Linguistic influences in adult perception of non-native vowel contrasts

Linda Polka
School of Communication Sciences and Disorders, McGill University, 1266 Pine Avenue West, Montreal, Quebec H3G 1A8, Canada

(Received 25 June 1993; revised 7 July 1994; accepted 6 September 1994)

Perception of natural productions of two German vowels contrasts, /y/ vs /u/ and /ʏ/ vs /ʊ/, was examined in monolingual English-speaking adults. Subjects were tested on multiple exemplars of the contrasting vowels produced in a dVt syllable by a native German speaker. Discrimination accuracy in an AXB discrimination task was well above chance for both contrasts. Most of the English adults failed to attain "nativelike" discrimination accuracy for the lax vowel pair /ʊ/ vs /ʏ/, whereas all subjects showed nativelike performance in discriminating the tense vowel pair /u/ vs /y/. Results of a keyword identification and rating task provided evidence that English listeners' mapping of the German vowel to English vowel categories can be characterized as a category goodness difference assimilation, and that the difference in category goodness was more pronounced for the tense vowel pair than for the lax vowel pair. The results failed to support the hypothesis that the acoustic structure of vowels consistently favors auditory coding. Overall, the findings are compatible with existing data on discrimination of cross-language consonant contrasts in natural speech and suggest that linguistic experience shapes the discrimination of vowels and consonants as phonetic segmental units in similar ways.

PACS numbers: 43.71.Hw, 43.71.Es

INTRODUCTION

Interest in the role of language experience in speech perception has been growing steadily since the first studies of cross-language speech perception were reported in the early 1970s. The earliest and most extensive research investigating linguistic influences on speech perception has centered on studies of consonant perception. The goal of the present study was to examine cross-language vowel perception in adult listeners.

Cross-language speech perception studies have repeatedly shown that adults often have difficulty discriminating and identifying syllables that differ in a single non-native consonantal feature when tested with both natural speech and synthesized speech continua (e.g., Flege and Hillenbrand, 1986; Goto, 1971; Jamieson and Morosan, 1986; MacKain *et al.*, 1981; Mochuzuki, 1981; Miyawaki *et al.*, 1975; Pisoni *et al.*, 1982; Polka, 1991, 1992; Werker and Tees, 1984. For reviews, see also Pisoni *et al.*, 1994; Strange and Jenkins, 1978; Werker, 1994). This perceptual advantage for native (phonemic) over non-native (nonphonemic) contrasts demonstrates that adult speech perception is organized to process the native language with the greatest efficiency and least effort.

Although adults often have difficulty perceiving foreign consonantal contrasts, they also experience quite varying degrees of difficulty in response to different non-native consonant contrasts even when they are tested using natural speech and identical testing conditions. Studies directed at this issue have shown discrimination accuracy to vary from chance to nativelike performance for different non-native consonant contrasts that were presented to English-speaking adults (Polka, 1991, 1992; Best, 1993, 1994). Furthermore, this variability is not readily eliminated through perceptual training, as a substantial amount of laboratory training is required to yield small improvements in some difficult non-native contrasts (Pruitt *et al.*, 1990; Logan *et al.*, 1991). Thus it is clear that language-specific influences are more evident in perception of some non-native contrasts than in others.

Recently, Best has proposed a perceptual assimilation model to account for variability in discrimination of non-native contrasts (Best, 1993, 1994). Best maintains that adults perceptually assimilate non-native contrasts to the phonemic categories of their native language whenever possible, with the direction and degree of assimilation being determined by phonetic similarities between native and non-native phones.[1] Within this framework five basic patterns of assimilation have been outlined along with some predictions regarding their relative discrimination difficulty for adults. Discrimination is expected to be excellent for *two-category* contrasts, in which each non-native phone is assimilated to a different native phoneme category, whereas poor discrimination is predicted for *single-category* contrasts, in which both non-native phones are equally assimilable to the same native phoneme category. *Uncategorizable* contrasts, in which both phones are heard as speech but neither can be assimilated to a native phonemic category, are also expected to be poorly discriminated, but somewhat better than single-category contrasts. Good to moderate discrimination is predicted for contrasts that are assimilated as a *category goodness difference* in which both phones are assimilated to the same native phoneme category but differ in the goodness of fit to the category, and also for *nonassimilated* contrasts, in which both phones fall entirely outside of the native phonetic space and are not even heard as speech. Some variation among cat-

 0001-4966/95/97(2)/1286/11/$6.00

egory goodness difference and nonassimilated contrasts is also predicted, but for different reasons. The relative difficulty among category goodness contrasts depends on the degree of category goodness difference between contrasting phones whereas the relative difficulty among nonassimilated contrasts depends on the acoustic distinctiveness of the contrasting phones as nonspeech sounds. Examples of consonant contrasts that fit each type of assimilation have been presented elsewhere (Best *et al.*, 1988; Best, 1994; Polka, 1991, 1992; Werker, 1991).

Comparisons of discrimination performance among consonant contrasts conforming to different assimilation patterns have provided support for Best's model (Best *et al.*, 1988; Best, 1994; Best and Strange, 1992; Polka, 1991, 1992). However, to date, no studies have examined discrimination of non-native vowel contrasts within this framework. Further studies of cross-language vowel perception are needed to build a comprehensive description of experiential influences on speech perception. The present study contributes to this objective.

Comparing cross-language vowel and consonant perception is also meaningful given the differences in the acoustic and perceptual properties of vowels and consonants that have been documented (Ladefoged, 1982; Borden and Harris, 1980). Categorical perception studies focusing on native language contrasts have shown that discrimination of within-category differences tends to more accurate for vowel than for consonant stimulus series. This description is most accurate with respect to the perception of isolated vowels which is not strongly categorical, although there is often a contribution of phonetic categorization (Repp and Crowder, 1990). For vowels occurring in a syllabic context listeners often show a category boundary effect, though within-category discrimination performance is usually still better than identification performance would predict (for a review see Repp, 1984). These differences (and many other findings) have been taken as evidence for a dual coding of speech in both an auditory code and a phonetic category code (e.g., Ades, 1977; Fujisaki and Kawashima, 1970; Pisoni, 1973; Sawusch *et al.*, 1980). Although both codes are available in processing vowels and consonants the typical acoustic properties of vowels in natural speech (long duration, large and slow spectral change) favor auditory coding whereas the inherent acoustic properties of consonants (brevity, small and rapid spectral change) do not (Studdert-Kennedy, 1993). Thus comparing cross-language vowel and consonant discrimination can contribute information regarding how language experience interacts with these coding processes.

Some previous cross-language vowel perception studies have examined the perception of synthesized speech series using the categorical perception paradigm. In such work, language-specific influences have been repeatedly demonstrated in vowel identification performance, consistent with effects found in many consonant studies (Beddor and Strange, 1982; Bohn and Flege, 1990; Flege and Bohn, 1989; Gottfried and Beddor, 1988; Stevens *et al.*, 1969). However, the few cross-language vowel studies that have also examined discrimination do not conform to findings from cross-language consonant studies. Stevens *et al.* (1969) compared Swedish-speaking and English-speaking adults' perception of a continuum of synthesized isolated vowels in which the end points contrasted rounded vowels that are phonemic in Swedish, but not in English. In this study, language-specific effects were not apparent in discrimination performance; discrimination functions were similar and were not categorical (i.e., predictable from identification performance) for either Swedish or English adults. A later study by Beddor and Strange (1982) compared English-speaking and Hindi-speaking adults' perception of a series synthesized to simulate the Hindi oral versus nasal vowel contrast, /ba/–/bã/. A category boundary effect was more clearly evident in discrimination functions for the Hindi listeners than for the English listeners'. English adults' discrimination of between-category differences was quite accurate, but they were more accurate than the Hindi listeners in discriminating within-category differences. The finding suggests that the effects of language experience may be expressed differently in perception of consonant and vowel contrasts presented in a natural syllabic context (see also Mack, 1989).

Although effects of linguistic experience have been shown using natural speech stimuli for consonant contrasts, cross-language vowel perception has rarely been examined using natural speech stimuli. Bohn and Flege (1990) reported linguistic influences in German adults' identification of natural exemplars of several English vowels. Gottfried (1984) found that English-speaking adults were less accurate than French-speaking adults in categorizing natural CV syllables produced by several talkers according to vowel contrasts that are phonemic in French but not in English. Thus Gottfried's vowel data replicated the basic finding reported in cross-language consonant studies using a different discrimination paradigm.

The present study was undertaken to assess further whether language experience is expressed differently in vowel and consonant perception by examining cross-language vowel perception in natural speech using discriminations tasks that have been frequently employed in cross-language consonant studies. If cross-language vowel and consonant perception show the same patterns of language-specific effects, non-native listeners should be less accurate than native listeners in discriminating some non-native vowel contrasts. Discrimination performance would also be expected to vary for different non-native vowel contrasts and this variability should be associated with differences in the assimilation of non-native vowel contrasts to native vowel categories as outlined by Best (1993). If such parallel findings are observed for discrimination of vowel contrasts and consonant contrasts in natural speech, it would then indicate that the prominent effect of native phonemic structure evident in previous consonant studies is also evident for vowels.

On the other hand, it is possible that the acoustic structure of vowels in natural speech consistently favors auditory coding in certain perceptual tasks. Previous studies using synthetic speech are consistent with this hypothesis in that discrimination of between-category vowel differences was good in both native and non-native listeners whereas language effects were evident only for within-category discrimination performance (Beddor and Strange, 1982). If this hypothesis is correct, adults would be expected to show uniformly high levels of discrimination for non-native vowel contrasts, regardless of how the non-native vowels map onto native vowel categories. This outcome would indicate that linguistic influences are not evident for vowels in some perceptual tasks (e.g., discriminating a pair of natural syllables), even though they may be quite clear in others (e.g., identification or categorization across multiple talkers).

The research reported here was designed to evaluate these hypotheses by examining English listeners' perception of two German vowel contrasts. Data were gathered to answer three specific questions regarding language-specific influences on adult vowel perception. First, do English adults have difficulty discriminating non-English vowel contrasts in natural productions spoken by a single talker as has been shown repeatedly in consonant studies? Second, do English adults perceive the German vowels as similar to their native English vowel categories? Third, do these two contrasts conform to particular assimilation patterns outlined by Best, and if so, is the relative discriminability of the two contrasts consistent with Best's model?

Perception of the two German vowel contrasts by monolingual English speaking adults was evaluated using two perceptual tasks, an AXB discrimination task and a vowel identification and rating task using English keywords. The AXB task provided data on the ability of inexperienced adult English listeners to discriminate the non-native vowel categories. Specifically, the AXB discrimination task was structured to assess the listeners' ability to recognize different vowel categories produced by the same talker over and above within-category variation among multiple natural exemplars. This task or variations of it have been used frequently in studies that examined the relative discriminability of non-native consonant contrasts in adults using natural speech stimuli (Best *et al.*, 1988; Polka, 1991, 1992; Werker and Tees, 1984).

The identification and rating task was designed to show whether (and to what degree) English listeners perceive the German vowels as being similar to specific English vowel categories. The results will describe the assimilation of the German vowel contrasts to English phonology and will be interpreted in terms of the assimilation patterns and discrimination predictions outlined by Best (1993).

I. METHOD

A. German vowels

Phonetic descriptions characterize vowels in terms of at least three corresponding articulatory and acoustic features: (1) tongue position in the oral cavity specified in terms of height (high, mid, low) and front versus back, with both dimensions influencing the location of spectral peaks (formants), (2) tense versus lax, which corresponds to differences in timing and extent of tongue movement and is acoustically specified in the spectral and temporal structure of formant patterns, and (3) lip shape as unrounded versus rounded where lip rounding tends to lower the frequencies of all spectral peaks (Ladefoged, 1982).

German vowel contrasts were chosen for this research

TABLE I. German vowel contrasts.

Lax contrast		Tense contrast	
ʏ	ʊ	y	u
high	high	high	high
front	*back*	*front*	*back*
rounded	rounded	rounded	rounded
lax	lax	tense	tense

because German has vowel categories and contrasts that are not used in English. The phonetic features specifying the German vowels examined in the present study are outlined in Table I, with the distinctive feature for each contrast in italic. As Table I shows, both German vowel pairs are front–back minimal contrasts between high lip-rounded vowels. One pair contrasts two lax German vowels, /ʊ/ and /ʏ/, and the other pair contrasts corresponding tense vowels, /u/ and /y/. Thus German /y/ vs /u/ and /ʏ/ vs /ʊ/ contrast in the same phonetic feature (front versus back) but the vowels in each contrast combine different articulatory features.

The English vowels /u/ (as in "ooze") and /ʊ/ (as in "foot") are described using the same articulatory phonetic features as German /u/ and /ʊ/, respectively. German /y/ and /ʏ/ are also articulated with a high tongue position and rounded lips, like English /u/ and /ʊ/, but with a front tongue position similar to the English vowels /i/ (as in "tea") and /ɪ/ (as in "pit") which are not lip rounded. In English, only high back vowels have the lip-rounding feature; front vowels are not lip rounded. Therefore English listeners may have difficulty discriminating these German front-rounded versus back-rounded vowel contrasts because English does not have a phonemic distinction between high front versus back vowels that is independent of lip-rounding differences.

B. Stimuli

A male native speaker of German produced the sample of German vowels used in this study. He was a 31-year-old native of Karlsruhe, Germany and was, at the time of the recording, a graduate student in German at McGill University. He arrived in Canada at age 27 and speaks German daily with family and friends.

The talker was recorded producing multiple instances of six German vowels (/ʏ/, /ʊ/, /y/, /u/, /ø/, /ɔ/) in a /dVt/ context.[2] According to phonological rules of German these six vowels are produced as monophthongs in this phonetic environment. The /dVt/ tokens were produced in citation form and recorded in a sound attenuated chamber using a Revox A77 reel-to-reel tape recorder and a Sennheiser MD-441-U microphone. Five native German speakers (including the talker) listened to and identified the vowel in each /dVt/ syllable. A few items that were not identified consistently by all five German listeners were eliminated from the pool of stimulus items.

Next, the /dVt/ syllables were converted to digital waveforms (10-kHz sampling rate, 12-bit resolution, low-pass-filtered at 4.54 kHz) and stored as separate files on a Macintosh II computer using MACSPEECH LAB II software. Some acoustic measures of the /dyt/, /dut/, /dUt/, and /dYt/ syl-

TABLE II. Duration, amplitude, and $f0$ measures for the selected dVt syllables with German vowels.

	German lax vowels				German tense vowels			
	/ʊ/		/ʏ/		/u/		/y/	
	mean	range	mean	range	mean	range	mean	range
Duration (ms)								
VOT	12.9	11.9–14.4	13.6	12.3–15.2	15.1	11.1–19.2	13.3	12.3–15.3
vocalic	88.8	79.7–105.1	84	76.4–94.4	186.1	173.4–202.8	187.2	169.4–206.0
closure	115.7	106.0–136.1	114.6	105.8–128.3	100.6	92.1–104.4	89.1	76.0–96.9
final burst	58	36.8–75.7	57	33.3–72.6	52.4	37.7–66.1	48.6	37.2–63.0
total	270.8	243.5–307.7	269.2	240.8–285.8	354.3	333.7–387.9	338.2	302.7–372.1
Amplitude (dB)								
mean	65.7	64.7–66.5	65.9	65.1–66.3	65.9	65.5–66.4	66.7	66.5–67.0
Average $F0$ (Hz)								
beginning	102.5	97–109	104.8	99–112	106	100–110	105.3	102–109
middle	105.1	92–114	104.8	99–118	107.3	102–115	100.7	93–106
end	105.5	89–117	101.5	99–113	109.8	102–122	111.2	101–119

lables were gathered to assess differences that are not contrastive segmental cues for distinguishing front-rounded and back-rounded vowels, including amplitude, duration, fundamental frequency, and several measures related to stop consonant production. Using these measures, six /dVt/ tokens of each vowel were selected to avoid systematic differences in these noncontrastive acoustic dimensions within each minimal vowel pair. Thus the selected tokens provided some within-category variability in the stimulus set, but also insured that noncontrastive acoustic differences could not be used to discriminate the contrasting vowel categories. Overall amplitude of the selected stimuli was adjusted so that output levels of all of the stimuli peaked a VU meter within a very narrow range. Amplitude measures were very similar among tokens within each contrasting vowel pair; the adjustments served to equate the two contrasts to each other with respect to loudness, not to alter inherent within-contrast differences in vowel amplitude. A summary of the duration, amplitude, and $f0$ measures which describe the final set of stimulus items is provided in Table II.[3]

To conduct computer-controlled presentation of stimuli for the perceptual testing, the selected stimuli were redigitized (10-kHz sampling rate, 12-bit resolution) using BLISS software (Mertus, 1990) on a Compaq 286 computer. Stimuli were presented on-line from the computer over a single BR26 loudspeaker in a Tracoustics sound attenuated room. Stimulus presentation level was 69 dBA, measured at the approximate position of the subject head using a General Radio Precision sound level meter (type 1561).

C. Formant analysis

Analyses of the formants patterns in the six tokens of each German vowel were conducted using MACSPEECH LAB II software. Formant frequencies corresponding to the first three oral formants were measured using an LPC algorithm (13 coefficients) with a 24-ms Hamming window centered at successive 12.5-ms locations within each syllable beginning at the onset of voicing until voicing offset. Table III summarizes the formant frequency measures for the six tokens of each syllable. Formant measures corresponding to three locations within each syllable are presented including (1) onset of voicing, (2) offset of voicing, and (3) a mid vowel location which was defined as the earliest location within the syllable where $F2$ reached a frequency minimum. The latter location was chosen because the coarticulation with preceding and following alveolar stops resulted in a characteristic falling $F2$ transition into the vowel nucleus and rising $F2$ transition from vowel nucleus to final stop closure. Also, $F2$ is the primary acoustic cue that distinguishes front–back tongue position. Figure 1 presents the midvowel formant measurements plotted in an $F1$ by $F2$ vowel space. For comparison, the average formant values for English /u/, /ʊ/, /i/, and /ɪ/ produced by five male Western Canadian English speakers taken from Assmann (1979) are also shown.

Several additional formant values are presented in Table III. For formant measures taken at the midvowel location, $F1$-$F2$ differences and $F2-F3$ differences were computed to describe the frequency relationships among the first three formants. In addition, to characterize the extent of overall $F2$ movement, change in $F2$ frequency from onset of voicing to the $F2$ midvowel location (i.e., $F2$ minimum), and from $F2$ midpoint to voicing offset are also reported.[4]

In Table III, the formant measures that distinguish the contrasting vowel categories are highlighted in the Cue Status column. Those formant measures which show completely nonoverlapping distributions for the contrasting phones are designated with an asterisk. These nonoverlapping formant cues provide highly reliable information for distinguishing the front versus back vowels. Those formant measures which differed in mean value (by at least 100 Hz) but showed some overlap in the range of formant values corresponding to front versus back vowel are designated with a plus sign. These overlapping formant cues tend to differ between the vowels within a contrastive pair and thus also contribute some information for distinguishing the vowel categories. The remaining formant measures (those without either an asterisk or

TABLE III. Formant measures for dVt syllables containing German vowels. * =nonoverlapping values; + =overlapping values.

Formant frequency (Hz)	German lax vowels /ʊ/ mean	range	/ʏ/ mean	range	Cue status	German tense vowels /u/ mean	range	/y/ mean	range	Cue status
onset:										
$F1$	302	244–326	300	285–315		273	254–285	259	244–275	
$F2$	1583	1404–1719	1643	1607–1678		1597	1495–1688	1768	1739–1831	*
$F3$	2238	2126–2390	2234	2177–2268		2219	2095–2329	2412	2187–2879	+
midvowel:										
$F1$	322	325–336	300	264–326		270	244–295	251	234–285	
$F2$	1226	1099–1353	1614	1566–1658	*	970	742–1170	1699	1658–1739	*
$F3$	2126	2055–2177	2217	2146–2350		2117	2075–2156	2060	2014–2116	
$F1-F2$	904	774–1017	1314	1292–1343	*	693	498–916	1448	1424–1475	*
$F2-F3$	900	804–987	604	519–783	*	1154	936–1333	361	305–397	*
offset:										
$F1$	317	295–336	297	254–326		271	254–285	246	224–275	
$F2$	1466	1363–1536	1689	1617–1760	*	1119*	1038–1200	1865	1800–1984	*
$F3$	2104	2024–2207	2360	2207–2930	+	2012	1851–2065	2158	1984–2278	+
$F2$ change:										
onset to mid	358	183–509	29	10–71	*	643	478–773	70	10–163	*
offset to mid	241	142–346	74	41–132	*	285[a]	132–458	166	61–315	+

[a]Based on three syllables.

plus sign) were quite similar (in mean and range values) across the two contrasting vowels and thus provide little or no information for distinguishing these vowel contrasts in this stimulus set.

As shown in Table III and Fig. 1, several nonoverlapping $F2$ cues distinguished the front and back vowels within both contrasts. Overall, the same $F2$ parameters distinguished the front and back vowels within each contrast, but the magnitude of difference for each $F2$ cue was consistently larger for the tense contrast than for the lax contrast. As expected, the front vowels had higher $F2$ frequencies at the midvowel location relative to the back vowels in both contrasts. $F2$ frequency was also higher for the front vowels at voicing offset in both contrasts and at voicing onset for the tense vowels. The $F1$-$F2$ and $F2$-$F3$ differences at midvowel location also distinguish the front and back vowels in each contrast. For both the tense and lax front vowels, $F2$ was closer in frequency to $F3$ than to $F1$. For the tense back vowel /u/, $F2$ was closer to $F1$ than to $F3$; for the lax back vowel /ʊ/ $F2$ was roughly an equal frequency distance between $F3$ and $F1$. The extent of $F2$ movement also distinguished both contrasts, with less extensive change in $F2$ from voicing onset to midvowel observed for the front vowel than for the back vowel. The extent of $F2$ frequency change from mid- to voicing offset was also an overlapping cue for the tense contrast and a nonoverlapping cue for the lax contrast.

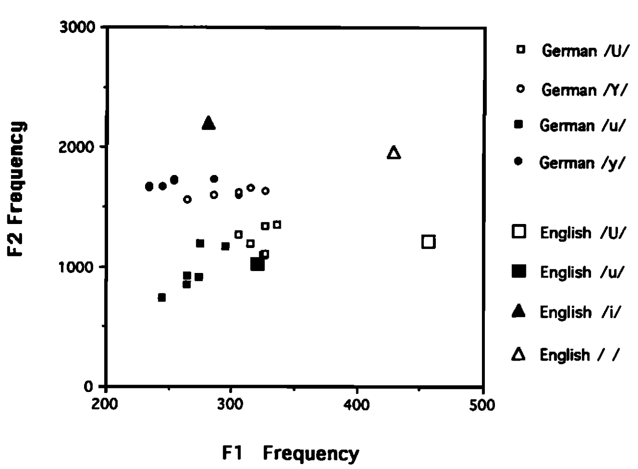

FIG. 1. $F1$ by $F2$ plot of German vowels and mean values for English vowels (for male talkers) from Assmann (1979).

Several differences were observed between the tense and lax vowels. First, durational measures (presented in Table II) show that the vocalic portions of the syllables were approximately twice as long for the tense vowels relative to the lax vowels. Second, differences between the tense and lax vowels in extent of tongue movement were evident in the measures of $F2$ change shown in Table III with more extensive $F2$ frequency change observed for the tense vowels than for the lax vowels. These tense/lax differences are consistent with acoustic studies that have shown tense and lax monophthongs in Northern German dialects to differ primarily in duration, but also to show some differences related to tongue movement and tongue height (Bennett, 1968). Finally, the tense vowel pair was more acoustically distinct than the lax vowel pair in that a larger number of nonoverlapping and overlapping cues were observed for the tense vowels, the formant differences between front and back vowels were larger for the tense contrast, and, being longer, the tense vowels also potentially provided greater opportunity for detection of these differences.

D. Subjects

Ten adult monolingual speakers of Canadian English, five males and five females, served as subjects. Subjects ranged from 18 to 27 years of age with a mean age of 20.8 years. Nine of the subjects were natives of Western Canada or had resided in Western Canada for most of their life. One subject was from Ontario, but had resided in Western Canada for eight years.

Eight native German-speaking adults were also tested in the identification and rating task, but not the categorial AXB task. This group included five males and three females and had a mean age of 33 years (range 24–55). All of the German subjects also spoke English and most of them had been residing in Canada between 3 and 9 years; one subject had resided in Canada for 1 year and another for 29 years. The subjects were from the following broad dialectal regions of Germany: Northern German ($n=3$), Northwestern German ($n=2$), and Southwest German ($n=3$).

E. Procedure

All English subjects completed the AXB discrimination task followed by the keyword task.

1. AXB discrimination

In the AXB task subjects were told that they would hear syllable triads in which the first syllable and the last syllable in a triad contained different vowel sounds and the middle syllable had the same vowel as either the first or the last syllable. Subjects were instructed to indicate which syllable, i.e., first or last, matched the middle syllable by pressing one of two buttons. The subjects were told that their responses would be timed. The middle syllable was not physically identical to the syllable which it matched in the triad; it was a different instance of the same vowel category. Thus subjects were required to make a vowel category match rather than an exact acoustic match.

Subjects were presented 36 AXB trials with the lax contrast and 36 AXB trials with the tense contrast with a brief pause between contrasts. The order in which the two contrasts were tested was counterbalanced across subjects. Eight familiarization trials (four per contrast) were presented in random order at the beginning of the AXB test session. Feedback was not provided during familiarization or test trials. An ISI of 2000 ms was used because it has been shown that language-specific effects are clearly observed in discrimination tasks when the ISI is relatively long (Werker and Logan, 1985; Carney *et al.*, 1977). The ITI was 3500 ms. For each German contrast, all 36 possible AB pairs were presented with an equal number of AAB, BAA, BBA, and ABB trials across the 36 trials. Across the 36 AXB trials, the six different tokens of each category were presented six times in both the A and the B position, and three times in the X position.

2. Vowel identification and rating using English keywords

This task was modeled after a vowel identification procedure developed by Strange and her colleagues and which has been used extensively in studies of English vowel perception (see Strange and Gottfried, 1980). In this procedure subjects use specific English words as responses in identifying English vowel categories. The 12 keywords designated for identifying English vowel categories are: eat (/i/), it (/ɪ/), ate (/e/), Ed (/ɛ/), at (/æ/), odd (/ɔ/), hawk (/ɑ/), oat (/o/), hook (/ʊ/), ooze (/u/), up (/ʌ/) and heard (/ɝ/). In the present study English subjects identified German vowels by matching them to English vowels using the same keyword responses and then rated the quality of the match they had just made.

The task was implemented in two stages. In the first stage, subjects were familiarized with the keywords and completed a brief English vowel identification test. The purpose of this stage was to establish that subjects were facile in identifying English vowels using keywords. In the first stage of the keyword task, subjects were presented 12 different English vowels produced in isolated /dVt/ syllables by a native female speaker of English. The stimuli were digitized using the same procedures as described above for the German stimuli. On each identification trial a /dVt/ syllable was played twice with an ISI of 1000 ms. The ITI was 5000 ms. Thirty-six identification trials were presented in three blocks of randomized trials to provide the subject three opportunities to identified a single exemplar of each English vowel.

In the second stage of the task subjects were presented /dVt/ syllables containing six different German vowels. The two extra vowels (/ø/ and /ɔ/) were included in case subjects perceived all four German high vowels as similar to the same English vowel category which would lead them to select the same keyword response on every trial. In this event it would be difficult to identify subjects who did not understand instructions or were not motivated to perform the task. To avoid this problem, the two extra German vowels, which were not likely to be perceived as similar to the same English vowels, were also included in the task. Subjects were asked to match each German vowel sound to the most similar English vowel category using the keywords and then to rate the quality of the match on a five-point scale (1=poor to 5=very good). As in the first stage, the syllable was played twice on each identification trial with an ISI of 1000 ms and the ITI was 5000 ms. Six blocks of 12 identification trials each were presented. Across the 72 trials, two trials were presented for each of the six tokens of each German vowel category.

The German vowels were also presented in the same way to the native German listeners for identification and rating responses with respect to German vowel categories. Although the /dVt/ syllables are not words in German, German vowels can be unambiguously specified in different spellings. Native German speakers were presented 14 written syllables to select from in identifying the German vowels in a /dVt/ context. After identifying each syllable, the subjects also rated the quality of the match to the selected vowel using a 1–5 rating scale.

II. RESULTS

A. AXB discrimination

Overall, English adults were highly accurate in their discrimination of both German contrasts. Mean percent correct

TABLE IV. English listeners' responses to the German vowels in the keyword identification and rating task. Left number=% of responses (120 possible); right number=average rating (1=poor, 5=good).

	German lax vowels				German tense vowels			
English keyword/vowel	back ʊ		front ʏ		back u		front y	
Back vowels:	**83.2**		**73.5**		**95.9**		**93**	
ooze /u/	38.3	3.3	15	2.7	79.2	4.4	85.5	2.5
hook /ʊ/	37.5	3.4	52.5	3.2	15	4.3	5	3.8
oat /o/	1.6	1.5	1	1	0	0	0.8	5
odd /a/+hawk /ɔ/	5.8	2.9	5	3.7	1.7	2	1.7	5
Front vowels:	**6.6**		**14.1**		**2.5**		**4.1**	
at /æ/	0	0	0	0	0.8	3	0.8	3
Ed /ɛ/	1.6	2.5	1.7	2	0	0	0.8	3
ate /eɪ/	0	0	0	0	0	0	0	0
it /ɪ/	5	1.7	11.6	3.3	1.7	4	1.7	4
eat /i/	0	0	0.8	1	0	0	0.8	2
Other:	**10.5**		**12.5**		**2**		**2.4**	
up /ʌ/	5.8	2.3	6.7	2.2	2	3.5	0.8	3
heard /ɝ/	3	2.7	5	2.4	0	0	0.8	3
no response	1.7		0.8		0		0.8	

on the AXB task was 98.6% for the German tense vowels (s.d.=2.4) and 86.9% for the German lax vowels (s.d.=8.2). For each contrast, AXB performance was compared to chance performance of 50% using single mean t tests. Performance on both contrasts was well above chance predictions ($t=14.277$, $p<0.0001$ for the lax contrast; $t=65.118$, $p<0.0001$ for tense vowel contrast).

To compare individual subject's performance more directly to that of native listeners, a criterion for nativelike performance was established as greater than 90% correct.[5] All ten subjects met this criterion for the tense German vowels (seven subjects had perfect performance), but only two subjects attained this level of accuracy for the lax German vowels.

To evaluate differences in perceptual difficulty of the two German contrasts, performance in the AXB task was also analyzed in a mixed factor ANOVA with contrast (tense versus lax) as the within-subjects factor and order (tense/lax versus lax/tense) as the between-subjects factor. This analysis was conducted for three dependent measures: (1) number of errors, (2) average response time on correct trials, and (3) standard deviation of response times on correct trials. In all three ANOVAs only the contrast effect was significant [$F(1,8)=18.186$ $p<0.0027$ for analysis of errors, $F(1,8)=35.385$, $p<0.003$ for analysis of average response time, and $F(1,8)=16.737$, $p<0.0035$ for analysis of variation in response time]. These results show that English subjects made more errors and had longer and more variable response times in discriminating the German lax vowel contrast than in discriminating the German tense vowel contrast.

B. Keyword identification and rating task

The native German speakers' identification of the intended German vowels was essentially perfect (99%–100% accuracy among the four vowels). Ratings were also consistently high with close to a unanimous assignment of a 5 rating to all instances of the four German vowels. Overall mean ratings for the four vowels varied from 4.8 to 4.9. Thus German adults readily identified all four vowels to be good instances of their native vowel categories.

Overall, English listeners were able to perform the first stage of the keyword task with little difficulty. Among the ten subjects, eight identified the English vowels using keywords with 100% accuracy; two subjects made a single error on the first two trial blocks.[6] Data from only one additional subject were discarded because the subject could not reliably perform the keyword task with English vowels.

The results of the keyword identification and rating task with the German vowels are summarized in Table IV. In each column, the left number is the percentage of times (of possible 120) in which the English keyword listed in the row was selected to match the German vowel listed in the column. The right number is the overall average rating for all the matches designated by the percentage on the left.

As shown in Table IV, the mapping to English vowel categories was quite restricted for the German tense vowels with the English vowels in keywords "hook" and "ooze" accounting for over 90% of responses. For the German lax vowels, English keywords ooze and hook also represent a very large proportion of the responses although the mapping to English vowel categories was more distributed. Thus all four German vowels were matched most often to keywords containing English vowels that are high, back, and lip rounded. However, English listeners showed some sensitivity to the front tongue position of German /y/ and /ʏ/ in that there were some matches of the German front vowels, especially /ʏ/, to English high front vowels.

To analyze the mapping of the German vowels to English vowel categories, the selection and rating data were combined into a composite similarity score by summing each subject's ratings for each keyword/German vowel pair (yielding a value between 0 and 60[7]) and multiplying by a constant so that the similarity score had a maximum value of 100. For each contrast an ANOVA was conducted on the

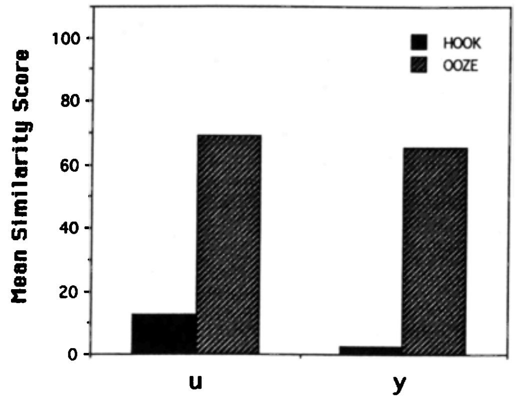

FIG. 2. Similarity scores for German tense vowels.

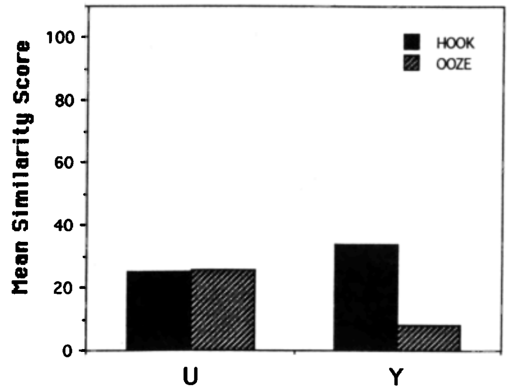

FIG. 3. Similarity scores for German lax vowels.

similarity scores with German vowel (front versus back) and keyword (ooze, hook) as within-subject factors.[8] Mean similarity scores (averaged across the ten English listeners) for keywords ooze and hook are shown in Figs. 2 and 3 for the tense and lax vowels, respectively.

For the analysis of the German tense vowels, only the two main effects were statistically significant. The significant keyword factor [$F(1,9)=34.564$, $p<0.0002$] showed that both German /u/ and /y/ were perceived to be more similar to ooze than to hook. The significant German vowel factor [$F(1,9)=9.491$, $p<0.013$] indicated that similarity scores were higher for German /u/ than for German /y/ in matches to both ooze and to hook.

For the German lax vowels, the ANOVA yielded a significant main effect for German vowel [$F(1,9)=29.824$, $p<0.0004$] showing that overall similarity scores were higher for the German back vowel /ʊ/. The German vowel by keyword interaction was also significant [$F(1,9)=11.586$, $p<0.0078$]. Tukey pairwise comparisons ($p<0.05$) of all possible keyword/vowel pairs were conducted. These comparisons showed that German /ʊ/ was perceived to be equally similar to both ooze and hook, whereas German /ʏ/ was more similar to hook than to ooze. As observed for the German tense vowels, similarity scores were higher for German /ʊ/ than for German /ʏ/ in matches to ooze. Similarity scores were not significantly different for German /ʊ/ and /ʏ/ in matches to hook.

III. DISCUSSION

This study was directed at three questions. First, do adult monolingual speakers of English have difficulty discriminating non-English vowel contrasts in natural syllables produced by a single talker? Although performance was above chance predictions for both contrasts, English listeners had more difficulty discriminating the front–back distinction for the lax vowel contrast, /ʏ/ vs /ʊ/, than for the tense vowel contrast, /y/ vs /u/. Most of the English adults failed to achieve nativelike performance in discriminating the German lax contrast, /ʊ/ vs /ʏ/, whereas all displayed nativelike discrimination accuracy for the tense contrast, /u/ vs /y/.

Results of the vowel identification and rating task using English keywords provided data to address the second question: do adult English listeners perceive the German vowel categories to be similar to English vowel categories? The identification data clearly demonstrated that English adults recognize similarities between these German vowels and the vowel categories in their native vowel inventory.[9] Consistent with English phonology, English adults perceived German /u/ and /ʊ/ to be similar, but not identical to, the English vowels in ooze and hook which are also described as high, back, lip-rounded vowels. Likewise, the German front vowels /y/ and /ʏ/ were also perceived as most similar to the English vowels in ooze and hook, respectively. Thus in matching the German high front-rounded vowels to English vowel categories, the lip-rounding feature was more salient than front tongue position.

For the vowel pair /u/–/y/, English listeners' mapping of the German vowels to English vowel categories can be characterized as a *category goodness difference* assimilation pattern. Matches to English ooze accounted for a large proportion of identification responses for both /u/ and /y/ and the similarity scores were higher for German /u/ than /y/. Thus German /u/ and /y/ appears to be perceived as a "good" versus "less good" exemplar of English /u/. Subjects also matched both German tense vowels to the English vowel in hook, though with much lower frequency. Here also, the similarity scores were higher for German /u/ than for German /y/, consistent with a category goodness difference assimilation.

The mapping of the German lax vowels was more complex because subjects consistently matched both German lax vowels to two English vowels. These data show that English adults subjects map the German lax vowels to a small number of English vowels and thus this contrast does not appear to be either *uncategorizable* or *nonassimilated.* Moreover, there were differences in the identification and rating results for the two German lax vowels. Although the data do not exactly fit a category goodness difference pattern with respect to a single category, similarity scores were significantly higher for German /ʊ/ than for German /ʏ/ when these vowels were matched to ooze and also when matches to hook and ooze are combined. Thus English listener's perception of German /ʊ/ and /ʏ/ can be described as differing in the degree of fit to the English /u/ or /ʊ/ vowel categories.

A comparison of the discrimination and identification results addressed the third question of whether variability in discrimination of non-native contrasts depends on the extent to which non-native phones can be assimilated to native phonemic contrasts, as predicted by Best's model (Best, 1993). Findings in the present study uphold this general prediction in that English adults had greater difficulty discriminating the German lax vowel contrast which showed a less pronounced pattern of category goodness assimilation relative to the German tense contrast. Thus differences in assimilability of these two German vowel contrasts to native phonemic categories appear to be reflected in the relative discriminability of these two non-native vowel contrasts.[10]

The present study sought to evaluate hypotheses regarding language-specific influences in discrimination of vowel contrasts. Overall, the findings are in line with studies that have examined perception of non-native consonant contrasts using similar tasks and stimulus materials. Discrimination accuracy was quite good for both contrasts, but nevertheless fell short of nativelike performance for one vowel contrast and varied significantly between the two non-native vowel contrasts. The relative difficulty on the two contrasts was also broadly consistent with the assimilation model which has been proposed by Best and which is supported by data from cross-language studies of consonant perception (Best *et al.*, 1990; Best and Strange, 1992; Best *et al.*, 1988; Polka, 1991, 1992; Werker, 1991). The identification and rating tasks showed that, for both contrasts, a category goodness assimilation pattern was evident in the English keywords that were selected most frequently and given the highest quality ratings. Furthermore, as Best's model would predict, relative discriminability of the two German contrasts was reflected in assimilation patterns in that adults had more difficulty with the contrast that they were less able to relate consistently to English vowel categories. The compatibility of these results with findings from comparable studies of consonant perception point to a common pattern of language-specific effects for discrimination of vowel and consonant contrasts in natural speech.

The present findings argue against the hypothesis that the acoustic structure of vowels in natural speech consistently favors auditory coding in discrimination tasks. However, the contribution of auditory coding in cross-language vowel discrimination may nevertheless be greater than is observed in cross-language consonant perception. Once discrimination of a more varied sample of non-native vowel contrasts has been examined, we can also assess whether discrimination of non-native vowel distinctions is generally better and less variable (compared to consonants) in adult listeners owing to the acoustic structure of vowels or to the patterns of assimilation that characterize cross-language vowel perception.[11] The relatively high levels of discrimination performance observed in this study are consistent with this notion but further studies with additional non-native vowel contrasts are needed to test this hypothesis.

Best's model emphasizes the role of category coding processes in the perception of speech segments; acoustic properties become relevant only in accounting for differences among nonassimilated contrasts. While the present study provides some additional support for Best's model, the potential role of auditory coding processes in accounting for the present findings cannot be overlooked. Acoustic analyses of the German vowels showed that formant differences were more pronounced for the tense contrast than for the lax contrast. Thus the differences in discrimination performance between these two contrasts may be accounted for by auditory coding processes which are sensitive to the degree of physical acoustic differences. In light of this, the present results are also entirely consistent with the speech learning model that has been put forth by Flege (e.g., Flege *et al.*, 1994; in press). Flege has proposed that, in vowel perception, category coding is prominent only when auditory coding is disfavored, for example, when listeners must attend to small formant differences. The present findings support this hypothesis in that English adults performed as well as native listeners in discriminating the more physically distinct tense vowel pair. German adults, who have acquired a phonology in which they can readily categorize these vowels, outperformed the English adults only on the less acoustically distinct lax vowel pair. The present study does not provide data to assess the role of acoustic differences in cross-language vowel perception independent of differences in mapping of the German vowels to English vowel categories. In this regard, a comparison of discrimination performance between two groups of non-native listeners who are likely to have different assimilation patterns (e.g., English and Japanese) for these two German vowel pairs would be informative.

English adults' identification and rating of the German vowels in this study is also interesting in light of Kuhl's thesis that vowel categories are internally organized around language-specific prototypes (Kuhl, 1992). In recent work, Kuhl and her colleagues (Grieser and Kuhl, 1989; Kuhl, 1991) have provided evidence for language-specific influences on the internal structure of a single vowel category which are described in terms of a perceptual magnet effect. The magnet effect is observed in listeners' responses to a set of synthesized steady-state vowels that vary along $F1$ and $F2$ dimensions in equal mel steps, forming a ring of vowel tokens that are equal mel distance from a central vowel stimulus which is a prototypic instance (ideal exemplar) of the vowel category. Adults identify the entire set of stimuli as the same vowel, but show an asymmetry in their discrimination of the vowel set. Adults had greater difficulty discriminating a change from the central stimulus to a more peripheral (less prototypic) stimulus than a change from a peripheral stimulus to a more central (more prototypic) stimulus, even when the change to be discriminated involves the same pair of stimuli. Thus the central "prototypic" stimulus appears to act like a perceptual magnet, effectively reducing the perceptual distance between itself and more peripheral (and less prototypic) members. Kuhl *et al.* (1992) also examined the magnet effect for stimuli synthesized around a prototypic Swedish /y/ and around a prototypic English /i/ in 6-month-old infants from English and from Swedish families. The magnet effect was stronger in infants' discrimination of their native vowels, suggesting that language experience operates to shape the internal structure of vowel categories.

English adults' identification and rating data for the German vowel contrasts in the present study are also in line with the notion of a language-specific prototype structure. The identification and rating results indicate that the German back vowels /u/ and /U/ were perceived as more prototypic instances of English high back vowels while the corresponding German front vowels, /y/ and /Y/ were perceived to be less prototypic instances of the same English vowel categories. This observation raises the question of whether a perceptual magnet effect is associated with this typicality difference. If so, then English adults should have more difficulty in discriminating a change from the prototypic German /u/ to less prototypic German /y/ than a change from German /y/ to German /u/. Studies investigating possible magnet effects are currently underway.

In summary, the pattern of findings in the present study is compatible with existing data on discrimination of cross-language consonant contrasts in natural speech. The results contribute additional support for the assimilation model put forth by Best and are also completely compatible with Flege's speech learning model. A general conclusion drawn from these results is that linguistic experience shapes the discrimination of vowels and consonants as phonetic segmental units in similar ways. This conclusion is a tentative one. Further research is required before it can be firmly concluded that linguistic influences on vowel and consonant perception are best explained within a common conceptual framework. In particular, it will be useful to compare language-specific effects on vowel and consonant perception under different processing demands and to examine cross-language comparisons that isolate the contribution of acoustic factors and assimilation differences to performance in cross-language speech perception tasks.

The present study was focused on cross-language vowel perception in adults. The development of cross-language vowel discrimination during infancy has also been investigated using these same German vowel stimuli (Polka and Werker, 1994). The results of this first study are consistent with earlier studies of cross-language consonant discrimination showing a shift from a language-general toward a language-specific pattern during the first year of life. However, this shift appears to begin earlier in development for vowels than for consonants. Further studies that explore cross-language vowel perception in both infants and adults are currently underway.

ACKNOWLEDGMENTS

This study was conducted while the author was as postdoctoral fellow working in the Infant Studies Laboratory at the University of British Columbia and was supported by a postdoctoral fellowship from the National Institute of Deafness and other Communication Disorders (No. F32 DC00021-01) and a grant from the Natural Science and Engineering Research Council of Canada (No. OGP01205397). Thanks are extended to Jennifer Fraser for assistance in data collection, and to Diane Hobday for assistance in recruiting and scheduling subjects. The author is also grateful to Shari Baum, Catherine Best, Jim Flege, and Janet Werker for their helpful comments on earlier drafts of this manuscript.

[1]Best defines phonetic similarities with reference to criterial articulatory gestures.

[2]Although only four German vowels were of interest in this study, two extra vowels (/ɔ/ and /ø/) were recorded and, for reasons explained below, were presented as extra items in the keyword task.

[3]Amplitude values reported were measured on the digital waveforms after amplitude adjustments had been made.

[4]$F2$ frequency measures corresponding to voicing offset were not observable for three of the six syllables with tense back vowels.

[5]In previous work an A' score of 0.95 or greater was used as a criterion for nativelike performance (Polka, 1991). An A' score of 0.95 is roughly equivalent to 90% correct.

[6]In both stages of the keyword task, odd and hawk keyword responses were collapsed into a single response category because not all dialects of English have different vowels for these two keywords. In stage 1 of the task, confusions between these two vowels were not scored as errors.

[7]Each subject had 12 opportunities to select and rate (from 1 to 5) each German vowel category; thus the resulting similarity score could vary from 0 (if the keyword was never chosen) to 60 (if the keyword was chosen 12 times and always given the highest rating possible).

[8]Analyses were also conducted with the five most frequently chosen keywords as levels of the keyword factor, thus including over 90% of the identification responses to all four vowels. The pattern of significant differences was not altered when this larger array of keywords was included in the analysis. For each contrast, similarity scores for the additional keywords were significantly lower than scores for ooze or hook and did not differ significantly between contrasting front and back German vowels.

[9]In a subsequent study (Polka and Bohn, 1994) English adults were presented a different sample of the same German vowels in the same identification task except that "none" was also provided as a response option. The none response was never chosen. These results further strengthen the conclusion that English adults readily perceive these German vowels as being similar to certain English vowels.

[10]Other vowel studies have demonstrated that both categorical and noncategorical aspects of the internal representation of vowels are influenced by changes in stimulus context (Repp and Crowder, 1990). Thus conclusions drawn here from comparing results across the identification and ABX task research would be stronger if both the labeling and the ABX data had been obtained from presentation of identical stimulus sequences.

[11]It is not known whether assimilation patterns that describe perception of non-native vowel contrasts will include all or only those patterns outlined by Best to characterize perception of non-native consonants contrasts.

Ades, A. E. (**1977**). "Vowels, consonants, speech and non-speech," Psychol. Rev. **84**, 524–530.

Assmann, P. (**1979**). "The role of context in vowel perception," Master's thesis, University of Alberta, Edmonton, Alberta.

Beddor, P. S., and Strange, W. (**1982**). "Cross-language study of the oral–nasal distinction," J. Acoust. Soc. Am. **71**, 1551–1561.

Bennett, D. C. (**1968**). "Spectral form an duration as cues in the recognition of English and German vowels," Lang. Speech **11**, 65–85.

Best, C. T. (**1994**). "The emergence of native-language phonological influences in infants: A perceptual assimilation model," in *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, edited by H. Nusbaum and J. Goodman (MIT, Cambridge, MA).

Best, C. T. (**1993**). "Emergence of language-specific constraints in perception of non-native speech perception: A window on early phonological development," in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, edited by B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton (Kluwer, Dordrecht).

Best, C. T., McRoberts, G. W., and Sithole, N. N. (**1988**). "Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants," J. Exp. Psychol. Hum. Percept. Perform. **14**, 345–60.

Best, C. T., and Strange, W. (**1992**). "Effects of phonological and phonetic factors on cross-language perception of approximants," J. Phon. **20**, 305–330.

Bohn, O.-S., and Flege, J. E. (**1990**). "Interlingual identification and the role of foreign language experience in L2 vowel perception," Appl. Psycholinguist **11**, 303–328.

Borden, G. J., and Harris, K. S. (**1980**). *Speech Science Primer* (Williams and Wilkins, Baltimore).

Carney, A. E., Widin, G. P., and Viemiester, N. F. (**1977**). "Non-categorical perception of stop consonants differing in VOT," J. Acoust. Soc. Am. **62**, 961–970.

Flege, J. E. (**in press**). "Second-language speech learning: Theory, findings, and problems," In *Speech Perception and Linguistic Experience: Issues in Cross-Language Speech Research*, edited by W. Strange (York, Timonium, MD).

Flege, J. E., and Hillenbrand, J. (**1986**). "Differential use of temporal cues to the /s/–/z/ contrast by native and non-native speakers of English," J. Acoust. Soc. Am. **79**, 508–517.

Flege, J. E., and Bohn, O.-S. (**1989**). "The perception of English vowels by native speakers of Spanish," J. Acoust. Soc. Am. Suppl. 1 **85**, S85.

Flege, J. E., Munro, M. J., and Fox, R. A. (**1994**). "Auditory and categorical effects on cross-language vowel perception," J. Acoust. Soc. Am. **95**, 3623–3641.

Fujisaki, H., and Kawashima, T. (**1970**). "Some experiments on speech perception and a model for the perceptual mechanism," Ann. Rep. Eng. Res. Inst. Univ. Tokyo, **29**, 207–214.

Goto, H. (**1971**). "Auditory perception by normal Japanese adults of the sounds "*L*" and "*R*", Neuropsychologia **9**, 317–323.

Gottfried, T. L. (**1984**). "Effects of consonant context on the perception of French vowels," J. Phon. **2**, 91–114.

Gottfried, T. L., and Beddor, P. S. (**1988**). "Perception of temporal and spectral information in French vowels," Lang. Speech **31**, 57–75.

Grieser, D., and Kuhl, P. K. (**1989**). "Categorization of speech by infants: Support for speech-sound prototypes," Dev. Psychol. **25**, 577–588.

Jamieson, D., and Morosan, D. (**1986**). "Training non-native speech contrasts in adults: Acquisition of the English /θ/–/ð/ contrast by francophones," Percept. Psychophys. **40**, 205–215.

Kuhl, P. J. (**1991**). "Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories; monkeys do not," Percept. Psychophys. **50**, 93–107.

Kuhl, P. J. (**1992**). "Psychoacoustics and speech perception: Internal standards, perceptual anchors, and prototypes," in *Developmental Psychoacoustics*, edited by L. A. Werner and E. W. Rubel (American Psychological Association, Washington, D.C.).

Kuhl, P. J., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (**1992**). "Linguistic experience alters phonetic perception in infants by 6 months of age," Science **255**, 606–608.

Ladefoged, P. (**1982**). *A Course in Phonetics* (Harcourt Brace Jovanovich, New York).

Logan, J. S., Livley, S. E., and Pisoni, D. B. (**1991**). "Training Japanese listeners to identify English /r/ and /l/: A first report," J. Acoust. Soc. Am. **89**, 874–886.

MacKain, K. S., Best, C. T., and Strange, W. (**1981**). "Categorical perception of English /r/ and /l/ by Japanese bilinguals," Appl. Psycholinguist **2**, 369–390.

Mack, M. (**1989**). "Consonant and vowel perception and production: Early English–French bilinguals and English monolinguals," Percept. Psychophys. **46**, 187–200.

Mertus, J. (**1990**). BLISS User's Manual (Brown University, Providence, RI).

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and Fujimura, O. (**1975**). "An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English," Percept. Psychophys. **18**, 331–340.

Mochizuki, M. (**1981**). "The identification of /r/ and /l/ in natural and synthesized speech," J. Phon. **9**, 283–303.

Pisoni, D. B. (**1973**). "Auditory and phonetic codes in the discrimination of consonants and vowels," Percept. Psychophys. **13**, 253–260.

Pisoni, D. B., Aslin, R. N., Perey, A. J., and Hennessy, B. L. (**1982**). "Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants," J. Exp. Psychol. Hum. Percept. Perform. **8**, 297–314.

Pisoni, D. B., Logan, J. S., and Lively, S. E. (**1994**). "Perceptual learning of nonnative speech contrasts: Implications for theories of speech perception," in *Development of Speech Perception: The Transition From Recognizing Speech Sounds to Spoken Words,* edited by H. C. Nusbaum and J. Goodman (MIT, Cambridge).

Polka, L. (**1991**). "Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions," J. Acoust. Soc. Am. **89**, 2961–2977.

Polka, L. (**1992**). "Characterizing the influence of native experience on adult speech perception," Percept. Psychophys. **52**, 37–52.

Polka, L., and Werker, J. F. (**1994**). "Developmental changes in perception of non-native vowel contrasts," J. Exp. Psychol. Hum. Percept. Perform. **20**, 421–435.

Polka, L., and Bohn, O.-S. (**1994**). "A cross-language comparison of vowel perception in English-learning and German-learning infants," paper presented at the International Conference on Infant Studies, Paris (unpublished).

Pruitt, J. S., Strange, W., Polka, L., and Aguilar, M. (**1990**). "Effects of category knowledge and syllable truncation during auditory training on American's discrimination of Hindi retroflex-dental contrasts," J. Acoust. Soc. Am. Suppl. 1 **87**, CC8.

Repp, B. H. (**1984**). "Categorical perception: Issues, methods, and findings," in *Speech and Languages: Advances in Basic Research and Practice*, edited by N. L. Lass (Academic, New York), Vol. 1.

Repp, B. H., and Crowder, R. G. (**1990**). "Stimulus order effects in vowel discrimination," J. Acoust. Soc. Am. **88**, 2080–2090.

Sawusch, J. R., Nusbaum, H. C., and Schwab, E. C. (**1980**). "Contextual effects in vowel perception II: Evidence for two processing mechanisms," Percept. Psychophys. **27**, 421–434.

Stevens, K. N., Liberman, A. M., Studdert-Kennedy, M., and Ohman, S. E. G. (**1969**). "Crosslanguage study of vowel perception," Lang. Speech **12**, 1–23.

Strange, W., and Gottfried, T. L. (**1980**). "Task variables in the study of vowel perception," J. Acoust. Soc. Am. **68**, 1622–1625.

Strange, W., and Jenkins, J. (**1978**). "Role of linguistic experience in the perception of speech," in *Perception and Experience*, edited by R. D. Walk and H. L. Pick (Plenum, New York).

Studdert-Kennedy, M. (**1993**). "Discovering phonetic function," J. Phon. **21**, 147–155.

Werker, J. F. (**1994**). "Cross-language speech perception: Developmental change does not involve loss," in *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, edited by H. Nusbaum and J. Goodman (MIT, Cambridge, MA).

Werker, J. F. (**1991**). "The ontogeny of speech perception," in *Modularity and the Motor Theory of Speech Perception*, edited by I. G. Mattingly and M. Studdert-Kennedy (Erlbaum, Hillsdale, NJ).

Werker, J. F., and Tees, R. C. (**1984**). "Phonemic and phonetic factors in adult cross-language speech perception," J. Acoust. Soc. Am. **75**, 1866–1878.

Werker, J. F., and Logan, J. S. (**1985**). "Cross-language evidence for three factors in speech perception," Percept. Psychophys. **37**, 35–44.