International Journal of English Studies

IJES

UNIVERSITY OF MURCIA

www.um.es/engphil/ijes

Allophonic Splits in L2 Phonology: The Question of Learnability

FRED R. ECKMAN*
University of Wisconsin–Milwaukee

ABDULLAH ELREYES
United Arab Emirates University

GREGORY K. IVERSON
University of Wisconsin–Milwaukee

ABSTRACT

The research reported in this paper is intended as a contribution to the understanding of several well-known problems relating to the learning of phonemic contrasts in second language (L2) phonology. The paper describes a series of ongoing studies examining what Lado (1957) hypothesized to represent maximum difficulty in second language pronunciation, namely, a phonemic split. This is the process involved when an L2 learner must split native language (NL) allophones into separate target language (TL) phonemes. Two core principles of phonological theory are described and evaluated for their relevance in explaining the series of well-defined, implicationally-related stages involved in a phonemic split. Finally, the paper reports the results of an empirical study designed to test the explanatory adequacy of these principles, and concludes with a discussion of the implications of these studies for second language phonology in general.

KEYWORDS: Second-language phonology; interlanguage phonology; pronunciation difficulty; phonemic split; stages of second-language acquisition; learnability; structure preservation; derived environment constraint.

* *Address for correspondence*: Fred R.. Eckman, Department of Foreign Languages and Linguistics, University of Wisconsin-Milwaukee,PO Box 413, Milwaukee, Wisconsin 53201, USA. E-mail: eckman@uwm.edu

INTRODUCTION

Over the last few years there has been a resurgence within second language acquisition (SLA) theory and instruction in the amount of attention that has been devoted to the teaching of pronunciation, though by common concession this aspect of language learning is still poorly understood, and often poorly taught (Celce-Murcia et al. 1996; Morley 1987, 1991, 1994). The research reported in this paper is intended as a contribution to the understanding of several well-known problems relating to the learning of phonemic contrasts in second language (L2) pronunciation. In particular this paper focuses on some of the effects that the competing influences of similarity and difference between native and target language sound systems have on the learning of (L2) phonology (Wode 1983a, Flege 1980,1987; Major & Kim 1999). The purpose of the present paper is to report on a series of ongoing studies examining the role of phonological theory in the explanation of L2 pronunciation; in particular, the paper seeks to evaluate two core principles in phonological theory for their relevance in explaining what Lado (1957) hypothesized to represent maximum difficulty in second language pronunciation, namely, the splitting of native language (NL) allophones into separate target language phonemes.

The paper is structured as follows. Reprising discussion in Eckman and Iverson (1997, 1999), we first describe two linguistic constructs that we believe are crucial in learning the pronunciation of a target language, and review the issues that are involved in splitting native language allophones into separate target language phonemes. We then outline the phonological principles which are relevant to our investigation and follow this by reporting the results from a study designed to test these principles. We frame the discussion in terms of conventional "rules" rather than optimality theoretic "constraints", primarily for clarity of exposition, but we believe that the general principles at play (which emerged from work in the theory of lexical phonology) will hold for any version of phonology in which issues such as these are addressed.

I. PRONUNCIATION DIFFICULTY

We start with the assumption that, in order to acquire a target language (TL), the L2 learner must acquire a lexicon (a set of words and their affixes) along with a set of rules (or equivalent constraints) for combining the lexical items into larger utterances, and then pronouncing them. Potential impediments to this leaming arise from two areas: 1) from certain inherent difficulty in learning the various TL lexical items and rules, and 2) from areas of the NL that may interfere with this acquisition.

Given this, and focusing on the area ofpronunciation, we can identify at least two aspects of the NL and TL where differences may cause difficulty: differences in inventory, in which the TL contains sound segments that do not exist in the NL, and positional differences, such that the

TL may have a contrast between two sounds that are allophones of the sarne phonerne in the NL. Phoneme inventory differences have long been recognized as a source of leaming difficulty, at least as far back as Lado (1957), and as recently as Flege (1987) and Major & Kim (1999), but a special status has been accorded to positional differences in which the allophones of an NL phonerne represent separate phonemes in the TL (Lado 1957, Hammerly 1982). The task of the leamer in such cases is to split the NL allophones into separate TL phonemes.

Two examples of an allophonic split, both relevant to the arguments in this paper, are: (1) a native speaker of Spanish leaming the English distinction between /d/ and /ð/, and (2) a native speaker of Korean acquiring the English contrast between /s/ and /š/. In Spanish, [d] and [d] are allophones of the phoneme /d/, because [ð] occurs after continuant segments and [d] occurs elsewhere; in Korean, [s] and [S] are allophones of syllable-initial /s/, because [S] occurs only before the vowel [i], [s] elsewhere. In English, of course, all of these sounds are separate phonemes, and thus a Spanish speaker leaming English must learn to factor the allophones [d] and [d] into separate phonernes, and a Korean-speaking ESL leamer must acquire the contrast between /s/ and /š/. In what follows, we will argue that the splitting of NL allophones into TL phonemes potentially involves two stages which are explained by established phonological principles.

II. THE PHONOLOGICAL CONTEXT

In this section, we summarize the rnotivation for two general principles which have ernerged out of the theory of lexical phonology (Kiparsky 1973), Structure Preservation and the Lexical Derived Environment Constraint.

(1) STRUCTURE PRESERVATION
Representations within the lexicon may be cornposed only of elements drawn from the phonemic inventory.

(2) LEXICAL DERIVED ENVIRONMENT CONSTRAINT
Lexical rules apply only in derived environments; postlexical rules apply across-the-board.

These principles presuppose that phonological rules are divided into two groups: those that apply within the lexicon of the language as words are being formed, i.e., the lexical rules, and those that come into play after words have been entered into sentences, the postlexical rules. Lexical rules exhibit two special properties that are of concem to us: (1) they apply only to "derived" forms (i.e., to words whose relevant portions have been modified by previous rule, or which are built up out of separate meaningful elements); (2) they are constrained to produce only segrnents

which are found in the phonemic inventory, or, more generally, to produce just those kinds of structures which exist in the lexicon. Postlexical rules, on the other hand, do not require the form to which they apply to be derived, or composite, and are not constrained to be structure preserving, hence they may produce segments which are not part of the phonemic inventory.

A frequently cited example of a typical lexical rule in English is Trisyllabic Laxing, so named because it has the effect of making a stressed or accented vowel short if it is in the third syllable from the end of the word. This rule accounts for altemations in vowels such as those in the word pairs listed in (3).

(3)	sane	[sén]	sanity	[sǽnəri]
	divine	[dəváɪn]	divinity	[davínari]
	serene	[sarín]	serenity	[səréńəri]

The stressed vowel in each of the unsuffixed words in (3) is tense, but that same vowel is pronounced as lax when the word it is in consists of a stem followed by the two-vowel suffix *-ity*. The words in (4a, b), on the other hand, illustrate that this rule applies only in so-called derived environments (i.e., when an affix has been appended, not when the word itself consists ofjust the stem), and the word in (4c) exemplifies that only particular suffixes (e.g., *-ity* but not *-able*) will trigger Trisyllabic Laxing.

(4)	a.	stevedore	[stívədɔr]	*[stívédɔr]
	b.	nightengale	[náɪtəngel]	*[nítengel]
	c.	notable	[nórabal]	*[nɔ́rəbəl]

An example of a postlexical rule in (American) English is Flapping, which accounts for the pronunciation altemations in (5).

(5)	a. bet	[bet]	betting	[bɛrɪŋ]
	b. ride	[raid]	nding	[raɪrɪŋ]

Flapping must be a postlexical rule because it is not structure preserving in that it produces the sound [4], which is not part of the phonemic inventory of English. Unlike lexical rules such as Trisyllabic Shortening, moreover, Flapping may apply between words (e.g., to the first [t] in Hit it) as well as within single lexical entries (e.g., the noun matter may be pronounced the same as the comparative adjective madder, both with medial flaps). The distinction is thus one between lexical rules that apply strictly within words as they are being created, preserving structure in the sense of (1), and postlexical rules that may apply within as well as between words after they have been created, without regard for any limitations on the inventory of speech

sounds.

The other core principle, the Lexical Derived Environrnent Constraint as stated in (2), overlaps substantially with Structure Presewation inasrnuch as lexical rules are structure presewing, and by (2) are restricted to apply only to configurations that are derived through processes of affixation or word formation, or the application of another rule, i.e., they rnay not affect basic lexical entries. If such rules were to apply to unmodified lexical iterns without affixes, there would be no trace lefi in terms of crucial alternations which support the recovery of underlying representations. As Kiparsky illustrated with respect to Finnish, for exarnple, the structure-presewing rule in that language converting /t/ to [s] before /i/ crucially applies only in derived contexts, as in (6a), where processes of word formation have brought the (stern) /t/ and the (suffix) /i/ into juxtaposition.

(6) Finnish assibilation

(a)	/halut+i/	→	[halusi]	'want-ed'
(b)	/koti/	→	[koti]	'horne'
(c)	*[kosi]			
(d)	/halut+a/	→	[haluta]	'to want'

If the /t/ plus /i/ sequence is already on hand in the basic lexical listing, on the other hand, the rule does not apply, as shown in (6b). Of course, if the rule were to apply here, producing (6c), there would be no basis for "recovery" of the underlying /t/: Finnish speakers would never be able to figure out that the word for 'house' is [koti] if it were always pronounced as *[kosi]. The /t/ in /halut/ 'want', conversely, does undergo the change to [s] when a (suffix) /i/ follows, because this /t/ rernains in other instances of the form that do not undergo the rule, as exernplified in (6d). Sirnilarly, if the lexical Trisyllabic Laxing rule in English were to apply in nonderived contexts, i.e., within single-rneaning structures like ***nightengale,*** there would be no basis for recovery of the fact that the first vowel in this word is /ay/, not /ɪ/, since the forrn would always be pronounced with the incorrect lax vowel.

Thus, Structure Presewation requires that lexical rules produce segrnents which are phonernes of the language, and the Lexical Derived Environrnent Constraint holds that (structure presewing) lexical rules rnay apply only to configurations that are crucially derived, as through a process of affixation. The relationship between these two notions has been argued to be even tighter than this, however. Based on the analysis of primary language data relating to rules with lexical as well as postlexical functions, Iverson (1993) rnakes the more general case that not only are lexical rules constrained to apply just in derived environments, as in conventional lexical phonology, but so are the applications of all structure presewing rules, whether functioning lexically or postlexically. The effect of this narrower lirnitation, which we adopt here as the operative version of the Derived Environrnent Constraint (cf. also Kiparsky 1973), is that neutralizing rule applications in any part of the grarnrnar rnay not affect basic lexical iterns:

(7) DERIVED ENVIRONMENT CONSTRAINT
Stmcture preserving rule applications are restricted to derived environrnents.

Both Stmcture Preservation and the Derived Environrnent Constraint have irnplications for leamability. The Derived Environrnent Constraint is fundamentally a condition on the recoverability, or leamability, ofwords and their parts. Applying neutralizing rules to nonderived forrns would rnake the lexical form of the word essentially unleamable, because there would be no altemations frorn which the leamer could acquire the phonernic representation. Likewise, Stmcture Presemation, which associates chiefly with lexical mles and is not applicable in the postlexical component, correlates generally with the distinction between phonernic and allophonic distribution. Since postlexical mles are typically (though not exclusively) allophonic, and since lexical rules alrnost always result in the loss of contrast between sounds in specific environments, the long-standing distinction between distributional staternents defined on phonernes and those defined on allophones is accornmodated rather directly, reflecting the presurned prirnary cognitive status of the traditional phonerne. That is, a language's inventory of phonernes is part of what rnust be actually learned in leaming the language, along with other essentially arbitrary inforrnation encoded in the lexicon, including the particular rneanings of lexical entries and their individual syntactic properties. Postlexical material, by contrast, is cognitively less prominent, presurnably precisely because it lies outside the arena where meaningful contributions to word forrnation take place. i.e., the lexicon.

These two principles have interesting irnplications for the developrnent of L2 leamers' sound patterns.

III. SECOND LANGUAGE ACQUISITION

Hypothesizing that Structure Presemation and the Derived Environrnent Constraint also govern interlanguage gramrnars, we predict the existence of progressive stages of leaming associated with the influence of an NL allophonic rule on the acquisition of the TL pronunciation. To illustrate, we reconsider the two exarnples of an allophonic split rnentioned above (and discussed in Eckrnan & Iverson 1997, 1999), namely, that in Spanish [d] and [ð] are allophones of the phonerne /d/, and in Korean, [s] and [S] are allophones of /s/.

In a language-contact situation in which the NL grarnmar incorporates a postlexical (allophonic) rule relating segrnents already contained in the phonemic inventory of the TL, the transfer of the NL rule to the IL would not result in any change in the rule's applicational status for a leamer who has not yet acquired the TL contrast. That is, the rule still is not structure preserving, and so will continue to apply postlexically in the IL, with the leamer consequently erring across-the-board on TL words containing the contrast in question. In the Spanish exarnple,

the prediction is that the learner, at stage 1, would err consistently on English words with intervocalic /d/, producing forms such as [læðər] 'ladder' and [rɛðər] 'redder' rather than [lædər] and [rɛdər].[1] A first-stage Korean learner of English would be predicted to err consistently on TL words containing a /si/ sequence, pronouncing *receive* as [rišiv] and the words *messy* and *meshy* both as [mɛši].

Once the learner begins to acquire the TL contrast, however, the status of the NL (postlexical) rule becomes structure preseming in the IL grammar, and thus subject to the Derived Environrnent Constraint. This means that the rule now may no longer apply in all contexts, but rather is restricted to derived environrnents, i. e., across a morpheme boundary. In our Spanish-English example, the learner would continue to make errors contrasting /d/ and /ð/, but would make them only in derived contexts, now pronouncing *ladder* with [d] ([lædər], non-derived context), but still producing *redder* with [ð] ([rebar], derived context). At some later point, if the learner continues to progress, we might expect this rule to be eliminated from the IL altogether.

This scenario reduces to the claim that an NL postlexical rule which produces as output a TL phoneme will, if incorporated into the IL grammar, observe the principles of Structure Presemation and the Derived Environment Constraint. We state this claim explicitly as the hypothesis in (8).

(8) Interlanguage phonological mles conform to the principles of phonological theory.

According to (8), the predicted stages of acquisition, using a Korean learner as an example, are these:

(9) The three predicted possible stages for a learner:
Stage I, NO CONTRAST: not to make the relevant target language contrast, applying the native language rule in both derived and nonderived contexts (e.g., a Korean ESL learner says the pairs *sea–she* and *messing– meshing* homophonously, as [Si] and [mɛšɪŋ]);

Stage II, PARTIAL CONTRAST: to make the relevant contrast in some words, applying the native rule only in derived contexts (a Korean ESL learner says *sea–she* correctly but errs by producing *messing– meshing* homophonously);
Stage III, CONTRAST: to make the relevant contrast in all words, applying the native rule in neither derived nor nonderived contexts (a Korean ESL leamer says the pairs *sea–she* and *messing– meshing* correctly);

Excluded: to make the relevant contrast in some words, applying the native rule only in nonderived contexts (a Korean ESL leamer says the pairs ***sea–she*** homophonously, but says ***messing– meshing*** correctly).

In our view, then, universal principles of grammar place learnability constraints on the kinds of IL grammars that can be acquired. If we are correct about this, it would be possible for a Spanish leamer of English to first acquire the contrast between [d] and [ð] in only non-derived environments (words consisting of only a single morpheme), but it would never be possible for a learner to acquire this contrast only in derived environments. In other words, our hypothesis reduces ultimately to a learnablility claim: IL grammars in which [d] and [ð] are contrasted only in derived environments will never be learned.

To test these predictions empirically, we conducted both a cross-sectional and instructional study.

IV. THE STUDIES

The purpose of the cross-sectional study was to test for the existence of the three predicted stages outlined in (9), and the absence of the excluded stage. Accordingly, for the hypothesis to be supported by the data from the cross-sectional study, we should attest only three kinds of leamers: those who make the relevant contrast (between [d] and [ð] for Spanish speakers, and between [s] and [S] for Korean speakers) in both derived and nonderived contexts; those who make the relevant contrast in nonderived environments, but who may not make the contrast in derived environments; and finally, those who have not yet acquired the relevant contrast in either context. We should not find, according to the hypothesis, a leamer who has the contrast in derived environments but lacks it in basic words.

The purpose ofthe instructional study was to test the two pedagogical irnplications of the hypothesis. It is predicted that a leamer who is taught to make a phonemic split between NL allophones only in a derived environment will generalize this leaming to the nonderived environment, but a leamer who is trained to make the contrast in a nonderived context will not necessarily extend it to derived environments. To support these claims, it must be the case that a learner who initially lacks the contrast in both derived and basic environments and who is trained to make the contrast in only derived environments either will leam the contrast also in nonderived words, or will leam it in both derived and nonderived words. Such a leamer, however, will not learn the relevant contrast only in derived words. But a learner who is trained on the contrast in only nonderived contexts may acquire that contrast without generalizing it to derived contexts.

IV.1. The cross-sectional study

IV.1.a. Subjects

We elicited pronunciations of English words from sixteen ESL leamers, nine native speakers of Spanish, and seven native speakers of Korean. Leamers with these two NL backgrounds were chosen because, as outlined above, their NL includes an allophonic distribution of segments which are contrastive in English. All of the subjects were in the process of learning English as a second language. These leamers ranged in age from 17 to 31, each had been in the United States for less than six months, and each was from one of the two lower modules in the University of Wisconsin–Milwaukee ESL Intensive Program. All of the subjects were paid for their participation.

IV.1.b. Methodology

The first step was to establish a baseline on each of the subjects to determine whether their IL exhibited the relevant contrast: /d/ vs. /ð/ for Spanish-speaking subjects, /s/ vs. /š/ for Korean speakers. In order to accomplish this, the subjects met individually with one of the authors and/or one of the research assistants appointed to the project. The subjects' pronunciations of words containing the sounds in question were elicited using pictures accompanied by definitions. Pictures were used to avoid the subjects basing their pronunciation on the spelling of the words. The subjects were given directions and examples for an exercise in which they were presented with a loose-leaf notebook containing drawings depicting a word on one page, and a definition of the word on the facing page. The subjects were instructed to pronounce the word that was depicted.

The exercise was designed to elicit English words exhibiting the relevant contrast in both a derived and nonderived environment. Words exhibiting the contrast in a nonderived environment were basic, monomorphemic lexical items. The words exhibiting the contrast in a derived environrnent contained a suffix, either the progressive "ing" or the adjectival "y" suffix. The exercise was constructed so that the pictures contained a cue indicating which of the two suffixes was to be added to the word being pictured. For example, if the subject was shown a picture of some grass on one page, and a definition of grass on the facing page, the subject was to produce the word ***grass.*** If the picture and definition presented to the subject also contained the cue "adjective" on the page below the picture and the definition, then the subject was to produce the adjectival form of ***grass,*** namely, *grassy.* Thus, the subjects produced two kinds of baseline words, those containing the sounds in question in a nonderived context, i.e., without a suffix added, and those with the sound in a derived context, i.e., with the addition of a suffix. Some examples of the pictures and definitions used in this elicitation are contained in Appendix A.

To ensure that the subjects understood the exercise, they were given written directions along with a set of practice words. All of the subjects were able to complete the practice words satisfactorily and move on to the baseline words. During the elicitation of the baseline, subjects were prompted on the words they did not recognize from the pictures and definitions. All of the subjects were able to produce all of the baseline words elicited by the pictures and definitions by the end of the first session. The lists of words used for each NL group along with the directions used for this exercise are given in Appendix B[2].

Baselines were established on all of the subjects over two to five sessions held as close together as the subjects' schedules permitted, in most cases within one or two weeks. All of the sessions were tape recorded. Two transcriptions were done for each session: one was made during the session itself, whereby the interviewer transcribed only the segments relevant to the contrast in question (i.e., the [d] and [ð] for the Spanish speakers and the [s] and [S] for the Korean subjects) on a score sheet; the other was transcribed at a later date by one of the research assistants. Two reliability checks were then done on the transcriptions. The live transcription of the segments in question was checked against the transcription of those segments based on the tape. Where the two transcriptions differed, which occurred in only 0.88% of the cases, those segments were not scored as part of the data.' Additionally, randomly selected, five-minute portions of the tapes were later re-transcribed by a research assistant who had not performed the original transcription. A reliability figure was computed by making a point-to-point comparison between the two transcriptions and then dividing the number of agreements (2,520) between the transcriptions by the number of agreements and disagreements (2,778). This yielded a figure of .91, which was deemed adequate[4].

IV.1.c. Scoring

We now tum to a description of how the subjects' productions were scored. Because the focus of the study was to determine whether the subjects could make a contrast between two segments which occurred in the NL, albeit as allophones, the question was not whether the subjects could produce the segments in question, but whether they could produce them in the appropriate environment[5]. Accordingly, subjects were scored on their ability to produce the relevant segments in TL positions where the segment did not occur in the NL. For example, [S] in Korean occurs only before the vowel [i], whereas [s] never occurs before [i], but does occur before all other vowels. Consequently, we were interested for scoring purposes in a subject's ability to produce [s] before [i] in TL words, and, conversely, their ability to produce [S] before vowels other than [i]. A subject's score, therefore, is the percentage of relevant segments produced in the appropriate TL contexts, where that context is different from where that segment occurs in the NL. For example, Korean subjects were given credit for exhibiting the /s/–/S/ contrast in nonderived contexts only if the subjects reached criterion (see below) producing [s] in words where [s] occurred before [i], and also reached criterion producing [S] in words where this sound

occurred before some vowel other than [i]. We did not score, in other words, Korean subjects' productions of [s] before vowels other than [i], or their pronunciation of [S] before [i], because this is where these segments occur in the NL. In short, we scored only those productions of the relevant sounds that were in a non-NL position; had we scored segments in the environments where they occurred in the subject's NL, the scores would have been artificially inflated.

One other point needs to be made about scoring. Only the features that were relevant to the particular contrasts in question were scored. In some cases, this meant that the subject was given credit for a "correct" production, even though the segment the subject produced may not have been entirely target-like. For example, virtually all of our Spanish subjects devoiced final obstruents to some extent, causing them to render words such as ***head*** variably, at times as [hɛt] and on other occasions as [hɛd]. Because voicing was not the focus of this study, the subject was given credit in these cases for producing a /d/, despite the fact that a voiceless alveolar stop was produced. Likewise, if the subject spirantized the final stop and produced variably [hɛð] as well as [hɛθ], the subject was scored as producing a word-final /ð/, despite the fact that it was realized as its voiceless counterpart. To do othenvise would have artificially inflated the error rates on this contrast as well.

The data were then analyzed to determine whether the subjects exhibited the relevant contrasts in both the derived and nonderived contexts. The criterial threshold used to determine the presence of a contrast was successful production of the contrast in at least 80% of the attempts in two consecutive sessions[6]. This criterion was chosen because we observed that any subject whose performance exceeded 80% for two straight sessions did not subsequently fall below the 80% threshold. Thus it seemed that 80% performance represented a systematicity from which the subject did not later retreat.

Those subjects who lacked the relevant contrast in both derived and nonderived environments were entered into the instructional study. Those that evidenced the contrast in at least some positions were not eligible for the instructional study, and were therefore designated for the cross-sectional study, the results of which we now outline.

IV.1.d. Results of the cross-sectional study

As it turned out, there were no Stage I Korean subjects; therefore, the cross-sectional results include those from all seven of the Korean subjects, plus two Spanish-speaking subjects who were Stage II learners.

The protocol stipulated that only subjects who lacked the contrast in both the nonderived and derived environments were to be entered into the instructional study. Accordingly, any subject who had the contrast in question in at least one of the environments, became part of the cross-sectional study, the purpose of which was to attest only the predicted stages in (9)[7].

Figures 1 through 7 show that all of the Koreans exhibited the contrast between /s/ and /š/ in at least the nonderived context. More specifically, the facts represented in Figures 1

through 3 show that subjects K1, K2 and K3 were Stage III learners who evinced the contrast in both derived and nonderived environments. The results in Figures 4 through 7 depict Korean learners who, during the initial baseline measures, showed the contrast only in the nonderived contexts, but shortly thereafter evidenced the contrast also in the derived environment.

Figure 1. Baseline Probes for Subject K1

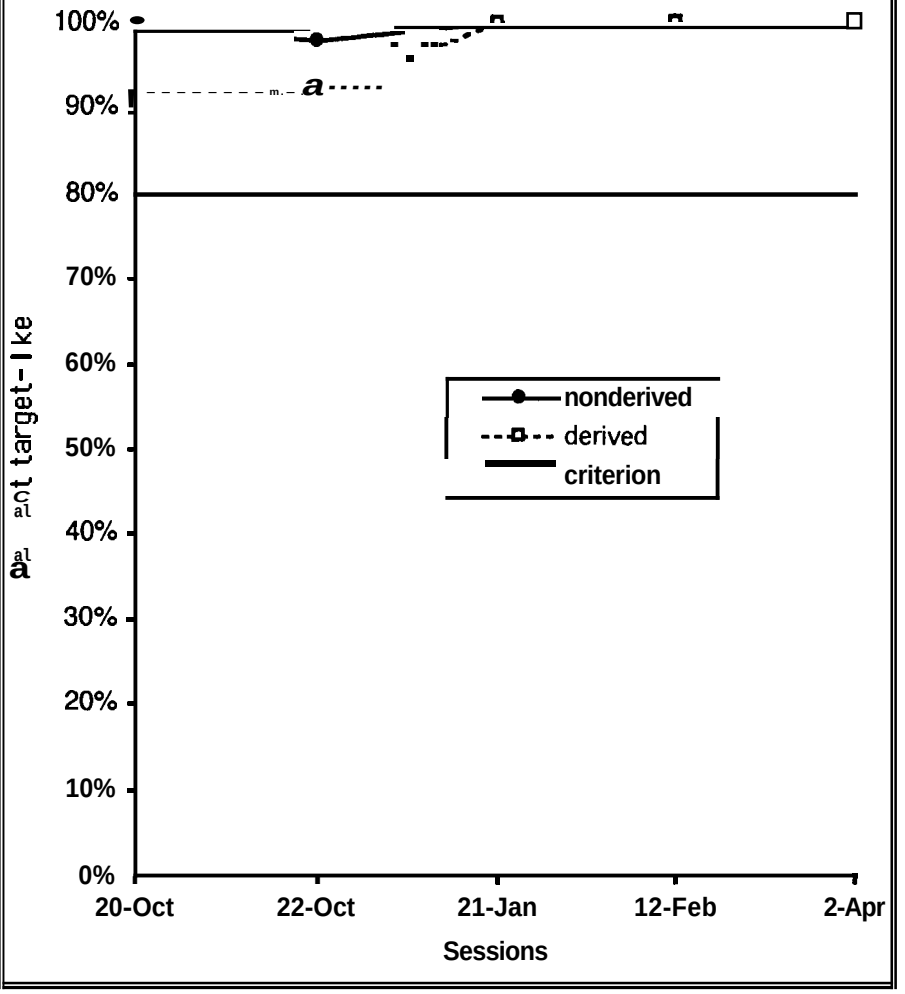


Figure 2

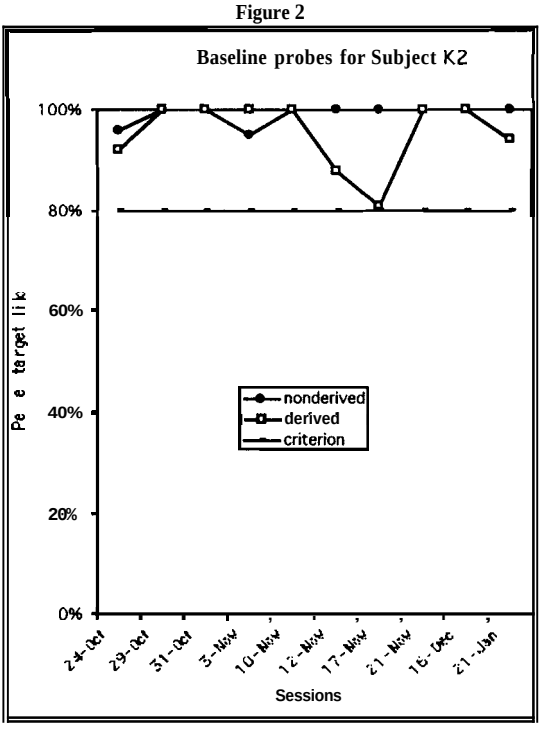


Figure 3

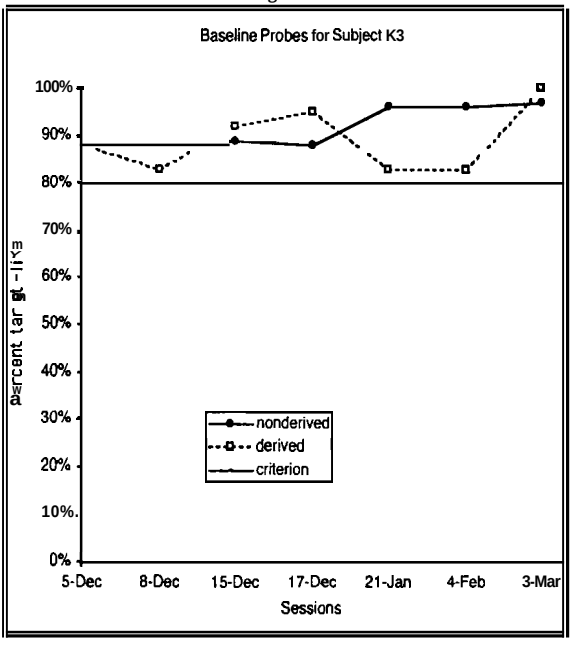


Figure 4. Baseline probes for Subject K4

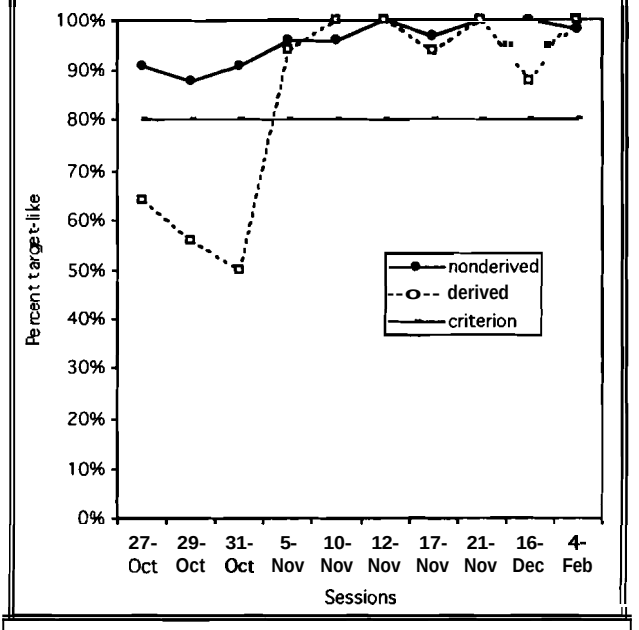


Figure 5. Baseline probes for Subject K5

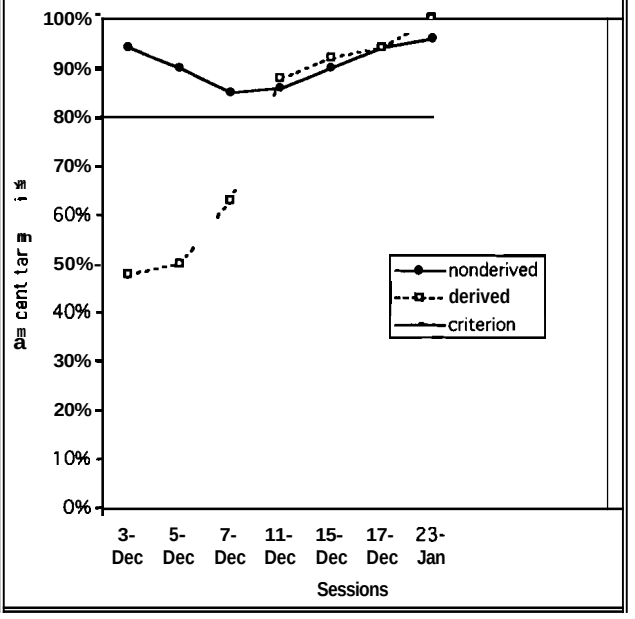


Figure 6. Baseline Probes for Subject K6

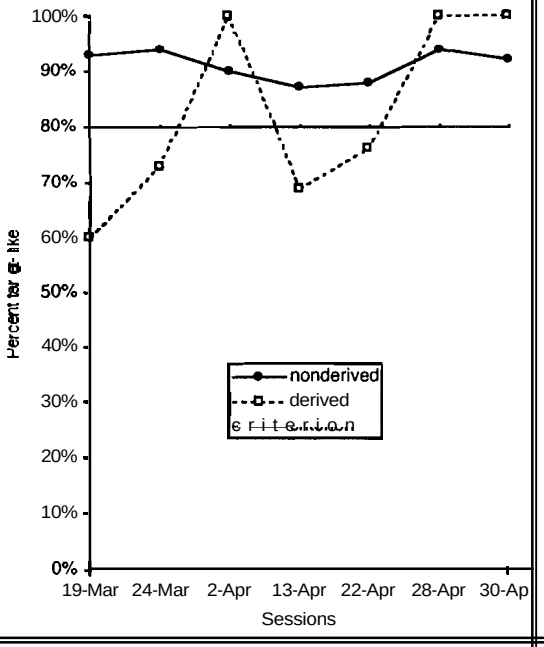
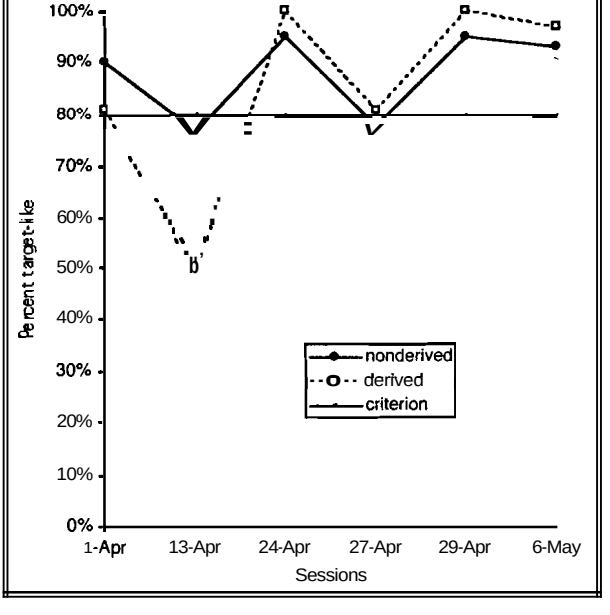
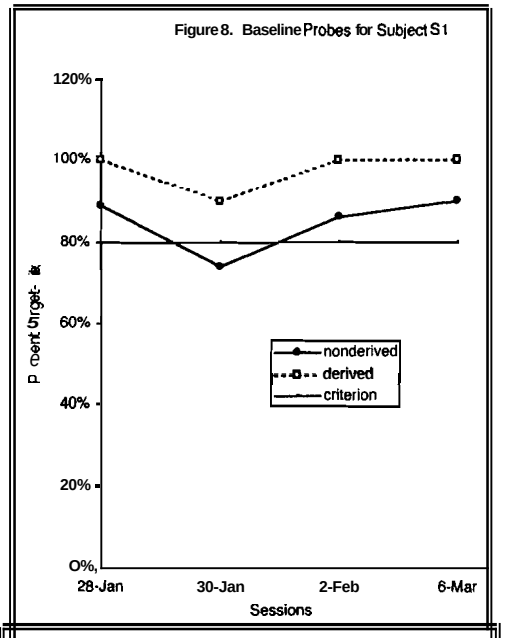
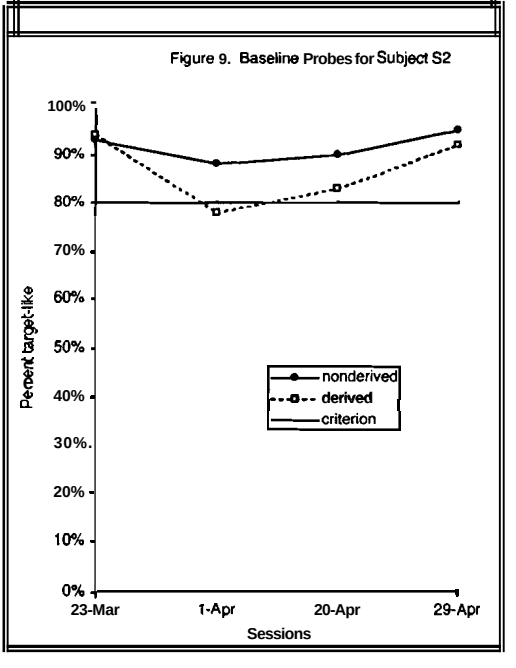


Figure 7. Baseline Probes for Subject K7



There were two Spanish-speaking subjects who also were entered into the cross-sectional study. Figures 8 and 9 represent the baseline results for subjects S1 and S2, both Stage III learners who exhibited the /d/-/ð/ contrast in both derived and nonderived environments.





In sum, all of the results from the cross-sectional study depict IL grammars that are at either Stage II, having the relevant contrast in only nonderived environments, or Stage III, evincing the contrast in both derived and nonderived contexts. None of the IL grammars we analyzed had the contrast only in derived environments. Therefore, all of the results from the cross-sectional study are in conformity with the hypothesis. We now tum to the instructional study.

IV.2. The instructional study

IV.2.a. Subjects

All of the subjects who lacked the relevant contrast in both derived and nonderived contexts, based on the baseline probes, were entered into the instructional study. As there were no Stage I Korean subjects, all seven of the subjects in the instructional study were Spanish speakers.

IV.2.b. Methodology

The subjects who were entered into the instructional study were trained on the relevant contrasts using a single-subject design (also called a within-subject design, McReynolds and Keams (1983)). Because there has been little or no discussion of such designs in the SLA literature, it would be worthwhile for us to describe this methodology in more detail. Much of what follows is based on the discussion in McReynolds and Keams (1983).

In any experimental situation, the goal is to show it was the treatment applied in the course of the experiment that caused the observed change in the subjects' behavior. Because the subjects are exposed to a variety of input and stimuli outside the experiment room during the course of the study, however, it is important for the experimenter to control for these extraneous variables, and the design of the experiment must be structured accordingly. The vast majority of experiments in the L2 literature are group designs, and although these can take several forms, the standard design is to identify a large set of subjects from which two groups are formed: an experimental group and a control group. Both groups are measured on the dependent variable (in our case, the relevant L2 contrast) at the beginning of the experiment and again at the end. In the interim, the independent variable (in our case, training on the relevant contrast in either a derived or a nonderived context) is administered to the experimental group, but not to the control group. Data from the subjects in each group are pooled and a mean is computed. The mean of the experimental group is compared with the mean of the control group, and if a difference is found, it is submitted to a statistical test to see if the difference is significant, or reliable. Extraneous variables in group designs are controlled for by randomly drawing both the experimental group and the control group from the same population, and exposing the control group to the pre-treatment and post-treatment measures, but not to the treatment itself. The

control group's performance is an indication that factors outside the experimental conditions do not have an effect on the subjects' responses. In other words, the less change the control exhibits between the pre- and post-treatment measures, the more control has been exercised during the experiment. The assumption is that the same external factors are operating on both the control group and the experimental group. If the control group's behavior does not change during this time and the experimental group's behavior does, the conclusion is that this change must be due to the treatment and not to the external variables.

In single-subject designs, by contrast, there is no control group; instead, the control is within the subject. Each subject goes through both a non-treatment and a treatment period. In other words, each subject in a single-subject design goes through all phases of the experiment, whereas in a group design the control group never receives the treatment (the experimental group goes through a treatment period but never through a time where there is no treatment). The assumption underlying single-subject designs is that although external stimuli could affect the subjects' responses, these factors are present during the non-treatment phase of the experiment as well. Thus, if the subjects' performance on the dependent variable changes during the period of treatment, the conclusion is that this change was caused by the treatment.

For our purposes, however, the clear advantage of a single-subject design as described by McReynolds and Kearns (1983) is that it enables the particular question we are posing to be addressed in the first place, and directly so: Will a learner who acquires a TL contrast in derived environments necessarily generalize it to nonderived environments, as implied by the hypothesis in (8)? As Eckman (1994) has argued in detail, questions bearing on whether IL grammars will adhere to universal principles must be addressed by studying individual IL grammars, not by using group designs in which the data are pooled. It would not even be possible, in our view, to investigate this question using a group design because the answer revolves around whether there are ***any*** IL grammars that violate the hypothesized relationship between derived and nonderived environments, not whether the mean performance of a group of subjects supports the hypothesis.

In a single-subject design, then, one subject can serve to falsify the hypothesis. In a group design, this is not the case, as there may be—and usually are—subjects whose performance runs counter to the hypothesis. Yet because the data from all subjects in the group are pooled, there may be enough subjects whose behavior is in conformity with the hypothesis to counterbalance that of a few whose performance contradicts the hypothesis. In our study, on the other hand, data from a single, recalcitrant subject are sufficient to falsify our claim. Thus, the hypothesis we are testing is claimed to hold for all learners, not just for the mean of a group.

This point, we believe, needs to be emphasized for another reason, also pointed out by McReynolds & Kearns (1983). A single-subject design allows for the recording of individualized data, whereas individual patterns may well be masked in group studies. For example, as will be seen in the results reported below, there are several ways in which a subject's performance can be in compliance with the hypothesis. Subjects, regardless of whether they were trained on the contrast in derived environments only or nonderived environments only, would support the

hypothesis if they (a) acquired the contrast only in the nonderived environment; (b) leamed the contrast in both derived and nonderived contexts; or (c) did not acquire the contrast in either environrnent. Pooling such data frorn a group study, on the other hand, may well obscure the fact that the data support the hypothesis, especially if the data reflect all three of these situations.

And finally, a single-subject design gives us the freedom to conduct studies with relatively srnall numbers of subjects. If we were to conduct a group design, we would be forced to find large nurnbers of subjects who lacked the relevant contrast before we would be able to apply the treatrnent. We would, in other words, have to wait until we could recruit nurnerous appropriate subjects before we could conduct the study. In a university-level ESL program, of course, this is not practical, because it is unlikely that there would be a sufficient number of students in the prograrn at that time who would also be at that level of proficiency.

We retum now to the description of the methodology of the instructional study. As our first step we established a baseline on each subject to determine which of them evinced the relevant contrast according to the criteria discussed above. Generally speaking, in single-subject designs, the baseline consists of the scores on the first severa1 sessions. For this study, however, we did not score the first session for the purposes of establishing the baseline, because, in the initial session, many of our subjects did not always recognize which words were being depicted by the pictures and the definitions. In these cases, the subjects were given prompts until they learned which word went with which picture. The initial sessions, therefore, elicited rnany pronunciations of the baseline words that were based on imitations. But because all of the subjects had learned which baseline word went with which picture by the second session, and no longer had to be prompted, we established our baselines beginning from the second session in which the baseline words were elicited.

For the instructional study, the baseline established the starting point for each subject with respect to the relevant contrast. As indicated, only those subjects who did not reach criterion on the relevant contrast on the baseline words were entered into the instructional study. Subjects were randomly assigned to one of two training conditions: either the subject was trained using nonce words exhibiting the contrast only in nonderived environments, or the subject was trained on nonce words showing the contrast only in derived environments. Nonce words were used for training to ensure that all subjects were equal with respect to their knowledge of the training words; that is, none of the subjects knew any of the training words at the outset. The subjects were given directions at the beginning of training that the exercise required them to produce words on the basis of a picture and a definition, as was the case with the baseline words. However, in the instructional study, the directions informed the subjects that the words used in the exercise were not real words of English, but had been made up for the purposes of this exercise.

There were twelve training words in all —six minirnal pairs— each of which was associated with a fabricated definition and a picture. An example of a picture used for the instructional study is shown in Appendix A, and the list of the training words is given in

Appendix B. Since the training words were not real words, the subjects were prompted during the initial sessions on which word went with which picture. The subjects were told in the directions that they were to try to leam the words and their associated pictures as quickly as possible. To prevent the subjects from becoming bored with the exercise, they were told that after they had learned the words on the basis of the pictures and definitions given together, they would be asked to name the words on the basis ofjust the pictures alone, or just the definitions alone. During each training session, the subjects went through eight to ten trials of the words[8]. All of the subjects had leamed which training words went with which picture and definition by the end of the second training session. The subjects were taught to make the relevant contrast through the investigators' describing and modeling the correct pronunciation, and then correcting the subjects' productions[9]. All of the subjects' pronunciations were recorded during the sessions and later transcribed by research assistants who were experimentally blind as to the intent of the study.

The specific type of single-subject design used for the instructional study was a staggered, multiple baseline design in which three subjects were entered into one training condition, and four subjects were entered into the other (McReynolds and Kearns, 1983). Each successive subject in a given condition was administered one additional baseline measure. More specifically, subjects S3, S4, and S5 received instruction on the /d/–/ð/ contrast in only derived environments, while subjects S6, S7, S8 and S9 were instructed on the contrast in only nonderived environments. Subjects S4 and S5 are considered direct replications of S3's treatment. Therefore, S3's baseline was established over two sessions, while the baselines for S4 and S5 were established over three and four sessions, respectively. The procedure was identical with the other treatment group: S6's baseline was established over two sessions, with an additional baseline measure added to the baseline of each additional, replicating subject, meaning that S9's baseline consisted of five measures.

From time to time during the training, the baseline words were elicited from the subjects. It was hypothesized that the subjects would generalize the contrast learned on the basis of the training words (i.e., the nonce words) to the baseline words (i.e., the real words). In fact, it is the subjects' performance on the baseline words that provides the test of the hypothesis: it was predicted that subjects who were trained only on nonce words exhibiting the contrast in derived environments would generalize this contrast to the baseline words and evince the contrast in both nonderived and derived environments; it was further hypothesized that subjects trained only on nonce words exhibiting the contrast in nonderived environments would not necessarily generalize this contrast to derived environments in the baseline words.

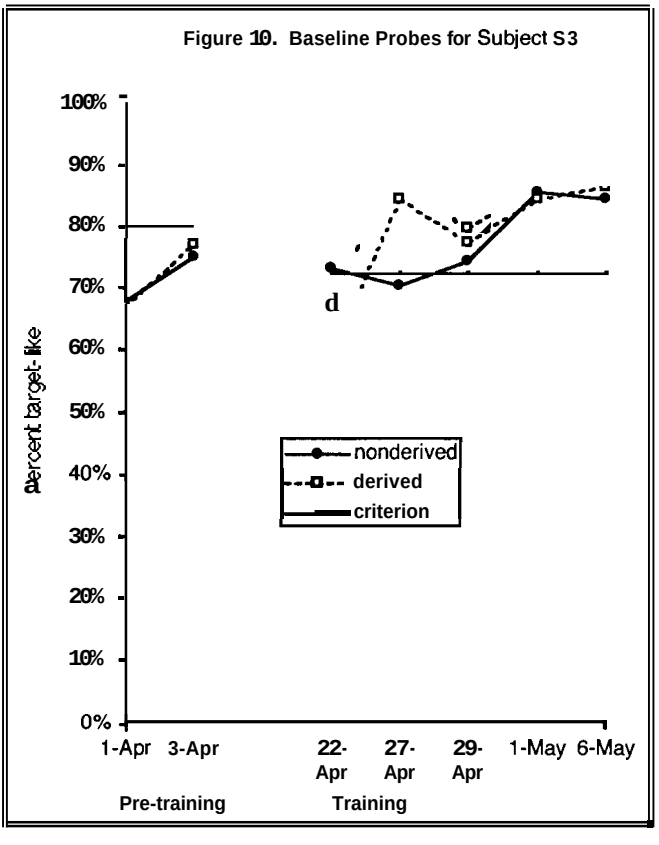
IV.2.c. Results of the instructional study

Figures 10 through 16 represent the results from the Spanish-speaking subjects entered into the instructional study. As can been seen from the graphs, none of the subjects had the contrast

between /d/ and /ð/ during the baseline, or pre-training sessions.

S3, S4 and S5 were trained on words showing the contrast only in derived environments, while S6 through S9 were trained using words containing the contrast only in basic environments. Figure 10 shows that S3 acquired the contrast in both basic and derived environments at about the same time. Figures 11 and 12 present results which are particularly interesting. S4, although trained on words with the contrast only in derived contexts, generalized this training first to baseline words with the contrast in nonderived positions, and then subsequently to derived environments, while S5, who was also trained in the derived context condition, implemented this contrast in nonderived environments, but not in derived contexts.

Stated differently, S3 responded to the treatment by quickly becoming a Stage III leamer. S4 first passed through Stage II, where she had the contrast only in basic contexts, before becoming a Stage III leamer. S4 became a Stage II leamer, and did not generalize the contrast to derived environments in the baseline words despite having been instructed only on derived-environment training words. All three of these outcomes are permissible under the hypothesis.



Subjects S6 through S9, whose results are depicted in Figures 13 through 16, respectively, were trained in the non-derived condition. As shown in Figure 13, S6 generalized the contrast from basic to derived contexts, an outcome which, while not expected, is nevertheless allowed by the hypothesis. The results from S7 are particularly interesting. She acquired the contrast in the non-derived environment on the baseline words by the 5th (February 25th) baseline session, but did not acquire the contrast in derived environrnents until the 10th baseline elicitation (May 8th). Thus, S7 clearly evidences an acquisition sequence in which she acquired the contrast first in lexically basic environments and then, more than two months later, also in morphologically composite environments. Subject S8 acquired the contrast in the basic environments in which she was trained, but did not generalize the contrast to derived environments. And S9 acquired the contrast in both environments at the same time, as was the case with S6.

Figure 11. Baseline probes for Subject S4

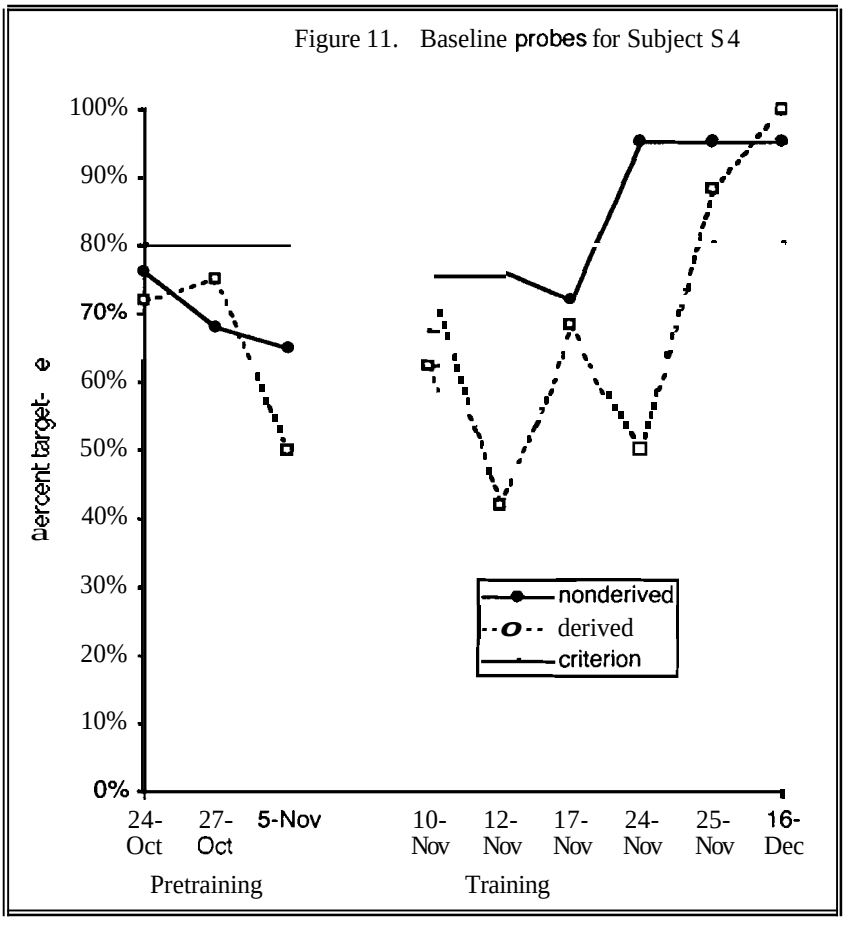


Figure 12. Baseline Probes for Subject S5

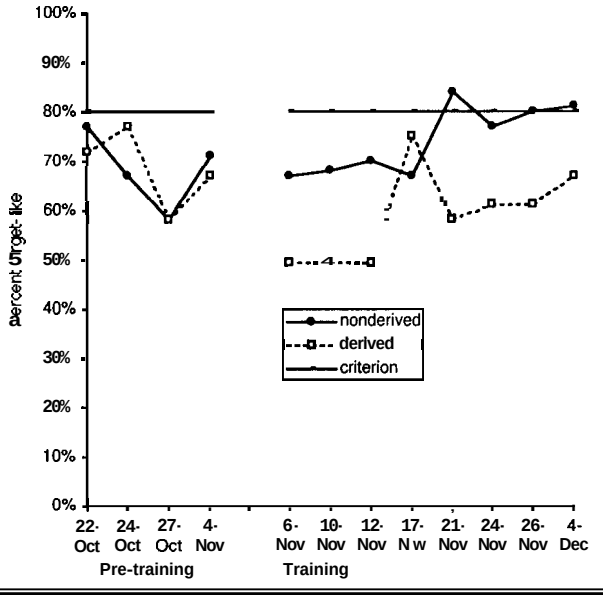
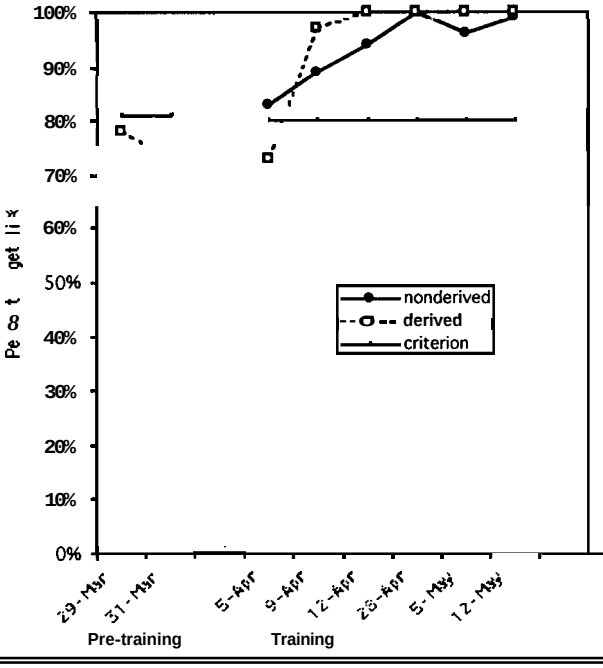


Figure 13. Baseline probes for Subject S6



Our training of Stage I subjects, then, produced leamers who were either Stage II or Stage III, while not producing any leamers whose IL grarnrnar is excluded by the hypothesis in (8). All of these outcornes confirrn our claims, with the results frorn S4, S5, S7 and S8 being supportive in particularly interesting ways.

To surnrnarize this section, results frorn our training study suggest that splitting NL allophones into separate TL phonernes entails significantly more than learning to pronounce new sounds. The acquisition of a TL contrast where none exists in the NL is, as our results support, govemed by phonological principles which constrain the acquisition to proceed through only sorne of the logically possible stages of learning

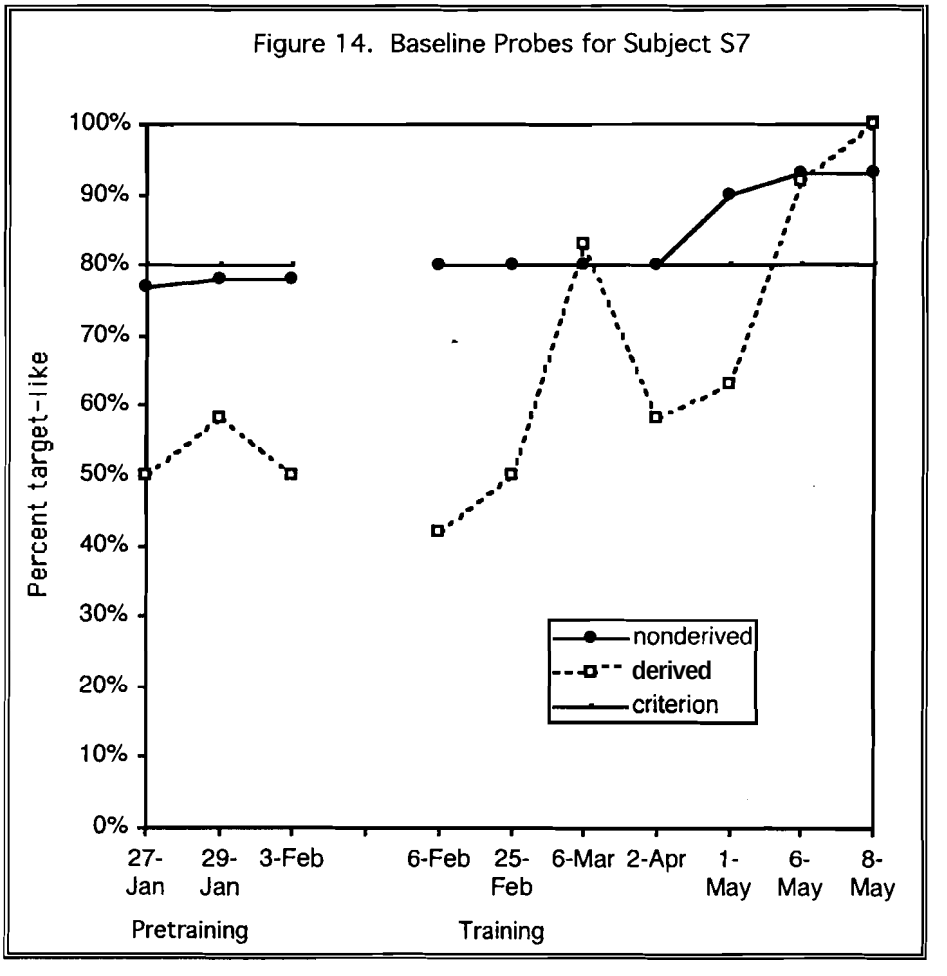


Figure 15. Baseline probes for Subject S8

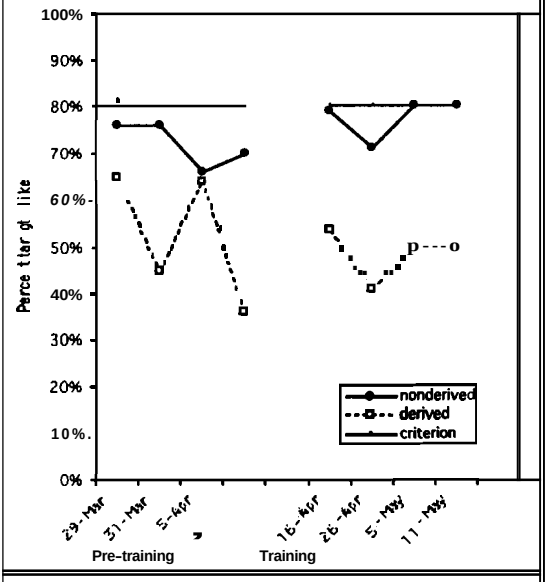
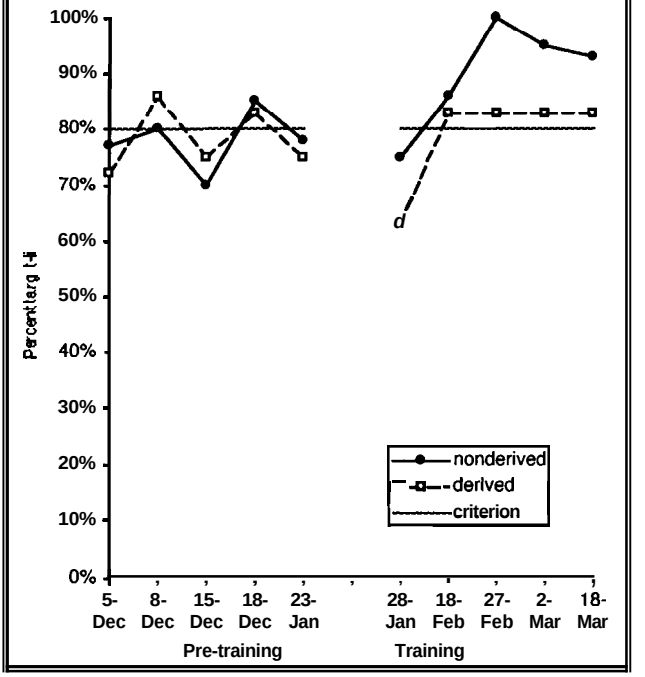


Figure 16. Baseline Probes for Subject S9



V. DISCUSSION

We focus here on three points: (1) the fundamentally abstract nature of IL phonology, (2) the fact that we encountered no Stage I Korean subjects, and (3) the implications of our findings for pronunciation pedagogy.

Results from the experimental study reported here support the claim that certain facts about the pattem of IL phonological development and interference can be accounted for through interaction of the principles of Structure Preservation and the Derived Environment Constraint. We have argued that these principles, which can be explicitly linked to conditions of learnability, provide an explanation for why one type of phonological leaming—splitting NL allophones into TL phonemes—takes place as it does.

The leaming of L2 pronunciation thus amounts to more than the simple mimicking of TL sounds. Rather, in the cases that we have considered, it is clear that acquisition of TL pronunciation involves incorporating contrasts as part of a general system that is constrained by universal principles of phonology. In our view, here as elsewhere (e.g., Eckman & Iverson 1996), second language phonology is a fundamentally abstract enterprise, parallel (though obviously not always identical) to the organization of sound structure which is characteristic of natively leamed languages. We have tried to show in this paper that the perhaps most basic of abstractions in phonology, the familiar notion of contrast, is incorporated into interlanguages in a progressive way that conforms to principles that have been uncovered in the analysis of primary languages.

The fact that we encountered no Korean-speaking leamers who lacked the contrast between /s/ and /š/ in both environments perhaps needs some comment, and two possible explanations come to mind. First, there is a possibly confounding variable among Korean leamers of English in that their NL contrasts two strident alveolar fricatives: one of these phonemes is a glottally tense /s'/ (e.g., [s'al] 'uncooked rice'), produced with increased vocal fold constriction, the other is a lax /s/ (e.g., [sal] 'skin'), produced with the breathy quality of a substantially more open glottis (Iverson 1983). Of these two phonemes, at least in the standard Seoul dialect, only lax /s/ palatalizes before /i/; thus, we have [Si] 'city', but [s'i] 'seed', i.e., we do not get *[š'i] for 'city' (Ahn, 1998). It is therefore possible that the Korean subjects were implementing the TL contrast between /s/ and /š/ before high front vowels by substituting the NL glottally tense /s'/ for English /s/ and the NL plain /s/, which palatalizes before [i], for English /š/. Indeed, many of the Korean subjects' productions of TL [s] did seem to be equivalent to NL [s']. Thus, it is possible that Korean ESL learners who have had suffícient English exposure to matriculate in an ESL program at an American university will probably already be aware of the TL contrast between /s/ and /š/, and they may well realize that this contrast can be successfully implemented using NL phones. The second explanatory factor, as implied in the work with Chinese and Japanese leamers by Brown (1998), is that it can also be the case that the Korean subjects are rather easily able to implement a plain vs. palatalized

contrast in fricatives because Korean already contrasts plain anterior versus palatalized corona1 segments, e.g., /t/ vs. /č/. Still, nothing in this observation would account for the stages of acquisition which are hypothesized in (9) and attested in our studies.

The final issue we address concems the implications of our findings for second language pedagogy, and here we have two points to make. The first reflects back to the claim we made above, namely, that learning L2 pronunciation involves far more than simply mimicking TL sounds. IL phonology, in other words, is abstract in that it invokes higher-order principles of phonological theory while incorporating phonemic contrasts into a system. And as L2 pronunciation takes place in stages, instruction and assessment of pronunciation must take these stages into account.

To be more specific, let us ask what might be indicated by systematic learner errors relating to an allophonic split made in monomorphemic lexical items versus errors made in words that are morphologically composite. According to the framework we have proposed, systematic errors made on the contrast in basic, monomorphemic lexical items indicate that the learner is at Stage I. If mistakes are made here, according to our findings, the leamer will err in morphologically composite items as well. Errors made only in derived contexts, on the other hand, indicate progress in leaming the contrast. In our framework, this indicates a Stage II leamer, the point at which the contrast has been learned only partially (in terms of the contexts in which it has been acquired). Conversely, the absence of errors in monomorphemic forms does not mean that the contrast is completely mastered, as the leaner may still err in derived contexts. Our point, simply stated, is that not all errors involved in splitting NL allophones are "equal"—some errors (derived contexts) are "better" than others (monomorphemic contexts) in that they indicate progress in acquisition.

And finally, these points can be applied to pronunciation instruction as well. We note that recent methodological principles in pronunciation pedagogy (Celce-Murcia, et al. 1996) stress that pronunciation teaching cannot focus only on words, but must also take larger domains such as the sentence and discourse into account. The results from our studies support these claims, for the added reason that the distinction between derived and nonderived contexts, in the sense expressed by the Derived Environment Constraint, is crucial to a learner's fully acquiring the TL contrast between noncontrasting NL sounds.

VI. CONCLUSION

In this paper we have reported and atiernpted to explain the stages and pattems involved in the acquisition of a split between NL allophones. We have argued, on the basis of both cross-sectional and instructional data, that the principles of phonological theory, which can be linked to leamability, govem the way in which this acquisition takes place. We have tried to show, in particular, that TL contrasts between NL allophones are incorporated into interlanguages

progressively, not at once, and that the progression follows a path which is laid out by the interaction of two very general phonological considerations: the Derived Environment Constraint and Structure Preservation.

NOTES

1. Although we use the term "err consistently", we do not want to imply that there is no variation here, as variation in the /l/s of L2 learners has been well documented and is clearly present in our own data.

2. We chose the baseline words, as much as possible, according to how easy they were to picture and how likely it would be that the subjects were familiar with the words. For the Spanish speakers, we chose words that had the targeted contrast in onset position before a vowel, in coda position following a vowel, and in the middle of a word between vowels. For the Korean subjects, we invoked the same considerations, but in addition we chose words instantiating the contrast before the high front vowels [i] and [I], as well as before other vowels. In the lists, any word with the suffix *–y* or *–ing* is a derived context.

3. The percentage of agreement varied from subject to subject, and from group to group, though the percentage of disagreement between the live transcription (which included only the consonants in question) and the tape transcription never exceeded 0.97%. The higher disagreement percentages occurred, in general, with the Spanish-speaking subjects more than with the Korean-speaking subjects, as it was more difficult to distinguish [d] and [ð] on the tape than it was to differentiate [s] and [S].

4. The reliability figure based on the re-transcription of randomly-selected portions of the tape is lower than that computed between the live transcription and the tape transcription because the former was based on a point-to-point comparison between transcriptions of the entire word, whereas the latter was based on a comparison of the just the consonants in question. The research assistants transcribed the subject's pronunciation of the whole word, on both the original transcription and the re-transcription, so that the assistants could remain experimentally blind as to what the focus of the study was.

5. One of the anonymous reviewers questioned why we did not conduct spectrographic analyses of the subjects' utterances, citing that this could have pointed out cases of "covert contrast" or "near merger" in which subjects may be making a contrast, but in a way that does not phonetically match how the contrast is implemented in the TL (Flege 1980). While we agree that it is reasonable to ask whether there are instances of our subjects' making a covert contrast between the segments in question, we also believe that, within an L2 context, it is interesting to investigate whether the subjects are producing the appropriate phonetic categories as perceived by native speakers of the TL. Given this as the goal of our study, it is rather beside the point whether the subject is making a covert contrast or near merger.

6. An anonymous reviewer pointed out that the 80% criterion is often used without discussion in the SLA literature, and further suggested that instead of using such a threshold, we should report the scores in terms of percentages and statistical levels of significance. We believe, however, that establishing a meaningful criterial threshold is the most insightful way to report the data, and further, that employing levels of statistical significance does not obviate the need for the criterial threshold. First, we consider that performance at the 80% level on two successive sessions is meaningful because, as we stated in the text, this represents a level of systematicity below which the subjects did not fall at a later date. And second, simply reporting percentages and levels of significance, as the reviewer suggested, does not address the questions we

are posing. To test our hypothesis, we must be able to say whether or not a given learner has the contrast in question. The basis for this decision, it seems to us, is whether the subject evidences enough systematicity with respect to that contrast for one to confidently conclude that the contrast is present. Suppose that a given subject performs at the 40% level in the nonderived context and at the 20% level in the derived context, and suppose, further, that it can be shown statistically that those two levels are significantly different. This result still does not provide an answer to the question as to whether the subject *has* the contrasts in the specified contexts, because one still has to decide whether 40% and 20% are systematic enough to warrant the conclusion that the contrast is present. Consequently, the use of statistical levels of significance does not remove the need for a criterial threshold.

7. The subjects who were entered into the cross-sectional study were, while the instructional study was being conducted, held in an extended baseline phase, during which time the investigators continued to meet with the subjects and to elicit the baseline words. This is why there are as many as ten baseline measures on some of the cross-sectional subjects.

8. The number of tokens of both the baseline words and the training words varied for each subject, which is why we report the scores in terms of percentages. In the initial sessions of the baseline words, the subjects went through four or five trials of the words; in the later baseline sessions, as the pictures and definitions became much more familiar, the subjects went through only two or three trials. In any given baseline session, however, the subject performed at least two trials of the baseline words. The number of tokens of the training words also varied from subject to subject and from session to session. In the earlier training sessions, subjects went through the words more slowly, producing on average five or six trials of each word. In the later sessions, subjects often produced up to ten trials of each word. In the later sessions, to prevent the subjects from becoming bored with the exercise, the training words were also elicited on the basis of only the pictures or only the definitions.

9. The training given to the subjects dealt only with the consonants in question ([d] and [ð] for the Spanish speakers, [s] and [S] for the Koreans), and thus did not focus on the pronunciation of vowels or on the production of other consonants. Moreover, there was nothing innovative or "exciting" about the training: the pronunciations were modeled, at times as single words and at other times as part of a minimal pair, and the subjects were then given feedback on their productions. In short, the focus of the study was not to investigate the effects of learning a contrast based on different teaching methods, but rather to identify the grammatical implications of learning a contrast in a given environment.

REFERENCES

Ahn, S-C. (1998). *An Introduction to Korean Phonology.* **Seoul: Hanshin Publishing Co.**

Brown, C. (1998). The Role of the L1 Grammar in the L2 Acquisition of Segmental Structure. *Second Language Research,* **14, 136-193.**

Celce-Murcia, M., D. M. Brinton & J. M. Goodwin. (1996). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages.* **Cambridge: Cambridge University Press.**

Dinnsen, D. A. & Barlow, J. A. (1998). On the Characterization of a Chain Shift in Normal and Delayed Phonological Acquisition. *Journal of Child Language,* **25, 61-94.**

Eckman, F. (1994). Local and Long Distance Anaphora in Second Language Acquisition. In E. Tarone

et al. (Eds.), *Research Methodology in Second Language Acquisition,* (pp. 207 – 225). Mahwah, NJ: Lawrence Erlbaum Associates.

Eckrnan, F. R. & G. K. Iverson. (1996). Onset and Coda Preferences in Korean-English Interlanguage. In Jae-Kee Shirn, et al. (Eds.), *Essays in Honor of Ki-Moon Lee,* (pp. 916-938). Seoul: Shin-gu Publishing Co.

Eckrnan, F. R. & G. K. Iverson. (1997). Structure Preservation in lnterlanguage Phonology. In S. J. Hannahs and M. Young-Scholten (Eds.), *Focus on Phonological Acquisition,* (pp. 183-207). Arnsterdarn & Philadelphia: John Benjarnins Publishing Co.

Eckrnan, F. R. & G. K. Iverson. (1999). Learnability in Second Language Pronunciation: Splitting Native-Language Allophones into Target-Language Phonernes. In *Representation and Process: Proceedings from The Third Pacific Second Language Research Forum.* Tokyo: Aoyarna Gakuin University.

Flege, J. (1980). Phonetic approxirnation in second language acquisition. *Language Learning,* 30, 117–134.

Flege, J. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics,* 15, 47-65.

Hamrnerly, H. (1982). Contrastive Phonology and Error Analysis. *International Review of Applied Linguistics* 20, 17-32.

Iverson, G. K. (1983). Korean /s/. *Journal of Phonetics,* 11, 191-200.

Iverson, G. K. (1993). (Post)lexical Rule Application. In E. M. Kaisse & S. Hargus (Eds.), *Studies in Lexical Phonology.* (pp. 255-275). New York: Acadernic Press.

Kenstowicz, M. (1994). *Phonology in Generative Grammar.* Carnbridge, MA: Blackwell Publishers.

Kiparsky, P. (1973). Abstractness, Opacity, and Global Rules. In O. Fujimura (Ed.), *Three Dimensions of Linguistic Theory.* (pp. 57-86). Tokyo: TEC.

Kiparsky, P. (1982). Lexical Phonology and Morphology. ln I. S. Yang (Ed.), *Linguistics in the Morning Calm.* (pp. 3-91). Seoul: Hanshin.

Kiparsky, P. (1985). Some Consequences of Lexical Phonology. *Phonology Yearbook,* 2, 85-138.

Lado, R. 1957. *Linguistics Across Cultures: Applied Linguistics for Language Teachers.* Ann Arbor: University of Michigan Press.

Major, R. & E. Kim. (1999). The Similarity Differential Rate Hypothesis. In J. Leather (Ed.), *Phonological Issues in Language Learning: A supplement to Language Learning* 49, Supplement 1, (pp. 151-183). Ann Arbor, MI: Basil Blackwell Publishers.

McReynolds, L. V. & K. P. Keams. (1983). *Single-Subject Experimental Designs in Communicative Disorders.* Austin: Pro-Ed.

Morley, J. (Ed.) (1987). *Current Perspectives on Pronunciation.* Washington, D.C: Teachers of English to Speakers of Other Languages.

Morley, J. (1991). The Pronunciation Component in Teaching English to Speakers of Other Languages. *TESOL Quarterly.* 25,481-520.

Morley, J. (1994). A Multidimensional Curriculum Design for Speech-Pronunciation Instruction. In Joan Morley (Ed.), *Pronunciation Pedagogy and Theory.* (pp. 66-91). Alexandria, VA: Teachers of English to Speakers of Other Languages, Inc.

Vennemann, Theo, gen. Nierfeld. (1972). Rule Inversion. *Lingua,* 29, 209-42.

