgments 'Good', 'Acceptable', and 'Not Acceptable', were written as heads of columns. The chart was divided into learning trials and testing trials. Informants could not see each others' judgments. Any specific part of a tape was replayed as soon as an informant signaled his desire for a replay. The judgments given in the first session were transferred to a master chart, compared with one another, and checked against the experimenter's phonetic transcriptions of each S's renditions, recorded at the time of the S's actual performances. These comparisons revealed slightly more than 4% disagreement between informants or between the informants and the experimenter on the 6,720 trials recorded. All judgments of individual responses which included contradictory ratings of 'Good' and 'Not Acceptable' were considered instances of disagreement.

(2) Every category on which there had been disagreement was then abstracted from the Ss tapes and re-recorded so that all Arabic items were assembled on one tape, all French items on another, and all Vietnamese items on a third. The two informants for each specific language then met in the recording chamber with

the experimenter. The re-recorded tapes were played, one language at a time. Only the informants needed for the language concerned were present in the chamber with the experimenter during the playing of these "initial disagreement tapes". At this time both of the informants and the experimenter made judgments and compared ratings. The disagreement was reduced to slightly less than 1%.

(3) For the remaining 1%, "pooled judgments" were used in which the informants and the experimenter discussed the reasons for a particular rating and then reached, where possible, a final judgment satisfactory to all. The disagreement was now reduced to 12 specific items.

(4) Broad band spectograms were made of the remaining twelve items. Vowels were analyzed at 4.5 kcs and 8 kcs and consonants were analyzed at 8 kcs. The additional acoustic information coupled with the results just described enabled the experimenter to make a final decision on all consonants. However, it was still impossible, even with the spectrograms, to reach a decision on six specific vowels: two examples of /ï/ and four examples of /ü/. These remaining six items were randomly assigned as 'Acceptable' or 'Not Acceptable'.[10]

Method of Determining Hierarchy. — Arbitrary numerical values were assigned to all informant judgments (1 for 'Good', 2 for 'Acceptable', and 3 for 'Not Acceptable') and then transferred to the appropriate squares on a master chart for each S. The averages in each square were then taken and all averages of 1-1.5 were counted as correct and all averages greater than 1.5 were counted as incorrect. Learning and testing trials were counted separately.

A mean number of correct trials was compiled for every category for the total sample population. The resulting hierarchy is based on the mean number correct of test trials, the largest number of

[10] The number of items randomly assigned represents approximately .0008 of the total number of responses.

correct responses representing the 'easiest' category and the smallest number of correct responses representing the 'hardest' category. Learning curves were plotted for each sound category by plotting the number of correct responses made by the total population on any *nth* trial.

VIII

RESULTS

The resulting hierarchies of difficulties for both learning and testing trials are given in Table I below. It will be noted that there were very slight changes in the order of individual categories between learning trials and testing trials, but that the changes occurred between items that were not significantly different from each other. For example, in the learning trials, /x/ scored slightly higher than /ey/ (.1 of one learning trial), but in the testing trials /ey/ scored slightly higher than /x/ (.25 of one testing trial). The discussion below is based on the hierarchy as established by the testing trials only. It is felt that the testing trials (items presented in random order rather than in minimal pair sections) permitted less inter-item associative connections, permitted less serial order expectancy, and resembled more closely a language learning situation as experienced in everyday life. It is also interesting to note that (even though it was not statistically significant) except for /ŋ/, all categories received the same or slightly higher scores in the learning trials.

A test for significant differences between individual categories, based on the hierarchy established by the test trials, was made. The test used was the Duncan "multiple range test"[1] and any ensuing discussion of significant differences between individual categories is based on the Duncan "Significant Studentized Ranges" with *alpha* = to the .05 level of significance.[2]

It should be noted that there has been some question regarding

[1] D. B. Duncan, "Multiple Range and Multiple *F* Tests", *Biometrics*, 11 (1955) pp. 1-42.

[2] A chart of critical values at the .05 level, as compiled by H. Leon Harter, can be found in Allen Edwards, *Experimental Design in Psychological Research* (New York, Rinehart and Company, Inc., 1960), p. 373.

TABLE I

Hierarchy of Difficulties

LEARNING		TESTING	
ž	12.00	ž	12.00
ε	11.90	ε	11.70
x	11.10	ey	10.75
ey	11.00	x	10.55
e	10.80	e	10.00
h	10.55	t	9.95
t	10.10	h	9.75
ε̃	9.45	ε̃	9.35
ü	9.00	ŋ	9.20
ŋ	8.90	ü	8.50
γ	8.90	γ	8.45
t'	8.90	t'	8.25
ḥ	4.25	ī	4.35
ī	4.05	ḥ	3.20

Hierarchy of difficulties, ranging from the easiest to the hardest, based on mean number of correct learning and testing trials.

the principles underlying the sampling distributions which Duncan used in obtaining the critical values for the tests (cf., for example, Scheffé).[3] More conservative tests of significance between multiple means have been devised (e.g., by Newman-Keuls, Tukey, and Scheffé).

A table which compares the different tests directly is given in Winer[4] (reproduced below) in which it is clearly shown that the higher the number of degrees of freedom, the higher is the critical value required as the tests become more and more conservative.

Finney points out that a researcher is perfectly justified in using a less conservative test "if he is particularly anxious not to miss any indications of departure from hypothesis".[5] Finney's statement serves as the rationale for choosing the Duncan test for this paper. In other words, in light of the 'pioneer' aspect of this experiment,

[3] H. A. Scheffé, *The Analysis of Variance* (New York, Wiley, 1960), p. 78.
[4] B. J. Winer, *Statistical Principles in Experimental Design.* (New York, McGraw Hill Company, Inc., 1962), p. 88.
[5] D. J. Finney, *Experimental Design and Its Statistical Basis* (Chicago, The University of Chicago Press, 1955), p. 13.

TABLE II

Comparison of Different Tests of Significance of Multiple Means and the Required Values for Significance at the .01 Level per Number of K Treatments[6]

Method k:	2	3	4	5	6	7
Scheffé	13.02	13.02	13.02	13.02	13.02	13.02
Tukey (a)	10.90	10.90	10.90	10.90	10.90	10.90
Tukey (b)	9.36	9.93	10.29	10.54	10.74	10.90
Newman-Keuls	7.82	8.96	9.68	10.18	10.56	10.90
Duncan	7.82	8.16	8.36	8.56	8.68	8.78
Individual Comparisons	7.82	7.82	7.82	7.82	7.82	7.82

it is believed that the more statistically significant categories there are to talk about, the more very much needed research will be generated in the following critical areas: (1) the different concepts of establishing hierarchies of difficulties of learning different phonological categories (e.g., research should be conducted to determine the optimum primes to be used in a contrastive analysis); (2) the concept of a hierarchy of importance of features to encoder and decoder from various languages; (3) the possibilities of establishing a more empirically determined set of 'distinctive' features for specific sound categories; (4) the possible programming techniques for teaching that should be developed from the empirical findings.[7]

Table III indicates the categories whose scores are significantly different from each other at the .05 level of confidence.

[6] B. J. Winter, *Statistical Principles in Experimental Design* (New York, McGraw-Hill Book Company, 1962), p. 88.

[7] Some people (e.g., Stanley Sapon, personal communication) feel that the "hierarchy of difficulty" is in the "teacher", therefore a category, if properly taught, would present no hierarchy of difficulty for the student at all. In other words, it could well be that the more we know about two conflicting sound categories, the less of a "hierarchy of difficulty" would be presented to the student. Although the results of this experiment would not seem to bear this out, the possibility does exist that any significant differences may be due to a lack of understanding, on the teacher's part, of the underlying distinctive parameters of specific sound categories.

TABLE III

Individual Comparisons

		ḥ	ī	t'	γ	ü	ŋ	ɛ̃	h	t	e	x	ey	ɛ	ž
		64	87	165	169	170	184	187	195	199	200	211	215	234	240
ḥ	64	0	23	101	105	106	120	123	131	135	136	147	151	170	176
ī	87		0	78	82	83	97	100	108	112	113	124	128	147	153
t'	165			0	4	5	19	22	30	34	35	46	50	69	75
γ	169				0	1	15	18	26	30	31	42	46	65	71
ü	170					0	14	17	25	29	30	41	45	64	70
ŋ	184						0	3	11	15	16	27	31	50	56
ɛ̃	187							0	8	12	13	24	28	47	53
h	195								0	4	5	16	20	39	45
t	199									0	1	12	16	35	41
e	200										0	11	15	34	40
x	211											0	4	23	29
ey	215												0	19	25
ɛ	234													0	6
ž	240														0

Differences between individual categories in terms of raw scores of total number of correct responses, on testing trials, for the entire sample population. Significantly different scores, based on the Duncan multiple range test, are underlined.

Before describing the results of the comparisons of the multiple means, it should first be noted that there was no significant difference between subjects ($F = 1.47$, $df = 1/19$, $P > .10$), and that the difference between the sound categories themselves was highly significant ($F = 24.75$, $df = 1/13$, $P \leqslant .001$).

From Table III, it can be seen that the following categories are statistically significant:

(1) /ḥ/ and /ï/ are significantly different from (harder than), all other sounds but not significantly different from each other.

(2) /t'/ is significantly easier than /ḥ/ and /ï/ but significantly harder than /t/, /e/, /x/, /ey/, /ɛ/, and /ž/.
(3) /γ/ and /ü/ are significantly easier than /ḥ/ and /ï/, significantly harder than /x/, /ey/, /ɛ/, and /ž/, but not significantly different from each other.
(4) /ŋ/, /ɛ̃/, /h/, /t/, and /e/ are significantly easier than /ḥ/ and /ï/, significantly harder than /ɛ/ and /ž/, but not significantly different from each other.
(5) /x/, /ey/, /ɛ/, and /ž/ are significantly easier than /ḥ/, /ï/, /t'/, /γ/, and /ü/, but not significantly different from each other.

In addition to the multiple comparisons on Table III, orthogonal comparisons were made of the following sounds, taken as a class, versus the rest of the sounds, taken as a class, to test for significant differences between the classes. Three classes (two not yet mentioned), were tested: reasons for testing these specific groups will be discussed below.

(A) An 'OUT' class, consisting of /ḥ/, /x/, /γ/, /t'/, /ï/, /ü/, and /e/, versus an 'IN' class consisting of /h/, /t/, /ŋ/, /ey/, /ɛ̃/, /ɛ/, and /ž/. The 'OUT' group was highly significantly different from the 'IN' group ($F = 101.58$, $P \ll .001$).
(B) /h/ and /ḥ/ taken as a class were highly significantly different from the rest of the sounds taken as a class ($F = 59.27$, $P \ll .001$).
(C) /t'/ and /t/ as a class were not significantly different from the rest of the sounds taken as a class ($F = .47$, $P \gg .25$).

The discussion of the results already shown is as follows: (1) significant groups shown in Table III; (2) groups tested for significance by means of orthogonal comparisons; (3) productive versus perceptual phenomena.

Group 1 contains a sound /ḥ/, requiring a completely new articulatory feature (pharyngealization), and a sound /ï/, requiring a regrouping of existing features (high-back-unrounded). Since there was no significant difference between /ḥ/ and /ï/, one can assume that learning the new feature of pharyngealization and learning how to make back-unrounded vowels are equally difficult for the AE speaker. Since /ü/ was significantly easier than /ï/, one can recognize a clear basis for the assumptions of many teachers (cf. p. 42 above) that teaching front-rounded vowels to AE speakers is easier than teaching back-unrounded vowels. The fact that there is a clear difference in the degree of difficulty involved in the learning of the two members of the class labelled "regrouping of existing features" and that there is no significant difference between a sound category requiring the learning of a completely new feature and one of the members of the 'regrouping' class makes one feel that such a labelling is artificial and not very useful in predicting a hierarchy of difficulty. It is felt here that in any contrastive analysis descriptions in terms of articulatory features on the phonetic level, rather than descriptions in terms that are abstractions quite removed from the phonetic reality of the sounds, will produce a higher degree of accuracy of prediction of specific difficulties.

On the basis of the description given at the phonetic level in the contrastive analysis in chapter VI, it is not surprising that /t'/ (Group 2) turns out to be a difficult sound for the AE speaker. Since the AE speakers' substitutions of a fortis allophone required the inhibition of the feature fortis and the substitution of the feature lenis, one would expect some difficulty with this sound. What is rather surprising is that /t'/ was significantly more difficult than /t/, an allophone never found in syllable initial position in AE. This would indicate that transplanting an allophone from the N system to a new position in the T system is easier than learning to produce and perceive completely new features.

One word of caution is necessary. The features fortis-lenis *seem* to be extremely important for V speakers in determining whether a response is acceptable or not acceptable. Since fortis-lenis have, so far, not been clearly defined in specific articulatory or acoustic parameters to everyone's satisfaction, a great deal of research is necessary to determine the specific features being used by the V encoders in classifying their sounds. Length of aspiration was considered and checked on spectrograms. Many responses made by the subjects and rated "near native proficiency" by the V informants, had widely varying lengths of aspiration (as measured on spectrograms), whereas many responses containing the same length of aspiration as the stimulus 'word' were rated 'not acceptable'. The problem may lie in onset of aspiration versus the closure, or in some yet undetermined features. Therefore, the difference in degree of difficulty between /t'/ and /t/ might simply be due to a lack of complete understanding of the precise underlying parameters of the two sounds, since this would seriously affect the efficiency of the teaching sessions concerned with them.

Group 3, /γ/ and /ü/, once again combines a "new feature" category with a "regrouping-of-existing-features" category. /ü/ has already been discussed in relation to /ï/. The fact the /γ/ is significantly harder than /x/ is an interesting phenomenon. Both sounds are made in the same manner except that the former is voiced and the latter is voiceless. The difference in difficulty between /γ/ and /x/ is further explored in the discussion of group A.

Group 4 /ŋ/, /ε̃/, /h/, /t/, and /e/ is primarily interesting from the standpoint that it combines into one class, showing no significant difference among the members, sounds that were predicted to involve different kinds of learning problems. /ŋ/ and /t/, phonemic and allophonic, respectively, in the AE system, must now be learned in a new context. /ε̃/ and /h/ require articulatory changes from the partially similar allophones in AE and, in addition, they also must be learned in a new context. /e/ is a 'completely' new sound. Expected differences of difficulty among these sounds simply did not appear. All were significantly more difficult than /ε/ and /ž/, which might be expected, since the latter two sounds exhibited perfect to almost perfect positive transfer, respectively, from the AE system.

Group 5 /x/, /ey/, /ε/, and /ž/ (not significantly different from each other) includes a completely new sound /x/ and three sounds permitting complete or nearly complete positive transfer from the AE system to the T system. Except for /x/, one would expect the latter three sounds to be significantly easier than /ḥ/, /ï/, /t'/, /γ/, and /ü/.

Group A: On the basis of the groupings found to be significant from other groupings in Table III, but whose individual members were not significantly different from each other, this comparison was made to see if an 'OUT' group, consisting of completely new sounds in the T system, would be significantly different from an 'IN' group composed of sounds that are 'present' in the AE system, i.e., relatively similar to partially corresponding sounds in the T system, whether the sounds were classified as allophonic or phonemic in either system. The 'IN' group was highly significantly easier than the 'OUT' group. In other words, regardless whether a sound is in the AE system as a phoneme (e.g., /ŋ/) or as an allophone (e.g., /t/) these sounds were learned much more rapidly than sounds that are not in the AE system at all (e.g., /γ/).

The only notable exception was /x/, which is neither allophonic, nor phonemic, in AE but was learned with amazing rapidity. Since careful control was exercised to exclude native language backgrounds that contained this sound, one can only conclude that: (1) there are certain articulatory features of /x/ which have not been discussed in the literature that have partially corresponding features in AE, permiting a great deal of positive transfer for the voiceless /x/ but not for the voiced /γ/, or (2) the Arabic tolerance for the voiceless, so called, uvular fricative is extremely wide and varies from a pre-velar articulation to the usually cited uvular articulation.[8] It would seem that the same tolerance does not exist for /γ/. The degree of length of aspiration was considered a possible factor but, as in /t'/ and /t/, the Ss' aspirations were timed and widely varying lengths of aspiration were accepted by informants as "near native". A great deal of research on the phonetic level is necessary to determine the precise classificatory features attended to by the Arabic speaker when confronted with either of the stimuli /x/ or /γ/.

Before discussing the comparisons of groups B and C, the

[8] During his practice sessions with the Arabic informants, the experimenter tried varying points of articulation ranging from pre-velar to uvular. The Arab informants rated as "near native proficiency" sounds that were produced over an amazingly large number of different points of articulation.

introduction and explanation of the following tables are necessary.

Tables IV and V are confusion matrices for learning and testing trials, respectively. They represent the sounds actually occurring as Ss' responses (shown on the abscissa) for a particular given T stimulus (on the ordinate). The charts are divided into a T system response, on the left side, and an AE system response, on the right side. Included with the AE system responses is the exception /R/, uvular trill, which is not from the AE system but more typically French or German. Sixteen additional symbols are required for AE substitutions on the learning trials and eighteen additional symbols for the testing trials. There were 578 AE substitutions made on the learning trials and 596 substitutions made on the testing trials.

Of the symbols shown on the chart, the following should perhaps be described: /hh/, a heavily fricative, voiceless, AE /h/; /ʔ/, glottal stop; /ø/, no response whatsoever to the syllable initial consonantal stimulus; /*t*'/, AE voiceless fortis alveolar aspirated stop; /ny/ palatalized alveolar nasal; /ng/ alveolar nasal followed by a voiced velar stop; /ö/, mean mid front rounded vowel. The other transcriptions are the common symbols for the specific sound referents. The discussion of the substitutions made on the chart are in terms of the test trials (Table V) only.

From Table V, it can be seen that:

(1) /h/, /ḥ/, /x/, and /hh/ formed a perceptual confusion group with /x/ rapidly dropping out as a perceptual or productive difficulty but with /h/, /ḥ/, and /hh/ persisting for some time. /ḥ/ caused the greatest difficulty from the standpoint of both perception and production and /hh/ was most commonly used as a substitution when an attempt at producing /ḥ/ was made. /ḥ/ was substituted 20 times for /h/.[9]

(2) Though the AE /g/ served as the most frequent substitution for /γ/, it is evident that the prediction of the precise sound to be substituted by a speaker of X when learning Y can not always be

[9] Not apparent from the chart was the constant complaint by the subjects that they could "make" the sounds but that they were unable to distinguish between them on the test trials.

TABLE IV

Confusion Matrix for Learning Trials

	T SYSTEM														AE SYSTEM															
	h	ḥ	x	γ	t'	t	ï	ŋ	ey	ε	ü	ž	ẽ	e	hh	?	ø	g	r	t'	d	uw	I	ng	n	ö	æ		iy	Ey
h	**211**	10	4												15															
ḥ	43	**85**	39												75															
x	6	2	**222**												10															
γ	2			**178**												5		48	7											
t'					**178**	29														33										
t					5	**202**														6	27									
ï							**81**				1											134	21			3				
ŋ								**178**								1	1	19						20	21					
ey									**220**	14																				6
ε									1	**238**																	1			
ü											**180**											38					22			
ž												**240**																		
ẽ													**189**														37	14		
e									1	9				**216**									10						2	2

TABLE V

Confusion Matrix for Testing Trials

	T SYSTEM														AE SYSTEM																	
	h	ḥ	x	γ	t'	t	ï	ŋ	ey	ε	ü	ž	ε̃	e	hh	?	ø	g	r	R	*t'*	d	uw	I	n	ny	ng	ö	æ		iy	Ey
h	**195**	20	11												14																	
ḥ	66	**64**	40												70																	
x	10	4	**213**												13																	
γ	3			**169**												6	4	49	5	4												
t'					**164**	60															16											
t					13	**199**															16	21										
ï							**86**																133	14				7				
ŋ								**186**								2	2	30							12	3	5					
ey									**215**	6																						19
ε									1	**235**				1															3			
ü											**174**												39					27				
ž												**240**																				
ε̃													**187**																36	17		
e									3	15				**200**										9							4	9

made in terms of the conflicting structures alone. AE speakers also substituted /ʔ/, /ø/, /r/, and /R/, the latter of which is not even in their system. In other words, the 'psychoacoustics' of the encoder's perception of a given phonological stimulus need not be completely governed by the N system alone and many other factors need to be explored to explain a particular substitution for a specific subject.

(3) One rather surprising substitution pattern was that which existed for T /t'/. Although AE /t'/ was substituted 16 times, unaspirated T /t/ was substituted 60 times. Two possibilities might account for these substitutions: (a) AE speakers had difficulty distinguishing T /t'/ from T /t/ but the bulk of the perceptual confusion was experienced with /t'/;[10] (b) a degree of hypercorrection could have been occurring (which is also a possible explanation for the /ḥ/ substitution for /h/).

(4) After the very first period of instruction, there was no perceptual problem in distinguishing /ï/ from /uw/ but the production problem persisted for an extremely long period throughout the experiment.

(5) The substitutions for /ŋ/ were quite varied. Though the most frequent substitution was /g/, there were examples of /ʔ/, /ø/, /n/, /ny/, and /ng/ substitutions.

(6) Although there were 16 substitutions of diphthongs for T /e/ (4 substitutions of /iy/ and 9 of /ɛ/ from AE and three of /ei/ from the T system) there were 24 substitutions of monopthongs (15 of /ɛ/ and 9 of /I/). This makes the usual analysis of the learning problem involved (AE speaker learning French /e/) that it is concerned with reduction of the second part of the AE diphthong /ɛy/ most questionable. Previous assumptions that the AE speaker would consistently substitute /ɛy/ as the closest corresponding sound[11] are not substantiated by these data and Delattre's observa-

[10] This would seem to be borne out by the comments made by the subjects. The most frequent remark was, "I have a hard time telling them apart, especially when the aspirated sound comes first. I can usually get both of them if the unaspirated comes first."

[11] For example, Politzer, *ibid.*, p. 52.

tion of the close similarity between French /e/ and AE /I/[12] and Pierre Léon's report of an experiment in which AE speakers substituted /I/ for French /e/ 43.16% of the time[13] seem to correspond much more closely to the data gathered in this experiment.

On the basis of the actual substitutions made by the subjects, groups B and C were compared for significant differences from the rest of the sounds.

Group B (/h/ and /ḥ/) was tested as a group versus the rest of the sounds, as a group, and found to be highly significantly different. In addition to the problem of production of /ḥ/, there was the additional burden of the problem of perceiving a difference between these two sounds. We can hardly call these two sounds a 'split' category on the basis of the AE system alone, since /ḥ/ does not even exist in the AE system. It seems safe to posit, however, that the AE encoder formed a "perceptual-confusion-pair" category solely on the basis of the distinctive cues within the T system itself. In addition, the perceptual confusion was not cleared by the occurrence of one or the other as the first stimulus on any given test tape as was the case when /t/, most frequently recognized, occurred before /t'/, thereby automatically assuring the correct identification of the second sound. /h/ and /ḥ/ seemed to be confused regardless of which sound occurred first on any testing tape.

Group C /t'/ and /t/ was compared with the rest of the sounds as a group to help determine whether the predicted 'split' really did exist. There was no significant difference between /t'/ and /t/ taken as a class versus the rest of the sounds taken as a class. In light of the significant difference of difficulty between /t'/ and /t/ (/t'/ being the harder) the fact that taken together as a class they were not significantly different from the rest and the fact that /t/ was a much more frequent substitution for T /t'/ than was AE /t'/, it is concluded here that the two members do not belong together

[12] Pierre Delattre, "Les indices acoustiques de la parole, Premier Rapport", *Phonetica*, vol. 2, nos. ½ (Sept., 1958), p. 246.

[13] Pierre Leon, "Perception of French Vowels by American Listeners", *Etudes de Linguistiques Appliquée*, 2 (Paris, Didier, 1963,) pp. 145-146.

in a class called 'divergent' on the basis of the AE system. On the phonetic level of production, /t/ is treated as a familiar allophone being learned in a new context (which is relatively rapid) and /t'/ is a new sound, more difficult to learn. The fact that /t'/ and /t/ form a "perceptual confusion pair" is not dependent upon the allophonic membership of the AE phoneme /t/ but rather upon the phonetic cues in the T system which the AE speaker must now perceive as distinguishing two sounds that seem very similar to him.

One fact not readily apparent from the data is that subjects, given proper instruction, were able to produce the members of the perceptual confusion pairs, /h/ and /ḥ/ and /t'/ and /t/, in isolation long before they were able to perceive accurately which particular sound was being given as a stimulus when the items occurred in random order on a test tape. There is a strong implication that there are at least two levels of learning involved: (1) production of the sound in isolation, which can be quickly achieved through proper physiological descriptions; and (2) perception of the distinction between relatively 'highly' similar sounds (from the point of view of the speaker of a particular language), which requires a much longer time to differentiate within the T system. One might say that there is the problem of learning to produce a given token first and then learning the rule that determines the class membership of that token. Put differently, one might assume that a subject can produce the criterial attributes of a category without being perceptually aware that they are criterial and, consequently, treat the criterial attributes as a noisy adjunct to a category when it becomes necessary to perceptually distinguish one category from another at the perceptual level.

The attempt to establish /ey/ and /ɛ/ as a convergent class was a failure. As with the concept of divergent categories, it is felt here that mapping psychological parameters over linguistic parameters in terms of phonemic status is not the best method of predicting difficulties. From the results of this paper, there is serious doubt whether the V speaker learning French really 'collapses' his /t/ and /t'/ to French /t/ or whether he is simply transferring his /t/

to the new system and the V /t'/ fails to enter into the learning structure at all. If concepts of divergence and convergence are to be meaningful for language learning, they probably should be defined in articulatory terms at the phonetic level of analysis, encompassing information concerning the production and the perception of the particular categories.

One final check was made to see whether there was any correlation between the frequency of occurrence of specific phonemes in the speech of the adult AE speaker and their position in the hierarchy of difficulty established in this paper. A recent experiment by Carterette and Jones[14] was conducted in which conversations of first, third, and fifth grade students were recorded, analyzed, and compared with a similarly analyzed recording of adult conversations. A frequency of occurrence of phonemes was determined in terms of a percentage of the total corpus.

Some of the sounds in this experiment were compared with the frequency of their occurrence in everyday adult speech. /ŋ/, significantly harder than /ž/ and /ɛ/ (which were not significantly different from each other) were checked. According to Carterette and Jones' report, /ŋ/ had a frequency of .011 per cent of the total corpus, which is considerably less than the .032 per cent recorded for /ɛ/ and considerably more than the .000 recorded for /ž/. /ž/ showed the highest degree of positive transfer from AE yet showed the lowest frequency of occurrence in everyday speech. /ɛ/, showing the next highest degree of transfer from AE, showed the highest frequency of occurrence of the sounds compared, yet showed no significant difference from /ž/ on the hierarchy. Presumably, there is no correlation between frequency of occurrence and hierarchy of difficulty and no prediction of transfer or interference can be based on the frequency of occurrence of phonemes within the N or T systems.

[14] E. C. Carterette and Margaret Hubbard Jones, *Contextual Constraints in the Language of the Child* (Office of Education, U.S. Dept. of Health, Education, and Welfare, Cooperative Research Project no. 1877, 1965).

IX

SUMMARY

The following summary states the expected and the unexpected results obtained in this experiment followed by the general conclusions drawn therefrom.

I. EXPECTED RESULTS

(1) /x/ was learned quite rapidly although this sound is non-existent in the AE phonological system. The observations made by Nancy Kennedy (cited on page 43 above) are confirmed but no explanation is apparent from the experimental procedure. (2) AE speakers did learn front rounded vowels significantly faster than back unrounded vowels. Once again, there is no apparent explanation for this phenomenon. (3) [ey] and [ɛ] were easy to learn and not significantly different from each other. The concept of overdifferentiation seemed to be a factor in the rapidity with which these two sounds were learned. (4) More important, perhaps, then the previous three results is the fact that /ž/ (which indicated complete positive transfer from AE to T) was significantly easier than /ŋ/. Previous analyses of /ž/ as an AE phoneme of restricted distribution were rejected in this paper and the prediction was made that /ž/ would be easy for an AE speaker. The use of the syllable as the prime unit of analysis (rather than the word) is mentioned below under 'conclusions'. (5) Previous analyses of T /t/ and /t'/ as a divergent learning structure for AE speakers were rejected in this study and the prediction that /t/ and /t'/, as a class, would not be significantly different from the rest of the sounds was corroborated in this experiment. (6) Though /t/ and /t'/ did not form a divergent learning

structure from the standpoint of production, the sounds did form a perceptual confusion pair for the AE speaker when he was in the role of decoder.

II. UNEXPECTED RESULTS

(1) /t/ was significantly easier to learn than /t'/. Linguists would not have expected this or the following result. (2) /t/ was the most frequent substitution for /t'/. Linguists would have expected that /t/ would be more difficult than AE /t'/ and that AE /t'/ would be the most frequent substitution for T /t'/. (3) Previous analyses have assumed that /h/ would be easier than /ɛ̃/ but this study rejected these assumptions and predicted that /ɛ̃/ would be easier than /h/. Both predictions were incorrect and /h/ and /ɛ̃/ were not significantly different from each other. (4) The most amazing observation of all was that poduction of sounds in isolation always preceded perception of sounds within the T system. Although this was especially noticeable in the case of perceptual confusion pairs, production in isolation preceded perception within the system for all sounds.

III. GENERAL CONCLUSIONS

(1) From the results obtained when /ž/ and /ŋ/ were compared, one can conclude that the syllable is a better prime on which to base a contrastive analysis of AE with any T language and any ensuing prediction of the hierarchy of difficulty of learning involved should be based on the syllable rather than on the word. (2) T sounds which are close equivalents of N system sounds, whether phonemic or allophonic, are easier to learn than T sounds without such equivalents. (3) Perceptual confusion pairs (e.g., /h/ and /ḥ/), though in part determined by the N system, are equally dependent upon the articulatory features being used as classificatory by the native speakers of T. (4) Previous descriptions of convergent and

divergent categories in terms of allophonic membership of phonemes in N and T were not corroborated in this study. From the results of this experiment it is concluded that, when analyzing the problems involved in learning any T phonological system, it is extremely difficult (if not impossible) to classify any of the underlying parameters as 'convergent' or 'divergent' structures, since (a) at the level of phonetic reality, the stimuli and the responses in N are never truly equal to the stimuli or the responses in T, and (b), more generally, the linguistic parameters are rarely the same as the psychological parameters. (5) Perhaps the most general conclusion that can be drawn from this study is that any prediction of a hierarchy of difficulty of learning phonological categories must be based on descriptions of these categories in terms of exhaustive information at the phonetic level rather than on description solely in terms of distinctive features or allophonic membership of phoneme classes.

A great deal of research is necessary to determine precisely (a) the underlying parameters of Arabic, Vietnamese, AE, and other sound categories; (b) the specific classificatory features attended to by the native speakers of the respective languages being investigated; (c) the underlying learning problems that cause difficulty in distinguishing between two members of a perceptual confusion pair; (d) the psychoacoustics of phone substitutions made by individuals.

APPENDIX I

Total Number of Correct Learning Trials per Subject per Category

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Σ i
h	10	12	12	12	11	10	12	12	7	6	10	10	12	10	8	11	12	12	10	12	211
ḥ	6	5	3	3	0	5	3	0	7	11	6	5	0	7	10	2	4	0	4	4	85
x	3	12	12	12	12	12	7	11	11	11	12	11	12	12	12	12	12	12	12	12	222
γ	6	11	9	12	3	9	7	10	9	9	11	7	8	6	7	11	12	12	10	9	178
t'	4	8	9	11	7	9	9	7	10	9	10	11	10	7	11	10	6	12	8	10	178
t	5	10	10	12	11	12	12	9	7	12	8	11	12	11	9	8	10	12	10	11	202
ï	7	7	3	6	1	7	1	0	7	6	2	10	2	1	5	3	5	0	5	3	81
ŋ	5	12	11	11	11	8	6	10	10	5	11	8	11	9	9	9	9	9	6	8	178
cy	12	12	11	9	7	12	9	12	12	12	12	11	8	10	12	12	11	12	12	12	220
ε	12	12	12	12	12	12	12	12	12	12	12	12	12	12	11	12	11	12	12	12	238
ü	10	12	8	8	8	7	11	4	10	11	8	7	10	6	12	9	9	9	11	10	180
ž	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	240
ε̃	8	12	12	12	12	0	10	11	12	11	0	12	10	12	0	12	9	12	12	10	189
e	10	12	12	12	11	10	11	2	12	12	11	12	10	12	10	12	12	12	10	11	216
Σj	110	149	136	144	118	125	122	112	138	139	125	139	129	127	128	135	135	137	134	136	2618

APPENDIX II

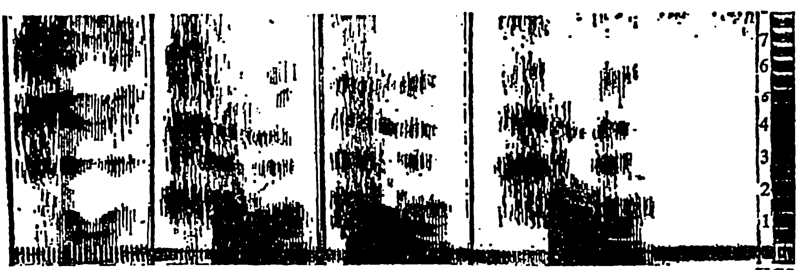
Total Number of Correct Testing Trials per Subject per Category

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Sum *i*
h	10	11	12	11	10	7	9	12	8	7	12	9	12	9	6	11	10	9	9	11	195
ḥ	6	5	3	1	1	4	3	0	5	7	5	4	1	5	4	3	2	1	1	3	64
x	5	10	11	11	12	11	7	10	11	8	12	12	11	11	12	10	12	12	11	12	211
γ	6	10	8	11	3	8	7	9	8	9	11	7	11	5	2	12	11	12	10	9	169
t'	3	9	6	12	7	5	5	9	10	5	10	11	11	6	12	9	6	12	7	10	165
t	8	12	12	11	9	8	10	11	5	12	11	10	12	9	8	9	10	12	7	12	198
ï	7	4	0	7	5	9	0	1	7	4	5	9	6	1	7	2	4	1	5	3	87
ŋ	5	11	10	9	11	7	9	10	9	5	11	9	11	9	11	10	8	9	11	9	184
ey	12	12	10	11	7	11	10	10	9	12	12	12	9	8	12	12	11	12	12	11	215
ε	12	12	12	12	12	12	11	12	11	12	11	12	12	12	11	12	12	11	12	12	235
ü	9	9	4	11	6	10	10	6	11	9	7	5	10	7	10	7	8	10	12	9	170
ž	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	240
ε̃	9	12	12	12	10	0	10	9	12	12	0	12	10	12	1	11	10	12	12	9	187
e	10	11	12	12	11	10	7	1	8	12	10	12	10	10	12	12	9	12	7	12	200
Σj	114	140	124	143	116	113	110	112	126	126	129	136	138	116	120	132	126	137	128	134	2520

APPENDIX III

Spectogram Comparison of /h/,/ḥ/,/x and AE/h/

I II III IV



h a w a x a w a ha wa˅ ḥ a wa KCS

American /h/; Voiceless. Heavy concentration of energy at 6kcs and above. F3 distinct; F4 and F5 blend together. Very slight energy below 2.5kcs. Laryngeal.

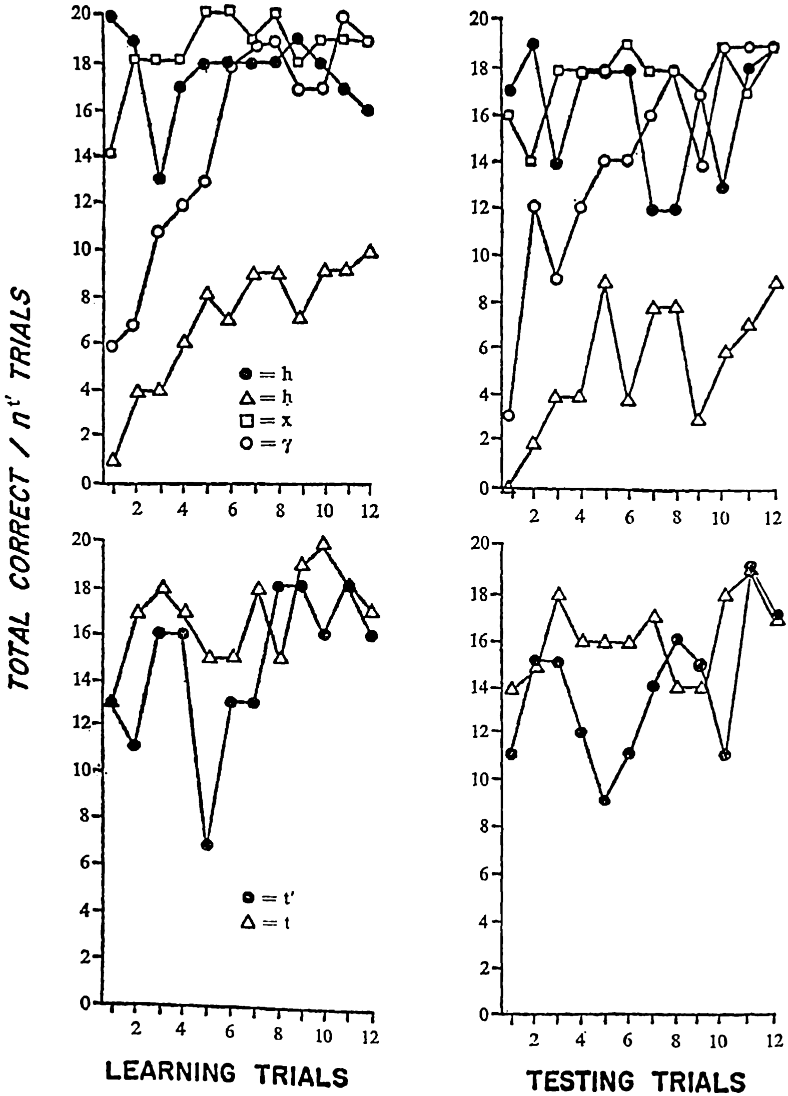
Arabic /x/; Voiceless. Clear F2. Concentration ofenergy blending around F3 and F4. Concentration of energy from 5kcs above. Velar.

Arabic /h/; breathy voice, Clear formant structure similar to following vowel from F1 through F5 Arabic Laryngeal

Arabic /ḥ/; indeterminate energy around 1-2kcs. Clear F structures only at F3 F4. Arabic pharyngela

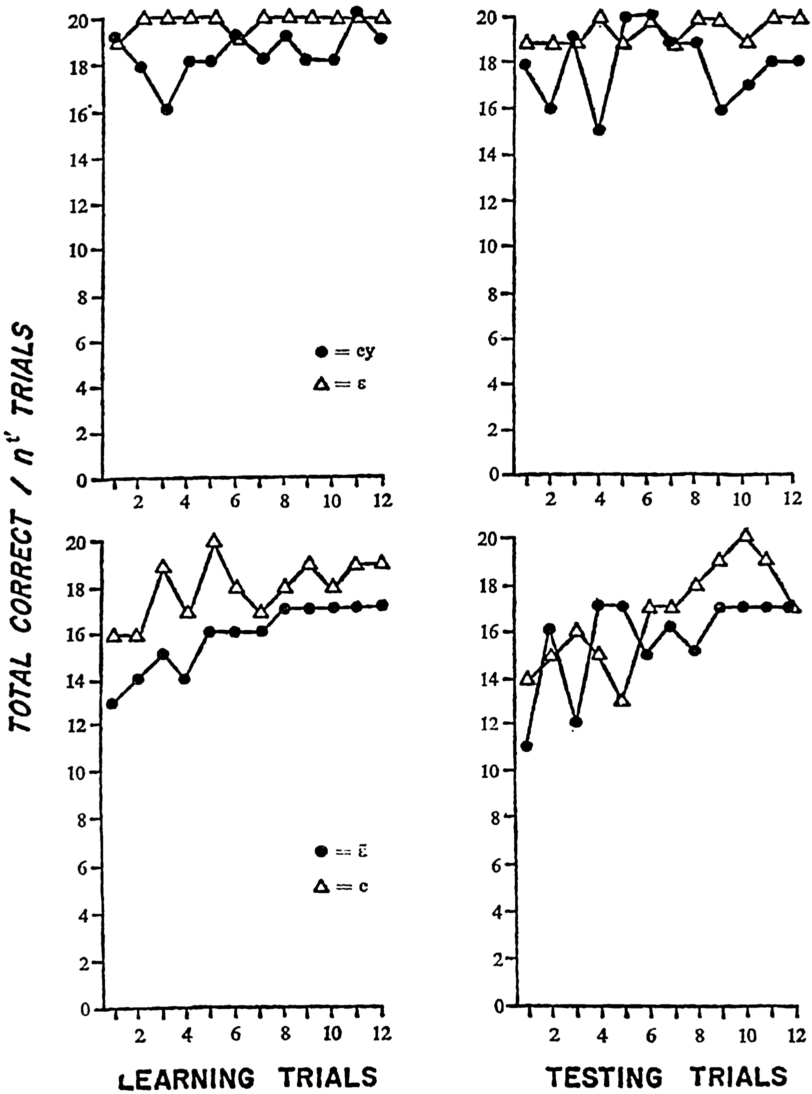
APPENDIX IV

Learning Curves /h/, /ḥ/, /x/, /γ/, /t/, /t'/



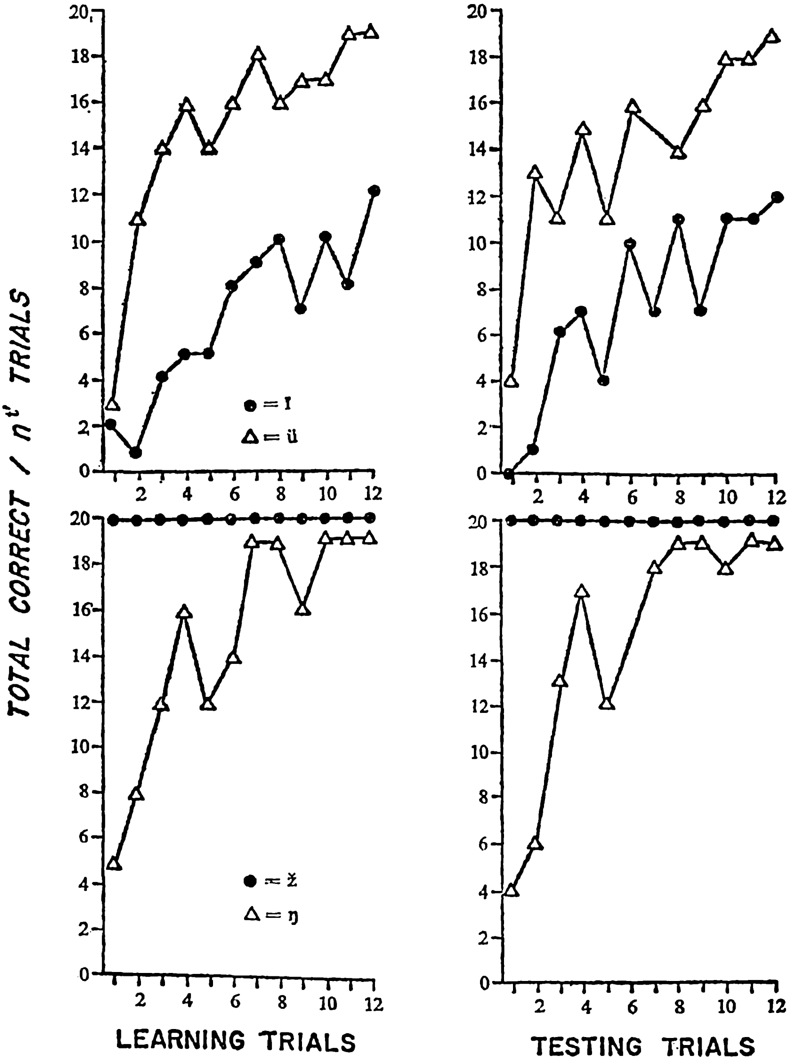
APPENDIX V

Learning Curves /eγ/, /ε/ and /ε̃/, /e/



APPENDIX VI

Learning Curves /ï/, /ü/ and /ž/, /ŋ/



REFERENCES CITED

Armstrong, Lilias E., *The Phonetics of French* (London, G. Bell and Sons, Ltd., 1932).

Bloch, Bernard, "Phonemic Overlapping", *American Speech*, 16 (1941), pp. 278-284, reprinted in Martin Joos (ed.), *Readings in Linguistics* (Washington, D.C., American Council of Learned Societies, 1957), pp. 93-96.

——, and George L. Trager, *Outline of Linguistic Analysis* (Baltimore, The Waverly Press, Inc., 1942).

Bloomfield, Leonard, *Language* (New York, Henry Holt and Company, 1933).

Bursill-Hall, Geoffrey L., "Frequency of Consonant Clusters in French", *The Journal of the Canadian Linguistic Association*, 2 (1956), pp. 66-78.

Carterette, E. C., and Margaret Hubbard Jones, *Contextual Constraints in the Language of the Child* (Office of Education, U.S. Department of Health, Education, and Welfare, Cooperative Research Project no. 1877, 1965).

Chomsky, Noam, "The Logical Basis of Linguistic Theory", in *Proceedings of the Ninth International Congress of Linguists* (The Hague, Mouton & Co., 1964).

Cofer, Charles (ed.), *Verbal Learning and Verbal Behavior* (New York, McGraw-Hill Book Company, Inc., 1961).

Delattre, Pierre, "Les Indices acoustiques de la parole", Premier Rapport, *Phonetica*, 2, nos. ½ (Sept., 1946), pp. 242-248.

Duncan, D. B., "Multiple Range and Multiple *F* Test", *Biometrics*, 11 (1955), pp. 1-42.

Edwards, Allen, *Experimental Design in Psychological Research* (New York, Rinehart and Company, Inc., 1960).

Emeneau, M. B., *Studies in Vietnamese (Anamese) Grammar* (Berkeley and Los Angeles, University of California Press, 1951).

Finney, D. J., *Experimental Design and Its Statistical Basis* (Chicago, The University of Chicago Press, 1955).

Gibson, Eleanor J., "A Systematic Application of the Concepts of Generalization Differentiation to Verbal Learning", *Psychological Review*, 47 (1940), pp. 196-229.

——, and J. J. Gibson, "Retention and the Interpolated Task", *American Journal of Psychology*, 46 (1934), pp. 606-610.

Harrell, Richard S., *The Phonology of Colloquial Arabic* (New York, American Council of Learned Societies, 1957).

Hockett, Charles F., *A Course in Modern Linguistics* (New York, The MacMillan Company, 1958).

Jakobson, Roman, "Mufaxxama — The "Emphatic" Phonemes in Arabic", in Ernst Pulgram (ed.) *Studies Presented to Joshua Whatmough* ('s-Gravenhage, Mouton & Co., 1957), pp. 105-117.

——, Gunnar Fant and Morris Halle, *Preliminaries to Speech Analysis* (Cambridge, Massachusetts, The MIT Press, 1963).

Jenkins, James J. "The Learning Theory Approach", in Charles Osgood and Thomas Sebeok (eds.), *Psycholinguistics* (Baltimore, Waverly Press, Inc., 1954), pp. 20-34.

Jones, Daniel, *An Outline of English Phonetics* (New York, E. P. Dutton and Co. Inc., 1956).

Jones, R. B., Jr., and Huynh Sanh Thong, *Introduction to Spoken Vietnamese* (Washington D.C., American Council of Learned Societies, 1957).

Joos, Martin, "Description of Language Design", *Journal of the Acoustical Society of America*, 22 (1950), pp. 701-708, reprinted in Martin Joos (ed.), *Readings in Linguistics* (Washington D.C., American Council of Learned Societies, 1957), pp. 349-356.

——, (ed.), *Readings in Linguistics* (Washington, D.C., American Council of Learned Societies, 1957).

Kennedy, Nancy, *Problems of Americans Mastering the Pronunciation of Egyptian Arabic* (Washington, D.C., Center for Applied Linguistics, 1960).

Ladefoged, Peter, "Some Possibilities in Speech Synthesis", *Language and Speech*, Dec. 1964.

Lado, Robert, *Linguistics Across Cultures* (Ann Arbor, The University of Michigan Press, 1957).

Lehn, Walter and William Slager, "A Contrascove Study of Egyptian Arabic and American English: the Segmental Phonemes", *Language Learning*, IX (1959), pp. 25-34.

Leon, Pierre, "Perception of French Vowels by American Listeners", in B. Quemada (ed.), *Etudes de Linguistique Appliquée*, 2 (Paris, Didier, 1963), pp. 135-146.

Malmberg, Bertil, *Le Système consonantique du français moderne* (Copenhague, CWK Gleerup, Lund, 1943).

——, *Structural Linguistics and Human Communication* (New York, Academic Press Inc., 1963).

Mandelbaum, David G. (ed.), *Selected Writings of Edward Sapir* (Berkeley and Los Angeles, University of California Press, 1949).

McGeoch, J. A., "The Influence of Degree of Interpolated Learning Upon Retroactive Inhibition", *American Journal of Psychology*, 44 (1932), pp. 695-708.

——, "Studies in Retroactive Inhibition: VII. Retroactive Inhibition as a Function of the Length and Frequency of Presentation of the Interpolated Lists", *Journal of Experimental Psychology*, 19 (1936), pp. 674-693.

——, *The Psychology of Human Learning: An Introduction* (New York, Longmans, 1942).

Moulton, William G., *The Sounds of English and German* (Chicago and London, The University of Chicago Press, 1962).

——, "Toward a Classification of Pronunciation Errors", *The Modern Language Journal*, XL, no. 3 (March, 1962), pp. 101-109.

Osgood, Charles, "The Similarity Paradox in Human Learning: A Resolution", *Psychological Review*, 56 (1949), pp. 132-143.

——, and Thomas Sebeok (eds.), *Psycholinguistics* (Baltimore, Waverly Press, Inc., 1954).

Pike, Eunice, "A Test for Predicting Phonetic Ability", *Language Learning*, IX, 1 and 2 (1959), pp. 35-43.

Pike, Kenneth L., *Phonemics* (Ann Arbor, University of Michigan Press, 1947).

Politzer, Robert, *Teaching French: An Introduction to Applied Linguistics* (Boston and New York, Ginn and Company, 1960).

Postman, Leo, "The Present Status of Interference Theory", in Charles Cofer (ed.), *Verbal Learning and Verbal Behavior* (New York, McGraw-Hill Book Company, Inc., 1961), pp. 152-196.

Pulgram, Ernst (ed.), *Studies Presented to Joshua Whatmough* ('s-Gravenhage, Mouton & Co., 1957).

Sapir, Edward, "Sound Patterns in Language", *Language*, 1 (1925), pp. 37-51; reprinted in David Mandelbaum (ed.), *Selected Writings of Edward Sapir* (Berkeley and Los Angeles, University of California Press, 1949); and in Martin Joos (ed.), *Readings in Linguistics* (Washington, D.C., American Council of Learned Societies, 1957), pp. 19-25.

——, "La Réalité psychologique des phonèmes", *Journal de Psychologie Normale et Pathologique*, 30 (1933), pp. 247-265, reprinted in David Mandelbaum (ed.), *Selected Writings of Edward Sapir* (Berkeley and Los Angeles, University of California Press, 1949), pp. 46-60.

Saporta, Sol, "Frequency of Consonant Clusters", *Language*, 31 (1955), pp. 25-31.

Scheffé, H. A., *The Analysis of Variance* (New York, Wiley, 1960).

Stockwell, Robert, *A Contrastive Analysis of English and Tagalog*, unpublished fascicle, UCLA, 1957.

——, and J. Donald Bowen, *The Sounds of English and Spanish* (Chicago, University of Chicago Press, Dec., 1964).

Swadesh, Morris, "The Phonemic Principle", *Language*, 10 (1934), pp. 117-129, reprinted in Martin Joos (ed.), *Readings in Linguistics* (Washington, D.C., American Council of Learned Societies, 1957), pp. 32-37.

Thompson, Laurence C., "Saigon Phonemics", *Language*, 35 (1959), pp. 454-478.

Trager, George L. and Henry Lee Smith, Jr., *An Outline of English Structure* (Washington, D.C., American Council of Learned Societies, 1957).

Troubetskoy, N. S. (J. Cantineau, translator), *Principes de Phonologie* (Paris, Librairie Klincksieck, 1955).

Twaddell, W. Freeman, "On Defining the Phoneme", *Language Monograph*, no. 16 (1935), reprinted in Martin Joos (ed.), *Readings in Linguistics* (Washington, D.C., American Council of Learned Societies, 1957), pp. 55-80.

Underwood, B. J., "The Effect of Successive Interpolations on Retroactive and Proactive Inhibition", *Psychological Monographs*, 59, No. 3 (1945).

——, "Interference and Forgetting", *Psychological Review*, 64 (1957), pp. 49-60.

Valdman, Albert, *Applied Linguistics, French*, Simon Belasco (Gen. ed.) (Boston, D. C., Heath and Company, 1961).

Weinreich, Uriel, *Languages in Contact* (New York, Publications of the Linguistic Circle of New York, 1953).

Wells, Rulon, "De Saussure's System of Linguistics", *Word*, 3 (1947), pp. 1-31, reprinted in Martin Joos (ed.), *Readings in Linguistics* (Washington, D. C., American Council of Learned Societies, 1957), pp. 1-18.

Whitely, P. L. and A. B. Blankenship, "The Influence of Certain Conditions Prior to Learning Upon Subsequent Recall", *Journal of Experimental Psychology*, 19 (1936), pp. 496-504.

Winer, B. J., *Statistical Principles in Experimental Design* (New York, McGraw Hill Book Company, Inc., 1962).

Wolff, Hans, "Partial Comparisons of the Sound System of English and Puerto Rican Spanish", *Language Learning*, III, 1 and 2 (1958), pp. 38-41.