13

Explicit Training and Implicit Learning of L2 Phonemic Contrasts

FRED R. ECKMAN
University of Wisconsin–Milwaukee

GREGORY K. IVERSON
University of Wisconsin–Milwaukee and University of Maryland Center for Advanced Study of Language

ROBERT ALLEN FOX AND EWA JACEWICZ
The Ohio State University

SUE ANN LEE
Texas Tech University

THE PURPOSE OF THIS PAPER is to report preliminary findings of an ongoing investigation into constraints on the acquisition of L2 phonemic contrasts. We elicited production and perception data in two of the three logically possible ways in which a NL and a TL can differ with respect to a two-way phonemic contrast, as listed in (1).

(1) NL–TL Differences in a Two-Way Phonemic Contrast
  (a) The NL lacks sounds corresponding to either of the two TL phonemes.
  (b) The NL has sounds corresponding to one, but not both, of the two TL phonemes.
  (c) The NL has sounds corresponding to both of the TL phonemes, but in complementary distribution as allophones of the same phoneme.

This chapter considers only the latter two language-contact situations, those depicted in (1b) and (1c), and reports on the elicitation of both production and perception data to investigate these two NL–TL combinations.

A language-contact situation that illustrates (1b) as well as (1c) arises with respect to Korean as the NL and English as the TL. Thus, exemplifying (1b), Korean has [p], as does English, but Korean lacks [f] altogether, a sound that stands in phonemic contrast to /p/ in English. At the same time, both Korean and English have [s] and [š], but whereas these sounds contrast in English, they are in complementary distribution in Korean, because [š] occurs only before a (phonological) high front vowel

or glide, and [s] occurs elsewhere. In Korean these two sounds are related by a principle or rule such as that in (2).

(2) Korean Allophonic Rule

/s/ is realized as [š] before the high front vowel or glide, elsewhere as [s].

The task of a Korean learner of English in acquiring these two contrasts seems straightforward: on the one hand, the learner must acquire the phoneme /f/ to differentiate words such as *pan* and *fan,* and, on the other hand, the learner must suppress the application of the NL allophonic rule so as not to render *see* the same as *she.* If the early stages of the interlanguage (IL) grammar are tied closely to the NL phonological patterns, then the learner will err on TL words containing /f/, most likely substituting /p/ (*fan* = *pan*), the phonetically closest segment in NL inventory, and the learner will also transfer (2) into the IL grammar, erring on TL words containing [s] before a high front vowel (*see* = *she*).

However, general principles of phonology, to be discussed below, constrain the application of (2) in the IL and thus restrict the errors that the second-language learner makes, depending on whether the learner is at a stage in which [š] represents the phoneme /š/ rather than /s/ in the IL, at least for some words. Through the hypotheses developed in detail below, we predict that L2 learners who acquire a contrast such as that in (1c) in what we term morphologically derived environments will necessarily generalize that contrast to morphologically basic environments but not vice versa. Employing these general phonological principles as the basis for an intervention strategy, we attempted to manipulate the learning and generalization of the /s/–/š/ and /p/–/f/ contrasts.

The remainder of the chapter is structured as follows. The next section outlines the background for the study, setting the theoretical basis for the work and connecting it to previous research in the area. The section concludes with the statement and rationale of the hypotheses. We then lay out the methodology used to elicit the data. The two sections following deal with the findings, reporting them in the results section and interpreting them in the discussion section. The final section concludes the chapter.

Background

This section sets the context for the study by reviewing the literature in the two areas where the findings impinge on previous work. The discussion is followed by a description of the theoretical grounding for, and the statement of, the hypotheses.

The results of this research can be seen as a contribution to the discussion of two areas of L2 phonology, the first a long-standing issue dating back to the times of the Contrastive Analysis Hypothesis (CAH) (Lado 1957), the second a more-recent question on the relationship between a second-language learner's perception versus production of a TL phonemic contrast. We consider each in turn.

As far back at least as the CAH, allophones have played a significant role in hypotheses about L2 learning difficulty. The central claim of the CAH is that differences between the NL and TL are the major source of difficulty in L2 acquisition (Lado 1957, 2). A corollary of this claim is that the degree of difficulty associated

with any given NL–TL difference is a function of the degree of difference between the NL and TL. Thus, the area of maximum difficulty in Lado's terms would result from structures where the NL and TL are maximally different. Although it may be possible to interpret the notion of maximal difference between the NL and TL in several ways, a reasonable interpretation in terms of the present discussion would be that of an NL–TL combination as in (1a) or (1b), where the NL is lacking one or both of the TL phonemes in question. However, instead of adopting this conclusion as characterizing maximum difficulty, Lado states that "when one significant unit or element in the native language equates bilingually with two significant units in the foreign language we have maximum learning difficulty" (1957, 15).

This statement depicts the language contact situation exemplified for (1c), in which Korean learners of English must split the allophones [s] and [š] into separate phonemes. In fact, the example that Lado uses to illustrate his claim about maximum difficulty is that of a learner whose NL is Spanish, which has the sounds [d] and [ð] as allophones of /d/, who then must split the categorization of these sounds into two phonemic units in English.

In addition to the seminal work by Lado, allophones have played an important role in the description and explanation of L2 phonological difficulty up through recent work. Considerations of space allow discussion on only one such work, viz., that of Hammerly (1982), who conducted a pronunciation study in which he proposed a six-level hierarchy of difficulty. Allophones figured into the four highest degrees of difficulty, the highest of all being NL allophones that fall into a different (including contrastive) distribution in the TL, supporting on empirical grounds the above claim about maximum difficulty.

Allophonic distribution is also part and parcel of more recent proposals regarding L2 phonological difficulty. The first two hypotheses of Flege's (1995) Speech Learning Model (SLM) can be construed in terms of Lado's claim. Specifically, the SLM's first hypothesis asserts that NL and TL sounds are related to each other at the allophonic level. The second hypothesis claims that, in the process of L2 acquisition, new phonetic categories are set up if a phonetic difference is perceived between the sounds in question. As allophones are outside of the lexicon and therefore unlikely to be perceived by the learner, the hypothesis predicts that TL sounds that correspond to NL allophones of the same phoneme are unlikely to be perceived as different and thus not likely to be set up as different categories, that is, as distinct phonemes.

Whereas it is clear that the role of NL and TL allophones has been prominent in various proposals regarding L2 phonological difficulty over the decades, what is equally apparent is that no one has been able to suggest an explanation for this fact. The findings of the present study have a bearing on this question, and we will speculate on a possible place to look for an explanation.

We now turn to the second area of L2 phonology that forms part of the context for this study, specifically, the relationship between a L2 learner's perception of a TL phonemic contrast and the production of that contrast.

One of the overarching questions in this respect is whether learners have to perceive contrasts before they can produce them. Although it would seem to be intuitively clear that a learner must perceive any given contrast before being able to implement

it in production, the literature in fact attests all four logical possibilities: a contrast can be neither perceived nor produced; it can be both perceived and produced; it can be produced but not perceived; and it can be perceived but not produced.

Two of these possibilities are straightforward and require little or no elaboration: there are numerous documented examples of learners who can neither perceive nor produce a TL contrast that is absent in the NL, as well as instances in which L2 learners can both perceive and produce TL contrasts that happen to match up with NL distinctions (e.g., Bion et al. 2006). The other two logical possibilities are not as straightforward and therefore require discussion.

The first of these two, the one that is more intriguing, is that L2 learner productions of certain contrasts can exceed their ability to perceive those contrasts, which has been reported in Sheldon and Strange (1982), who replicated and extended earlier work. Sheldon and Strange tested native speakers of Japanese learning English on their ability to perceive and produce the distinction between /r/ and /l/, a contrast lacking in Japanese. Specifically, it was found that native speakers of English, when listening to recordings of the subjects' productions of minimal pairs containing /r/ and /l/, could successfully distinguish /r/ from /l/ better than the subjects themselves could distinguish this contrast in their own productions.

It is these results from Sheldon and Strange (1982) that clearly render the relationship between phonological perception and production not straightforward and that raise the question of how L2 learners can produce a contrast unless they know that a contrast has to be produced.

The answer, it seems, lies in the fact that the L2 subjects in the Sheldon and Strange study cited above received written input on the contrasts in question. The subjects did not have to hear the difference between /r/ and /l/ in order to know that they had to produce this contrast; they could discern that a contrast existed between /r/ and /l/ from the spelling of the words. Therefore, in language acquisition by preliterate children, it can still be maintained that perception of a contrast will precede its successful production; however, in L2 acquisition, if the learner is provided with written access to the contrast, or at least with some nonauditory way to discern the distinction, it is possible that production approximating a contrast may precede its production.

The final logical possibility with respect to perception and production of a TL contrast is the one in which the learners can perceive the contrast in question but are unable successfully to produce it. As outlined above, Flege's (1995) SLM is predicated on the notion of "equivalence classification," according to which an L2 learner sets up phonetic categories for TL phones on the basis of the learner's perception of the TL segments in terms of the established NL categories. In a review of the literature on the relationship between L2 learners' perception and production of TL contrasts on both consonants and vowels, Flege (1999) found that perception and production are correlated but only weakly so. The works examined by Flege included selected studies on the perception and production of consonants as well as vowels; however, due to space limitations, we will limit ourselves to the studies on consonants.

Aoyama et al. (2004) investigated Japanese-speaking learners of English on the perception of English /r/ and /l/. Because English /r/ (a rhotic vocoid without

tongue contact) is perceptually less similar to Japanese /r/ (an alveolar tap) than is English /l/ (an alveolar lateral, also with tongue contact), the SLM predicts that English /r/ should be acquired by Japanese learners faster than English /l/. The results supported this. More recently, Kluge et al. (2007) studied the production and perception of English /m/ and /n/ in coda position by native speakers of Brazilian Portuguese, for whom these are merged in the NL. The results showed a positive correlation between the perception and production tests. Given this background, we now turn to the theoretical context for the four hypotheses forming the basis for this study.

The assumption underlying the general research program to which this study belongs is that IL grammars are the way they are, in part, because they are constrained by general grammatical principles. Specific to the research being reported here is the hypothesis that two of these principles, listed below in (3) as adapted from work by Kiparsky (1982, and elsewhere), can also form the basis for intervention strategies regarding the IL grammars.

(3) Phonological Principles

(a) Structure Preservation

Representations within the lexicon consist only of elements drawn from the phonemic inventory.

(b) Derived Environment Constraint

Structure-preserving rule applications are restricted to derived environments (i.e., rule applications that involve phonemes of the language apply only across morpheme boundaries).

Structure Preservation states that words and morphemes in the lexicon of a grammar comprised only phonemes; no allophones are part of the lexicon. The Derived Environment Constraint claims that rules that apply to, or produce, phonemes must apply in environments that arise from putting two morphemes together. That is, such rules must produce morphemic alternations. Rules that produce allophones are not restricted in this way and so can apply everywhere, in morphologically simple and morphologically composite environments.

As outlined in Eckman, Elreyes, and Iverson (2003), Structure Preservation and the Derived Environment Constraint have important implications for learnability, in general, and for the acquisition of the English contrast between /s/ and /š/, in particular. It follows from these principles that the acquisition of a TL phonemic distinction whose contrasting segments correspond to allophones of the same phoneme in the NL will take place in stages. At the beginning, when the IL grammar lacks the contrast, the transferred NL rule will apply across the board in both basic and derived environments. As the learner begins to acquire the contrast in question, the two segments take on the status of phoneme, and therefore become part of the IL lexicon. As a consequence, the Derived Environment Constraint permits the rule to apply only in derived contexts, that is, only across morpheme boundaries. The last stage would be one in which the contrast has become acquired to the point where the rule is suppressed altogether.

Thus, if a native speaker of Korean learning English transfers the NL allophonic rule to the IL grammar and is subject to the two constraints in (3), then the following stages of acquisition for the /s/–/š/ contrast are predicted.

(4) Acquisition sequence

Stage I, No Contrast: not able to make the relevant target language contrast, applying the NL rule in both derived and basic environments (e.g., a Korean learner says the pairs *sea–she* and *messing–meshing* homophonously, as [ši] and [mɛšɪŋ]);

Stage II, Partial Contrast: able to make the contrast in some words, applying the NL rule only in derived environments (a Korean learner says *sea–she* correctly but errs by producing *messing–meshing* homophonously);

Stage III, Contrast: able to make the contrast in all words, applying the NL rule in neither derived nor basic environments (a Korean learner says the pairs *sea–she* and *messing–meshing* correctly);

Excluded stage: able to make the contrast in some words, applying the NL rule only in basic contexts (a Korean learner says the pair *sea–she* homophonously, but says *messing–meshing* correctly).

Of importance is that the stages of acquisition in (4) would not be predicted for learning a contrast such as (1b) because there is no allophonic rule involved.

Within this context, we now posit the hypotheses in (5).

(5) Hypotheses

a. Acquisition of the production of a contrast such as (1b) will not be sensitive to morphological structure.

b. Acquisition of the perception of a contrast such as (1b) will not be sensitive to morphological structure.

c. Acquisition of the production of a contrast such as (1c) will exhibit a derived environment effect sensitive to morphological structure.

d. Acquisition of the perception of a contrast such as (1c) will not exhibit a derived environment effect sensitive to morphological structure.

The rationale for (5a) is that there is no rule involved in the contrast; the rationale for (5b) and (5d) stems from the fact that virtually all (if not truly all) generalizations about L2 grammars have been made on the basis of data from production, not perception. The hypothesis in (5c) has its roots in the two general principles of phonology, Structure Preservation and the Derived Environment Constraint, given in (3).

Before concluding this section, we wish to emphasize the claims underlying the hypotheses in (5). As the consonantal phonemic inventory in (6) shows, Korean has lax, tense, and aspirated contrasts among its bilabial, alveolar, and velar stops, the same contrast in its postalveolar affricates, and a lax versus tense contrast in its coronal fricatives. And in addition to a three-way point of articulation contrast in nasals, Korean also has a liquid phoneme (rendered here as /l/, though it has both central

and lateral allophones) and /h/. Yet the claim embodied in the hypotheses in (5) is that, despite the possibility that a Korean learner of English may be able to perceive a number of different contrasts on the basis of the inventory in (6), the explanation for the staged learning outlined in (4) for the /s/–/š/ contrast—as opposed to /p/–/f/—is that the principles in (3) constrain the application of the NL allophonic rule after it is transferred to the IL grammar.

(6) Phonemic Inventory of Korean Consonants

p p^h p' t t^h t' k k^h k' c c^h c'
s s'
m n ŋ
l
h

Given this background along with the above hypotheses, we turn now to a description of the elicitation methods used to gather our production and perception data.

Methods

In a test of the hypotheses in (5), we elicited baseline productions and perception judgments from 10 participants on the /p/–/f/ contrast and from twenty different participants on the /s/–/š/ contrast and elicited perception judgments from the thirty participants on their respective contrast. We then trained the participants on their respective contrast using nonce words that showed the contrast either in the basic environment or in the derived environment. After the training, the baseline productions were elicited again to serve as a posttest. Of the thirty native speakers of Korean serving as participants for the study, ten were students at the City University of Incheon, Korea, ranging in age from nineteen to twenty-five years, and twenty were students in the Intensive English Program at the University of Wisconsin–Milwaukee, ranging in age from eighteen to thirty-six. A group of eight native-speaker controls was used for the perception stimuli for each contrast.

Two sets of stimuli were used in the study, one for the production of the target sounds and one for their perception. These two sets were used twice to collect subject responses at two points in time: (1) as a pretest (or baseline), at the beginning of the study before each subject entered the training phase, and (2) as a posttest, after the training had been completed. For the baseline pretest, sixty target words and thirty fillers was selected, all existing lexical items in English, each target word containing either /p/ or /f/, or /s/ or /š/, in three different positions in a morphologically basic word: initial before a high front vowel (e.g., *sip/ship*), medial (e.g., *lesson/ocean*), final (e.g., *pass/crash*) and in one additional position (medial, at the juncture with another morpheme) in morphologically composite words containing either the suffix *-ing* or *-y,* (e.g., *passing/brushing* or *messy/bushy*).

The perception stimuli consisted of naturally produced, single words recorded by a male native speaker of American English. All words were existing minimal pairs in English, in which either /p/ or /f/, or /s/ or /š/, occurred in initial and final positions in basic words and in medial position followed by the suffix *-ing* or *-y,* e.g.,

fan/pan, laugh/lap, cuffing/cupping; seep/sheep, plus/plush, classing/clashing. The stimulus set consisted of seventy-two items (four stimuli × two contrastive segments × three positions in a word × three repetitions).

Experimental Procedures

Several custom programs were written in MATLAB for the purposes of the present study. For the production of the baseline pretest and the posttest, a program controlling the recordings displayed on a computer screen a set of pictures, clues, and commands such as "Wait" or "Speak" designed to guide the subject and the experimenter in order to elicit the word in question. The stimuli were presented in a random order, recorded directly onto a hard disc drive at the sampling rate of 44.1 kHz. Subjects spoke into a head-mounted microphone at a distance of one inch from the lips.

A second program controlled the perception experiment, employing a single-interval two-alternative forced choice (2AFC) identification procedure, with the two response choices, /s/ and /š/, displayed on the computer monitor. After hearing the stimulus word, the subject indicated with the press of a mouse button whether the word contained a /p/ or an /f/, or an /s/ or an /š/, depending on the contrast being tested. The stimuli were presented in a random order over Sennheizer HD600 headphones at a comfortable listening level (~70 dB HL). Each subject was tested individually. To make the perception task more demanding, the stimuli were presented in masking white noise at two different levels of sound-to-noise (S/N) ratios: 0 dB, and −4

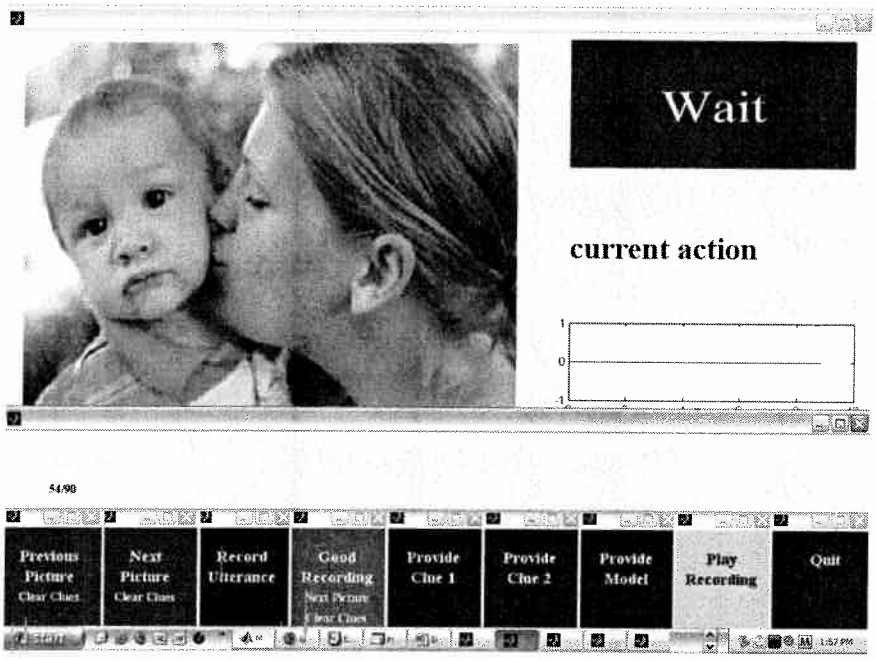

Figure 13.1 Computer Screen Display for Baseline for Eliciting the Baseline Production of the Derived-environment Word, "Kissing."

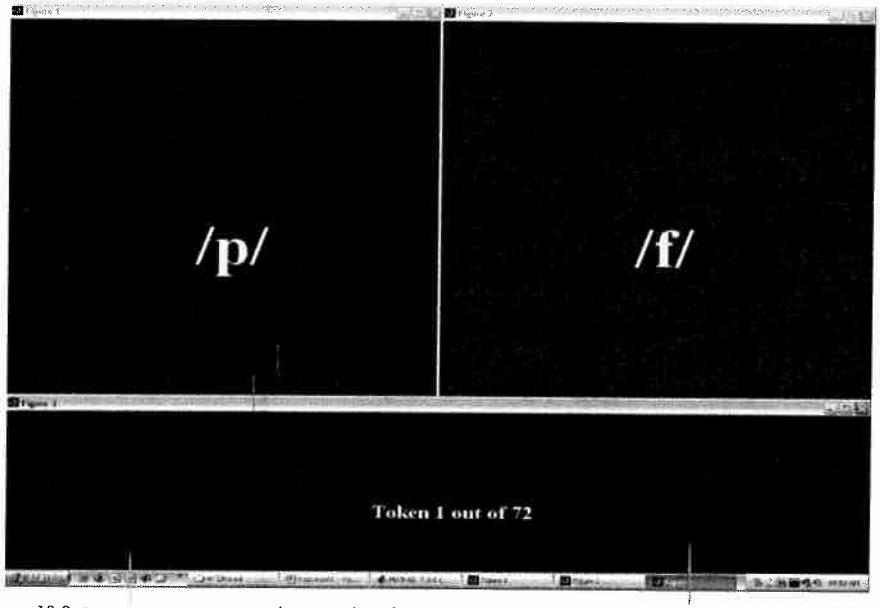

Figure 13.2 Computer Screen Display Used in the Perception Task for the /p/–/f/ Contrast.

dB. The use of masking noise is not uncommon in perception tasks because it is often necessary, as in the present study, to determine not simply whether the learner has acquired the contrast perceptually, but also the degree to which the perceptual contrast has been learned.

A third program guided the training of the participants on the production of the /p/–/f/ or /s/–/š/ contrast, depending on which baseline the person produced. The training program took the participants through a series of steps that were somewhat similar to the baseline production task in which pictures and verbal models that were presented to, and were to be learned by, the participant. Only nonce words (e.g., *nafe, kefing, hosing, hisi*) were used in the training phase.

After the training sessions the same production and perception tasks were conducted to elicit subjects' responses to the /p/–/f/ and /s/–/š/ contrast in the form of posttraining tests to assess the effects of learning.

The data were collected at Milwaukee and then transferred to Ohio State where they were transcribed by research assistants who were blind to the hypotheses. The transcriptions were then returned to Milwaukee, where they were scored. A subject's performance on a task had to reach the 80 percent criterion, the threshold that has been invoked in L2 research for several decades, in order for the subject's interlanguage grammar to be credited with having the contrast in question.

Results

In this section we describe the results according to how they bear on each of the hypotheses in (5). In production of the baselines and the posttest of the /p/–/f/ contrast, we observed no pattern of sensitivity to morphological structure (see fig. 13.3). On

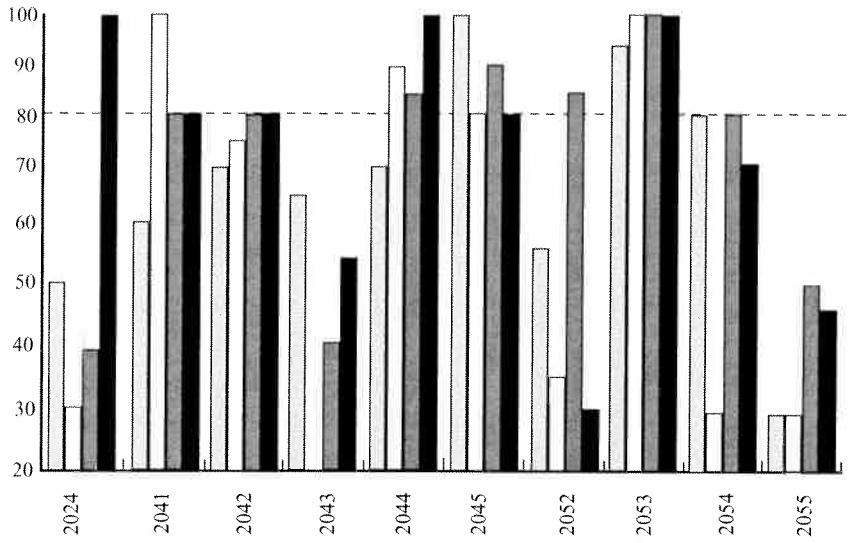

Figure 13.3 Results for the L2 Subjects on the Production of the /p/–/f/ Contrast on the Pretest and Posttest in Basic and Derived Environments

either the pretest or the posttest, the subjects' performance on the contrast evinced all four of the logical possibilities. Three of subjects (2024, 2041, 2044), on either the pretest or posttest, crossed the threshold in the morphologically composite words without doing so on the morphologically basic words; four of the subjects (2024, 2043, 2052, 2055) did not reach the criterial threshold in either environment on one or both of the tests; two of the subjects (2045 and 2053) showed that they had the contrast in both environments on both pre- and posttests; and finally, two of the subjects (2052 and 2054) showed the contrast in the morphologically basic environment without having the contrast in the morphological-composite words on the posttest. Thus, hypothesis (5a) was supported because there is no morphologically sensitive pattern in the participants' production of the /p/–/f/ contrast. Within our framework, this is because there is no NL allophonic rule associated with the /p/–/f/ contrast, and therefore the general phonological principles in (3) above do not predict the stages of acquisition shown in (4).

The results of the perception task involving the /p/–/f/ contrast are shown in table 13.1. The subject identification numbers are shown at the head of the rows in the tables, and the columns indicate whether the performance was on the pretest or posttest, whether the environment was basic or derived, and the amount of the signal degradation. As can be seen from table 13.1, the performance of only one subject (2044) reached the criterial threshold on either the pre- or posttest; the scores of all the others were much lower and showed no morphologically sensitive pattern. The control subjects performed better overall than did the L2 participants on the perception task; however, they also showed no morphologically sensitive pattern in their performance. Therefore, hypothesis (5b) was supported.

We now turn to the /s/–/š/ contrast, where we see different results than those for the /p/–/f/ contrast. In the production of the baselines and the posttest on this

Table 13.1
Results of Perception for Subjects on /p/–/f/

	Pretest 0dB		Pretest −4dB		Posttest 0dB		Posttest −4dB	
	Basic	Derived	Basic	Derived	Basic	Derived	Basic	Derived
2024	60.42	62.5	62.5	33.33	79.17	50	75	58.33
2041	54.17	45.83	62.5	62.5	60.42	37.5	54.17	45.83
2042	62.5	58.33	60.42	50	64.58	41.67	47.92	66.67
2043	54.17	25	60.42	45.83	58.33	50	60.42	33.33
2044	83.33	54.17	81.25	45.83	75	50	68.75	58.33
2045	68.75	37.5	56.25	62.5	64.58	62.5	60.42	58.33
2052	68.75	41.67	70.83	58.33	60.42	75	60.42	41.67
2053	66.67	50	58.33	58.33	58.33	58.33	62.5	54.17
2054	56.25	58.33	37.5	41.67	54.17	66.67	66.67	54.17
2055	62.5	62.5	50	62.5	83.33	66.67	70.83	62.5

contrast, five subjects lacked the contrast in both basic and derived environments (see fig. 13.4); the performance of eight subjects showed that their IL had the contrast by reaching the 80 percent criterial threshold in both basic and derived environments (shown in fig. 13.5); and seven subjects evinced a derived environment effect by having the contrast in basic environments only, on the pretest, the posttest,

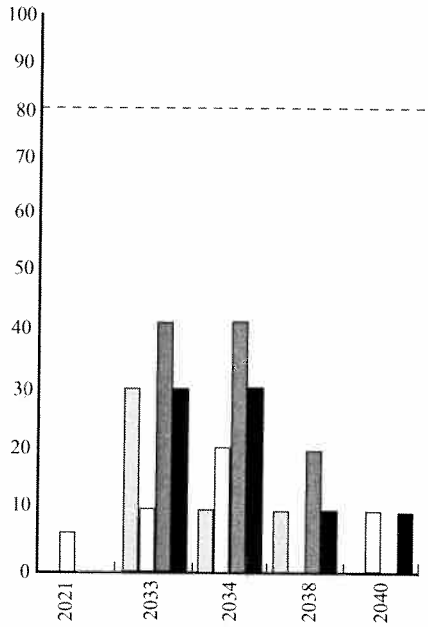

Figure 13.4 Results for the Production of the /s/–/š/ Contrast by the L2 Subjects That Lacked the Contrast in Both Basic and Derived Environments on both the Pretest and the Posttest

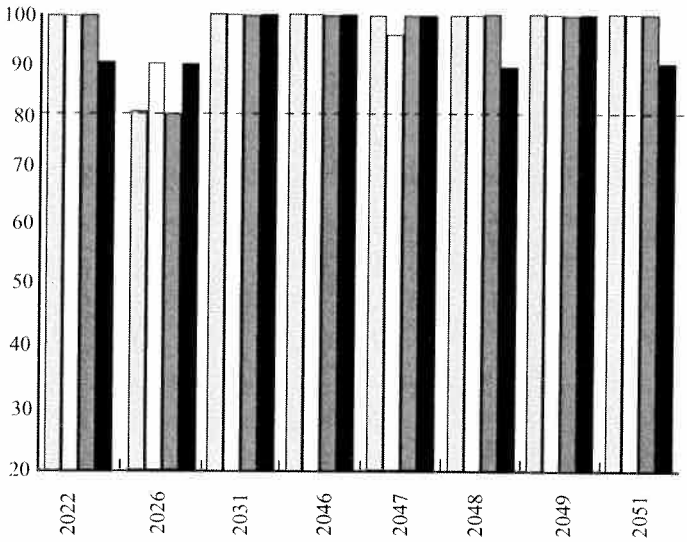

Figure 13.5 Results for the Production of the /s/–/š/ Contrast by the L2 Subjects That Exhibited the Contrast in Both Basic and Derived Environments on Both the Pretest and the Posttest

or both (depicted in fig. 13.6). No subjects evidenced an IL that had the contrast in derived environments but lacked it in basic environments. Therefore hypothesis (5c) was supported.

Table 13.2 presents the results for the L2 subjects on the perception task for the /s/–/š/ contrast. The rows and columns of the table show the same information as that

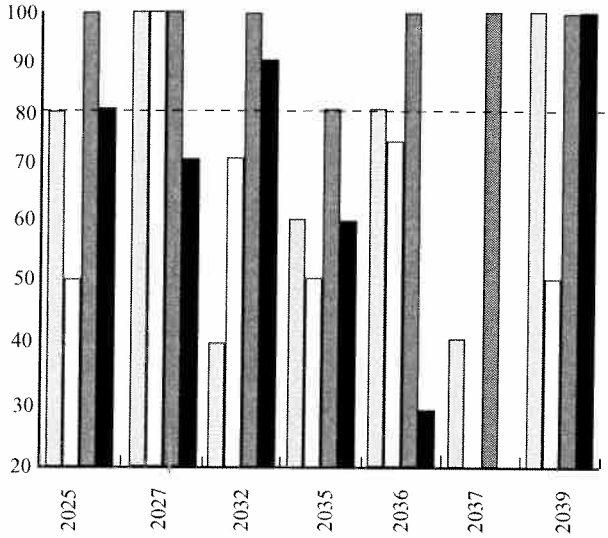

Figure 13.6 Results for the Production of the /s/–/š/ Contrast by the L2 Subjects That Evidenced the Contrast in the Basic Environment but Not in the Derived Environment on Either the Pretest or the Posttest

Table 13.2
Results of Perception for Subjects on /s/–/ʃ/

	Pretest 0dB		Pretest −4dB		Posttest 0dB		Posttest −4dB	
	Basic	Derived	Basic	Derived	Basic	Derived	Basic	Derived
2031	77.82	87.5	79.17	88	100	95.83	85.42	95.83
2032	77.08	91.67	70.83	87.5	77.08	75	68.75	83.33
2033	72.92	54.17	68.75	54.17	89.58	79.17	81.25	79.17
2034	97.92	87.5	83.33	95.83	95.83	100	89.58	95.83
2035	79.17	100	70.83	91.67	91.67	83.33	72.92	79.17
2036	83.33	83.33	87.5	91.67	97.92	100	95.83	100
2037	70.83	62.5	70.83	79.17	91.67	79.17	75	83.33
2038	75	54.17	58.33	54.17	87.5	75	72.92	75
2047	87.5	87.5	83.33	79.17	93.75	91.67	70.83	100
2048	66.67	87.5	70.83	95.83	83.33	75	79.17	70.83
2049	81.25	87.5	72.92	75	91.67	100	72.92	91.67

for table 13.1. Using once again the 80 percent threshold for acquisition of a contrast, we see that there was no patterned sensitivity to morphological structure in either the 0dB or −4 dB SNR degradation level on either the pretest or posttest. In fact, seven of the subjects performed contrary to what would be expected for a derived environment effect. Once again, the controls performed better overall on this task than did the L2 participants, but there was no pattern of morphological sensitivity. Therefore, hypothesis (5d) was supported.

To recapitulate this subsection, the results from the study support the four hypotheses. The acquisition of the /s/–/š/ contrast by Korean-speaking L2 learners of English attests the three permissible stages in (4) above, while showing no evidence of the excluded stage. Moreover, this derived environment effect was relevant only in the case where an NL allophonic rule was involved, that is, not in the acquisition of the /p/–/f/ contrast, where no allophonic rule was applicable. Thus, our general claim that IL grammars are subject to general constraints was supported, as were the hypotheses concerning the perception of the contrasts in question.

Discussion

Our discussion of the results will reprise each of the major points described in the beginning of the chapter. The first is that our findings suggest a possible line of investigation that could shed light on why it has been reported over several decades that the splitting of NL allophones into separate TL phonemic categories is so difficult. The second point of discussion is to suggest why it may be the case that the perception and production of L2 phonemic contrasts are only weakly, rather than strongly, correlated. The third point is that general principles of phonology constrain IL grammars in such a way that L2 contrasts are acquired according to different paths, some following staged development through basic and derived environments, and others

not. Finally, we take up the claim that these general principles can be the basis for intervention strategies. We begin with the question of maximum difficulty.

As the findings of our study show, acquiring a contrast between two sounds that are allophones of the same phoneme in the NL will take place in stages whereby the contrast is acquired first in morphologically simple words and then proceeds to include words that are both morphologically simple and morphologically complex. No such staging is predicted for the acquisition of contrasts where the NL lacks one or both of the sounds. Thus, past observations that splitting NL allophones into TL phonemes is more difficult than learning other kinds of phonemic distinctions may be due to the fact the learners have two stages to pass through to reach full acquisition in the case of the former, but not in the case of the latter. Of course, our claim at this point can be only suggestive until there is empirical work to address the question.

The explanation of the second point, the finding by Flege (1999) that perception and production of L2 contrasts are only weakly correlated, may also be related to the stages of acquisition associated with the production, but not the perception, of TL contrasts involving allophonic splits. Our findings suggest that the stages of acquisition in (4) hold for the production of a contrast such as in (1c) but not for the perception of that contrast. Thus, it would seem that the correlation between L2 learners' production and perception of contrasts involving allophonic splits would pattern in a way that is distinct from the correlation between the production and perception of other L2 phonemic contrasts. As with the first point above, the status of this suggestion as an explanation for the weak correlation is ultimately an empirical question and therefore must await further study.

The third point of discussion is that IL grammars are the way they are because they are constrained, at least in the case at hand, by general phonological principles. This was shown through the three stages of acquisition in (4) being attested by the production patterns of the L2 learners for the /s/–/š/ contrast, where an NL allophonic rule is involved but not in the case of the acquisition of the /p/–/f/ contrast, where no such NL rule is motivated. This point leads in turn to the final topic to be discussed, the general claim that the principles described in (3) can form the basis of a strategy for intervening in the IL; we therefore take these two point together.

We return to the results from the seven subjects whose IL grammar showed a derived environment effect (fig. 13.6), where we can observe some interesting results with respect to the effect of the training. Of these seven subjects, whose performance on the contrast in either the pretest or the posttest was at stage 2 in (4) above, four subjects were trained using nonce words with the contrast in the basic environment (2025, 2027, 2036, and 2039) and three were trained using nonce words with the contrast in a derived environment (2032, 2035, 2037). With the exception of 2027, all the subjects improved their performance between the pretest and posttest on the production of the contrast in at least one of the environments. Subject 2027 produced 100 percent of the contrasts in both basic and derived environments on the pretest, but then on the posttest scored 100 percent in the basic environment, but 70 percent in the derived environment. We have no explanation for why subject 2027 did worse on the posttest in the derived environment; however, we note that the IL pattern of contrast on the posttest nevertheless conforms to one of the predicted stages in (4).

Now let us consider the pattern of generalization of learning for the other subjects. Subjects 2025 and 2039 were trained using words with the contrast in the basic environment, and were able to generalize this training to the derived environment. This type of generalization from basic to derived environment is possible according to the principles in (3) but not necessary.

What is more interesting, we believe, in terms of the patterns of generalization of the contrast are the three cases in which the subjects were trained on words with the contrast only in derived environments. Subject 2032 lacked the contrast in both basic and derived environments on the pretest but, after training, evinced the contrast in both environments. Thus, this subject was able to generalize the contrast from the derived environment, on which training took place, to the basic environment, which was not trained. Similar results were obtained for subjects 2035 and 2037, except that on the posttest the contrast was not evidenced in the derived environment but was shown only in the basic environment, despite the fact that the contrast was not trained on the basic environment.

Thus, if a subject that lacks the contrast in both the basic and derived environments is trained on the contrast in only the derived environment and in fact learns the contrast in only that environment, the result would be an IL grammar that represents the excluded stage in (4) and is not licensed according to the principles in (3). According to our results, what happens in this case is that the contrast seems to be generalized to the basic environment, thereby producing an IL grammar that is allowed according to the principles in (3).

Conclusion

This chapter has reported findings that support the conclusion that the acquisition of L2 phonemic contrasts is constrained by general grammatical principles, which, in turn, lead to different paths of learning and can be used as the basis for intervention strategies.

ACKNOWLEDGMENTS

An earlier version of this paper was presented at the 2009 Georgetown University Round Table conference, March, 15, 2009. We would like to thank the members of the audience for their questions, comments, and general feedback. As always, any remaining errors or inconsistencies are our own. We also wish to express our appreciation to the following for their assistance with this research: Cara Campbell, Anne Hoffmann, Samantha Lyle, John Olstad, Heather Povletich, and Julia Sammet.

This work was supported in part by a grant from the National Institutes of Health 1 R01 HD046908-03. The positions expressed in this paper are those of the authors and do not necessarily reflect those of NIH.

REFERENCES

Aoyama, K., James Flege, S. Guion, R. Akahane-Yamada, and Y. Tsuneo. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 23:233–50.

Bion, R. A. H., Paula Escudero, Andreia Rauber, and Barbara Baptista. 2006. Category formation and the role of spectral quality in the perception and production of English front vowels. *Proceedings of Interspeech* 2006:1363–66.

Eckman, Fred, Abdullah Elreyes, and Gregory Iverson. 2003. Some principles of second language phonology. *Second Language Research* 19:169–208.

Flege, James. 1995. Second language speech learning: Theory, findings and problems. In *Speech production and linguistic experience: Issues in cross-language research,* ed. W. Strange, 233–77. Timonium, MD: York Press.

———. 1999. The relation between L2 production and perception. *14th International Congress of Phonetic Sciences* 99:1273–76.

Hammerly, Hector. 1982. Contrastive phonology and error analysis. *International Review of Applied Linguistics* 20:17–32.

Kiparsky, Paul. 1982. Lexical phonology and morphology. In *Linguistics in the morning calm,* ed. I. S. Yang, 3–91. Seoul: Hanshin.

Kluge, C. D., Andreia Rauber, M. S. Reis, and R. A. H. Bion. 2007. The relationship between perception and production of English nasal codas by Brazilian learners of English. *Interspeech* 2007:2297–2300.

Lado, Robert. 1957. *Linguistics across cultures.* Ann Arbor: University of Michigan Press.

Sheldon, Amy, and Winifred Strange. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech perception can precede speech production. *Applied Psycholinguistics* 3:243–61.