

# Characterizing the influence of native language experience on adult speech perception

LINDA POLKA

*McGill University, Montreal, Quebec, Canada*

Previous cross-language research has indicated that some speech contrasts present greater perceptual difficulty for adult non-native listeners than others do. It has been hypothesized that phonemic, phonetic, and acoustic factors contribute to this variability. Two experiments were conducted to evaluate systematically the role of phonemic status and phonetic familiarity in the perception of non-native speech contrasts and to test predictions derived from a model proposed by Best, McRoberts, and Sithole (1988). Experiment 1 showed that perception of an unfamiliar phonetic contrast was not less difficult for subjects who had experience with an analogous phonemic distinction in their native language than for subjects without such analogous experience. These results suggest that substantive phonetic experience influences the perception of non-native contrasts, and thus should contribute to a conceptualization of native language-processing skills. In Experiment 2, English listeners' perception of two related nonphonemic place contrasts was not consistently different as had been expected on the basis of phonetic familiarity. A clear order effect in the perceptual data suggests that interactions between different perceptual assimilation patterns or acoustic properties of the two contrasts, or interactions involving both of these factors, underlie the perception of the two contrasts in this experiment. It was concluded that both phonetic familiarity and acoustic factors are potentially important to the explanation of variability in perception of nonphonemic contrasts. The explanation of how linguistic experience shapes speech perception will require characterizing the relative contribution of these factors, as well as other factors, including individual differences and variables that influence a listener's orientation to speech stimuli.

Our understanding of the role of language experience in the development of speech perception has been advanced by cross-language studies of speech perception. Two clear findings have emerged from these comparative investigations. First, in tasks with processing demands relevant to linguistic communication, the perception of non-native speech distinctions is typically not as accurate or efficient as is the perception of speech contrasts that convey meaning in the listener's native language (Cara-

mazza, Yeni-Komshian, Zurif, & Carbone, 1973; Flege & Hillenbrand, 1986; Goto, 1971; Lisker & Abramson, 1970; MacKain, Best, & Strange, 1981; Miyawaki et al., 1975; Pisoni, Aslin, Perey, & Hennessy, 1982, Werker & Tees, 1984b; for reviews, see also Pisoni, Logan, & Lively, in press; Strange & Jenkins, 1978; Werker, in press). Second, considerable variation in the perceptual difficulty of different non-native distinctions has been reported (Best, McRoberts, & Sithole, 1988; Burnham, 1986; Eilers, Gavin, & Wilson, 1979; Pisoni et al., in press; Strange, 1972; Werker, Gilbert, Humphrey, & Tees, 1981). In addition, the phonetic and phonotactic context in which the same non-native distinction occurs affects perceptual difficulty (Henley & Sheldon, 1986; Logan, Lively, & Pisoni, 1991; Sheldon & Strange, 1982). The research reported here was directed at explaining this variability in perceptual difficulty, with the goal of providing a clearer conceptualization of the facets of speech perception that undergo change during the course of language acquisition.

In the cross-language studies that have been done so far, the perception of a relatively small number of non-native voicing or place-of-articulation (hereafter place) distinctions has been examined. In most of this research, the perception of one or more speech contrasts has been compared between subjects whose native languages differ in such a way that the distinction is phonemic for one group but nonphonemic for the other group. Discrepan-

---

This research was part of a doctoral dissertation completed in the Department of Psychology at the University of South Florida. Special thanks are extended to the following people for their invaluable assistance: Farah Khorsandian, for facilitating contact with Farsi students; Abbas Rakshani, for producing the Farsi stimuli; the Farsi subjects who contributed to the preparation of the stimulus materials; Janet Werker, for providing the Salish stimuli; Desi Saludes and Doug Whalen for their technical assistance; and Winifred Strange, James J. Jenkins, Catherine Best, and Janet Werker for their comments, questions, and advice during the course of this research. Financial support was provided by NINCDS Grant NS22568 and NIDCD Grant DC00323 to Winifred Strange and James J. Jenkins, co-principal investigators. Preparation of the stimulus materials was also supported, in part, by NINCHD Grant N01-HD-5-2910 to Haskins Laboratories. This project would not have been possible without the equipment funded by the Office of Sponsored Research, the Department of Communication Sciences and Disorders, and the Department of Psychology at the University of South Florida. Correspondence should be addressed to Linda Polka, School of Human Communication Disorders, McGill University, 1266 Pine Avenue West, Montreal, Quebec, Canada H3G 1A8.

cies in the performance of native and non-native listeners are typically observed in such comparisons and thus provide evidence for the influence of native language phonemic structure on speech perception in adulthood (Abramson & Lisker, 1970; Best, MacKain, & Strange, 1982; Briere, 1966; Flege & Hillenbrand, 1986; Mann, 1986; Marckwardt, 1946; Miyawaki et al., 1975; Singh & Black, 1966; Strange & Jenkins, 1978). Research documenting difficulties in training adult listeners to perceive various nonphonemic distinctions also supports this view (Flege, 1989; Flege & Wang, 1989; Jamieson & Morosan, 1986; Logan et al., 1991; McClaskey, Pisoni, & Carrell, 1983; Miranda, Underbakke, Strange, & Micceri, 1989; Mochizuki, 1981; Pisoni et al., 1982; Strange & Dittmann, 1984).

Although phonemic experience undoubtedly influences the development of speech perception, it has also been noted that some nonphonemic contrasts present greater difficulty than others. This nonuniformity indicates that phonemic experience alone does not provide a complete account of the effects of specific language experience. In previous cross-language studies of speech perception in adults, it has been hypothesized that three factors underlie differences in perceptual difficulty among non-native speech contrasts (Polka, 1991): phonemic status, phonetic familiarity, and acoustic salience. *Phonemic status* refers to speech as consisting of functional categories—that is, classificatory units that differentiate lexical items in a perceiver's native language. *Phonetic familiarity* refers to the experience with the substantive aspects of linguistic structure, including language-specific regularities in production corresponding to both within- and between-phoneme category variants (i.e., allophonic rules of pronunciation). The third factor, *acoustic salience*, refers to the acoustic structure that underlies different speech distinctions. This factor concerns differences in the salience of acoustic information that may contribute to variability in the perception of non-native contrasts.<sup>1</sup>

The influence of phonetic familiarity is supported by the observation that allophonic experience with the contrasting phones may decrease perceptual difficulty. For example, in an attempt to explain the finding that non-native place distinctions may be harder than non-native voicing contrasts, researchers have noted that subjects are more likely to have had allophonic experience with the phones of the non-native voicing contrasts (e.g., Best et al., 1982; MacKain, 1982; Pisoni et al., in press; Tees & Werker, 1984). The role of acoustic salience in the perception of non-native speech distinctions is also supported by the increased discriminability of voicing over place distinctions. Specifically, it has been hypothesized that the perception of contrasts that are distinguished primarily by spectral acoustic differences (e.g., place contrasts) is more dependent on linguistic experience than is the perception of contrasts that are distinguished primarily by temporal acoustic dimensions (e.g., voicing contrasts; see Burnham, 1986; Edman, 1981). A detailed discussion of

specific findings that provide support for phonetic familiarity and acoustic salience factors in explaining varying difficulty among diverse nonphonemic contrasts can be found in Polka (1991).

The importance of both acoustic and phonetic factors has also been highlighted in work on the /r/-/l/ contrast. For example, several studies have shown that Japanese listeners' difficulty in perceiving the English /r/-/l/ contrast varies as a function of context (Logan et al., 1991; Mochizuki, 1981; Sheldon & Strange, 1982). Sheldon and Strange (1982) and Dissoway-Huff, Port, and Pisoni (1982) proposed that differences in the acoustic cues for /r/ and /l/ across various contexts could account for the perceptual differences. However, Henley and Sheldon (1986) found that native speakers of Cantonese, another language without a phonemic /r/-/l/ distinction, showed a different pattern of relative difficulty than did Japanese listeners in the perception of English /r/-/l/ in different phonetic contexts. They concluded that contextual variability could not be explained by acoustic factors alone; the relationship between English /r/ and /l/ and the phonetic and phonotactic structure of Cantonese and Japanese must also be considered.

Recently, Best and her colleagues (Best, in press; Best et al., 1988) have presented a model that provides an integrated view of how phonemic, phonetic, and acoustic factors may underlie variation in the perception of non-native speech contrasts. Best et al. maintain that attention is normally focused at the phonemic (functional) level during speech perception in such a way that listeners perceptually assimilate non-native phones to their native phonemic categories whenever possible. The pattern of assimilation is a function of phonetic similarity between native and non-native phones. Best et al. propose four ways in which non-native phones may be assimilated, and they make several predictions regarding the relative perceptual difficulty of the four assimilation patterns.

According to this model (Best, in press), the greatest perceptual difficulty is associated with *single-category* assimilation, in which both non-native phones are perceived as instances of the same native phoneme category. The least perceptual difficulty is associated with *two-category* assimilation, in which the contrasting non-native phones are assimilated to two different native phoneme categories. Two other assimilation patterns give rise to intermediate levels of perceptual difficulty, *category-goodness* assimilation and *nonassimilation*. In category-goodness assimilation, both phones are assimilated to the single native phonemic category, but the phones are differentiable because one phone is a better perceptual fit to the category than is the other phone. Finally, in nonassimilated contrasts, both phones are sufficiently dissimilar from any native category so as to be perceived as nonspeech sounds, and, as a result, the listener differentiates them by attending to psychoacoustic differences between them. Thus, according to this model, psychoacoustic factors are relevant only when non-native phones are not perceived as speech.

Examples of each assimilation type have been presented elsewhere (Best, in press; Best et al., 1988; Polka, 1991; Werker, 1991).

The existing cross-language data do not provide an adequate basis for evaluating the role of phonemic, phonetic, and acoustic factors in the perception of non-native contrasts for several reasons. First, in most cross-language studies, the perception of a single contrast has been examined, so that any relevant comparisons must be made across experiments that differ methodologically in numerous ways. Second, in only a few studies have perceptions of more than one non-native contrast by the same listeners been compared in similar procedures (e.g., Best et al., 1988; Werker, 1991). Likewise, in not many studies have perceptions of the same contrast been compared among multiple groups of non-native listeners whose linguistic experience relates to the particular contrast in different ways (e.g., Flege & Wang, 1989; Henley & Sheldon, 1986). Previous cross-language studies also have not fully tested the predictions provided by Best's assimilation model. In particular, perception of contrasts that fit different assimilation patterns has been compared among contrasts involving different classes of phonetic features and thus often differ markedly in acoustic features.

In the two experiments reported here, the same testing procedures were used to systematically evaluate the influence of phonemic status and of phonetic familiarity on the perception of non-native place contrasts. Predictions derived from Best's assimilation model were also tested. In Experiment 1, the perception of a single place contrast from the Salish language by one group of non-native listeners was compared with that of another. The groups differed in their phonemic experience but not in their phonetic familiarity with the non-native contrast.<sup>2</sup> This comparison should provide a way to assess the influence of phonemic status without allowing phonetic familiarity or acoustic salience to vary. In Experiment 2, the perception of one non-native place contrast was compared with that of another, within a single group of English listeners. Specifically, the perception of the Salish contrast from Experiment 1 and that of another variant of the same distinction from another language were compared. Because these two nonphonemic contrasts vary in phonetic familiarity to native English listeners, this comparison provides a test of the influence of phonetic familiarity on the perception of non-native contrasts. Together, the results of these two experiments shed light on the influence of phonetic and phonemic factors in cross-language speech perception and also permit further evaluation of the utility of Best's assimilation model in predicting relative difficulty of non-native contrasts.

In both experiments, stimulus materials were developed with the use of a multiple-exemplar approach in which natural productions of the contrasting consonants were carefully selected to avoid systematic variation along non-phonetically relevant dimensions. Subjects' perception of these materials was then evaluated in two tasks: closed-set identification (hereafter referred to as the ID test) and

categorical AX discrimination (hereafter referred to as the AX test).

In the ID test, subjects are provided a set of labels to use in responding to stimulus items that are presented individually. In the AX test, subjects are asked to indicate whether a pair of physically different stimuli are two instances of the same category or instances from two different categories. In contrast to a simple AX discrimination (in which X either is or is not physically identical to A), performance in this task involves some degree of perceptual constancy, because the listener is required to recognize relevant between-category differences and to ignore irrelevant within-category variation.

The decision to use both an ID test and an AX test was made for both methodological and theoretical reasons. The ID task requires that subjects sort stimuli into categories, and thus it should reflect the demands of natural speech processing. However, the interpretation of the identification data in cross-language studies can be problematic, because accurate identification requires both the recognition of relevant category differences and the appropriate assignment of labels to category members. When subjects label stimuli accurately, we can conclude that they are able to carry out both components of the ID task. Likewise, when subjects are consistently inaccurate in labeling stimuli in a two-category ID task, we can conclude that they can differentiate the categories, but have reversed the assignment of category labels. However, when identification performance is not highly consistent, it is not evident whether subjects have failed to recognize category differences, assigned inappropriate labels, or both.

AX tasks can also be structured to assess perception of phonetic categories, and they have been used previously in cross-language research (Best et al., 1988; Gottfried, Jenkins, & Strange, 1985). An AX procedure was chosen for the present study because it had been used in previous studies on the perception of non-native contrasts. In contrast with ID tasks, AX tasks do not require that subjects assign category labels. Thus, as in the present study, the interpretation of identification results is appropriately supplemented by data gathered in an AX task structured around the same phonetic categories. In both experiments, subjects completed the AX test followed by the ID test.

## EXPERIMENT 1

The relationship between the phonological structure of two languages can be described by comparing their respective phonetic inventories (i.e., the specific combinations of phonetic features that occur) or by comparing the phonemic (i.e., distinctive) features whereby they convey word meaning. In the latter approach, although a particular *combination* of phonetic features may not be found in a given language as an allophone of some phoneme class, the distinctive feature for a given place contrast may be phonemic in combination with other features—that is, in a different manner or voicing class. In such a situa-

tion, phonemic experience with the contrastive dimension would be provided, but it would not be phonetically specific. As mentioned earlier, in previous cross-language studies involving place contrasts, phonemic relevance and phonetic familiarity have often been confounded. The role of phonemically relevant, but phonetically nonspecific, experience has not been examined in cross-language research. This issue was examined in Experiment 1.

The specific question addressed was whether experience with a particular place distinction facilitates perception of the same phonemic place distinction in a different manner context, or alternatively, whether the effects of native language experience are specific to particular phonetic categories or combinations of phonetic features. This was addressed by comparing the perception of the Salish uvular versus velar glottalized stop contrast by native speakers of English with that of native speakers of Farsi, a language spoken in Iran. Although glottalized stops are not found on the phonetic surface of either English or Farsi, the uvular-velar distinction has a different phonemic status in the two languages. This feature contrast is not phonemic in English, which does not contain phones produced with a uvular place of articulation. In Farsi, the uvular-velar place distinction is phonemic for voiced stops. Voiceless uvular fricatives also occur in the Farsi language, although not in contrast with voiceless velar fricatives. Thus, neither Farsi nor English includes the specific combination of phonetic features that constitutes the Salish contrasting phones. Likewise, the salience of the acoustic differences between the contrasting Salish phones is presumably the same to the ears of both English and Farsi listeners. However, from a phonemic perspective, Farsi is more similar to Salish than English, in that Farsi contains phones produced with both the uvular and the velar places of articulation, and the uvular-velar contrast occurs distinctively in Farsi, in a different manner and voicing context.

Previous research on the Salish uvular versus velar glottalized stops has shown that English listeners find it difficult to perceive this contrast, particularly in tests with procedural requirements similar to those of everyday speech processing (Werker & Tees, 1984b). If perception of the Salish contrast is less difficult for Farsi listeners, this would suggest that phonemic experience exerts a facilitative influence on perception of place feature contrasts that is independent of the specific phonetic context. Alternatively, if perception of the Salish contrast is found to be equally difficult for Farsi listeners and English listeners, this would suggest that substantive phonetic experience plays a role in characterizing the effects of linguistic experience.

Werker (1991) has described English listeners' perception of the Salish contrast as conforming to either a single-category or a nonassimilated pattern. If Farsi listeners, for whom this place distinction is phonemic, are able to assimilate the Salish phones to analogous native place categories, their perception of the Salish contrast would be expected to conform to a two-category assimilation pat-

tern. According to Best (in press), this characterization predicts that Farsi listeners would perform better than English listeners, because both single-category assimilation and nonassimilation are expected to present greater perceptual difficulty than two-category assimilation.

Error patterns and subjects' descriptions of the contrast were also evaluated to corroborate the assimilation strategies proposed to underlie perception of the Salish contrast by Farsi and English listeners. Subjects' descriptions of the contrast were expected to portray (1) perceived similarities to analogous uvular and velar sounds from Farsi (two-category assimilation), (2) perceived similarity to a single native category or reports that syllables begin with the same sound (single-category assimilation), or (3) a failure to relate the Salish sounds to one or two native categories in a consistent fashion or attention to acoustic features of the syllables (non-assimilation).

It is difficult to distinguish these three types of assimilation on the basis of error patterns that reflect differential accuracy in perception of /q'/ versus /k'/ categories. However, subjects who perceived the Salish contrast as a nonassimilated contrast were considered more likely to show an improvement in performance during the course of either test than subjects who perceived the contrast as either a single-category or a two-category assimilation. This was expected because, in the absence of native-language constraints on processing, subjects were expected to attend to a wider range of acoustic differences and to be more likely to shift attention to different stimulus features during the course of the tasks in attempting to isolate systematic differences.

## Method

**Stimuli.** The Salish stimuli were the same as those used by Werker and Tees (1984a, 1984b). They consisted of six consonant-vowel (CV) syllables, each with a voiceless velar, /k'/, or uvular, /q'/, glottalized stop as the initial consonant followed by the same vowel, produced by a native speaker of Salish. Three instances of each CV syllable were selected by using a multiple natural exemplar approach; they are described in detail in Werker and Tees. (Acoustic differences between the contrasting syllables are discussed further in Experiment 2.) Recorded test sequences were presented free-field in an IAC sound-attenuated chamber. The delivery system included a Teac (A-3300SX-2T) tape recorder and a Yamaha amplifier, which was located outside the chamber and was connected to a single ADS loudspeaker located inside the chamber. The stimuli were presented at a level of 62-68 dBA; the ambient noise level in the chamber was 23 dBA.

**Subjects.** Twenty-eight native monolingual English speakers and 16 native speakers of Farsi participated in the experiment. All subjects in both groups reported normal hearing and were not trained in phonetics. The English subjects were recruited through the psychology subject pool and received extra credit. The Farsi subjects were recruited through letters sent out by the campus center for foreign students and were offered \$10 for participating. Only subjects who had acquired Farsi (Tehran dialect) as a first language and also spoke English were tested.

The English subjects were 4 males and 24 females, who ranged in age from 18 to 32 years ( $M = 22$  years). The Farsi subjects were 14 males and 2 females, who ranged from 17 to 52 years ( $M = 29.4$  years). The Farsi subjects had spent between 4 and 19 years in the United States. All Farsi subjects reported that they spoke

Farsi with their families or friends regularly. Two Farsi subjects had acquired both English and Farsi beginning in early childhood, and a 3rd subject was trilingual, having acquired Armenian and Farsi from early childhood and having learned English in adulthood. Armenian does not have a uvular-velar place distinction (Ruhlen, 1976). The three Farsi subjects who had acquired two or more languages during childhood will hereafter be referred to as *early bilinguals*.

**Procedure.** Two similar, but not identical, sets of procedures were employed in testing subjects in this experiment. First, 10 Farsi and 10 English subjects were tested with the procedures outlined below for Group 1. Several changes in the procedures were employed in testing a second group of subjects (6 Farsi subjects and 18 English subjects), described below for Group 2.

**Group 1.** At the beginning of the test session, all subjects were told that they would hear CV syllables produced in a foreign language that began with two different consonants. Both the English and the Farsi subjects were told that one of the consonants was similar to /k/. The English subjects were told that the other consonant could not easily be compared to a consonant in English, whereas the Farsi subjects were told that the second consonant might or might not sound like a consonant in Farsi. For the AX test, the subjects were then told that they would hear pairs of syllables that contained either two different instances of the same consonant (*same* pairs) or instances of two different consonants. They were asked to respond by writing an S or a D on a response sheet. Before responding in each test, the subjects listened to a brief familiarization (12 AX or ID trials) while following a list that indicated the syllables that were presented in each trial. The 12 familiarization trials for the two tests provided equal numbers of presentations of each consonant. After the familiarization was presented, the subject began the AX or the ID test, without feedback.<sup>3</sup> The subjects were presented a random sequence of 180 pairs in the AX test, which included equal numbers of *same* and *different* pairs and equal numbers of presentations of each of the six stimuli. The interstimulus interval (ISI) was 1,500 msec, and the intertrial interval (ITI) was 3,000 msec.

In the ID test, each of the six stimuli was presented 27 times in random order, with an ISI of 3,000 msec. The subjects were instructed to identify the stimulus as either beginning with a consonant similar to /k/ (by circling a K) or other (by circling an X). Before beginning the ID test, the subjects heard six familiarization trials. After completing both tests, the subjects described (in writing) what the stimuli sounded like to them.

**Group 2.** For Group 2, two changes in procedure were introduced. First, all subjects were oriented to the experiment in a more comparable fashion. Both the Farsi and the English subjects were provided a brief description of the place-of-articulation difference for the two consonants using a cross-sectional diagram of the vocal tract. For the Farsi subjects, it was pointed out that several consonants in Farsi are produced with the use of the same place of articulation as that for each of the Salish consonants. Examples of Farsi words that begin with uvular and velar consonants were provided. For the English subjects, it was pointed out that English consonants are produced with the use of the velar, but not the uvular, place of articulation. The second change in procedure was introduced to equate the AX and ID tests in terms of the amount of auditory exposure provided in both the familiarization and the test trials. The subjects therefore received 192 trials in both the ID and the AX tests, with each stimulus presented twice per trial in the ID test. In both tests, the ISI was 1,500 msec and the ITI was 3,000 msec.

## Results

Each subject's performance in the ID and in the AX test was converted to an  $A'$  score.  $A'$  is a nonparametric unbiased index of sensitivity which ranges from 0 to 1.0,

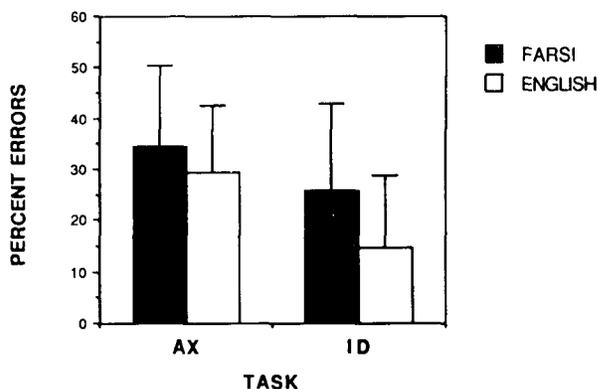


Figure 1. Overall errors on the Salish AX and ID tests for English and Farsi subjects.

with .50 corresponding to chance performance (McNichol, 1972). For each test, the  $A'$  scores for subjects in Groups 1 and 2 were compared in one-way analyses of variance (ANOVAs) conducted separately for English and Farsi subjects.<sup>4</sup> In analyses of both AX and ID performance, there were no significant differences between English Groups 1 and 2 or between Farsi Groups 1 and 2. Therefore, the subjects in Groups 1 and 2 were combined for further analysis.

Figure 1 presents the mean percent errors (expressed as percentage of opportunities) made by Farsi and English listeners on the AX and ID tests. The range of error scores corresponding to +1  $SD$  are also indicated. Mean  $A'$  scores for each test are presented in Table 1.<sup>5</sup>  $A'$  scores were analyzed in a mixed model native language (Farsi vs. English)  $\times$  task (ID vs. AX) ANOVA. The main effects of native language [ $F(1,42) = 4.977, p < .03$ ] and of task [ $F(1,42) = 28.17, p < .0001$ ] were significant. The results indicate that on both perceptual tests, the Farsi listeners made significantly more errors than the English listeners did. Also, subjects in both groups made more errors on the AX test than they did on the ID test.

$A'$  scores for the Farsi and the English subjects on each test were also compared with chance performance ( $A' = .50$ ) with single mean  $t$  tests. The obtained  $t$  values are also presented in Table 1. These analyses show that the performance of both the Farsi and the English listeners exceeded chance predictions on both tasks. Although the average performance on both tests in both groups was significantly better than that predicted by chance, the performance of most subjects fell below that of native listeners who identified the stimuli with 100% accuracy. A criterion  $A'$  score of .95 was defined as *native-like* perception. This score corresponds to an overall error rate of less than 10%. The number of subjects in each group who reached the native-like criterion on each test is also presented in Table 1. Results of a Fischer exact probability test indicated that the proportion of subjects who met criterion was not significantly different for the English and Farsi subjects on either the AX or the ID test. Thus, the perception of Farsi and English subjects was not differ-

**Table 1**  
**Mean *A'* Scores, Results of Single Mean *t* Tests,**  
**and Numbers of Subjects Who Reached**  
**Native-Like Criterion on ID and AX Tests in Experiment 1**

Test	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	No.
English Group					
AX	.772	.131	10.805	< .001	2/28
ID	.901	.022	18.09	< .001	14/28
Farsi Group					
AX	.702	.043	4.596	< .001	0/16
ID	.793	.160	7.09	< .001	4/16

ent when compared with that predicted by chance or with native-like performance for either task.

For each test, the distribution of *A'* scores for subjects in each group was compared with a theoretical distribution of *A'* scores corresponding to native-like perception, defined as obtaining an *A'* score of .95 or greater. These distributions were compared in the Kolmogorov-Smirnov one-sample test, which assesses whether the point of maximum divergence between a theoretical and an observed cumulative frequency distribution is greater than would be expected on the basis of chance.<sup>6</sup> For both Farsi and English subjects, the distribution of *A'* scores on each test deviated significantly from the distribution characterizing native-like perception. Therefore, neither Farsi nor English listeners as a group showed performance that was comparable to native perception.

The types of errors made in each task were also very similar for Farsi and English subjects who showed either no differences in perception of the two categories or fewer errors in response to the uvular category relative to the velar category. The only consistent difference between the error patterns of the Farsi and English listeners was that Farsi subjects failed to show improvement in the AX test, whereas English subjects showed significant improvement from the first to the second half of the AX test.

## Discussion

The results as a whole support three conclusions. First, the Farsi listeners, as a group, did not demonstrate a perceptual advantage by virtue of their experience with the place contrast in another context. This was the case even when instructions focused the subjects' attention on the specific place dimension that was distinctive in the contrast and how the place difference corresponded to phonemes in their native language. Second, neither the Farsi nor the English subjects, as a group, demonstrated the high level of accuracy that characterizes native perception. Third, substantial individual differences in performance on both tests were observed for both the English and the Farsi listeners. Less consistent support was obtained for a fourth conclusion—that perception of the Salish contrast was more difficult for the Farsi than for the English listeners.

Given the individual differences observed in both the Farsi and the English groups, the subjects' descriptions

at the end of the test session were of considerable interest. Three of the 5 Farsi subjects who met the native-like criterion reported that the consonants were similar to velar and uvular sounds in their language. Two of these subjects were early bilinguals—that is, they had acquired two languages from early childhood. The reports and performance of these subjects are consistent with the two-category assimilation pattern described by Best et al. (1988). These Farsi subjects appeared to have assimilated the Salish contrast to velar and uvular consonant categories in Farsi.

One of the other 2 Farsi listeners who attained native-like performance failed to provide a description of the contrasting phones. The other Farsi subject and the English subjects who met the native-like criterion tended to report one of three things: (1) that the one consonant was like a *t* or *d* and that the other sounded like a combination of several consonants, including *tr*, *tw*, *trv*, *ktr*, and *tgr*, (2) that they discriminated acoustic differences between the categories (e.g., *K* was shorter than *X*), or (3) that they listened to the vowels to differentiate the syllables. These subjects did not appear to assimilate either sound to a single native phonemic category, but rather seemed to be responding on the basis of phonetic or acoustic details that they heard.<sup>7</sup> The reports and performance of these subjects were most consistent with a nonassimilation pattern.<sup>8</sup> Error patterns for both the Farsi and the English subjects failed to suggest alternative assimilation patterns, showing either little difference in cohesiveness of the two categories or greater cohesiveness in the uvular category, which has phonemic status in Farsi but not in English.

The Farsi and English subjects who did not reach the native-like criterion of performance tended to report that the two consonants sounded very similar or were both very unclear. The reports and performance of these subjects were most consistent with a single-category assimilation pattern.

Given these patterns of individual performance, the conclusions outlined above merit the following elaboration. At least for some Farsi listeners, evidence was obtained that nonspecific linguistic experience had a facilitative influence. Two of the 3 Farsi subjects who provided support for this conclusion (i.e., Farsi subjects who met the native-like criterion and related the Salish contrast to Farsi

consonants) were early bilinguals. Thus, evidence for effects of nonspecific linguistic experience may be readily observed in multilingual subjects who have acquired at least one language with the uvular-velar distinction. The performance and descriptions of the remaining subjects provide further evidence that the Salish contrast can support several patterns of assimilation in individual listeners, as Werker (1991) has reported. The single-category assimilation pattern leads to a lower level of performance than does the nonassimilated pattern. Although both assimilation patterns were evident among Farsi and English listeners, error patterns and subjects' descriptions suggested that more Farsi subjects than English subjects perceived the Salish contrast as a single category assimilation. The superior overall performance of English subjects and the finding that only the English subjects showed improvement in the AX test are also consistent with this interpretation.

The better performance of the English subjects could also be interpreted as evidence that experience with Farsi interfered, rather than facilitated, perceptual categorization of the Salish contrast. However, evidence supporting interference effects was not available in the reports provided by the Farsi subjects. Every Farsi subject who reported that the Salish consonants sounded similar to Farsi consonants also achieved native-like accuracy in their perceptual performance. In addition, the Farsi subjects who received the more informative instructions explicitly stated that the Salish syllables did not sound similar to Farsi.

In conclusion, the group data suggest that the highly accurate and effortless perception characteristic of native listeners requires experience that is specific to particular combinations of phonetic features (i.e., particular phonetic categories). Native language experience with the Farsi uvular-velar place distinction in voiced stops did not, in general, facilitate perception of the Salish uvular-velar place distinction in voiceless glottalized stops. However, such facilitation was observed in individual Farsi listeners who were early bilinguals. Thus, at least for the Salish /k'/-/q'/ contrast, perception of a non-native contrast appears to be influenced by the phonetic substance of the stimuli as well as individual difference variables. The performance of the early bilinguals in this experiment suggests that both broad and specific language experience may contribute to variability in perception of non-native contrasts by individual listeners.

## EXPERIMENT 2

The results of Experiment 1 provide at least tentative support for the idea that the experience with specific phonetic categories influences perception of non-native place contrasts. The role of substantive phonetic experience in the perception of non-native place contrasts was evaluated further in Experiment 2. In this experiment, the perception of the Salish contrast was compared with the perception of another variant of the same phonemic contrast

from Farsi. In Farsi, the uvular-velar place distinction is minimal only for voiced stops [G] versus [g]. In English, consonants are not produced with a uvular place of articulation and the uvular-velar distinction is not phonemic in any phonetic context. Although the Salish and the Farsi contrasts convey the same non-native place distinction, the two contrasts differ in phonetic familiarity to monolingual English listeners. Specifically, the Farsi contrast pairs a familiar phone, [g], with an unfamiliar phonetic category, [G], whereas both Salish phones, [k'] and [q'], are phonetically unfamiliar to English listeners. Alternatively, the relationship of the two contrasts to English can be described in terms of combinations of phonetic features. For English listeners, the Farsi contrast involves a non-native place distinction within a familiar manner and voicing context (i.e., voiced stops), whereas the Salish contrast involves a non-native place difference within a non-native manner class (i.e., voiceless glottalized stops).

If phonetic familiarity facilitates the perception of non-native contrasts, the perception of the Farsi contrast should be easier for English listeners than the perception of the Salish contrast. Experiment 2 was designed to evaluate this prediction by comparing English listeners' performance on the Farsi and Salish contrasts in the AX and ID tests. Perception of the two contrasts was examined within the same set of English listeners in order also to observe the extent to which individual differences are stable across the perception of different non-native contrasts.

Several predictions provided by Best's assimilation model were also tested. According to this model, the Farsi /g/ versus /G/ contrast should conform to the category-goodness assimilation pattern and the Salish contrast should be assimilated as either a single-category or a nonassimilated contrast. The category-goodness pattern was predicted to be less difficult than the single-category pattern, and equal in difficulty to the nonassimilated pattern. Thus, performance on the Farsi contrast was expected to be equal to or better than perception of the Salish contrast, depending on which pattern would be observed for the Salish contrast. However, perception of the Salish contrast was not expected to be easier than perception of the Farsi contrast.

Because the assimilation patterns characterizing the two contrasts do not predict a definitive outcome regarding overall performance, further evidence that different assimilation patterns underlie perception of these two contrasts was also examined. Specifically, if English listeners' perception of the Farsi contrast fits the category-goodness pattern of assimilation, they would be expected to make fewer errors in response to the velar /g/ category than the uvular /G/ category, and their descriptions of the Farsi contrast would be expected to portray a perceived difference in fit to their native /g/ category. For the Salish contrast, no systematic difference in errors was expected in response to the two Salish phones, and, as in Experiment 1, subjects' descriptions were expected to portray either assimilation to a single category (i.e., reports that

both phones are perceived as the same speech sound) or a nonassimilated pattern (e.g., reports that reflect failure to relate to one or two native phonemic categories or attention to acoustic dimensions of the syllables).

## Method

**Stimuli.** In the Tehran dialect of Farsi, voiced velar stops contrast phonemically with both voiceless aspirated stops, /g/ versus /k<sup>h</sup>/, and with voiced uvular stops, /g/ versus /G/ (Lambton, 1961; Ruhlen, 1976). Although voicing is not contrastive among uvular stops in Farsi, Lambton (1961) reports that uvular stops "tend to be voiceless" except when they occur between back vowels, when they "tend to be voiced." Therefore, to avoid redundant voicing cues to the /g/-/G/ contrast, the perception of uvular and velar voiced stops produced in a back vowel context (i.e., in an /a/-C-/a/ context) was examined in Experiment 2.

About 100 instances of velar and uvular stops produced by an adult male native speaker of Farsi in the VCV context were recorded in a sound-attenuated chamber. Four native speakers of Farsi identified the recorded items; any unclear items were eliminated. Perceptual judgments were provided by several phonetically trained American English listeners, and preliminary acoustic analyses were conducted to evaluate differences in nonphonetically relevant dimensions (loudness, *F0*, rate, etc.). These measures were used to select four "clear-case" tokens of each consonant category that did not vary systematically on nonphonetic dimensions across the uvular and velar tokens. The eight tokens were converted into digital waveforms (20-kHz sampling, 12-bit resolution, 10-kHz lowpass filtering) on the Haskins PCM system. To eliminate several potential voicing cues that might have been confounded with place cues, the first vowel in each selected VCV syllable was deleted by editing the digitized wave form of each stimulus. This was accomplished without the introduction of distortion because a period of silence was observed following the first vowel in each stimulus.<sup>9</sup> Table 2 summarizes the results of the preliminary acoustic analyses by presenting the acoustic measures that did not vary systematically between the contrasting phones. Although vowel quality was perceived by trained listeners to be similar for the uvular and velar syllables, formant measures of the vocalic portions were difficult to equate across the two categories.<sup>10</sup> At the center of the vowel, the *F2* frequency tended to be higher and the *F3* frequency tended to be lower for /g/ items than for /G/ items. (See Table 5 and the Acoustic Analysis section for details and further discussion of acoustic parameters that differentiate /g/ and /G/.)

Test sequences for the discrimination and identification tests were created and recorded on audio tape. The test sequences were constructed to be comparable to the ID and AX tests presented in Experiment 1 (Group 2), allowing for the difference in the number of tokens in each category (i.e., four per category for Farsi, three per category for Salish). Both tests consisted of 192 trials. As in Experiment 1, each stimulus was presented twice per trial in the

ID test in order to equate the ID and AX tests for overall auditory exposure to the stimuli. In both tests, each of the eight stimuli was presented 48 times, the ISI was 1,500 msec, and the ITI was 3,000 msec. The ID test included 24 identification trials for each of the eight stimulus tokens. Trials were randomly presented in blocks of 12, with equal numbers of velar and uvular items per block. The test sequence for the AX test consisted of 96 *same* pairs and 96 *different* pairs. There were 48 uvular-*same* pairs, 48 velar-*same* pairs, 48 *different* pairs with the uvular category first, and 48 *different* pairs with the velar category first. Trials were randomized in blocks of 16, with equal numbers of *same* and *different* trials per block. The recorded test sequences were presented in a sound-attenuated chamber, using the delivery system described in Experiment 1. The stimuli were presented at a level of 62-68 dBA.

To determine whether the stimulus items had been distorted or mislabeled during stimulus processing, the tests were presented to two native Farsi listeners who had not previously heard the stimuli. Both native listeners demonstrated errorless performance on both tasks. One native listener also rated the intelligibility of each item (as "very clear," "clear," or "?"). None of the stimuli received a "?" rating from either subject.

**Subjects.** Two groups of 12 monolingual, native speakers of American or Canadian English served as subjects. All subjects reported having normal hearing and had received no phonetics training. Each group included an equal number of males and females; their age range was 16-29 years (*M* = 22 years). One group (S/F) was tested on the Salish contrast first and the Farsi contrast second; the other group (F/S) was tested on the contrasts in the reverse order.

**Procedure.** The subjects in both groups were tested on both Farsi and Salish contrasts, with the AX test followed by the ID test. In order to minimize carryover effects, testing on each contrast was conducted on consecutive days. At the end of each session, the subjects described what the consonants sounded like, and at the end of the second session, the subjects were asked whether the two contrasts differed in difficulty.

At the beginning of the first session, the subjects were told that they would hear syllables that began with two different consonants. A brief description of the difference in place of articulation for the two contrasts was provided using a cross-sectional drawing of the vocal tract. Familiarization and instructions preceding the Salish tests were the same as those described for the English listeners in Experiment 1 (Group 2).

Similar procedures were followed with the Farsi tests. Before the Farsi contrast was presented, the subjects were told that one of the consonants was used in English and that one was not used in English but was used in other languages. For the Farsi ID test, the subjects were instructed to respond by circling "G" for the English consonant or "X" for the foreign consonant on a response sheet. As with the Salish contrast, before responding in each test, the subjects listened to a brief familiarization (12 AX pairs or ID trials) while following a list of the syllables that were presented

Table 2  
Preliminary Acoustic Analysis of Farsi /g/ and /G/ Stimuli

	Duration (msec)				Vowel Intensity (dB)		Pitch H <sub>2</sub>			Formant Frequencies at Center	
	Total	Burst	VOT	Vocalic	<i>M</i>	Peak	Average <i>F0</i>	Starting <i>F0</i>	Ending <i>F0</i>	<i>F1</i>	<i>F4</i>
<i>/g/</i> Items											
Mean	338	7	16	314	43.3	60.0	98	105	91	600	3400
Range	328-353	6-9	7-21	303-325	41.7-45.3	57.6-63.7	95-99	101-107	86-96	600	3200-3500
<i>/G/</i> Items											
Mean	338	8	18	300	45.5	63.1	98	102	95	600	3338
Range	328-353	6-9	14-21	282-327	42.5-47.5	60.6-64.7	97-99	100-104	89-100	600	3200-3550

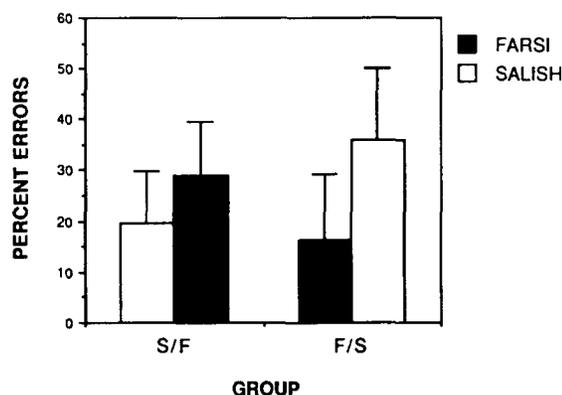


Figure 2. Overall errors on Farsi and Salish AX tests for English subjects (top of page).

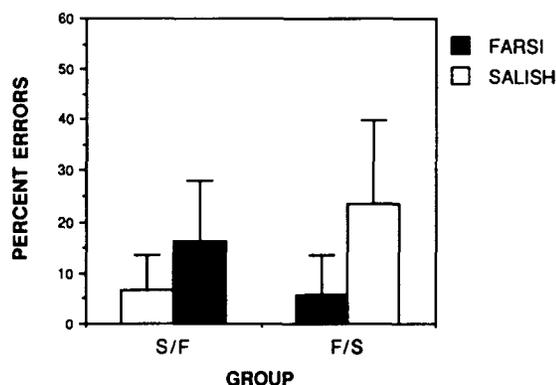


Figure 3. Overall errors on Farsi and Salish ID tests for English subjects (bottom of page).

in each trial. The 12 familiarization trials for the two Farsi tests provided equal numbers of presentations of each consonant. As in Experiment 1, the subjects followed a premarked response sheet while listening to the familiarization trials, but no feedback was provided during the AX or ID test.

## Results

**Overall performance.** Mean percent errors on each contrast are presented in Figure 2 for the AX test and in Figure 3 for the ID test. Performance is plotted separately for groups receiving the two contrasts in different orders, because large order effects were evident.<sup>11</sup> For each group, mean  $A'$  scores for each contrast are presented in

Tables 3 and 4 for the AX and ID tests, respectively.  $A'$  scores for the two contrasts were compared in a mixed model group (S/F vs. F/S)  $\times$  contrast (Farsi vs. Salish) ANOVA. This analysis was conducted separately for each task. The group  $\times$  contrast interaction was significant for both the AX test [ $F(1,22) = 21.758, p < .00001$ ] and the ID test [ $F(1,22) = 12.522, p < .0018$ ].

Next, for each test, the simple effects of group and contrast were evaluated. Analysis of the simple effect of group revealed that, within the S/F group, performance on the two contrasts was not significantly different for either test. However, within the F/S group, subjects did significantly worse on the Salish contrast than on the Farsi contrast on both the ID test [ $F(1,22) = 11.710, p < .01$ ] and the AX test [ $F(1,22) = 20.682, p < .01$ ]. Thus, the predicted difference between the two contrasts was significant in the F/S group but not in the S/F group.

Analysis of the simple effect of contrast (between the two groups) revealed a significant difference in performance (Salish poorer than Farsi) between the two groups on the second, but not on the first, contrast that they were presented. The difference between subjects on the second contrast was evident in both the ID test [ $F(1,44) = 7.944, p < .01$ ] and the AX test [ $F(1,42) = 14.323, p < .01$ ]. Thus, the pattern predicted on the basis of phonetic familiarity was observed between groups only for the second contrast that was presented.

$A'$  scores for each group on each contrast were also compared to chance performance ( $A' = .50$ ) by using single mean  $t$  tests; the results are shown in Tables 3 and 4. As in Experiment 1, both groups made significantly fewer errors than predicted by chance on both contrasts and both tasks. The number of subjects who reached the native-like criterion on each contrast is also presented in Tables 3 and 4. Simple effects of group and contrast were also evaluated for each perceptual test, with the number of subjects reaching criterion as the dependent variable. These analyses showed some, but not all, of the effects evident in the  $A'$  data. Results of a Cochran Q test showed that, in the F/S group, significantly more subjects reached criterion on the Farsi ID test than on the Salish ID test ( $p < .01$ ). Thus, as in the analysis of  $A'$  scores, the predicted difference in ID performance between the two contrasts was significant in the F/S group, but not in the S/F group. The contrast effect was not significant for the AX test. None of the between-group comparisons were significant.

Table 3  
Mean  $A'$  Scores, Results of Single Mean  $t$  Tests, and Numbers of Subjects Who Reached Native-Like Criterion on AX Tests in Experiment 2

Contrast	$M$	$SD$	$t$	$p$	No.
Group F/S					
Farsi	.888	.096	13.380	<.0001	5/12
Salish	.690	.155	4.060	<.0019	2/12
Group S/F					
Farsi	.784	.111	8.451	<.0001	0/12
Salish	.873	.078	15.861	<.0001	3/12

**Table 4**  
**Mean A' Scores, Results of Single Mean *t* Tests, and Numbers of Subjects**  
**Who Reached Native-Like Criterion on ID Tests in Experiment 2**

Contrast	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	No.
Group F/S					
Farsi	.966	.048	32.096	< .0001	9/12
Salish	.817	.167	6.309	< .0001	2/12
Group S/F					
Farsi	.895	.102	12.838	< .0001	4/12
Salish	.963	.040	37.913	< .0001	8/12

For each contrast, the distribution of *A'* scores for each test was compared with the distribution hypothesized to characterize native-like perception. This comparison was made separately for each group, using the Kolmogorov-Smirnov test. For both groups, the distribution of *A'* scores deviated significantly ( $p < .01$ ) from the native-like distribution on the AX test for both contrasts and on the ID test for the second contrast that subjects heard. For the contrast that subjects heard first, both groups failed to deviate from a native-like distribution on the ID test. Thus, both groups approximated the performance levels of native listeners in identifying the first, but not the second, contrast that they were presented. Relative to native-like accuracy, ID performance varied, depending on the order in which contrasts were presented.

For both tasks, Spearman rank order correlations between subjects' *A'* scores for the two contrasts failed to reach statistical significance. Thus, an insignificant portion of the variance in perceptual performance with each contrast was accounted for by individuals. However, the large order effect may have obscured the relationship between performance on the two contrasts.

In agreement with the order effects observed in overall performance, subjects often reported experiencing greater difficulty with the second contrast that was presented to them. At the end of both sessions, 8 subjects in the S/F group reported that the Farsi contrast was more difficult, but only 1 subject in the F/S group made this report. Likewise, 8 subjects in the F/S group reported that the Salish contrast was more difficult, whereas only 3 subjects in the S/F group made this report.

**Farsi contrast: Error patterns and subjects' descriptions.** Because of the large order effects found in overall performance, analyses of error patterns were conducted separately for the S/F and F/S groups. For the AX test, errors in response to *same* pairs for each category were analyzed in a place (/g/-*same* vs. /G/-*same*)  $\times$  half test ANOVA. This analysis revealed a significant main effect of place in both the F/S group [ $F(1,11) = 8.097, p < .02$ ] and the S/F group [ $F(1,11) = 6.731, p < .02$ ]. These results show that, in both groups, subjects were more accurate on /g/-*same* pairs than on /G/-*same* pairs. For the ID test, none of the factors in a category (/G/ vs. /g/)  $\times$  half test ANOVA reached significance in either group.

Almost all of the subjects in both groups reported that they heard "g." A number of subjects in both groups re-

ported that the uvular syllable sounded like an "ah" vowel without an initial consonant. One subject in the F/S group reported that the "g" sounded "clearer." This subject reached the native-like criterion on the ID test. Two of the subjects in the S/F group made similar comments. For example, one subject said that "the 'g' was more sharp and defined, easier to spot," and another subject said, "I listened for 'g' to distinguish the sounds." Both of these subjects reached the native-like criterion on the Farsi ID test. These comments support the notion that subjects in both groups perceived the Farsi contrast as a difference in goodness of fit to the native /g/ category.

**Salish contrast: Error patterns and subjects' descriptions.** Again, because of large order effects, error patterns were analyzed separately for each group. For the AX test, responses to *same* pairs failed to show any significant effects in either group when subjected to a place (/k'/-*same* vs. /q'/-*same*)  $\times$  half test ANOVA. Likewise, for the ID test, none of the factors in a category (/k'/ vs. /q'/)  $\times$  half test ANOVA were significant in either group. Thus, there was no evidence of differences in the perceptual cohesiveness of the two categories in either group.

The subjects in both groups reported a wide variety of descriptions of the Salish syllables. The subjects in both groups reported hearing "k" and "t" or a consonant blend with "t" (such as "tr," "ktr," "tw"), usually indicating that the latter description corresponded to the uvular consonant. One of the subjects who heard Salish first said, "I listened to the vowels" and "the X was longer than the K." This subject reached the native-like criterion on both tasks. Two other subjects in the S/F group provided similar descriptions. For example, one subject said that "the vowels are different," and another subject said that "X sounded longer than K." Both of these subjects reached criterion on the ID test. One of the subjects who heard Salish second also commented that the Salish "was easier to distinguish about the vowels." This subject also reached the native-like criterion.

#### Acoustic Analysis

Because acoustic and phonetic information co-occur in the speech stream, the perceptual differences observed may be attributed, at least in part, to differences in acoustic distinctiveness of the contrasting syllables differences. Although this study was not designed to systematically examine acoustic and phonetic factors, acoustic analyses of each contrast were examined in order to characterize the

Table 5  
Analysis of Burst and Formant Transition Cues for the Farsi /g/~/G/ Contrast

	DIFFERENTIATING ACOUSTIC PARAMETERS						
	F2 Formant Frequencies (Hz)			Burst Frequency (Hz)			
	Onset	At Center	Transition	1st Peak	2nd Peak	3rd Peak	4th Peak
/g/ Items							
Mean	1188	1122	falling	1200	4188		
Range	1100-1250	1100-1150		1150-1250	4150-4200		
/G/ Items							
Mean	975	962	flat	988	3700	5375	6325
Range	900-1000	900-1000		950-1000	3600-3800	4600-5900	6200-6700
	CORRELATED ACOUSTIC PARAMETERS						
	Burst Amplitude (dB)	F1 Formant Frequencies (Hz)		F3 Formant Frequencies (Hz)			
	M	Peak	Onset	Transition	Onset	At Center	Transition
/g/ Items							
Mean	34.3	44.6	350	slightly rising	2262	2475	slightly rising
Range	31.7-36.8	41.2-49.1	200-500		2200-2400	2400-2550	
/G/ Items							
Mean	32.1	44.1	525	flat	2550	2638	flat
Range	18.8-34.3	41.7-46.5	500-600		2300-2700	2550-2700	

role that acoustic factors may potentially play in explaining the perceptual results.

**Farsi contrast.** Both the release burst and *F2* and *F3* formant transitions have been shown to be important in differentiating velar stops from alveolar and bilabial stops in English. Acoustically, English velar stops are characterized as having a strong, compact burst in the midfrequency range, a falling *F2* transition, and either a flat or a rising *F3* transition. A characteristic feature of English velar stops in syllable-initial contexts is that *F2* and *F3* tend to be close together at onset and then diverge with transition into the vowel (Lieberman & Blumstein, 1988; Pickett, 1980). Acoustic or perceptual studies of Farsi velar and uvular stops have not been reported.

Spectrographs were made of each Farsi CV syllable, with a Kay digital sonograph (Model 7800). For each syllable, spectral sections were taken at successive 10-msec intervals from the onset of voicing until the steady state portion of the vowel was reached. From the spectral sections, formant frequency measures were obtained; these corresponded to the onset and asymptote of initial formant transitions. Frequency characteristics of the bursts were examined in a fast Fourier transform analysis with ILS software on a Vax Station II/GPX computer. Prominent peaks in the burst spectra were determined from the transform display for a window centered on the release burst. Duration and amplitude of release bursts were also measured from the digital wave forms of the syllables, using ILS software.

A summary of the acoustic analysis is presented in Table 5. Each measure was classified as either a differentiating or a correlated acoustic parameter. Differentiating parameters were defined as acoustic measures that did not overlap across /g/ and /G/ categories. Correlated acoustic parameters differed in central tendency but also showed some overlap across categories so that, as a sin-

gle cue, the parameter could not reliably support category differentiation.

As shown in Table 5, both *F2* onset frequency and *F2* frequency in the center of the vowel differentiated the Farsi syllables, with higher *F2* frequencies for the velar stops than for the uvular stops. *F2* formant transitions were observed to be flatter for the /G/ items than for the /g/ items, which contained slightly falling *F2* patterns. The peak frequencies of the release burst also differentiated uvular and velar stops. Two prominent spectral peaks in the burst were observed for the velar items, and three to four prominent peaks were observed for the uvular items. For both the uvular and velar items, the first peak in the burst spectrum was located near the onset frequency of *F2*, and the second peak was located above the onset frequency of *F4*. The first and the second peaks in the burst spectra occurred at lower frequencies for the /G/ items than for the /g/ items. Two additional peaks were prominent in the uvular burst spectra at frequencies above *F4*. In general, the uvular burst energy was distributed over a broader frequency range, extending into higher frequencies than the velar burst spectra.

Several acoustic dimensions also appeared to be correlated with place categories. Burst amplitude tended to be higher for the velar stops. The onset frequency of *F1* tended to be lower for the /g/ items than for the /G/ items. However, the frequency of *F1* in the center of the vocalic portion was the same for both categories. *F3* frequency at onset and at the center of the vocalic portion tended to be higher for the uvular stops than for the velar stops. In general, both *F1* and *F3* tended to be flat for /G/ items and gradually rising for the /g/ items.

**Salish contrast.** Detailed acoustic analyses of the Salish stimuli can be found in Werker and Tees (1984a, 1984b). Werker and Tees (and also Mayes, 1979) report that it was difficult to isolate acoustic parameters that dif-

differentiated these consonants. Syllables beginning with ejective stops have a high-amplitude ejective portion followed by a period of near silence before the onset of voicing. Werker and Tees (1984a) noted two differences between the uvular and velar syllables. First, amplitude and duration of the ejective portion were greater for the uvular consonant than for the velar consonant. However, the frequency range and location of spectral peaks for the ejective portions of the syllables were very similar for the uvular and velar stops. Second, in the vocalic portion, the third formant ( $F3$ ) was flat (around 2300 Hz) for the velar consonant, whereas  $F3$  rose slightly at the end of the vowel (from 2400 to 2900 Hz) for the uvular consonant. Thus, as in the Farsi contrast,  $F3$  in the vocalic portion of the syllables was higher in frequency for the uvular syllables than for the velar syllables.

Several other aspects of these stimuli are relevant in the present study. Because of difficulties encountered in obtaining the Salish sample, many CV tokens were recorded and analyzed, and only one pair of CV syllables was observed to have a similar vowel quality. Therefore, as reported in Werker and Tees (1984b), for each CV type, a single vocalic segment was spliced onto two other burst segments to obtain a set of three exemplars per consonant. This was possible because a period of silence preceded the onset of the vocalic portion of each syllable. As a result, within the set of three exemplars of each CV type, the vocalic segments were identical and were originally produced in a syllable beginning with the same consonant. After these modifications were made, the stimuli were presented (free-field) to native speakers of Salish and were readily identified as the appropriate consonant. Although these modifications did not change the intelligibility of the items for native listeners, the overall acoustic variation within each consonant category was greatly reduced.

### Discussion

English listeners' perception of the uvular-velar contrast in Salish was compared with their perception of the uvular-velar contrast in Farsi in order to examine the role of phonetic familiarity in perception of the same nonphonemic place distinction. If phonemic status takes precedence in accounting for experiential effects in speech perception, perception of the two contrasts should not differ. However, if phonetic familiarity is relevant in explaining experiential effects on perception of non-native contrasts, perception of the Farsi contrast should be easier than perception of the Salish contrast.

Overall, the perceptual results did not provide unalloyed support for the effects of phonetic familiarity. The data pattern predicted in support of this hypothesis was evident only among subjects who encountered the two contrasts in a particular sequence. Thus, for monolingual English listeners, whether or not the Farsi contrast was less perceptually difficult than the Salish contrast did not depend simply on phonetic familiarity. The acoustic dimensions that differentiate each contrast were not altered when the order of the contrast was varied. Therefore, the dif-

ferences in perceptual difficulty between these two contrasts also cannot be explained simply in terms of differences in the acoustic properties of the stimuli.

The results that emerged from statistical analysis of overall performance were consistent with predictions based on the particular assimilation patterns ascribed to the two contrasts. That is, the assimilation model predicted that perception of the Farsi contrast would be either as difficult as or less difficult than the perception of the Salish contrast, as the statistical analyses show. However, there was a nonsignificant trend in the opposite direction in the S/F group for both perceptual tests, which is clearly inconsistent with assimilation predictions. Furthermore, the assimilation model did not predict a priori the order effects that were clearly evident in both perceptual tests.

Although the analysis of error patterns and subjects' descriptions of each contrast cannot, by themselves, be taken as adequate support for the assimilation model, these data were consistent with different assimilation patterns for the two contrasts. The English listeners showed greater cohesiveness for the /g/ category, and their descriptions of the Farsi contrast support the idea that they perceived a difference in fit of Farsi /g/ and /G/ to their native /g/ category. Thus, perception of the Farsi contrast was consistent with a category-goodness difference assimilation pattern.

Error patterns in the perception of the Salish contrast did not show a consistent difference in perceptual cohesiveness of the Salish consonants. Subjects' descriptions of the Salish contrast in both experiments were consistent with either a single-category pattern or a nonassimilated pattern. As was observed in Experiment 1, the subjects' descriptions of the Salish syllables did not suggest that they related the Salish consonants to native phoneme categories. Instead, when successful, the subjects reported relying on vowel differences or particular acoustic details to differentiate the two syllables; when not successful, they typically reported that both sounded alike or unclear.

Several additional conclusions can be drawn from this experiment. First, although test sessions were conducted 24 h apart, both the perceptual results and the subjects' reports indicate that perception of the first contrast, in some way, disrupted perceptual performance on the second contrast. Furthermore, the asymmetry of the order effect reveals that the disruption was greater when subjects were tested on the Salish contrast after the Farsi contrast.

Second, because the order effect was asymmetric, it is unlikely that the observed disruption can be fully explained by the particular instructions that the subjects received. However, it is possible that the instructions encouraged them to listen in similar ways or to search for similarities in the two contrasts by pointing out similarities between the two contrasts.

Third, although most of the subjects performed better on the first contrast that they encountered, performance on the two contrasts was not significantly correlated in either group for either perceptual test. That is, in both groups, individuals' performance on the second contrast

was not related to their performance on the first contrast. Thus, the asymmetric order effect is not easily explained by individual differences.

Fourth, the acoustic properties that distinguish the Farsi and Salish contrasts differed in both kind and number. For the Farsi contrast, acoustic analyses revealed a number of systematic differences between the contrasted syllables, including parameters that are known to be important cues specifying stop place of articulation in English. In contrast, few systematic acoustic differences were evident between the Salish consonants, especially among parameters that are considered primary cues for place of articulation (i.e., differential burst spectra and initial formant transitions).

It is difficult to compare the relative salience of the acoustic dimensions for the two contrasts because the specific cues that differentiated each contrast do not correspond in a straightforward way. However, the two contrasts clearly differed with respect to variation within each contrasting category. The Salish contrast had a very narrow range of within-category acoustic variation relative to the Farsi contrast. This was the case because the vocalic portion of the Salish stimuli did not differ within the uvular or within the velar category and because there were fewer Salish stimuli than Farsi stimuli. Because there was no variation in the vowel within each Salish category, any between-category difference in the vowels, even a very subtle one, could serve as a reliable cue for differentiating the categories.<sup>12</sup> While systematic vowel differences were evident for the Farsi contrast, unlike the Salish contrast, these differences had to be extracted from the variation that existed within each place category if they were to support category differentiation.

Finally, the order effects indicate that the processes underlying perception of these two non-native contrasts are quite malleable, at least under some test conditions. However, it is not clear to what extent the malleability of perception observed in the present study is attributable to the perceptual instability of the Salish contrast or is a more general characteristic of the perception of non-native speech contrasts.

Overall, further research will be required to describe completely the nature of the disruption evident in the order effects. However, because both acoustic and phonetic properties of the two contrasts differed, three possibilities can be raised. First, the order effect may reflect an interaction between the assimilation processes (reflecting the interaction of phonetic and phonemic factors) evoked by the first contrast and the perceptual properties of the second contrast. The observation that perception of the Salish contrast can conform to two different assimilation patterns makes such an account reasonable. Second, the asymmetric order effect may reflect an interaction between perceptual processes evoked by acoustic properties of the first contrast and acoustic properties of the second contrast. The differences in number and kind of acoustic cues that distinguish the two contrasts suggest the plausibility of this sort of account. Though presented as alternative accounts, these two possibilities need not

be mutually exclusive. Hence, the third possibility exists that both acoustic and phonetic properties of the stimuli contribute to the perceptual processes evoked by these contrasts in ways that must be understood to explain the order effects. Recall that the assimilation model proposes that psychoacoustic properties play an important role only in the case of nonassimilated contrasts. Because perception of the contrasts presented in this study conforms to patterns of assimilation (i.e., single category or category goodness) as well nonassimilation, both phonetic and acoustic factors may well contribute to the observed order effects.

In summary, the lack of consistent differences in performance on the Salish and Farsi contrast does not clearly support the hypothesis that phonetic familiarity influences the perceptual difficulty of non-native speech contrasts. Likewise, an explanation that takes into account only differences in acoustic properties of the syllables falls short of fully explaining the clear order effects that were observed. Nevertheless, evidence was available to suggest that the two contrasts afford different assimilation patterns to native English listeners. Overall, the findings indicate that the perception of non-native contrasts involves perceptual processes that may be influenced by phonemic, phonetic, and acoustic properties of the stimuli as well as other variables, including orienting procedures, in ways that are not well understood. Because the assimilation model focuses on explaining how phonemic, phonetic, and acoustic factors interact to guide the processes that underlie the perception of non-native contrasts, this framework appears to hold much promise for explaining variability in the perception of non-native contrasts. However, the findings here challenge this model to specify the conditions under which assimilation processes result in least or most stable patterns of performance and what variables contribute to variation in assimilation patterns among individual listeners.

## GENERAL DISCUSSION

Cross-language studies have shown that some non-native speech distinctions present greater perceptual difficulty than others to the adult listener. Previous research has clearly demonstrated differences in the perception of phonemic versus nonphonemic contrasts. In the present study, this general finding was replicated, and it has also been shown that phonetic familiarity and acoustic factors may contribute to variability in perception of nonphonemic speech contrasts. In Experiment 1, the lack of a group difference in the perception of the Salish contrast by English and Farsi listeners suggests that the effects of linguistic experience on speech perception are specific to phonetic categories or feature combinations. However, neither between-subject nor within-subject comparisons of perception of Farsi and Salish contrasts provided consistent support for the role of phonetic familiarity. These negative results may be explained in terms of the interaction of divergent assimilation processes, the interaction of processes evoked by differential acoustic properties of

the Salish and Farsi stimuli, or the combined influence of these two factors.

Although this research has focused on the specification of the properties of speech contrasts that underlie the varying difficulty in the perception of non-native contrasts, other factors contributed to differences in performance. The tetrahedral model outlined by Jenkins (1979) provides a useful framework for more broadly assessing factors that affect the perception of non-native speech contrasts. According to this model, experimental outcomes in any domain of cognition are influenced by four classes of variables: (1) stimulus variables (e.g., linguistic properties, acoustic properties, natural versus synthetic speech), (2) criterial tasks (e.g., identification, discrimination), (3) orienting procedures (e.g., instructions, task configuration, cognitive strategies, feedback), and (4) differential characteristics of the subjects (e.g., language experience, age). (A discussion of each factor as it relates to training in the perception of non-native speech contrasts is found in Jenkins, Strange, & Polka, in press.)

In the present study, substantial individual differences in performance were observed in each experiment. Striking individual differences have been noted in previous cross-language studies (e.g., MacKain et al., 1981; Strange & Dittmann, 1984). Although possible bases for such differential abilities are not well understood, a study by Tees and Werker (1984) suggests that differences in early language experience may be important. They found that subjects who had been exposed to Hindi before 2 years of age (but who did not acquire Hindi) were able to improve their perception of the Hindi retroflex-dental place contrast through a brief laboratory training period. Subjects without such early exposure, including subjects with a year of college-level instruction in Hindi, failed to benefit from perceptual training with this contrast. The performance of individual subjects in Experiment 1 also points to early language experience as a potential source of individual differences. The performance of the early bilingual subjects in this experiment suggests that both broad and specific facets of early language experience may shape individual abilities.

Differences in perceptual performance related to testing procedures (e.g., ID vs. AX) and orienting procedures (e.g., order of contrast presentation) were also evident in the present study. Such manipulations are not likely to alter perception of naturally produced native speech contrasts. Thus, these findings reveal the fragility of the perception of foreign speech contrasts in highly inexperienced listeners, and they also indicate that researchers must be cautious in comparing absolute performance levels across experiments done with different testing procedures.

Overall, stronger conclusions could have been drawn from the present study if a more informative qualitative assessment of subjects' perception of the speech materials had been conducted. The subjects' descriptions of the non-native speech sounds were useful, but they failed to yield a consistent form of description across subjects and did not provide data on individual stimulus items. Thus, re-

searchers who wish to understand the processes involved in the perception of foreign speech sounds need reliable qualitative descriptions of subjects' perception of non-native contrasts. Conceptual progress may depend on the development of qualitative assessment tools that can clearly reveal what listeners perceive as they assimilate (or fail to assimilate) foreign speech sounds to native phonemic or phonetic categories.

In summary, phonetic and acoustic factors are potentially important in explaining why some non-native speech contrasts present greater perceptual difficulty than others. A clearer conceptualization of how these factors contribute to variability in adult cross-language speech perception will advance our understanding of how speech perception becomes shaped to meet the demands of processing a specific language. The development of methods to evaluate perceptual categories qualitatively and the growth of a better understanding of the bases of individual differences in speech perception present challenges for future studies of the role of experience in speech perception.

## REFERENCES

- ABRAMSON, A. S., & LISKER, L. (1970). Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the 6th International Congress of Phonetic Sciences* (pp. 569-573). Prague: Academia.
- BEST, C. T. (in press). The emergence of language-specific phonemic influences in infant speech perception. In H. C. Nusbaum & J. Goodman (Eds.), *Development of speech perception: The transition from recognizing speech sounds to spoken words*. Cambridge, MA: MIT Press.
- BEST, C. T., MACKAIN, K., & STRANGE, W. (1982, April). A cross-language study of categorical perception for semi-vowel and liquid contrasts. Paper presented at the 103rd meeting of the Acoustical Society of America, Chicago.
- BEST, C. T., MCROBERTS, G. W., & SITHOLE, N. N. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click perception by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception & Performance*, *14*, 345-360.
- BEST, C. T., MORRONGIELLO, B., & ROBSON, R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics*, *29*, 191-211.
- BRIERE, E. (1966). An investigation of phonological interference. *Language*, *44*, 786-796.
- BURNHAM, D. H. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics*, *7*, 207-240.
- CARAMAZZA, A., YENI-KOMSHIAN, G. H., ZURIF, E. B., & CARBONE, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*, *54*, 421-428.
- DISSOWAY-HUFF, P., PORT, R., & PISONI, D. B. (1982). Contextual effects in perception of /r/ and /l/ by Japanese. *Research on Speech Perception* (Progress Rep. No. 8). Bloomington, IN: Speech Research Laboratory, University of Indiana.
- EDMAN, T. R. (1981). Learning of intraphonemic discrimination for several synthetic speech continua. *Dissertation Abstracts International*, *41*, 4291B. (University Microfilms No. 81-09, 419)
- EILERS, R. E., GAVIN, W. J., & WILSON, W. R. (1979). Linguistic experience and phonemic experience in infancy: A cross-language study. *Child Development*, *50*, 14-18.
- FLEGE, J. E. (1989). Chinese subjects' perception of the word-final English /t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America*, *86*, 1684-1697.
- FLEGE, J. E., & HILLENBRAND, J. (1986). Differential use of temporal

- cues to the /s/-/z/ contrast by native and non-native speakers of English. *Journal of the Acoustical Society of America*, **79**, 508-517.
- FLEGE, J. E., & WANG, C. (1989). Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final /t/-/d/ contrast. *Journal of Phonetics*, **17**, 299-315.
- GOTO, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R." *Neuropsychologia*, **9**, 317-323.
- GOTTFRIED, T. L., JENKINS, J. J., & STRANGE, W. (1985). Categorical discrimination of vowels produced in syllable context and in isolation. *Bulletin of the Psychonomic Society*, **23**, 101-104.
- HENLEY, E., & SHELDON, A. (1986). Duration and context effects on the perception of English /r/ and /l/: A comparison of Cantonese and Japanese speakers. *Language Learning*, **36**, 505-521.
- JAMESON, D. G., & MOROSAN, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English /d/-/θ/ contrast by francophones. *Perception & Psychophysics*, **40**, 205-215.
- JENKINS, J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 429-466). Hillsdale, NJ: Erlbaum.
- JENKINS, J. J., STRANGE, W., & POLKA, L. (in press). Not everyone can tell a "rock" from a "lock": Assessing individual differences in speech perception. In R. V. Dawis & D. Lubinski (Eds.), *Assessing individual differences in human behavior: New methods, concepts, and findings*. Minneapolis, MN: University of Minnesota Press.
- LAMBTON, A. K. S. (1961). *Persian grammar*. London: Cambridge University Press.
- LIEBERMAN, P., & BLUMSTEIN, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.
- LISKER, L., & ABRAMSON, A. S. (1970). The voicing dimension: Some experiments in comparative phonetics. In *Proceedings of the Sixth International Congress on Phonetic Science* (pp. 563-567). Prague: Academia.
- LOGAN, J. S., LIVELY, S. E., & PISONI, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, **89**, 874-886.
- MACKAIN, K. S. (1982). Assessing the role of experience on infants' speech discrimination. *Journal of Child Language*, **9**, 527-542.
- MACKAIN, K. S., BEST, C. T., & STRANGE, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, **2**, 369-390.
- MANN, V. A. (1986). Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of English "l" and "r." *Cognition*, **24**, 169-196.
- MARCKWARDT, A. H. (1946). Phonemic structure and aural perception. *American Speech*, **21**, 106-111.
- MAYES, S. V. (1979). An acoustic analysis of Thompson velar /k/ and uvular /q/. *University of Hawaii Working Papers in Linguistics*, **11**, 11-22.
- MCCLASKEY, C. L., PISONI, D. B., & CARRELL, T. D. (1983). Transfer of training of a new linguistic contrast in voicing. *Perception & Psychophysics*, **34**, 323-330.
- MCNICHOL, D. (1972). *A primer of signal detection theory*. London: George Allen & Unwin.
- MIRANDA, S., UNDERBAKKE, M., STRANGE, W., & MICCERI, T. (1989, November). Training methods for the facilitation of Japanese students' perception of American English /r/ and /l/. *Journal of the Acoustical Society of America*, **86** (Suppl. 1), S102.
- MİYAWAKI, K., STRANGE, W., VERBRUGGE, R., LIBERMAN, A. M., JENKINS, J. J., & FUJIMURA, O. (1975). An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception & Psychophysics*, **18**, 331-340.
- MOCHIZUKI, M. (1981). The identification of /r/ and /l/ in natural and synthesized speech. *Journal of Phonetics*, **9**, 283-303.
- PICKETT, J. M. (1980). *The sounds of speech communication*. Baltimore: University Park Press.
- PISONI, D. B., ASLIN, R. N., PEREY, A. J., & HENNESSY, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 297-314.
- PISONI, D. B., LOGAN, J. S., & LIVELY, S. E. (in press). Perceptual learning of non-native speech contrasts: Implications for theories of speech perception. In H. C. Nusbaum & J. Goodman (Eds.), *Development of speech perception: The transition from recognizing speech sounds to spoken words*. Cambridge, MA: MIT Press.
- POLKA, L. (1987). Perception of Persian uvular and velar stops by speakers of American English. *Journal of the Acoustical Society of America*, **81**, (Suppl. 1) S68.
- POLKA, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America*, **89**, 2961-2977.
- REPP, B. H. (1981). Two strategies in fricative discrimination. *Perception & Psychophysics*, **30**, 217-227.
- RUHLEN, M. (1976). *A guide to the languages of the world*. Unpublished manuscript, Stanford University, Language Universals Project.
- SHELDON, A., & STRANGE, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, **3**, 243-261.
- SIEGEL, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- SINGH, S., & BLACK, J. W. (1966). Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *Journal of the Acoustical Society of America*, **20**, 372-387.
- STRANGE, W. (1972). *The effects of training on the perception of synthetic speech sounds: Voice onset time*. Unpublished doctoral dissertation, University of Minnesota.
- STRANGE, W., & DITTMANN, S. (1984). Effects of discrimination training on the perception of /r/-/l/ by Japanese adults learning English. *Perception & Psychophysics*, **36**, 131-145.
- STRANGE, W., & JENKINS, J. J. (1978). The role of linguistic experience in the perception of speech. In R. D. Walk & H. L. Pick, Jr. (Eds.), *Perception & Experience* (pp. 125-169). New York: Plenum.
- TEES, R. C., & WERKER, J. F. (1984). Perceptual flexibility: Maintenance or recovery of ability to discriminate non-native speech sounds. *Canadian Journal of Psychology*, **38**, 579-590.
- WERKER, J. F. (1991). The ontogeny of speech perception. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 91-109). Hillsdale, NJ: Erlbaum.
- WERKER, J. F. (in press). Cross-language speech perception: Developmental change does not involve loss. In H. C. Nusbaum, & J. Goodman (Eds.), *Development of speech perception: The transition from recognizing speech sounds to spoken words*. Cambridge, MA: MIT Press.
- WERKER, J. F., GILBERT, J. H. V., HUMPHREY, K., & TEES, R. (1981). Developmental aspects of cross-language speech perception. *Child Development*, **52**, 349-355.
- WERKER, J. F., & TEES, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, **7**, 49-63.
- WERKER, J. F., & TEES, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, **75**, 1865-1878.

## NOTES

1. Phonetic and acoustic structure are correlated in the speech signal. However, these two forms of substance characterize alternative kinds of information for a perceiver, and the acoustic information that is perceptually relevant depends on the listener's orientation to the stimulus array. Specifically, several studies have clearly demonstrated that subjects perceive the dimensions of a complex stimulus differently, depending on whether or not they perceive the stimulus as speech (Best, Morrongiello, & Robson, 1981; Repp, 1981). Such observations provide strong evidence in support of the existence of alternative modes of perception for speech and nonspeech. Specifically, it has been shown that when a signal is not perceived as speech, subjects can attend selectively to individual acoustic dimensions or divide attention among several acoustic dimensions (Repp, 1981). However, when the same signal is perceived as speech, subjects integrate the same acoustic dimensions in a phonetically relevant manner.

2. Note that, in other publications, different names for the Salish language have been used, including Thompson and Inslekampx.

3. The instructions and familiarization were intended to encourage subjects to treat the stimuli as speech categories. This may also have biased subjects to attend to phonetic or phonemic information and may have optimized absolute performance levels.

4. For several subjects, ID performance was significantly below chance predictions in a binomial test ( $p < .05$ ), whereas AX performance was well above chance level predictions. It was assumed that these subjects had reversed the assignment of category labels in the ID task. Therefore, for such subjects, ID data were rescored with the category labels reversed before their data were combined for group analyses. Data were rescored in this manner for 2 Farsi and 3 English subjects.

5. Werker and Tees (1984b) reported mean  $A'$  values of .654, .757, and .877 across three experiments on English listeners' perceptions of the Salish contrast in an AX task. Although other procedural differences exist between each of these experiments and the present study, the mean  $A'$  values reported fall within the range of values that have been reported in previous studies done with the AX procedure.

6. The distribution predicted for native-like perception is discontinuous, which makes this test more conservative (Siegel, 1956).

7. The fact that several subjects consistently reversed the category labels on the ID test provides further support for this interpretation.

8. However, it is not clear that subjects perceived the Salish as non-speech, which the assimilation model would presume for nonassimilated contrasts.

9. Stimuli modified in this fashion were also employed in a previous unpublished study done with VCV syllables spoken within sentences (Polka, 1987). The modification did not alter the phonetic identity of the consonants for native listeners.

10. Werker and Tees (1984b) also observed this with respect to the Salish contrast.

11. Before the data were combined for group analyses, the data for 1 subject in the S/F group were rescored with the assignment of category labels reversed, because this subject's performance was significantly below chance on the ID test and well above chance on the AX task.

12. In fact, English-speaking listeners have been shown to discriminate the isolated vocalic portions of the Salish syllables with a very high level of accuracy, even when no prior familiarization trials are provided (Werker and Tees, 1984b, Experiment 1).

(Manuscript received May 13, 1991;  
revision accepted for publication January 16, 1992.)