Chinese subjects' perception of the word-final English /t/–/d/ contrast: Performance before and after training

James Emil Flege
Department of Biocommunication, University of Alabama at Birmingham, University Station, Birmingham, Alabama 35294

(Received 8 August 1988; accepted for publication 20 July 1989)

Chinese words may begin with /t/ and /d/, but a /t/–/d/ contrast does not exist in word-final position. The question addressed by experiment 1 was whether Chinese speakers of English could identify the final stop in words like *beat* and *bead*. The Chinese subjects examined approached the near-perfect identification rates of native English adults and children for words that were unedited, but performed poorly for words from which final release bursts had been removed. Removing closure voicing had a small effect on the Chinese but not the English listeners' sensitivity. A regression analysis indicated that the Chinese subjects' native language (Mandarin, Taiwanese, Shanghainese) and their scores on an English comprehension test accounted for a significant amount of variance in sensitivity to the (burstless) /t/–/d/ contrast. In experiment 2, a small amount of feedback training administered to Chinese subjects led to a small, nonsignificant increase in sensitivity to the English /t/–/d/ contrast. In experiment 3, more training trials were presented for a smaller number of words. A slightly larger and significant effect of training was obtained. The Chinese subjects who were native speakers of a language that permits obstruents in word-final position seemed to benefit more from the training than those whose native language (L1) has no word-final obstruents. This was interpreted to mean that syllable-processing strategies established during L1 acquisition may influence later L2 learning.

PACS numbers: 43.71.Es, 43.71.Hw

INTRODUCTION

Much second language (L2) acquisition research has been motivated by the desire to learn whether a biologically (rather than cognitively) based "critical period" exists for human speech learning. Lenneberg (1967, p. 176) claimed that "foreign accents cannot be overcome easily after puberty." Rating-scale experiments have amply comfirmed the presence of foreign accents in the speech of L2 learners, even those who began learning L2 long before the onset of puberty (Asher and Garcia, 1969; Fathman, 1975; Flege and Eefting, 1987; Flege, 1988a,b; see also Seliger *et al.*, 1982).[1] The mispronunciation of particular phones, which contributes to foreign accent (see Johansson, 1978; Flege, 1984, 1988a; Schneiderman *et al.*, 1988), may derive in part from insufficient motor learning. This, in turn, may be impeded by an inability to perceive L2 phones or phonetic contrasts in a nativelike manner. Difficulty distinguishing phonemes in an L2 may also impede lexical development (see Leonard *et al.*, 1982). The purpose of the present study was to provide further insight into how phones in an L2 are perceived by examining the perception of the voicing feature in word-final English obstruents.

Prelinguistic infants are able to discriminate many, if not all, of the phones used contrastively in human languages (see, e.g., Jusczyk, 1979; Kuhl, 1987), but this ability may be attenuated near the end of the first year of life for certain phonetic contrasts. Werker and her colleagues found that young children lost the ability to discriminate phones not used contrastively in the language spoken around them, even though they were able to do so earlier as infants (Werker *et al.*, 1981; Werker and Tees, 1983, 1984a,b). An apparent consequence of this lost ability is the difficulty adults may have in comprehending an L2 later in life (Johansson, 1978; van Balen, 1980; Oyama, 1982b; Florentine, 1985; Koster, 1987; Mack and Tierney, 1987). Such comprehension difficulty may derive, at least in part, from difficulty perceiving phones in the L2.

Echoing Trubetzkoy's (1939) observation that the phonology of L1 acts like a "sieve" through which L2 sounds are processed, Borden *et al.* (1983) suggested that adult L2 learners are "constrained" to categorize the phones of an L2 "according to the phoneme contrasts" of their native language (p. 500). This might explain why, for example, Japanese speakers have difficulty in discriminating English /r/'s and /l/'s (Miyawaki *et al.*, 1975). It does not explain other experimental evidence so well, however. Flege and Hillenbrand (1985) found that Swedes and Finns, whose L1 has no /z/, were able to differentially identify the members of a continuum ranging from *peace* (/pis/) to *peas* (piz). They used only vowel duration, however, whereas native speakers of English used a combination of vowel and fricative duration cues.

The present study examined the identification of final /t/ and /d/ by Chinese subjects. The L1's of the subjects examined possess a contrast between phonologically voiced and voiceless stops, but only in word-initial position (Cheng, 1973; Howie, 1976; Li and Thompson, 1981; Shinn, 1985; Heyer, 1986). The word-initial phonemic contrast in Chinese is implemented as a phonetic contrast between voiceless unaspirated stops with short-lag voice onset time (VOT) values and voiceless aspirated stops with long-lag VOT values (Lisker and Abramson, 1964; Clumeck *et al.*, 1981).

Little is known about how Chinese subjects perceive final stops *in English, but much is known about how they* *produce* word-final stops. As predicted by the contrastive analysis hypotheses (Lado, 1957), Chinese learners of English are often unable to produce a perceptually effective contrast between voiced and voiceless stops in word-final position. Transcriptional data from several studies have shown that Chinese speakers devoice English /b,d,g/, delete voiced and voiceless stops, and/or add vowels following word-final stops (Tarone, 1980; Eckman, 1981; Anderson, 1983, 1987; Flege and Davidian, 1985; Heyer, 1986; Weinberger, 1987; Flege *et al.*, 1987). Detailed analyses of the production of /p/ and /b/ in word-final position have shown that Chinese subjects produce a much smaller (albeit significant) duration difference between vowels preceding /p/ vs /b/, are significantly less likely than native speakers to actively enlarge the oral cavity to sustain voicing in /b/, and do not produce /p/ with a greater force of bilabial constriction than /b/ (Flege *et al.*, 1987; Flege, 1988c).

These findings show that the ability to produce voiced and voiceless stops in word-initial position does not guarantee the ability to produce an effective voiced–voiceless contrast in the word-*final* position. This suggests that speech production is not based on freely commutable phonemes (see Flege and Port, 1981) and that speech production skills must be learned on an allophone-by-allophone basis (see also Brière, 1966). This inference is consistent with the observation that sounds may be more vulnerable to disruption in initial than final position following injury to the brain (Kent and Rosenbek 1983; Kelso and Tuller, 1981). It is also consistent with the conclusion, based on performance errors in normal individuals, that word-initial consonants are processed separately from subsequent consonants (Shattuck-Hufnagel, 1987).

The question addressed by experiment 1 was whether speech *perception* skills also need to be learned on an allophone-by-allophone basis. To determine this, native and Chinese subjects were asked to identify the final stop in unedited multiple natural tokens of English words such as *beat* and *bead*, and in copies from which the final release bursts and/or closure voicing had been removed. In conversational speech, English /t/ and /d/ are frequently produced without audible release bursts, and /d/ is "devoiced" (that is, produced without closure voicing through some part or all of the closure interval; see, e.g., Smith, 1979). Perhaps because of the unreliability of these cues, native speakers of English can identify stops without closure voicing and/or final release bursts nearly as well as stops with all acoustic cues available (see below).

Two alternative possibilities existed concerning how the Chinese subjects would perform in the identification task. Even though the acoustic cues to the /t/–/d/ contrast differ in the initial versus final position of English words (Strange, 1986), the Chinese subjects might attempt to identify the final /t/'s and /d/'s using the perceptual cues that distinguish /t/ from /d/ in the initial position of Chinese words.[2] Such a strategy would probably work well for the unedited words in which the final release bursts were intact, because they contained some of the information that cues the voiced–voiceless contrast in word-initial position. Such a strategy could be expected to work less well for stops from which release bursts had been removed, however. Alternatively, the Chinese subjects might regard word-final tokens of /t/ and /d/ as representing a "novel" phonetic contrast, and begin seeking cues other than those used for word-initial stops, such as preceding vowel duration. If so, the Chinese subjects might perform more poorly than the native English subjects for the tokens with release bursts, but the effect of removing release burst (and closure voicing) cues should be no greater for them than for native English subjects.

The results of experiment 1 supported the first alternative, so two feedback training experiments were carried out to better enable the Chinese subjects to perceive those cues to the /t/–/d/ contrast that remained after the closure voicing and release burst cues had been removed. The first training experiment involved subjects who had participated in experiment 1. A greater amount of feedback training was presented in the second training experiment to a new group of Chinese subjects. A "neural atrophy" hypothesis (see Werker and Tees, 1984b) implies that it would be difficult to increase Chinese subjects' sensitivity to the /t/–/d/ contrast through short-term laboratory training. The notion that auditory sensitivity is never really lost, but is simply "difficult to apply" to novel phonetic contrasts as the result of a "language based reorganization" of contrastive sound units (Werker, 1989) suggests a different outcome, however. Training might be effective if it succeeds in enabling subjects to pay attention to relevant acoustic cues (Strange, 1986), or if it helps subjects engage a phonetic, as opposed to a phonemic, mode of processing (Werker and Tees, 1984b).

I. EXPERIMENT 1

A. Introduction

The closure intervals of English /b,d,g/ and /p,t,k/ are distinguished by the presence versus absence of voicing during oral constriction in isolated, clearly spoken words, but as a rule voicing does not extend through the entire constriction of /b,d/g/. As discussed later (see also the Appendix), the removal of closure voicing may or may not lead to an increase in voiceless stop responses by native English-speaking listeners. English word-final stops are often not released audibly (Wardrip-Fruin, 1982), so it is hardly surprising that removing release bursts may have little perceptual effect.[3]

Previous research has shown that native speakers of English can identify the voicing feature of word-final tokens of /b,d,g/ and /p,t,k/ fairly well even when both closure voicing and release bursts have been removed. They do so on the basis of temporal and spectral cues residing in the preceding vowel and in the final consonant transitions (Halle *et al.*, 1957; Ohde and Sharf, 1977; Revoile *et al.*, 1982; Hillenbrand *et al.*, 1984). For example, Wolf (1978) found that /t/'s and /d/'s were identified correctly nearly as often in the final position of words from which release burst and closure voicing cues had been removed as in unedited words (89% vs 100% correct).[4] The aim of experiment 1 was to determine how well native and Chinese subjects could identify /t/'s and /d/'s in the final position of English words that were unedited and in copies from which closure voicing and/or release burst cues had been removed.

A secondary aim was to examine the effect of removing acoustic cues for native English-speaking children. Speech production research has shown that Chinese adults are less skillful in producing a contrast between word-final /p/ and /b/ than native English children who, in turn, are less skillful in doing so than native English adults (Flege *et al.*, 1987; Flege, 1988c). Parnell and Amerman (1978) found that 4 year olds identified the place of articulation of consonants in CV syllables less often than 11 year olds or adults when the steady-state portion of the vowel was removed (76% vs 98% for the two older groups). It appears that both children (Cole and Perfetti, 1980) and nonnative adults may require longer stretches of speech than mature native speakers to recognize words (Truin, 1981, reported by Koster, 1987).

If speech perception skills undergo a gradual attunement to the multiple coarticulated cues used to signal phonetic contrasts (Parnell and Amerman, 1978; Eguchi and Hirsch, 1969; Flege and Eefting, 1986), then removing closure voicing and/or release burst cues may affect native English children and the Chinese adults similarly. Alternatively, the children might resemble native English adults in showing little effect of editing. This could be expected if the children based their voicing judgments on the vowel duration cues that distinguished the edited words ending in /t/ and /d/ (see Krause, 1982a,b).

B. Methods

1. Subjects

Three groups of listeners participated as paid subjects, all of whom were able to detect four pure tones in the 500–4000-kHz range at 20 dB HL. The native English adult subjects ($N = 8$) and the nine English children ($N = 9$) were monolingual speakers of English from Birmingham, Alabama, who were recruited through an advertisement in the University newspaper. The adults were in their late twenties and early thirties; the mean age of the children was 9.2 years.

Most of the Chinese subjects (19 males, 11 females) were graduate students or faculty members at the University of Alabama at Birmingham who spoke English with foreign accents (see Flege, 1988b). As summarized in Table I, the Chinese subjects had a mean age of about 30 years and had lived in the U.S. for a little more than 2 years on average. The Chinese subjects reported that they spoke English about half the time, but some of them seemed to have difficulty comprehending English. This suggested that their estimated daily use of English might have been too high (see Gras, 1983). To test this, each subject was administered the listening comprehension subtest of the Michigan English Language Assessment Battery (English Language Institute, University of Michigan, 1986). The Chinese subjects received a 29th percentile score, which confirmed the inference concerning relatively limited native-speaker input.[5]

For the purpose of regression analysis (see below), the Chinese subjects were subdivided into three subgroups based on native language. Of the 30 subjects, some ($N = 16$) were native speakers of Mandarin from mainland China, others ($N = 5$) were native speakers of Taiwanese, and the remaining subjects were individuals from southern China who spoke either Shanghainese ($N = 8$) or Hakka ($N = 1$). None of the ANOVAs testing for differences between the subgroups in chronological age, age of arrival in the U.S., length of residence in the U.S., percentage daily use of English, or comprehension ability reached significance at the 0.10 level.

2. Stimuli

An adult native speaker of American English (the author) read real words formed by inserting seven vowels (/i,ɪ,e$^{\text{ɪ}}$,ɛ,æ,ɑ,u/) into /b–t/ and /b–d/ frames in the carrier phrase "Now I will say —." The first four tokens of each

TABLE I. Characteristics of the Chinese subjects in three native-language subgroups who participated in experiment 1. POB is place of birth, AOA is age of arrival in the U.S., MOS indicates months lived in the U.S., COMP is the score on an English comprehension test, and USE is self-estimated percentage daily use of English.

POB	Sex	AOA	MOS	COMP	USE
	Mandarin				
Beijing	M	36	19	74	80
Beijing	M	35	8	56	20
Beijing	M	32	12	85	50
Beijing	F	17	84	91	···
Xian	M	30	7	72	50
Xian	F	28	28	76	60
Tientsin	M	29	5	72	60
Lanzou	M	26	23	70	20
Harbin	F	27	16	87	80
Changchun	M	35	4	83	80
Shenyang	M	32	8	56	35
Chengdu	M	27	1	83	60
Nanchang	M	26	1	76	20
Taipei, ROC	F	25	132	91	85
Taoyuan, ROC	M	28	5	78	20
Kaohsiung, ROC	F	22	36	83	80
	M	28	24	77	49
		(5)	(35)	(11)	(25)
	Taiwanese				
Taipei, ROC	F	29	32	74	60
Taipei, ROC	F	24	3	79	40
Taichung, ROC	M	26	4	70	20
Taichung, ROC	F	31	27	46	20
Kaoshung, ROC	M	27	10	81	20
	M	27	15	70	32
		(3)	(13)	(14)	(18)
	Shanghainese				
Shanghai	M	27	18	89	40
Shanghai	F	25	2	53	20
Shanghai	M	24	72	87	95
Shanghai	M	35	47	79	20
Shanghai	M	27	52	70	60
Shanghai	M	29	4	72	50
Shanghai	M	42	27	78	20
Shanghai	F	25	36	91	50
Pingtong, ROC	F	25	25	78	50
	M	29	31	77	45
		(6)	(23)	(12)	(24)

word were low-pass filtered at 4 kHz, digitized at 10 kHz with 12-bit resolution, then normalized for overall rms amplitude. The initial /b/'s of the 56 test words were all prevoiced. Each final stop had an audible release burst, and the final /d/'s all had at least some voicing. As shown in Table II, the vowels averaged 133 ms (59%) longer before /d/ than /t/.

Copies of the "unedited" words were modified as follows. A set of "voicing-removed" words was made by attenuating to zero any energy present in the closure interval of the final stops. Stop closure duration information was preserved in these stimuli.[6] "Burst-removed" words were made by removing all energy in the release bursts. Both closure voicing and release bursts were removed from the "V + B-removed" words by zeroing any energy present after the constriction of the final stops.

3. Procedures

The words were presented binaurally via TDH-49 headphones at a comfortable level (74-dB peak syllable intensity). The subjects were told to push a button marked "t" or "d," and to guess if uncertain. Each new word was presented 1.0 s after a response was received for the preceding word. Three randomizations of the 56 words (seven vowels × two final stops × four replicate tokens) were presented in the four editing conditions (unedited, voicing-removed, burst removed, V + B-removed). The order of the four conditions was only partially counterbalanced. Subjects always heard the unedited words first and the V + B-removed words last. This ordering was used because pilot experiments suggested that the unedited words were relatively easy for the Chinese subjects to identify, whereas the V + B-removed words were difficult. The voicing-removed and burst-removed words were presented in counterbalanced order as the second and third conditions, which afforded the opportunity to determine the relative importance of closure voicing and release burst cues.

C. Results

1. Effects of editing

The mean percentage of correct identifications of the final stop in the 14 test words was calculated for each listener.[7] As shown in Table III, the overall rates for /d/ and /t/ differed little (85% vs 86% correct). The English adults identified stops in the burst- and V + B-removed words correctly at nearly the same rates (99%, 98%) as the unedited and voicing-removed words (100%, 99%). Similarly, the native English children identified the burst- and V + B-removed words correctly at about the same rates (96%, 96%) as the unedited and voicing-removed words (97% and 98%). The Chinese adults, on the other hand, identified the burst- and V + B-removed words far less well (64% correct in both instances) than the unedited and voicing-removed words (94%, 93%).

An unbiased estimate of perceptual sensitivity to the /t/–/d/ contrast was determined by calculating A' scores for each minimal pair (Grier, 1971). The A' scores were submitted to a group (English adults, English children, Chinese adults) × vowel (seven levels) × editing condition (four levels) ANOVA with repeated measures on the last two factors.[8] As shown in Fig. 1, the A' scores were only slightly lower in the unedited and voicing-removed conditions for the Chinese adults than for the English adults or children. However, in the burst- and V + B-removed conditions, the Chinese subjects' scores were substantially lower than the native speakers', which resulted in a significant group × condition interaction [$F(6,132) = 19.4, p < 0.05$].

The simple main effect of editing condition was nonsignificant for the English adults and also for the English children [$F(3,21) = 0.96$, $F(3,24) = 2.77$] but highly significant for the Chinese adults [$F(3,87) = 75.5, p < 0.05$]. *Post hoc* tests (Newman–Keuls, $\alpha = 0.05$) revealed that the Chinese subjects' A' scores were lower in the burst- and V + B-removed conditions than in the unedited and voicing-removed conditions. The simple main effect of group was significant for stops in the voicing-, burst-, and V + B-removed conditions [$F(2,44) = 4.73, 28.0, 21.9, p < 0.05$] and narrowly missed reaching significance in the unedited condition [$F(2,44) = 3.16$, $p = 0.052$]. *Post hoc* tests revealed that, for stops in the voicing-removed condition, the English adults' but not the English children's scores were significantly greater than the Chinese adults' scores. In the burst- and V + B-removed conditions, both the English adults' and children's A' scores were significantly higher than the Chinese adults' scores. The English adults and children did not differ significantly in any condition.[9]

2. Predictors of sensitivity

Some Chinese subjects succeeded better than others in identifying stops without release bursts. For example, the correct identification rate for stops in the burst-removed condition ranged from 35%–96%. Two forward stepwise multiple regression analyses were carried out to help account for intersubject variability among the 30 Chinese subjects.

a. Methods. The criterion variable in the regression analysis was sensitivity to the /t/–/d/ contrast in the burst- and V + B-removed conditions, as measured by d' (rather than A') scores.[10] The results of previous research suggested five of the six predictor variables examined. It is known that ability to comprehend L2 sentences may be correlated positively with the age of learning (Oyama, 1982b), length of residence in an L2-speaking country (Gras, 1983; Underbakke

TABLE II. The mean duration of vowels in the English /bVt/ and /bVd/ words used in experiment 1, in ms. Each mean is based on four tokens; standard deviations are in parentheses.

Word final /t/		Word final /d/	
beat	198 (12)	bead	335 (28)
bit	178 (19)	bid	304 (10)
bait	236 (8)	bade	364 (13)
bet	201 (8)	bed	327 (15)
bat	264 (21)	bad	410 (9)
boot	220 (13)	booed	353 (15)
bought	261 (15)	baud	396 (9)
M	223	*M*	356

TABLE III. Mean percent correct identifications of word-final stops by native English adults (EA), native English children (EC), and Chinese adults (CA) in English words that were unedited (UNED), had closure voicing removed (RVOI), had final release bursts removed (RBUR), or had both voicing and bursts removed (RV + B). The means for group EA are based on eight subjects × seven minimal pairs = 56 scores. Means for group EC are based on data for 9 subjects, those for CA on 30 subjects. Standard deviations are in parentheses.

	/d/				/t/				
	UNED	RVOI	RBUR	RV + B	UNED	RVOI	RBUR	RV + B	*M*
EA	99.6 (1.9)	98.4 (5.1)	99.1 (3.0)	97.2 (8.0)	99.6 (1.9)	99.6 (1.9)	99.6 (1.9)	99.4 (2.2)	99.0 (3.9)
EC	98.0 (3.9)	97.8 (4.8)	96.5 (5.6)	95.6 (7.8)	96.8 (6.8)	97.9 (4.7)	95.4 (7.4)	96.8 (4.8)	96.9 (5.9)
CA	91.5 (15.6)	88.1 (17.0)	70.0 (26.6)	69.0 (28.7)	96.5 (7.9)	97.9 (5.8)	58.5 (27.8)	58.7 (29.8)	78.5 (26.7)
M	94.1 (13.0)	91.7 (14.7)	80.3 (25.2)	78.9 (26.8)	97.1 (7.1)	98.2 (5.2)	72.5 (29.3)	72.9 (30.5)	

et al., 1988), overall experience with and proficiency in the L2 (van Balen, 1980; Mack and Tierney, 1987), and the extent to which English is used both actively and passively on a daily basis (Gras, 1983; Underbakke *et al.*, 1988; see also MacKain *et al.*, 1981). Since sentence comprehension may be related to the ability to perceive specific phonetic contrasts, the following predictor variables were examined: chronological age, age of arrival in the U.S., length of residence in the U.S., percentage of daily use of English, and scores on the English comprehension test. L1 background served as the sixth predictor variable because preliminary analyses of the correct identification scores revealed significant differences between the subjects who spoke Shanghainese, Taiwanese, and Mandarin (see Table I).[11]

b. Results. Analysis of data from the burst-removed condition indicated that L1 background and ability to comprehend English were independent predictors of sensitivity to the English /t/–/d/ contrast. A model wtih these two factors accounted for 31% of variance [$F(2,26) = 5.79$, $p < 0.05$]. The d' scores correlated significantly with the comprehension test scores, even when the L1 factor was partialed out [$r(28) = 0.441, p < 0.05$]. The d' scores correlated significantly with the L1 factor, even when the comprehension scores were partialed out [$r(28) = 0.420, p < 0.05$].

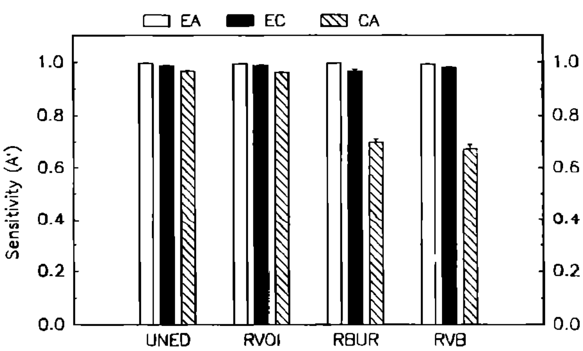

FIG. 1. Mean sensitivity (A') to the word-final English /t/–/d/ contrast by native English adults (EA), native English children (EC), and Chinese adults (CA) in experiment 1. The stimuli were unedited (UNED), had closure voicing removed (RVOI), the final release bursts removed (RBUR), or had both closure energy and release bursts removed (RV + B). The brackets enclose ± 1 standard error.

In the analyses of stops in the V + B-removed condition, L1 background and the English comprehension scores were again identified as significant predictors of the d' scores [$F(2,26) = 3.66$, $p < 0.05$], but the two-factor model accounted for only 22% of the variance in d' scores. The simple correlation between the d' scores and the English comprehension scores was significant [$r(28) = 0.372$, $p < 0.05$], but not the correlation between the d' scores and L1 background [$r(28) = 0.287$]. The correlation between the d' scores and the comprehension scores remained significant when variation due to L1 was partialed out [$r(28) = 0.393$, $p < 0.05$]. All subjects subjects participated in the burst-removed condition before the V + B-removed condition, which may have leveled differences between subjects in the three Chinese subgroups in the V + B-removed condition.

D. Discussion

1. Perceptual sensitivity

As expected, native English adults identified tokens of final /t/ and /d/ (from which closure voicing and/or final released burst cues were edited out) at high rates. Taken together with the result of previous research, this indicated that acoustic information in the vowel and consonant transitions of English CVC words, including the spectral quality and duration of the preceding vowel and F_1 offset frequency, are sufficient to cue accurate voicing judgments.

The native English children who were examined performed much like the English adults, but this does not mean necessarily that their perception was identical to that of adults (see Flege and Eefting, 1986, for word-initial stops). Bernstein Ratner and Luberoff (1984) reported that mothers speaking to their children were perceived to have deleted 40% of /t/'s and /d/'s (see also Shockey and Bond, 1980). These same mothers exaggerated vowel duration cues, so it is hardly surprising that other research (e.g., Krause, 1982a,b) has shown that vowel duration may provide a sufficient perceptual cue to the distinction between voiced and voiceless stops for young children. Thus any adult–child difference that did exist may have been obscured by a ceiling effect brought about by the use of vowel duration cues. The children may also have used F_1 offset frequency as a perceptual cue (see Simon and Fourcin, 1978, for word-initial stops).

One might infer that the Chinese subjects were unable to make the same use of vowel duration as the native English children because they performed more poorly than the children when release burst cues were removed. This inference is consistent with the observation that Chinese speakers of English produce only a relatively small (albeit significant) vowel duration difference between English words ending in /p/ and /b/ (Flege, 1988c).

The removal of closure voicing had little effect on the native English subjects, whereas it significantly lessened the Chinese subjects' sensitivity to the English /t/–/d/ contrast (it lowered the correct identification scores from 95% in the unedited condition to 93% in the voicing-removed condition). This does not mean necessarily that the Chinese subjects were more sensitive to closure voicing than the English adults. The lack of a decrease in sensitivity for the English adults was probably the result of a ceiling effect resulting from the use of other cues such as vowel duration. A short experiment presented in the Appendix showed that, in a paired-comparison task, native English adults significantly preferred final /d/ tokens with closure voicing over those without closure voicing.

Removing release burst had a much greater effect on the Chinese subjects than removing closure voicing. Their rate of correct identifications plummeted to 64% correct in the two conditions in which the release bursts were removed. This suggests that the Chinese subjects were identifying the word-final /t/'s and /d/'s on the basis of acoustic cues that distinguish stops in word-initial position, where a contrast between a voiceless unaspirated /d/ and a voiceless aspirated /t/ exists in the subject's L1's (Lisker and Abramson, 1964; Clumeck *et al.*, 1981; Shinn, 1985).

An alternative explanation is that the effect of altering cues might simply be greater for listeners whose speech perception skills are not mature than for mature native speakers. Previous research has shown that nonnative speakers recognize words more poorly than native speakers, especially in non-ideal listening conditions (Johansson, 1978; van Balen, 1980; Oyama, 1982b; Florentine, 1985; Koster, 1987; Mack and Tierney, 1987). Like children, nonnatives may require longer stretches of speech than mature native speakers to recognize words (Truin, 1981, reported by Koster, 1987). This explanation is probably not correct, however, because removing the closure voicing and release burst cues did not significantly lower the native English children's sensitivity to the /t/–/d/ contrast.

2. Factors predicting sensitivity

The multiple regression analyses indicated that L1 background and L2 comprehension ability were significant predictors of sensitivity to the English /t/–/d/ contrast. The native speakers of Shanghainese were somewhat more sensitive to the /t/–/d/ contrast than the native Mandarin subjects; and the better the Chinese comprehended English, the more sensitive they were to the /t/–/d/ contrast.

The basis for the L1 effect is uncertain. The Shanghainese subjects may have been relatively good at identifying word-final English stops because their L1 permits an obstruent (a glottal stop) in word-final position. One might speculate that this caused them to focus attention on phonetic information found at the end of syllables to a greater extent than the Mandarin subjects, whose L1 permits no final obstruents. However, there is indirect evidence that Mandarin speakers of English may pay greater attention to release bursts in word-final English stops than Shanghainese subjects. In examining the English speech production of Chinese subjects, Heyer (1986) found that Mandarin subjects "hyperaspirated" English /p,t,k/ more frequently than native speakers of southern Chinese dialects (46% vs 13% of errors observed).[12]

The relationship between the comprehension test scores and sensitivity to the /t/–/d/ contrast suggests, not surprisingly, that the ability to identify phonetic segments is important to comprehension. Variations in the comprehension scores may have been related to other factors such as amount of English-language input, but neither of the questionnaire variables relating directly to amount of English-language experience (length of residence in the U.S., percent daily use of English) was a significant predictor of perceptual sensitivity to English stops.

It must be recognized, of course, that comprehending an L2 depends on more than just bottom-up segmental processing (van Balen, 1980; Koster, 1987). Moreover, some of the intersubject variability not accounted for by the regression analysis may have derived from differences in overall listening skill, which was not assessed. A recent study examining word recognition revealed a great deal of variability in subjects' ability to perceive speech in noise and to exploit contextual constraints (Boothroyd and Nittrouer, 1988). In a study examining Dutch high school students, van Balen (1980) noted that students with high scores on an English listening proficiency test also tended to receive relatively high scores on a comparable Dutch listening test. More research will be needed to establish the basis for individual perceptual differences in both L1 and L2 speech processing.

II. EXPERIMENT 2

The near-perfect performance of the native English subjects in experiment 1 showed that sufficient acoustic cues remained in the V + B-removed stimuli to support accurate voicing judgments. The aim of this training experiment was to enable the Chinese subjects to make better use of acoustic cues in the V + B-removed stimuli. The reason for focusing on the V + B-removed stimuli is that phonologically voiced stops such as /d/ are often produced partially or completely without closure voicing in conversational speech, and both /d/ and /t/ often do not have an audible release burst (e.g., Smith, 1979; Bernstein Ratner and Luberoff, 1984).

It was uncertain whether the feedback training would improve sensitivity to the English /t/–/d/ contrast. If phonetic contrasts are not learned (or maintained) in early childhood, they may be difficult to learn (or train) later in life. There is evidence, however, that the perception of at least some nonnative phonetic contrasts can be improved through appropriate stimulation. MacKain *et al.* (1981) found that relatively inexperienced Japanese subjects performed at near-chance levels in identifying and discriminating synthetic /r/'s and /l/'s, whereas more experienced Japanese subjects (28 months residence in the U.S., 55% daily use of English) identified and discriminated /r/'s and /l/'s in a nativelike fashion.

The success of a training experiment, or the rapidity of learning in naturalistic acquisition, may depend to some extent on the acoustic nature of the L2 contrast being learned/trained. Certain nonnative contrasts, such as those between Zulu clicks, may require little or no training because they are salient for auditory reasons alone (Best *et al.*, 1988; see also Burnham, 1986; Polka, 1987). Others may prove to be more difficult. Werker and her colleagues (Werker and Tees, 1984a,b; Werker and Logan, 1985) found that English adults could not discriminate a relatively rare place of articulation contrast between uvular and velar stops in a category-change task, although they could easily discriminate the release burst portions of such stops, and could discriminate full stops, when a more sensitive testing procedure was used. In other research, Werker showed that English adults learned to discriminate a Hindi contrast between voiceless aspirated and breathy voiced dental stops more readily than a Hindi place of articulation contrast between dental and retroflex stops (Werker *et al.*, 1981; Werker and Tees, 1983; Tees and Werker, 1984).

This last finding suggests the possibility that voicing contrasts are relatively "robust" because of a "strong psychoacoustic basis" (Burnham, 1986, p. 222) and thus will be easy to train. However, the English subjects' success on the Hindi voicing contrast may have been due to previous phonetic experience. Even though the primary contrast between English /b,d,g/ and /p,t,k/ is not between prevoiced and long-lag stops, native speakers of English are nevertheless *exposed* to prevoiced realizations of /b,d,g/, which may have helped the Werker *et al.* (1981) subjects to perceive the Hindi voicing contrast.

In recent years, the technology used in speech perception research has been adapted to train nonnative phonetic contrasts. A number of training studies have focused on the English /r/–/l/ contrast (Cochrane, 1977; Gillette, 1980; Borden *et al.*, 1983; Strange and Dittmann, 1984; Logan *et al.*, 1989). The training techniques and the amount of training administered have varied widely.[13] Each study has resulted in some improvement in discrimination and/or identification, although performance has generally remained below native-speaker levels. Other studies have focused on word-initial stop consonants. Most but not all of the stop studies have met with success, and several provided evidence for transfer of training (Lisker, 1970; Kalikow and Swets, 1972; Strange and Jenkins, 1978; Pisoni *et al.*, 1982, McClasky *et al.*, 1983). No previous study has focused on the voiced–voiceless contrast in word-final stops, however.

A. Methods

Some previous studies have used discrimination training. One drawback of this technique is that it may highlight differences at the boundaries between the phonetic categories being trained rather than helping to define category *centers* (Repp, 1984).

The present experiment used methods similar to those of Jamieson and Morosan (1986; see also Logan *et al.*, 1989), who employed identification with feedback to train French subjects on the word-initial English /θ/ vs /ð/ contrast. The 56 V + B-removed words (seven vowels × two final stops × four repetitions) from experiment 1 were presented to Chinese subjects who had participated previously in that experiment. Seven subjects each were randomly assigned to experimental and control groups. All 14 participated in three blocks, in which they were told to identify the final stops as "t" or "d." Three randomizations of the 56 stimuli were presented without feedback to both groups in the first and third blocks (designated "pre-training" and "post-training" blocks). Four randomizations of the stimuli were presented in the second block, in which feedback was presented to the experimental but not the control subjects.

B. Results and discussion

Table IV presents the rate at which the /t/ and /d/ tokens were identified correctly. The mean identification rates in the pre-training block resembled those obtained in experiment 1, which shows that the Chinese subjects' difficulty with the V + B-removed stimuli persisted. The experimental subjects' overall identification rate was somewhat higher after than before training (74% vs 65% correct), whereas there was little difference for the control subjects (59% vs 60%).

The 9% increase observed here for the experimental subjects is comparable to the 11% increase in correct identifications obtained by Jamieson and Morosan (1986) for French learners of English. As in that study, subjects in the present study continued to identify the voicing feature in obstruents at rates that were far lower than those of native English speakers. (Recall that, in experiment 1, native English adults identified the V + B-removed stimuli in 99% of instances.) Feedback training resulted in fairly large increases in percent correct identifications for two experimental subjects, but the increases noted for most of the other subjects were quite small.

Table IV also presents the A' scores computed to estimate listeners' sensitivity to the /t/–/d/ contrast. The A' values were calculated based on the average proportion of

TABLE IV. Mean percentage of correct identifications of /t/ and /d/ before, during, and after training. The Chinese subjects in the experimental group received feedback training in the second of three blocks (FB), whereas those in the control group did not. The A' scores represent an unbiased measure of sensitivity to the /t/–/d/ contrast in seven minimal pairs. Standard deviations are in parentheses.

	Before	FB	After
		Experimental group	
/d/	70.4% (24)	74.8% (25)	77.4% (23)
/t/	60.0% (26)	62.8% (30)	71.6% (32)
A'	0.716 (0.126)	0.751 (0.151)	0.802 (0.161)
		Control group	
/d/	67.3% (29)	58.4% (33)	63.2% (32)
/t/	52.8% (28)	59.7% (31)	54.5% (34)
A'	0.658 (0.125)	0.636 (0.150)	0.628 (0.163)

/d/ responses given by each subject to the seven words ending in /d/ and the seven words ending in /t/ in the three blocks. The A' scores were submitted to a group (experimental, control) × block (pre-training, feedback, post-training) ANOVA, which yielded a significant two-way interaction [$F(2,24) = 3.53$, $p < 0.05$]. The two-way interaction occurred because the experimental but not the control subjects' sensitivity to the /t/–/d/ contrast was somewhat greater in the third than the first block. The block main effect was nonsignificant, however, for both the control and the experimental subjects [$F(2,12) = 1.06, 2.56, 1.06$, $p > 0.10$]. The experimental and control subjects did not differ significantly in the first (pre-training) or the second block [$F(1,12) = 0.76, 2.03$, $p > 0.10$]. In the third (post-training) block, the experimental subjects' sensitivity was greater than the control subjects' at a marginally significant level [$F(1,12) = 4.04$, $p = 0.067$].

A larger overall effect of training might have been obtained had more subjects from southern China been included in the experimental group, for the subjects showing the largest effects were from southern China. Of the three Mandarin subjects, two showed small effects of training, and one showed no effect. The training effects might also have been larger had more training trials been presented. These possibilities were tested further in experiment 3.

III. EXPERIMENT 3

A. Introduction

The training administered in experiment 2 might be regarded as effective, but just barely so. A second training experiment was therefore carried out in an attempt to find a better method for training the contrast between /t/ and /d/ in the final position of English words. Identification with feedback was again used as the training modality, but a larger number of training trials were given for a fewer number of words. The experiment also explored issues pertaining to generalization of training, acoustic variability of training tokens, and L1 background.

Little empirical evidence exists concerning the generalization of new speech perception abilities for adult learners of an L2. Neither applied behaviorists nor Piagetian structuralists have viewed generalization of training as an important research issue because they see it as an inherent part of the learning process. While it may be true that generalization is a "natural property" of human intelligence that can be "taken for granted" for certain aspects of language learning (Johnston, 1988, p. 325), it is nevertheless known (see the Introduction) that Chinese subjects have difficulty producing word-final English stops. Evidently, their prior experience producing /t/ and /d/ in the initial position of Chinese words does not automatically enable them to produce a perceptually effective contrast between these stops in the *final* position of English words (Flege and Davidian, 1985; Flege *et al.*, 1987; Flege, 1988c). Moreover, Strange and Dittmann (1984) assessed transfer of training following speech perception training. They found incomplete transfer to new synthetic materials, and no transfer to natural speech.

The subjects in the present experiment were trained using the four tokens of *beat*, *bead*, *bet*, and *bed* used previously in experiments 1 and 2. One question addressed was whether improvements in identifying the final stops of these words would generalize to tokens of other words (*bit*, *bid*, *bought*, *baud*) that had not been trained. If feedback training affects only the phonological/phonetic specification of individual words, one would not expect to see generalization. However, if feedback training leads to an increase in subjects' tacit knowledge about the properties of /t/ and /d/ (or differences between them), one might expect generalization from trained to untrained words.

The multiple natural token approach to speech training assumes that exposure to the acoustic variation between tokens of a single category will induce subjects to derive a more general representation than they would derive had they been trained on just a single token. Another issue addressed was whether changes in vowel context would increase the effectiveness of training. Native English-speaking listeners use variations in vowel quality as a perceptual cue to the voiced–voiceless contrast (e.g., van Summers, 1987). Subjects who receive training on *beat–bead* and *bet–bed* in a single block might therefore show a larger effect of training and/or more generalization than subjects who receive training on the two minimal pairs separately.

In experiment 1, native speakers of Shanghainese tended to identify English stops better than native speakers of Mandarin. In experiment 2, there was some evidence that Shanghainese subjects benefited more from feedback training than native Mandarin subjects, perhaps because Mandarin has no obstruents in word-final position, whereas Shanghainese has words ending in a glottal stop. The presence of a word-final obstruent in their L1 may cause native speakers of Shanghainese to attend to the end of syllables to a greater extent than Mandarin subjects. Children's prior knowledge may affect to some extent the effectiveness of language training (Johnston, 1988). If the same holds true for speech, differences in how syllables are processed might therefore influence the results obtained from feedback training.

B. Methods

The stimuli used in this experiment were the "V + B-removed" versions of minimal pairs used in experiments 1 and 2. The experiment was divided into three blocks. In each, the subject's task was to identify the final stop as "t" or "d." The subjects received feedback training on just two minimal pairs (*beat–bead*, *bit–bid*) in the second block, but were tested on all four minimal pairs (*beat–bead*, *bit–bid*, *bet–bed*, *bought–baud*) in the first and third blocks (designated the "before-" and "after-training" blocks). In the before- and after-training blocks, three separate randomizations of the 32 stimuli (eight words × four tokens) were presented. In the second block, 15 randomizations of the 16 stimuli (four tokens each of *beat*, *bead*, *bit*, *bid*) were presented for identification with immediate feedback. Thus more feedback trials were provided for each word here than in experiment 2 (60 vs 16).

A total of 16 Chinese subjects participated, none of whom had taken part in experiment 1 or 2. Eight subjects each were randomly assigned to "blocked" and "unblocked" conditions. Those in the blocked group were trained on *beat–bit* and *bit–bid* separately in counterbalanced order, whereas those in the unblocked condition received training on both minimal pairs at the same time. Six subjects were native speakers of Mandarin, six were native speakers of Taiwanese, and four were native speakers of a southern Chinese dialect. As summarized in Table V, the subjects in experiment 3 resembled those from experiment 2 in terms of English-language experience. All had learned English as adults and spoke English with a foreign accent. Most of them had arrived in the U.S. in their late twenties and had lived there for several years. Subjects assigned to the blocked and unblocked conditions differed little according to chronological age, age of arrival in the U.S., or length of residence in the U.S.

C. Results and discussion

Table VI presents the rate of correct identifications before and after training for stops in words that either were or were not trained. The results suggest that presenting more training trials on fewer words was not very useful. The rate at which stops were identified correctly increased from 64% to 75%, which is only a slightly greater increase than the one obtained in experiment 2. Moreover, the Chinese subjects continued to identify stops much more poorly than the native English subjects examined in experiment 1.

The effect of training that *did* occur, however, seemed to generalize. The identification rates were 11% higher after than before training for the words that were trained, and 10% higher for the words that were not trained. Subjects who were assigned to the blocked condition showed about the same effects of training, and about the same amount of generalization, as subjects who received training on both minimal pairs at the same time. The overall rate of correct identification for subjects in the blocked and unblocked conditions was 70%.

Table VI also presents the A' values representing the subjects' sensitivity to the /t/–/d/ contrast before and after training. The A' scores were submitted to a condition (blocked versus unblocked) × time (before versus after training) × minimal pair type (trained versus untrained) ANOVA with repeated measures on the last two factors. The A' scores were significantly higher after than before training [$F(1,14) = 8.65, p < 0.05$]. The condition main effect was nonsignificant [$F(1,14) = 0.016, p > 0.10$]. The effects of training generalized, for the A' scores obtained for the trained versus untrained minimal pairs did not differ significantly [$F(1,14) = 1.28$]. No interaction involving any factor reached significance at the 0.05 level.

The size of the training effect was influenced by L1 background. The overall increase in correct identifications was 21% for the southern Chinese subjects, 11% for the Taiwanese subjects, and just 5% for the Mandarin subjects. The mean value for the southern Chinese subjects may have been an underestimation. One subject in that group showed no increase in sensitivity because he identified /t/ and /d/ quite well even before any training was administered. The mean value for the Mandarin subjects may have been an *over*estimation, since the one Mandarin subject who showed an effect of training had been exposed to Cantonese (a language that permits unreleased /p,t,k/ to occur in word-final position). Finally, the mean value reported for the Taiwanese subgroup is somewhat misleading because it represents an average for three subjects who showed training effects of 18%–30% and three subjects who showed little or no effect of training.

TABLE V. Characteristics of the Chinese subjects in experiment 3. POB indicates place of birth, AOA is age of arrival in the U.S., MOS is length of residence in the U.S., in months, L1 is native language. Subjects assigned to the blocked condition received training separately on *beat–bit* and *bit–bid*, whereas those in the unblocked condition received training on both minimal pairs at one time. Most subjects whose L1 is listed as "Southern" spoke Shanghainese.

Sex	Age	AOA	MOS	POB	L1
			Blocked		
M	27	24	37	Taipei, ROC	Mandarin
M	30	28	10	Taipei, ROC	Taiwanese
F	33	26	72	Chufai, ROC	Taiwanese
M	31	29	9	Taipei, ROC	Taiwanese
F	28	22	73	Taipei, ROC	Mandarin
F	26	26	1	Xian	Mandarin
M	32	29	35	Shenying	Mandarin
F	24	24	1	Zhangchou	Southern
M	29	26	30		
	(3)	(3)	(30)		
			Unblocked		
M	29	28	9	Tainan, ROC	Taiwanese
M	31	31	25	Tainan, ROC	Mandarin
M	33	26	84	Taipei, ROC	Taiwanese
M	23	23	1	Wuhu	Southern
M	30	30	1	Hangzhou	Southern
F	29	29	1	Changzhou	Mandarin
M	32	28	37	Taipei, ROC	Southern
M	31	25	72	Taipei, ROC	Taiwanese
M	30	28	30		
	(3)	(3)	(32)		

IV. GENERAL DISCUSSION

Previous speech production studies have shown that the existence of a phonological voicing contrast in the initial position of Chinese words, which is implemented as a contrast between voiceless unaspirated versus voiceless aspirated stops, does not guarantee success in producing word-final stops in English. Chinese speakers of English often have difficulty producing a perceptually effective contrast between voiced and voiceless stops in word-final position (Tarone, 1980; Eckman, 1981; Anderson, 1983, 1987; Flege and Davidian, 1985; Heyer, 1986; Weinberger, 1987; Flege *et al.*, 1987). Experiment 1 provided analogous speech perception results. Chinese subjects were able to identify the final stops in words like *beat* and *bead* as well as native speakers of English when all acoustic cues were present, whereas they performed much more poorly than native speakers when release burst cues were removed.

The results of experiment 1 suggest that the English /t/–/d/ contrast was not treated as a novel phonetic contrast by the Chinese subjects, and that phonemic perception in an L2 must be learned on an allophone-by-allophone basis. It appeared that the Chinese subjects were attempting to

TABLE VI. Mean percent correct identifications of /t/ and /d/ by subjects who received training separately or together on multiple tokens of two English minimal pairs (blocked versus unblocked conditions). Performance was assessed before and after feedback training both for words that were trained (*bit-bid, bet-bed*) or not (*bet-bed, bought-baud*). The *A*′ scores represent an unbiased estimate of sensitivity to the /t/–/d/ contrast. Standard deviations are in parentheses.

	Trained words		Untrained words	
	Before	After	Before	After
Blocked condition				
Percent correct	66% (26)	75% (23)	63% (30)	75% (26)
A′	0.736(0.117)	0.832(0.086)	0.667(0.210)	0.817(0.119)
Unblocked condition				
Percent correct	66% (32)	80% (27)	62% (29)	72% (30)
A′	0.702(0.184)	0.825(0.227)	0.654(0.289)	0.755(0.227)

identify the word-final English stops using cues they applied to the word-initial /t/–/d/ contrast in their native language (L1). The voiced and voiceless bursts that occurred at release of lingual occlusion for final stops in the unedited condition of experiment 1 are likely to be similar to the release bursts for word-initial /t/ and /d/ in Chinese words. This may explain the relatively good performance of Chinese subjects for the unedited stimuli. If the Chinese subjects relied on word-initial cues, it would also explain their poor performance on the stimuli without release bursts, since cues that remained after editing (e.g., preceding vowel duration and quality) are absent in word-initial stop voicing distinctions.

This interpretation is consistent with the finding that the Chinese subjects showed only a very small decrease in sensitivity to the /t/–/d/ contrast when closure voicing was removed. One would not expect the Chinese subjects to rely on closure voicing as a cue for word-final stops in English because word-initial tokens of /t/ and /d/ are not distinguished by closure voicing in Chinese. The very small effect of removing closure voicing was not due to the fact that the voicing was inaudible. A follow-up study (see the Appendix) revealed that native English speakers showed a significant preference for final /d/'s with closure voicing over those without closure voicing. The Chinese subjects' sharp drop in correct identifications when release burst cues were removed was not an artifact of the waveform editing, since stops in conversational speech often have little or no closure voicing and no audible release burst. Moreover, the English children who were tested closely resemble the native English adults in showing little effect of editing.

It is not known if the Chinese subjects' difficulty in identifying /t/'s and /d/'s without release bursts would affect their comprehension of English, nor is it known if they established central phonetic representations for word-final English stops. If they did, their central phonetic representations were probably less elaborated than native English speakers' (see Massaro and Oden, 1980; Samuel, 1982; Flege, 1974, 1988a), which presumably specify parameter values for a large number of dimensions such as preceding vowel duration, vowel quality, and F_1 offset frequency in addition to information associated with final release bursts, such as overall intensity.

Two training experiments were undertaken to improve Chinese subjects's ability to identify /t/ and /d/ in English words from which closure voicing and release bursts had been removed. The first resulted in a nonsignificant 9% increase in the rate of correct identifications. The second, which included a larger number of training trials on a smaller number of words, resulted in a slightly larger and significant increase (11%) that generalized to untrained words. The training effects obtained here were similar in magnitude to the one obtained by Jamieson and Morosan (1986), who used similar techniques to train French subjects on the word-initial English /θ/-/ð/ contrast (see also Logan *et al.*, 1989). Although the Chinese subjects' performance improved significantly, they continued to identify stops much less well than native speakers of English, who performed at near-perfect levels. Thus an important question to ask is: Why did the Chinese subjects benefit *so little* from the training?

There are two reasons to think that the small effect of training probably cannot be attributed to the nature of the phonetic contrast being trained. First, the English subjects identified stops in the V + B-removed stimuli quite accurately. Second, previous research suggests that nonnative phonetic contrasts which involve the voicing feature are relatively *easy* to train because they are relatively robust in psychoacoustic terms (Tees and Werker, 1984; Burnham, 1986; Best *et al.*, 1988; Werker *et al.*, 1981; Werker and Tees, 1983, 1984a,b; Werker and Logan, 1985).

The difficulty of the word-final English /t/–/d/ contrast may have derived instead from a syllable-processing strategy established during L1 acquisition. This inference was drawn from the observation that certain subgroups of Chinese subjects tended to perform better than others in experiment 1, and to benefit more from the feedback training in experiments 2 and 3. In experiment 3, three of four subjects from southern China whose L1 includes word-final obstruents (but not /d/) showed gains that were much greater than average (18%–30% correct). These subjects may have learned to focus greater attention on the syllable termination properties of words in their L1 than the native Mandarin subjects (whose L1 has no word-final obstruents) in order to distinguish between open syllables and syllables ending in a glottal stop. This hypothesis received support in a recent study by Flege and Wang (1989). In that study, it was shown that subjects whose L1 (Cantonese) permits unreleased /p,t,k/ in final position identified edited tokens of /t/ and /d/ significantly better than subjects whose L1 (Mandarin) permits no final obstruents.

Locke (1983) speculated that syllable offset properties are important for children learning their L1. Adults may delete final /t/'s and /d/'s, or replace them with glottal stop (Bernstein Ratner and Luberoff, 1984), which might explain English-speaking children's propensity for omitting final stops or replacing them with glottal stops. The hypothesis offered here is that, if the L1 does not distinguish words according to information specified in syllable offsets, then L2 learners will focus relatively little attention later in life to the offset properties of syllables (or words) in an L2. This assumes that syllable-processing strategies are indeed learned. Consistent with this assumption, Walley (1987) found that in contexts where meaning was not semantically constrained, 4 and 5 year olds did not detect more mispronunciations in initial than final position. They did so in constraining contexts, however, thereby resembling adults. Walley (1988) found that adults but not children gave lower ratings to word-initial than word-final consonants that had been degraded by noise.

In conclusion, Chinese subjects identified fully specified English /t/'s and /d/'s as well as native English speakers but performed more poorly when release bursts were removed. Feedback training increased significantly Chinese subjects' sensitivity to stops without release bursts, but their gains were small in absolute terms. This was especially true for native speakers of Mandarin, a language that does not permit obstruents in word-final position. Additional research will be needed to further test the hypothesis that the effectiveness of feedback training for word-final stops will be limited if subjects learn to focus relatively little attention on the end of syllables during L1 acquisition.

ACKNOWLEDGMENTS

The research was supported by NIH grant NS20963. The author thanks E. James for help with all phases of the research. Thanks are due also to O.-S. Bohn, W. Strange, and two anonymous reviewers for comments on a previous draft of the article.

APPENDIX: PERCEPTION OF CLOSURE VOICING

Massaro and Oden (1980) hypothesized that listeners evaluate all potential acoustic cues to a phonetic contrast independently and continuously, even though no one property is necessary to establish category identity. However, removing closure voicing did not significantly lower English subjects' sensitivity to the /t/–/d/ contrast in experiment 1. This raises the possibility that the subjects may not use closure voicing as a cue to the /t/–/d/ contrast.

The nonuse of closure voicing may occur because /d/ is not reliably distinguished from /t/ in English by closure voicing (Flege and Hillenbrand, 1987), especially in the speech of children (Smith, 1979; Flege *et al.*, 1987). Previous research has shown that the perceptual effect of removing closure voicing from /b,d,g/ varies considerably for individual native speakers of English (Wardrip-Fruin, 1982; Revoile *et al.*, 1982; Raphael, 1981). Alternatively, the native English subjects may not have shown a decrease in sensitivity to the /t/–/d/ contrast when closure voicing was removed because they were able to make accurate identifications using other cues such as vowel duration.

To test this, a short paired-comparison experiment was carried out to determine whether native speakers of English do or do not use closure voicing in a phonetically relevant manner. The stimuli consisted of one of the four tokens each of the /d/-final word from experiment 1, and a copy from which closure voicing had been removed. In agreement with previous reports (e.g., Revoile *et al.*, 1982), the closure voicing in the final /d/'s was less intense (by 18 dB) than the vowel and died out before the end of the closure intervals in six of the seven words selected. Closure voicing filled 75% of the closure intervals on average. Pairs of /d/-final words with and without closure voicing were presented, separated by 0.5 s, to 11 normal-hearing native English subjects with a mean age of 29 years. The subjects were to decide which member of each pair sounded "better." Each pair was presented randomly 20 times each (10 times with the voiced /d/'s in the first position, and 10 times in the second position).

The results suggested that removing closure voicing in experiment 1 did not affect the English subjects' performance because of a ceiling effect. Subjects chose /d/'s with closure voicing more often than those without (64% vs 36%). A significant vowel × voicing interaction was obtained [$F(6,60) = 7.34$, $p < 0.05$] because the simple main effect of voicing was significant in all seven words except *booed* ($p < 0.01$). As expected from previous research, intersubject variability was great. The rate at which /d/'s with voicing were chosen over /d/'s without voicing ranged from 55% to 73%.

[1]Oyama (1982a) noted a significant correlation between degree of foreign accent and age of learning for Italian men who began learning English as an L2 between the ages of 6 to 20 years. Tahta *et al.* (1981) reported no accent for nonnative speakers who began learning English by the age of 6 years, moderage accents for those who began learning English between the ages of 7–11 years, and relatively strong accents for individuals who began learning English after the age of 12 years. Amount of L2 experience has generally been found to predict degree of foreign accent less well than age of learning. Tahta *et al.* (1981) found that age of learning but not length of residence accounted for a significant amount of variance in foreign accent ratings, perhaps because subjects with less than 2 years' experience were not included. Oyama (1982a) found that a correlation between degree of accent and age of learning remained significant when differences in length of residence were partialed out (see also Fathman, 1975; Seliger *et al.*, 1982, but cf. Asher and Garcia, 1969).

[2]The author thanks W. Strange for pointing this out.

[3]The decrease in correct identifications that occurs when release bursts are removed is generally greater for /p,t,k/ than /b,d,g/ (Malcot, 1958; Winitz *et al.*, 1972; Raphael, 1981; Revoile *et al.*, 1982), perhaps because voiceless stops are released more consistently than voiced stops (Wang, 1959), or because they are more intense.

[4]Wolf's (1978) subjects had three response alternatives: voiced, voiceless, or no final stop. The percentages given here for the edited words were based only on responses for words identified as having a final stop.

[5]Comprehenson was examined using a 45-item, tape-recorded test that assesses syntactic knowledge as well as word recognition. The test, a portion of the English Language Assessment Battery (English Language Institute, University of Michigan, 1986) does not gauge directly the ability to identify phonemic categories. The Chinese subjects' average score was 34.2 correct out of 45. The percentile score was based on the performance of 1486 students who had taken the test while enrolled in English language classes in Ann Arbor prior to enrolling as full-time degree seeking students.

[6]The closure interval extended from constriction of the final stop, identified by a decrease in signal amplitude and a concomitant simplification in waveshape, to the beginning of the final release burst.

[7]Preliminary analyses for the unedited and the burst-removed words

showed that there was not a significant correlation between the text frequencies of the 14 words examined and the rate of correct identifications [$r(12) = 0.172, 0.375$], nor was there a difference in the rate of correct identifications for male and female Chinese subjects for /t/ (57% vs 62%) or /d/ (68% vs 71%) [$F(1,28) = 1.01, 0.18, p > 0.10$].

[8]The sum of squares of the between-subjects factors were adjusted in a hierarchical fashion because between-subjects factors tend to be correlated when sample sizes are unequal.

[9]A significant vowel×editing condition interaction was obtained [$F(18,792) = 2.66, p < 0.05$] because of a differing effect of the vowel factor for the four editing conditions. This seems to have occurred because the difference in sensitivity between the unedited and edited conditions was greater for some minimal pairs than others. For example, the decrease in A' scores from the unedited to the V + B-removed condition was greater for *bit–bid* (0.257) than for *bought–baud* or *bet–bed* (0.137, 0.166). Vowel-related effects will not be discussed further because the group×vowel×editing condition interaction was nonsignificant [$F(36,792) = 0.781, p > 0.10$] and because the exact acoustic basis for these differences is uncertain (see Hillenbrand *et al.*, 1984; van Summers, 1987).

[10]The d' values were calculated according to the formula $d' = z$(proportion of hits) $- z$(proportion of false alarms). The A' scores were examined in the ANOVA reported earlier because of the small number of trials and the requirement of normal distributions for d', which could not be met for many of the native English subjects owing to perfect scores. Even though A' and d' scores may be correlated significantly (see Snodgras *et al.*, 1972), the regression analysis examined d' scores because they seemed to be more revealing of between-group differences than the A' scores. An ANOVA examining d' scores showed the same significant main effects and interactions as the A' analysis, but *post hoc* tests revealed between-group differences not seen in the A' analysis (significantly less sensitivity for the Chinese than English adults in the unedited condition, and less sensitivity to the /t/–/d/ contrast by the English children than English adults in the burst-removed condition; $p < 0.05$). When the six-factor regression model was applied to A' scores, it did not account for a significant amount of variance for either the burst-removed [$F(1,27) = 3.67$, $p = 0.063$] or the V + B-removed [$F(1,27) = 2.96$, $p = 0.069$] condition.

[11]The effect of L1 background on the correct identification rate for /d/'s was significant for stops in the voicing-removed and burst-removed conditions [$F(2,207) = 15.4, 4.43, p < 0.05$]. *Post hoc* tests revealed that the Shanghainese subjects identified /d/'s correctly significantly more often than the native Mandarin or Taiwanese speakers. Nonsignificant trends in the same direction were evident for final /t/'s.

[12]Hyperaspiration was not noted in /b,d,g/. This suggests, in agreement with two recent speech production studies (Flege *et al.*, 1987; Flege, 1988c), that the Chinese subjects produced voiceless but not voiced word-final stops in English with a laryngeal devoicing gesture.

[13]These studies employed a wide variety of methods: discrimination training with immediate feedback (Strange and Dittmann, 1984); perceptual feedback concerning consonant identity together with feedback concerning the difference between correct and incorrect productions, and additional information relating to "visual" and tactile–kinesthetic" differences between /r/ and /l/ (Borden *et al.*, 1983); instructions concerning tongue placement and imitation of a native speaker's correct productions (Cochrane, 1977); and articulatory descriptions followed by individual and group practice in producing /r/ and /l/ that were evaluated by fellow students and a native English-speaking instructor (Gillette, 1980). The length of training varied widely, from two 45-min clinical sessions (Borden *et al.*, 1983) to 5 of clinical training (Cochrane, 1977) to 15 laboratory sessions (Strange and Dittmann, 1984) to 5 weeks of classroom training (Gillette, 1980).

Anderson, J. (**1983**). "The difficulties of English syllable structure for Chinese ESL learners," Lang. Learn. Commun. **2**, 53–61.

Anderson, J. (**1987**). "The markedness differential hypothesis and syllable structure difficulty," in *Interlanguage Phonology*, edited by G. Ioup and S. Weinberger (Newbury House, Rowley, MA), pp. 279–291.

Asher, J., and Garcia, R. (**1969**). "The optimal age to learn a foreign language," Mod. Lang. J. **53**, 334–341.

Bernstein Ratner, N., and Luberoff, A. (**1984**). "Cues to post-vocalic voicing in mother–child speech," J. Phon. **12**, 285–289.

Best, C., McRoberts, G., and Sithole, N. (**1988**). "Examination of perceptual reorganization for nonnative speech contrasts. Zulu click discrimination by English speaking adults and infants," JEP:HPP **14**, 345–360.

Boothroyd, A., and Nittrouer, S. (**1988**). "Mathematical treatment of context effects in phoneme and word recognition," J. Acoust. Soc. Am. **84**, 101–114.

Borden, G., Gerber, A., and Milsark, G. (**1983**). "Production and perception of the /r/–/l/ contrast in Koren adults learning English," Lang. Learn. **33**, 499–526.

Burnham, D. (**1986**). "Developmental loss of speech perception: Exposure to and experience with a first language," Appl. Psycholing. **3**, 207–241.

Brière, E. (**1966**). "An investigation of phonological interference," Language **42**, 769–796.

Cheng, C.-C. (**1973**). *A Synchronic Phonology of Mandarin Chinese* (Mouton, The Hague).

Clumeck, H., Barton, D., Macken, M., and Huntington, D. (**1981**). "The aspiration contrast in Cantonese word-initial stops: Data from children and adults," J. Chin. Ling. **9**, 210–224.

Cochrane, R. (**1977**). "The acquisition of /r/ and /l/ by Japanese children and adults learning English as a second language," unpubl. Ph.D. thesis, University of Connecticut, Storrs, CT.

Cole, R., and Perfetti, C. (**1980**). "Listening for mispronunciations in a children's story: The use of context by children and adults," J. Acoust. Soc. Am. **67**, 1323–1332.

Eckman, F. (**1981**). "On the naturalness of interlanguage phonological rules," Lang. Learn. **31**, 195–216.

Eguchi, S., and Hirsch, I. (**1969**). "Development of speech sounds in children," Acta Oto-Laryngol. Suppl. **257**, 1–55.

English Language Institute (**1986**). *Manual for the English Language Institute Listening Comprehension Test* (Ann Arbor, MI), 16 pp.

Fathman, A. (**1975**). "The relationship between age and second language productive ability," Lang. Learn. **25**, 245–253.

Flege, J. (**1984**). "The detection of French accent by American listeners," J. Acoust. Soc. Am. **76**, 692–707.

Flege, J. (**1988a**). "The production and perception of foreign language speech sounds," in *Human Communication and Its Disorders, A Review, 1988*, edited by H. Winitz (Ablex, Norwood, NJ), pp. 224–401.

Flege, J. (**1988b**). "Factors affecting degree of foreign accent in English sentences," J. Acoust. Soc. Am. **84**, 70–79.

Flege, J. (**1988c**). "The development of skills in producing word-final /p/ and /b/: Kinematic parameters," J. Acoust. Soc. Am. **84**, 1639–1652.

Flege, J., and Davidian, R. (**1985**). "Transfer and development processes in adult foreign language speech production," J. Appl. Psycholing. Res. **5**, 323–347.

Flege, J., and Eefting, W. (**1986**). "Linguistic and developmental effects on the production and perception of stop consonants," Phonetica **43**, 155–171.

Flege, J., and Eefting, W. (**1987**). "Cross-Language switching in stop consonant production and perception by Dutch speakers of English," Speech Commun. **6**, 185–202.

Flege, J., and Hillenbrand, J. (**1985**). "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., and Hillenbrand, J. (**1987**). "A differential effect of release bursts on the stop voicing judgments of native French and English listeners," J. Phon. **15**, 203–208.

Flege, J., McCutcheon, M., and Smith, S. (**1987**). "The development of skill in producing word-final English stops," J. Acoust. Soc. Am. **82**, 433–447.

Flege, J., and Port, R. (**1981**). "Cross-language phonetic interference: Arabic to English," Lang. Speech **24**, 125–146.

Flege, J., and Wang, C. (**1989**). "Native-Language phonotactic constrains affect how well Chinese subjects perceive the word-final English /t/–/d/ contrast," J. Phon. (in press).

Florentine, M. (**1985**). "Speech perception in noise by fluent, non-native listeners," J. Acoust. Soc. Am. Suppl. 1 **77**, S106.

Gillette, S. (**1980**). "Contextual variation in the perception of L and R by Japanese and Korean speakers," Minn. Papers Ling. Philos. Lang. **6**, 59–72 (University of Minnesota, Minneapolis).

Gras, S. (**1983**). "Integrative attitude and difficulties second language learners encounter in comprehending informal speech: A study of Korean graduate students attending the University of Texas," unpubl. Ph.D. thesis, University of Texas at Austin.

Grier, J. (**1971**). "Nonparametric indexes for sensitivity and bias: Computing formulae," Psychol. Bull. **75**, 424–429.

Halle, M., Hughes, G., and Radley, J.-P. (**1957**). "Acoustic properties of stop consonants," J. Acoust. Soc. Am. **29**, 107–116.

Heyer, S. (**1986**). "English final consonants and the Chinese learner," unpubl. MA thesis, Dept. of Linguistics, Southern Illinois University, Carbondale, IL.

Hillenbrand, J., Ingrisano, D., Smith, B., and Flege, J. (1984). "Perception of the voiced–voiceless contrast in syllable-final stops," J. Acoust. Soc. Am. 76, 18–28.

Howie, J. (1976). *Acoustical Studies of Mandarin Vowels and Tones* (Cambridge U.P., London).

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

Johannson, S. (1978). *Studies in Error Gravity*, Gothenburg Studies in English 44 (Dept. of English, Gothenburg University of Sweden).

Johnston, J. (1988). "Generalization: The nature of change," Lang. Speech, Hear. Serv. Schools 19, 314–329.

Jusczyk, P. (1979). "Infant speech perception: A critical appraisal," in *Perspectives on the Study of Speech*, edited by P. Eimas and J. Miller (Erlbaum, Hillside, NJ).

Kalikow, D., and Swets, J. (1972). "Experiments with computer-controlled displays in second-language learning," IEEE Trans. Audio Electroacoust. AE-20, 23–28.

Kelso, S., and Tuller, B. (1981). "Toward a theory of apractic syndromes," Brain Lang. 12, 224–245.

Kent, R., and Rosenbek, J. (1983). "Acoustic patterns of apraxia of speech," J. Speech Hear. Res. 26, 231–249.

Koster, C. (1987). *Word Recognition in Foreign and Native Language* (Foris, Dordrecht, The Netherlands).

Krause, S. (1982a). "Vowel duration as a perceptual cue to postvocalic consonant voicing in young children and adults," J. Acoust. Soc. Am. 71, 990–995.

Krause, S. (1982b). "Developmental use of vowel duration as a cue to postvocalic stop consonant voicing," J. Speech Hear. Res. 25, 388–393.

Kuhl, P. (1987). "Perception of speech and sound in early infancy," in *Handbook of Infant Perception, Vol. 2*, edited by P. Salapatek and L. Cohen (Academic, New York), pp. 275–382.

Lado, R. (1957). *Linguistics Across Cultures* (University of Michigan, Ann Arbor).

Leonard, L., Schwartz, R., Chapman, K., Rowan, L., Prelocke, P., Weiss, A., and Messick, C. (1982). "Early lexical acquisition in children with specific language impairment," J. Speech Hear. Res. 25, 554–564.

Lenneberg, E. (1967). *Biological Foundations of Language* (Wiley, New York).

Li, C., and Thompson, S. (1981). *Mandarin Chinese: A Functional Reference Grammar* (University of California, Berkeley, CA).

Lisker, L. (1970). "On learning a new phonetic contrast," Haskins Lab. Stat. Rep. Speech Res. 24, 1–15.

Lisker, L., and Abramson, A. (1964). "A cross-language study of voicing in initial stops: Acoustical measurements," Word 20, 384–421.

Locke, J. (1983). *Phonological Acquisition and Change* (Academic, New York).

Logan, J., Pisoni, D., and Lively, S. (1989). "Training Japanese listeners to identify /r/ and /l/," Res. Speech Percept. 14, 1–22 (Dept. of Psychology, Indiana University).

Mack, M., and Tierney, J. (1987). "The intelligibility of natural and vocoded semantically anomalous sentences: A comparative analysis of English monolinguals and German–English bilinguals," Tech. Rep. 792 (Lincoln Lab., MIT, Cambridge, MA).

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

Malécot, A. (1958). "The role of releases in the identification of released final stops," Language 34, 370–380.

Massaro, D., and Oden, G. (1980). "Evaluation and integration of acoustic features in speech perception," J. Acoust. Soc. Am. 67, 996–1013.

McClasky, C., Pisoni, D., and Carrell, T. (1983). "Transfer of training of a new linguistic contrast in voicing," Percept. Psychophys. 34, 323–330.

Miyawaki, K., Strange, W., Verbrugge, R. Liberman, A., Jenkins, 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.

Ohde, R., and Sharf, D. (1977). "Order effect of acoustic segments of VC and CV syllables on stop and vowel identification," J. Speech Hear. Res. 20, 543–554.

Oyama, S. (1982a). "A sensitive period for the acquisition of a nonnative phonological system," in *Child–Adult Differences in Second Language Acquisition*, edited by S. Krashen and R. Scarcella (Newbury House, Rowley, MA), pp. 20–38.

Oyama, S. (1982b). "The sensitive period and comprehension of speech," in *Child–Adult Differences in Second Language Acquisition*, edited by S. Krashen and R. Scarcella (Newbury House, Rowley, MA), pp. 39–52.

Parnell, M., and Amerman, J. (1978). "Maturational influences on perception of coarticulatory effects," J. Speech Hear. Res. 21, 682–701.

Pisoni, D., Aslin, R., Perey, A., and Hennessy, B. (1982). "Some effects of laboratory training on identification and discrimination of voicing contrasts," JEP:HPP 8, 297–314.

Polka, L. (1987). "Perception of Persian uvular and velar stops by speakers of American English," J. Acoust. Soc. Am. Suppl. 1, 81, S66.

Raphael, L. (1981). "Durations and contexts as cues to word-final cognate oppositions in English," Phonetica 38, 126–147.

Repp, B. (1984). "Categorical perception: Issues, methods, findings," in *Speech and Language: Advances in Basic Research and Practice, Vol. 10*, edited by N. Lass (Academic, New York), pp. 243–335.

Revoile, S., Picket, J., Holden, L., and Talkin, D. (1982). "Acoustic cues to final stop voicing for impaired- and normal-hearing listeners," J. Acoust. Soc. Am. 72, 1145–1154.

Samuel, A. (1982). "Phonetic prototypes," Percept. Psychophysics 31, 307–314.

Schneiderman, E., Bourdages, J., and Champagne, C. (1988). "Second-language accent: The relationship between discrimination and perception in acquisition," Lang. Learn. 38, 1–21.

Seliger, H., Krashen, S., and Ladefoged, P. (1982). "Maturational constraints in the acquisition of second languages," in *Child–Adult Differences in Second Language Acquisition*, edited by S. Krashen, R. Scarcella, and M. Long (Newbury House, Rowley, MA), pp. 13–19.

Shattuck-Hufnagel, S. (1987). "The role of word-onset consonants in speech production planning: New evidence from speech error patterns," in *Motor and Sensory Processes of Language*, edited by E. Keller and M. Gopnik (Erlbaum, Hillside, NJ), pp. 17–52.

Shinn, P. (1985). "A cross-language investigation of the stop, affricate, and fricative manners of articulation," unpubl. Ph.D. thesis, Brown University, Providence, RI (Univ. Microfilms, No. 8519906).

Shockey, L., and Bond, Z. (1980). "Phonological processes in speech addressed to children," *Phonetica* 37, 267–274.

Simon, C., and Fourcin, C. (1978). "Cross-language study of speech-pattern learning," J. Acoust. Soc. Am. 63, 925–935.

Smith, B. (1979). "A phonetic analysis of consonantal devoicing in children's speech," J. Child Lang. 6, 19–28.

Snodgras, J., Volvovitz, R., and Walfish, E. (1972). "Recognition memory for words, pictures, and words + pictures," Psychonom. Sci. 27, 345–347.

Strange, W. (1986). "Speech input and the development of speech perception," in *Otitis Media and Child Development*, edited by J. Kavanagh (York, Parkton, MD), pp. 12–26.

Strange, W., and Dittmann, S. (1984). "Effects of discrimination training on the perception of /r–l/ by Japanese adults," Percept. Psychophys. 36, 131–145.

Strange W., and Jenkins, J. (1978). "The role of linguistic experience in the perception of speech," in *Perception and Experience*, edited by R. Walk and H. Pick, Jr. (Plenum, New York), pp. 125–169.

Tahta, S., Wood, M., and Lowenthal, K. (1981). "Foreign accents: Factors relating to transfer of accent from the first language to the second language," Lang. Speech 24, 265–272.

Tarone, E. (1980). "Some influences on the syllable structure of interlanguage phonology," Int. Rev. Appl. Ling. 18, 139–152.

Tees, R., and Werker, J. (1984). "Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds," Canad. J. Psych. 38, 579–590.

Trubetzkoy, N. (1939). *Grundzge der phonologie* (Travaux du cercle linguistique de Prague, 7).

Truin, P. (1981). "Herkenning van fragmenten van gesproken Nederlandse woorden door Nederlanders en niet-Nederlanders," *Toegepaste Taalwetenschap in Artikelen* 9, 170–182.

Underbakke, M., Polka, L., Gottfried, T., and Strange, W. (1988). "Trading relations in the perception of /r/–/l/ by Japanese learners of English," J. Acoust. Soc. Am. 84, 90–100.

van Balen, C. (1980). *Intelligibility of Speech Fragments* (Drukkerijelinkwijk, Utrecht).

van Summers, W. (1987). "Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses," J. Acoust. Soc. Am. 82, 847–863.

Walley, A. (1987). "Young children's detections of word-initial and -final mispronunciations in constrained and unconstrained context," Cogn.

Dev. 2, 145–167.
Walley, A. (1988). "Spoken word recognition by young children and adults," *Cogn. Dev.* 3, 137–165.
Wang, W. (1957). "Transition and release as perceptual cues for final plosives," J. Speech Hear. Res. 2, 66–73.
Wardrip-Fruin, C. (1982). "On the status of temporal cues to phonetic categories: Preceding vowel duration as a cue to voicing in final stop categories," J. Acoust. Soc. Am. 71, 187–195.
Weinberger, S. (1987). "The influence of linguistic context of syllable structure simplification," in *Interlanguage Phonology*, edited by G. Ioup and S. Weinberger (Newbury House, Rowley, MA), pp. 401–418.
Werker, J., Gilbert, J., Humphrey, K., and Tees, R. (1981). "Developmental aspects of cross-language speech perception," Child Dev. 52, 349–355.
Werker, J. (1989). "Becoming a native listener," Am. Sci. 77, 54–59.
Werker, J., and Logan, J. (1985). "Cross-language evidence for three factors in speech perception," Percept. Psychophys. 37, 35–44.
Werker, J., and Tees, R. (1983). "Developmental changes across childhood in the perception of non-native sounds," Can. J. Psychol. 37, 278–286.
Werker, J., and Tees, R. (1984a). "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life," Infant Behav. Dev. 7, 49–63.
Werker, J., and Tees, R. (1984b). "Phonemic and phonetic factors in adult cross-language speech perception," J. Acoust. Soc. Am. 75, 1866–1878.
Winitz, H., Scheib, M., and Reeds, J. (1972). "Identification of stops and vowels from the burst portion of /p,t,k/ isolated from conversational speech," J. Acoust. Soc. Am. 51, 1309–1317.
Wolf, C., (1978). "Voicing cues in English final stops," J. Phon. 6, 299–309.