The Effects of Early English Learning on Auditory Perception of English Minimal Pairs by Taiwan University Students

Hui-Li Lin,[1,3] Hsing-Wu Chang,[1] and Hintat Cheung[2]

Auditory perception of English minimal pairs was tested with or without noise background. Each subject was interviewed after the test to collect information regarding their early experience on learning English as a foreign language. This study was designed to examine the differential effects of learning English at three age-starting points and two learning durations. This study hopes to determine how childhood experience of English learning (which is not mandatory in public elementary schools) has affected the auditory competence of university students in distinguishing English minimal pairs. Results showed that age effects were salient only under condition of noise background. Without the interference of background noise, most subjects performed well enough to obliterate any potential differences.

INTRODUCTION

For some second-language educators, the notion "earlier is better" is a safe and simple answer. However, is earlier actually better? In what sense is "better" meant? Age effects in second-language learning have been investigated by many researchers. Although some findings based on initial gain did not seem to support this notion (e.g., Snow & Hoefnagel-Höhle, 1978), many studies that focused on long-term/ultimate effects did seem to show "the earlier, the better" (e.g., DeKeyser, 2000; Johnson, 1992; Johnson & Newport, 1989; Johnson, Shenkman, Newport, & Medin, 1996; Newport, 1990; Slavoff & Johnson, 1995). Consequently, some

---

[1] Department of Psychology, National Taiwan University, Taipei, Taiwan.
[2] Graduate Institute of Linguistics, National Taiwan University, Taipei, Taiwan.
[3] To whom all correspondence should be addressed: Department of Psychology, National Taiwan University, Taipei 106, Taiwan. email: hui_li_lin@yahoo.com

recent studies have further examined age effects, and results from these studies caution us in drawing hasty conclusions. For example, a meta-study by Marinova-Todd, Marshall, and Snow (2000) pointed out three misconceptions about age effects in second language acquisition: 1. misinterpretation of the facts regarding acquisition speed, 2. misattribution of age differences in language performance to neurobiology, and 3. an underemphasis on cases in which adults master a second language to native-level proficiency. Moreover, the global notion "earlier is better" may need further specification and elaboration. For instance, Flege, Yeni-Komshian, and Liu (1999) analyzed the relationship between age of arrival (AOA) of native Korean speakers who emigrated to the United States and their performance on English grammatical judgments and foreign accent in a matched-group design. They found no age effects on morphosyntax judgment tasks, but some age effects on foreign accent ratings. Based on these facts, the belief of "the earlier, the better" ought to be examined more closely, especially when it is likely to have significant implications in education policy.

To illustrate, interaction between L1 (first language) and L2 (second language) is affected by various factors and can be extremely complicated (e.g., Flege *et al*., 1999; Wode, 1995). This complexity implies that a significant ultimate gain in L2 might have to be accomplished at the price of L1 retardation. If this is true, earlier is not necessarily better in an overall sense. Furthermore, as suggested by studies on initial gain, learning a second language at an earlier age does not guarantee faster speed of acquisition at the initial stage. For example, in Snow and Hoefnägel-Höhle (1978), older second-language learners performed significantly better than their younger counterparts both in the middle and at the end of their 1-year longitudinal study. Ervin-Tripp (1974) also concluded that older second-language learners are better performers due to their more developed general intelligence. Therefore, if there are no age effects that cannot be accounted for by total length of time spent on second-language learning, then it should be easy to have a late starter who acquires L2 at a faster speed and whose L1 is mature enough to sustain possible L2 interference. In other words, unless there is a genuine age effect, late starters might very likely be able to catch up with earlier starters if they can receive more intense and compact instruction in a shorter period of time.

In addition to the concerns mentioned above, data supporting age effects were mostly collected without properly controlling important confounding variables. Under English-as-a-second-language (ESL)[4] settings, attempts to control confounding variables have not been entirely successful, thus many variables (e.g., length of learning submersion duration, total years of education received in the

---

[4] In an EFL (English-as-a-Foreign-Language) environment, English is used mostly in classroom situations, as a subject being taught. In contrast to an ESL (English-as-a-Second-Language) environment, where English is not only a language learner's second language, but also used by most of the people living in the environment, English usage patterns in Taiwan better resemble EFL.

immigrant country, daily usage of first vs. second languages, chronological age of subjects at test, and so forth) are left confounded (see discussion in Guion, Flege, Liu, and Yeni-Komshian, 2000). Many researchers recognize this problem and attempt to better control these variables. For example, Flege *et al.* (1999) utilized a matched-group design in an attempt to control these confoundings; however, as is well-known, matched-group design is not an ideal method to control all possible confounding variables. By matching one of the variables, such as age-of-arrival (AOA), and manipulating another (e.g., U.S. education), some other unmatched variables (e.g., chronological age or general academic capabilities) are often left uncontrolled. In other words, subjects who arrived in the United States at the same age but received more years of education were either older at time of testing or did not continue with higher education for personal reasons. Moreover, if AOA was matched and L1/L2 daily usage was manipulated, individual language learning abilities became a potential confounding variable because both L1/L2 daily usage and task performance might be affected by individual language learning abilities.

Despite difficulties in experimental control in the studies of age effects in second-language learning, there are better ways of confronting this issue than what has been done so far. For example, one may study age effects in an English-as-a-foreign-language (EFL) environment. When the second language is also the dominant official language, some variables are inexorably confounded with age of learning onset. However, in EFL, English learning is highly constrained in the classroom and often not mandatory; therefore, it is possible to separate different effects such as length of learning and starting age.

Taiwan is an EFL environment. English education policy in Taiwan is undergoing rapid changes. Until quite recently, English was taught in state schools only from the first year of junior high school when students are generally around 12 years old. Because of the profit involved, somewhere between 10 to 20 years ago, private schools of English language instruction, aimed at elementary school or even younger children, were hastily established throughout the island. Some elementary-school children take courses in English language instruction from these private schools as their after-school activities. Some younger children even attend exclusively English-instruction preschool/kindergarten programs offered by these schools.

Because these instructions are private and not mandatory in regular schools, and because English is not used in everyday life, we can find adults who were initially exposed to English as a foreign language early in life either formally in private institutes or informally at home by parents, but did not then continue with further learning until they were older; for example, when they reached 12 years of age where formal instruction of English begins in secondary school. We can also find subjects who received English immersion instruction for a short period of time at childhood, but did not continue with additional instruction of English. Two variables, starting point and duration, which are usually highly

correlated in studies conducted in English-speaking countries, are thus separable for some subjects in Taiwan.

To evaluate age effects, many aspects of linguistic competence may be used. Among them, phonological knowledge may be one of the most basic ones, due to its essential role in verbal communication. Many studies concerning speech accent support strongly a correlation between AOA and speech production attainment (e.g., Flege, 1995; Patkowski, 1990; see also Harley & Wang, 1997, for summary). Some studies have even claimed that, on phones appearing in both native and target languages but with a slight difference, only learners starting younger than 6 years of age can produce native-like pronunciation (Oyama, 1976). However, near-native performance can be found in some late starters as well (Bongaerts, van Summeren, Planken, & Schils, 1997; Ioup, Boustagui, El Tigi, & Moselle, 1994). These findings imply that phonological competence may not be entirely a function of learning duration. Furthermore, studies of speech perception have also shown that, without special training (in some cases even with special training), adults are often unable to discriminate fine differences in speech sounds that are not phonemic in their native language (e.g., Best, McRoberts, & Sithole, 1988; Strange, 1995). There are two principal modes of human auditory perception: the continuous mode and the categorical mode (Brown, 2000; Jusczyk, 1997; Wode, 1994). The continuous mode detects acoustic signals in speech, such as loudness and pitch on a gradual scale; the categorical mode helps discriminate and identify linguistic/phonetic categories and "operates on an all-or-none basis" (Wode, 1994; p. 328). Cooperation of the two modes enables the auditory system to process efficiently speech input. Given that all infants must initially be equipped to differentiate all phonemes, their perceptual competence, or at least part of it, must either have declined or become less accessible at certain point(s) in development. Phonetic category boundaries are influenced by their ambient language and become more and more language specific. Infant studies on cross-linguistic speech perception generally support this view. Results from many studies (Logan, Lively, & Pisoni, 1991; Werker, 1995; Werker & Lalonde, 1988; Werker & Tees, 1983, 1984; Wode, 1994; Wode, 1995) show that by 10 months of age, infants begin to lose or become less able to access some perceptual competence on discriminating non-native phonemic contrasts.

As cited above, findings on adults' L2 speech accent are inconclusive, and L2 speech perception has not been assessed directly in relation to English learning history under EFL settings. In addition, there are many unresolved issues due to confounding variables in ESL environments. It should theoretically be important to examine directly the relationship between second-language learning history and adult language learners' speech perception in an EFL environment such as Taiwan.

This study aims to determine how different childhood English learning experience affects the auditory competence of Taiwanese university students in distin-

guishing English minimal pairs. Furthermore, one of the primary objectives of the current study is to take advantage of the special conditions under EFL to partition possible influences from various variables. Specifically, the different types of effects of learning English at three different types of starting age (1, 2, and 3) and two learning durations (4 and 5) were investigated: 1. age at which they were first exposed and were able to recognize the English alphabet (ALPHAGE)[5]; 2. age at which they were taught English formally (FORMALAGE); 3. age at which they were taught English in English language immersion settings (IMMERAGE); 4. numbers of years in learning English under immersion settings (YRSIMM); and 5. numbers of years of English learning in formal/classroom instruction (YRSFORMAL).

Attending English language immersion programs, starting English instruction earlier, and recruiting native speakers as instructors are financially costly. It is therefore sensible to clarify how these early learning experiences affect the speech perception competence of L2 learners. In addition, contrary to predictions of learnability, positive evidence in terms of phonology (i.e., mere exposure to speech of native speakers) does lead to acquisition. On the other hand, lack of phonological attainment is also quite common in cases of adult language learners (Young-Scholten, 1995). Thus, the role of native speaker input is worthy of closer examination with age at exposure taken into account. Another point of interest concerns how native language background of dialect-speaking English learners affects their competence in English minimal pair discrimination. It is logical to assume that the broader phonological representation repertoire of multilinguals should be beneficial to learning a new language. However, informal observation of local dialect speakers does not seem to agree with this assumption. Therefore, this issue will be tested empirically.

Regarding childhood experience with English, several issues are investigated in this study. Specifically, does language immersion experience affect the competence of Taiwanese university students in distinguishing English minimal pairs? Can early starters of English better distinguish English minimal pairs? Does instruction by native speakers significantly enhance the competence of Taiwanese university students in perceiving English phonemes? And finally, does the native language background of subjects affect their competence in distinguishing English minimal pairs?

Concerning the first issue, we would like to take a closer look at the relationships between early English immersion experience and its long-term effects. With the hope that by comparing the performance of subjects who either had or

[5] We chose the age at which subjects were first exposed to the English alphabet (ALPHAGE) as an index of mere English exposure because of the following two reasons: 1. in informal pre-experimental interview, some Taiwanese university students claimed that they could recognize the alphabet before receiving formal instruction due to various reasons, for example, from playing games with parents; 2. it is clear-cut and thus serves as another objective starting point, which can be independent of and parallel to the other two relatively more formal starting points.

had not learned English by immersion methods, the utility of "language immersion learning settings" for speech perception can be examined. Many parents in Taiwan send their youngsters to immersion English programs that charge exorbitant tuition; these parents believe that language learning via immersion is more effective. It is therefore important to investigate this issue in this study.

On the second issue, by utilizing variance partitioning methods, it is hoped that effects of starting age and total years of instruction can be separated. Previous studies have shown that younger language learners acquire a second language at a relatively slower speed (Snow *et al*., 1978). The findings imply that lowering the starting age of second/foreign language instruction may not be worth the cost, both in terms of possible interference with native language acquisition and educational expenses. On the other hand, if unique age-related advantage could be found, education policymakers might need to take this effect into consideration when formulating policies on English instruction in general.

Regarding the third issue, accent-free auditory input (e.g., native speaker's input) may be crucial for sharpening speech perception capabilities. As many parents and educators in Taiwan believe, native speakers of English can better teach English than non-native teachers. Hence, many native speakers of English have been employed to teach preschoolers, kindergarteners, and elementary school students. But many of them do not have proper teaching credentials to teach ESL/EFL and related areas, and some do a poor job in teaching these youngsters. This lack of proper training is believed by many local, formal preschool/kindergarten and elementary school teachers to be the cause of behavioral incompetence of many school-aged children who previously attended English immersion preschool/kindergarten. Therefore, examining the effects empirically in this study could provide concrete evidence to help parents to decide where their best interests lie.

As to the fourth issue, data concerning native linguistic background of subjects were collected and compared with their results. Although Mandarin Chinese is the official language in Taiwan, the dialects Taiwanese and Hakka are also commonly used in daily activities. Therefore, some of our subjects were multilingual Chinese speakers. However, some Taiwanese are raised in a Mandarin-only family and grow up in a Mandarin-only community, so they are monolingual Mandarin speakers. There are yet no documented findings regarding the relationship between local linguistic groups and proficiency in EFL. This question will be explored in our study in order to shed some light on this issue.

Instead of single phoneme detection, minimal pairs of English words were used in the discrimination task for two reasons. First, in minimal pairs, both the identity and the position of the phonemes could be examined because some consonants are phonemic in both Chinese and English in terms of identity, but not so in terms of position. To be able to manipulate separately position and identity is very useful. The importance of this issue can be seen in Flege's notion of

phonetic prototypes or prototypical representations (Flege, 1992), which have separate representations based on positions (e.g. initial/p-/ and final/-p/). Second, testing auditory perception at the word level can avoid unrelated effects due to background knowledge of subjects. To distinguish efficiently two words of a certain minimal pair in their auditory forms, subjects must focus on the contrastive phonemes of the pair rather than on their prior semantic knowledge. The key is that subjects cannot efficiently guess what the phonemes are just by knowing what the two words mean if these two words cannot be distinguished from each other in their auditory forms. For similar reasons, only the most frequently used words were included in our study (word familiarity/use was determined according to frequency information provided in a dictionary; to be discussed in the following section), so that no cues would be available other than phonological ones. Not all possible identity and position combinations were covered due to the stimulus control listed above.

EXPERIMENT 1

Methods

Subjects

Sixty-six Introductory Psychology students (33 females; 33 males) at National Taiwan University volunteered for the experiment. The age of subjects ranged from 17 to 24 years, with an average age at 20.30 years and a standard deviation of 1.24 years. None of the subjects were native English speakers. Twenty-four of our subjects were monolingual native Mandarin speakers; 42 were multilingual native Mandarin and Taiwanese or Mandarin and Hakka speakers.

Materials

Twenty-eight English minimal pairs were selected as stimuli for the study. These pairs were obtained by the following procedures:

(1) Choosing phonemes: All English phonemes were first listed, then some of these phonemes were chosen to form contrasts according to their phonetic features. Only contrasts with two or less phonetic features were chosen to construct minimal pairs.
(2) Constructing minimal pairs: According to procedures in (1), 28 minimal pairs were generated (see Appendix). Words in these pairs were all among the most frequently used (spoken/written) 3000 words based on information provided in *Dictionary of Contemporary English* published by Longman Group Ltd (1995). Each phonemic contrast was used only once in one position to form these minimal pairs.

The only exceptions involved the following pairs: fly versus fry, glass versus grass, and dry versus try. These three pairs contain a consonant cluster in initial position (C1-C2 in short; C1 refers to the first consonance in the cluster combination, while C2 refers to the second), with C2 all being/r/versus/l/. The three pairs were all included because there is no identical consonant cluster combination in Mandarin Chinese and also because there is no specific reason to exclude any one of these three pairs. In addition, 15 out of the 28 minimal pairs had phonetic contrasts as initial consonants (i.e., Vowel (V) of V-C, C of C-V, and C1 or C2 of C1-C2-V). This was designed to control for the possible effects of initial versus final position.

A 26-year-old native male English speaker from Ohio with a Midwestern American accent was recruited to record the words. An Acer Extensa 712TE computer and WAVE digitalized sound recording software system were used for the recording and storage of the sounds. Further sound file editing was done through Cool Edit 2000. The sound intensity of the stimuli for speakers was between 68 db to 77 db with an average of 72.05 db. If measured at a middle point between the two speakers, the intensity is approximately 5 db lower than the intensity measured at the speakers. A complete list of the 28 pairs and their individual amplitude information can be found in the Appendix.

These sound files were then compiled into 168 trials based on the following rules.

(1) Every pair of words was tested as six different tri-word trials. For, example, for a pair containing words A and B, the six tri-word trials would be AAB, ABA, BAA, BBA, BAB, and ABB. Because there were 28 pairs, the total trials amounted to 168.

(2) For each subject, these 168 trials were presented in a random order and Visual Basic was used to program the randomization and the test procedure.

(3) Within each trial, the three-word combination was read out loud word by word. At the end of the third word, the computer started to time response time (in seconds) until subjects hit one of the three keys (1, 2, or 3) on the keyboard. The subject's task was to identify the odd member of this trio of words. The between-trial interval was 1 second, which started right after the subjects hit one of the target keys.

Procedures

Every subject was first given verbal instructions concerning the test. Examples were also given to ensure a subject's comprehension of the instructions. After the test, which usually took around 15 minutes, an interview was conducted to collect information on the subject's English learning history and other related matters. During the interview each subject was asked to recall 1. at what age they were first

exposed to the English alphabet and able to recognize all of them, formally or informally (ALPHAGE); 2. at what age they were taught English formally, which included tutoring, private institute, and formal school (FORMALAGE); 3. how many years of formal/classroom English instruction they had received (YRS-FORMAL); 4. if applicable, at what age they were taught in English language immersion settings (IMMERAGE; the number was then inversed and labeled as INVIMMAGE because some subjects never had any immersion experience and their starting age of immersion experience was arbitrarily set as infinity; their scores of INVIMMAGE were thus zero); 5. how many years in total they were taught English under immersion settings (YRSIMM); and 6. whether they were taught by native English speakers (NS).

Results

Summary of Descriptive Statistics

A descriptive summary of the variables is given in Table I. The mean score of correct response by all subjects is 153.68 (out of 168; SD = 7.28). The mean age at which subjects could recognize the English alphabet is 8.6 years. However, on average, these subjects did not start formal English class until around 9.4 years of age, and only some of them did it continuously. Therefore, we now have a group of EFL learners whose starting age does not necessarily co-vary with formal learning duration, and thus variance partitioning can be executed.

However, the average score of correct response is fairly high. In fact, 21 out of the 28 minimal pairs were correctly distinguished more than 90% of the time. Therefore, ceiling effects must be a real possibility if no significant effects are found.

Performance Difference Between Subgroups

We performed a series of analyses concerning the following questions.

(1) Did students with experience in an immersion program perform better on the auditory perceptual tests?
(2) Did students who have had native speakers as instructors perform better on the test?
(3) Did students having started formal English instruction before junior high school perform better than those who began learning English at junior high?
(4) Did students with exposure to the English alphabet before junior high school perform better on the auditory tests?
(5) Did native Mandarin speakers of English perform differently from their multilingual peers (Mandarin/Taiwanese or Mandarin/Hakka speakers of English)?

Table I. Experiment 1 Summary Statistics of the Variables ($n = 66$)

Variables	Mean	SD	Frequency
Age at test (years)	20.30	1.24	
ALPHAGE (years)[a]	8.58	2.36	
YRSFORMAL (years)[b]	9.18	1.70	
FORMALAGE (years)[c]	9.39	1.94	
YRSIMM (years)[d]	1.56	1.93	
INVIMMAGE[e]	0.061	0.057	
SCORES[f]	153.68	7.28	
RT (reaction time, total)[g]	0.54	0.42	
CRRT (reaction time, only correct)[h]	0.44	0.25	
Native language			
Mandarin			24
Mandarin and Taiwanese/Hakka			42
Native speaker as instructor			
Yes			32
No			34

[a] Age at which subjects were able to recognize the English alphabet.
[b] Total years of formal English instruction, including schools and "cram" schools.
[c] Age at which subjects started formal English instruction, including schools and "cram" schools.
[d] Total years of immersion experience (i.e., English used in instruction or daily conversation exclusively).
[e] Inversed quantity of age at which subjects started immersion experience.
[f] Out of 168.
[g,h] In seconds.
[h] Including only reaction time of correctly answered trials.

Findings are listed in Table II. Results of these tests together seem to indicate that Taiwanese university students who were exposed to the English alphabet before junior high school or who were taught by native English speakers, or who were in English immersion programs, performed significantly better in our task. However, on average, those who started taking formal English courses before junior high school did not perform significantly better. This finding implies that some of the late starters of formal English courses, who, nevertheless, learned the alphabet early, performed well enough so that a subgroup difference was found in performance regarding ALPHAGE, but not FORMALAGE. To further explore this finding, we divided the subjects into three groups: 1. late FORMALAGE and late ALPHAGE; 2. late FORMALAGE but early ALPHAGE; and 3. early FORMALAGE and early ALPHAGE. Means and standard deviations of these three groups are listed in Table III. The average performance of group 1 was not as good as that of the other two groups [1 vs. 2: $t = -1.592$, d$f = 11$, $p \leqq 0.07$ (one-tailed); 1 vs. 3: $t = -2.124$, d$f = 58$, $p \leqq 0.019$ (one-tailed)] while groups 2 and 3 did not show significant difference in test ($t = 0.308$, d$f = 57$; $p \leqq 0.759$ (two-tailed)).

Table II. Experiment 1 Performance Difference Between Subgroups

Subgroups (Mean Scores, SD, n)	t score	df	p
(1) With vs. without immersion experience (155.29, 5.83, 34) vs. (151.97, 8.31, 32)	1.87	55.24	$\leqq$0.05 (one-tailed)
(2) With vs. without native speaker instructor(s) (155.78, 5.64, 32) vs. (151.72, 8.14, 34)	2.38	58.99	$\leqq$0.05 (one-tailed)
(3) FORMALAGE before vs. after junior high school (154.25, 6.95, 53) vs. (151.38, 8.42, 13)	1.28	64	0.14 (one-tailed)
(4) ALPHAGE before vs. after junior high school (154.34, 6.90, 59) vs. (148.14, 8.65, 7)	2.19	64	$\leqq$0.05 (one-tailed)
(5) Mono- vs. multilingual (155.33, 6.55, 24) vs. (152.74, 7.59, 42)	1.40	64	0.17 (two-tailed)

Table III. Experiment 1 Means and Standard Deviations of Three Groups

Groups	n	Mean	SD
(1) Late ALPHAGE and late FORMALAGE	7	148.14	8.65
(2) Early ALPHAGE but late FORMALAGE	6	155.17	6.97
(3) Early ALPHAGE and early FORMALAGE	53	154.25	6.95

Correlation and Variance Partitioning

Table IV presents Pearson's correlation coefficients between mean score (SCORES), subjects' current age (AGE), ALPHAGE, FORMALAGE, YRSFORMAL, INVIMMAGE, and YRSIMM. Although AGE is not significantly correlated with SCORES ($r = 0.13$, $n = 66$; $p \leqq 0.32$), it is moderately correlated with ALPHAGE ($r = 0.33$, $n = 66$; $p \leqq 0.006$). It implies that younger subjects in our group are more likely to be exposed to the English alphabet earlier in their life, indicating an increasing popularity in Taiwan of learning English at an early age.

The three starting points (ALPHAGE, FORMALAGE and INVIMMAGE) and the two durations (YRSFORMAL and YRSIMM) are, as expected, correlated with one another as shown in Table IV. This finding can be accounted for via variance partitioning. However, some other findings are somewhat unexpected. For example, out of the 24 native Mandarin speakers of English, there are no late starters. This may be related to the socioeconomic status of this group of subjects.[6] We also

---

[6] The majority of the residents in Taiwan are multilingual speakers (Mandarin and at least one dialect, mostly Taiwanese). Only a few are monolingual native Mandarin speakers. Most monolingual native Mandarin speakers in Taiwan are either originally from mainland China or raised by parents who are mainlanders. Although fewer in number, mainlander' socioeconomic status is on average higher than that of the rest of the residents in Taiwan (Huang, 1994).

Table IV. Experiment 1 Intercorrelations of the Variables ($n = 66$)

Variable	1	2	3	4	5	6	7
(1) Age	—						
(2) ALPHAGE	0.33^a						
(3) YRSFORMAL	-0.10	-0.43^a					
(4) FORMALAGE	0.14	0.76^a	-0.68^a				
(5) YRSIMM	-0.19	-0.45^a	0.65^a	-0.49^a			
(6) INVIMMAGE	-0.23	-0.43^a	0.49^a	-0.44^a	0.80^a		
(7) SCORES	0.13	-0.28^b	0.23	-0.33^a	0.28^b	0.29^b	

[a] Correlation is significant at the 0.01 level (two-tailed).
[b] Correlation is significant at the 0.05 level (two-tailed).

found that 17 of these 24 subjects had been taught by native speakers of English whereas only 15 of the 42 multilingual subjects had been so taught. It is not our main research interest in this study to propose explanations for these curious results, but these findings do indicate the fact that these variables are not entirely independent of each other.

Variance partitioning procedure is appropriate for dealing with correlated variables. The analysis procedure makes use of forced-entry regression. By entering the variable in concern (or group of variables) last, we can partition the variance into several parts. This method is conservative in the sense that if significant effects of the last-entered variable (or group of variables) can be shown, the effects are usually very robust because the effects of the last-entered variable or block of variables represent its independent contribution to the criterion variable.

Five blocks of variables were entered in the following order:

(1) Current age and gender
(2) Native language
(3) Whether or not being taught by native speakers of English (NS)
(4) YRSIMM and YRSFORMAL
(5) FORMALAGE, ALPHAGE, and INVIMMAGE

Results are given in Table V. The change of R^2 values indicate 7.6% of the variances uniquely were contributed by the last block of variables; that is, the three starting points combined. However, it is not statistically significant ($p \leqq 0.114$). We also conducted a few more forced-entry regressions, with various variable(s) entered in the last step (results not shown in tables). Results show that NS uniquely contributes only 0.8% of the total variance. When YRSIMM and INVIMMAGE were entered last as a group in another forced-entry regression procedure, only 3.1% of the variance was uniquely accounted for by these two variables together.

Table V. Experiment 1 Forced-Entry by Block Regression Analyses

	Model summary				
Predictors entered by block	R^2 change	F change	df1	df2	Significant F Change
Age and Sex	0.047	1.566	2	63	0.217
Native language	0.026	1.714	1	62	0.195
NS	0.066	4.697	1	61	0.034
YRSIMM, YRSFORMAL	0.032	1.145	2	59	0.325
FORMALAGE, ALPHAGE, INVIMMAGE	0.076	1.874	3	56	0.144

Discussion

Statistically significant age effects were not found in this experiment according to variance partitioning analyses. However, the three starting points in our study did uniquely contribute more than 7% of the variance. One possible reason for the absence of significant age effects could be that the minimal pair test in this study was too easy and thus not sensitive enough to detect any age effects.

Our data also show that although we found both a native-speakers-as-instructor (NS) effect and some starting age effects in subgroup *t* tests and/or correlation analyses, variance partitioning analyses found no significant independent effects. Obviously, significant difference from comparing the NS with the non-NS group may be due to the combined influence of several correlated factors. Therefore, it is misleading to consider NS alone as a valid factor in auditory perceptual competence. No firm conclusion can be drawn on this specific variable unless we can control for possible confounding effects. The results of correlation analyses show that some of the variables in our study (ALPHAGE, FORMALAGE, INVIMMAGE, and YRSIMM) are significantly correlated with speech perception performance under quiet testing conditions. It also shows that YRSFORMAL is just barely correlated with SCORES. These findings mostly converge with results of subgroup *t* test analyses.

Furthermore, descriptive data indicate that learning English at a younger age has become increasingly popular in Taiwan. Most of our subjects had some exposure to English before junior high school, and some of them even received English instruction continuously from a very young age. Because their English language input is very different from that of ESL settings (Flege & Liu, 2001; Yamada, 1995), our study reveals a less often explored realm of second-language learning by utilizing an EFL learning environment. Our ultimate goal is to test findings of studies conducted in ESL settings with a different group of language learners. If age effects can still be found with confounding variables controlled, then this effect should be taken seriously.

EXPERIMENT 2

One way of increasing overall level of test difficulty without changing stimulus content is to introduce artificial background noise. A study testing early and late Spanish–English bilinguals' speech perception under noise has indicated that early bilinguals did better than late bilinguals (Mayo, Florentine, & Buss, 1997). This study, though well conceived, did not control the age of subjects at the time of test. Thus, all the late starters, who were native Spanish speakers and have learned English as a second language, were much older than their early-starting counterparts. It would be interesting to see if early starters who have learned English as a foreign language in Taiwan would perform better than late starters of the same age, under artificial noise background.

Methods

Subjects

Sixty-six Introductory Psychology students (35 females; 31 males) who did not participate in Experiment 1 volunteered for this experiment. The age of the subjects ranged from 18 to 24 years, with the mean and standard deviation at 20.45 and 1.33 years, respectively. None of subjects were native English speakers. Fifteen of them were monolingual native Mandarin speakers, and 51 were multilingual native Mandarin and Taiwanese or Mandarin and Hakka speakers.

Materials and Procedures

All materials and procedures were identical to those of Experiment 1, except that a white noise generated with Cool Edit 2000 software with frequency ranging from around 40 to 220 Hz was tape-recorded and superimposed upon the test trials.[7] The noise was played by an AIWA stereo headset connected to the speakers. Amplitude of the white noise was set at 83 db.

---

[7] To ensure the equivalence of subjects of Experiments 1 and 2, 23 Introductory Psychology students (12 females; 11 males) at National Taiwan University volunteered for the experiment. The age of the subjects ranged from 18 to 22 years, with an average at 20.17 years and a standard deviation of 1.23 years. All the subjects were local native Mandarin/Taiwanese/Hakka speakers. All materials and procedures were identical to those of Experiment 1, except that a different computer, an IBM Thinkpad notebook, was used. The overall pattern of the results was very close to that of Experiment 1. The mean scores of the 23 local subjects tested under the transitional condition are 154.74, with SD of 7.33. The findings almost replicated the findings of Experiment 1 (i.e., mean = 153.28, SD = 7.68); both order of difficulty and test scores were very close to those of Experiment 1. Therefore, we continued with Experiment 2 in which white noise background was superimposed upon the testing setting of this transition experiment without changing any of the soft- and hardware.

Results

Summary of Descriptive Statistics

A summary of descriptive statistics of the 66 subjects is shown in Table VI. All values of the independent variables of the two samples (tested under noise and no-noise conditions, respectively) were compared to determine the comparability of these two samples. Chi-square test shows that proportions of bilingual versus multilingual and NS versus non-NS are different between these two samples (bilingual vs. multilingual: $\chi^2 = 5.304$, d$f = 1$, $p \leqq 0.05$; NS vs. non-NS; $\chi^2 = 6.066$, d$f = 1$, $p \leqq 0.05$). However, other tests did not reveal any additional significant differences between these two samples.

The mean response score of the subjects is 121.74 (out of 168; SD = 9.38), which is much lower than the mean score under the quiet testing condition. Around 75% of the subjects started learning English formally before junior high school, and more than half of these subjects have learned English in immersion settings. The mean age at which the English alphabet was recognized is around 8.5 years of

Table VI. Experiment 2 Summary Statistics of the Variables ($n = 66$)

Variables	Mean	SD	Frequency
Age at test (years)	20.30	1.24	
ALPHAGE (years)[a]	8.50	2.47	
YRSFORMAL (years)[b]	9.02	1.87	
FORMALAGE (years)[c]	9.33	2.31	
YRSIMM (years)[d]	1.33	1.58	
INVIMMAGE[e]	0.052	0.059	
SCORES[f]	121.74	9.38	
RT (reaction time, total)[g]	0.60	0.29	
CRRT (reaction time, only correct)[h]	0.48	0.23	
Native Language			
Mandarin			15
Mandarin and Taiwanese/Hakka			51
Native speaker as instructor			
Yes			42
No			24

[a] Age at which subjects were able to recognize the English alphabet.
[b] Total years of formal English instruction, including schools and "cram" schools.
[c] Age at which subjects started formal English instruction, including schools and "cram" schools.
[d] Total years of immersion experience (i.e., English used in instruction or daily conversation exclusively).
[e] Inversed quantity of age at which subjects started immersion experience.
[f] Out of 168
[g,h] In seconds.
[h] Including only reaction time of correctly answered trials.

age. However, most of them did not start formal English class before 9.3 years of age on average. We now again have a group of EFL learners whose starting age of learning English does not totally co-vary with formal learning duration.

Perceptual difficulties were sorted in order of hard to easy. The results showed that some pairs with more than 90% correct response rate in Experiment 1 are now correctly discriminated only around 80% of the time or less. Almost all of these pairs contain consonant contrasts. This finding seems to imply that perception of consonant contrasts is affected by noise to a larger extent than that of vowel contrasts.

Performance Difference Between Subgroups

Tests for differences between various subgroups were conducted as in Experiment 1. Results are presented in Table VII. Subjects who had been taught by native speakers of English did significantly better, as under quiet conditions; so did subjects who had been exposed to the English alphabet. However, unlike the t test results of Experiment 1, which indicates that immersion experience is a valid factor and formal instruction starting age is not, performance under noise revealed a different pattern regarding these two variables. Under noise background, the group with early immersion experiences did not significantly outperform the group without the same experience whereas the group with early formal instruction

Table VII. Experiment 2 Performance Difference Between Subgroups

Subgroups (Mean scores, SD, n)	t score	df	p
(1) With vs. without immersion experience (122.97, 9.64, 38) vs. (120.07, 8.91, 28)	1.25	64	0.11 (one-tailed)
(2) With vs. without native speaker instructor(s) (124.71, 9.43, 24) vs. (120.05, 9.02, 42)	1.99	64	$\leqq 0.05$ (one-tailed)
(3) FORMALAGE before vs. after junior high school (123.31, 9.34, 49) vs. (117.24, 8.15, 17)	2.38	64	$\leqq 0.01$ (one-tailed)
(4) ALPHAGE before vs. after junior high school (122.78, 9.28, 57) vs. (115.22, 7.55, 9)	2.32	64	$\leqq 0.05$ (one-tailed)
(5) Mono- vs. multilingual (121.13, 9.08, 15) vs. (121.92, 9.55, 51)	1.40	64	0.17 (two-tailed)

outperformed their late-starting counterpart. Due to various covariations between these variables, these findings require further analysis.

Matched-Group Testing

Two subgroups were selected from the sample based on their total years of formal instruction (YRSFORMAL) and years at which they were first exposed to the English alphabet (ALPHAGE), with matched YRSFORMAL but variable ALPHAGE. Six subjects with exposure to the English alphabet before 6 years of age were selected, but they had not received English instruction on a continuous basis. Therefore, their formal English instruction merely adds up to around 7 years in total (mean = 7.50, SD = 0.55). Another group of eight subjects who were not exposed to the English alphabet before 12 years of age was identified. They also received a total of about 7 years of formal English instruction by the time of this experiment (mean = 7.13, SD = 0.35). YRSFORMAL between these two groups are not statistically different ($t = 1.46$, d$f = 8.05$; $p \leqq 0.18$, two-tailed). Performance difference between the early and the late ALPHAGE group was analyzed. It was revealed by t test that the early ALPHAGE group did better on the discriminating task than the late ALPHAGE group ($t = 2.44$, d$f = 12$; $p \leqq 0.05$, one-tailed). This result suggests that, with total years of formal English instruction held constant, subjects who were exposed to the English alphabet early in life had a significant advantage in phonemic discrimination later on at university age.

Correlation and Variance Partitioning

Intercorrelations between variables are listed in Table VIII. The overall correlation pattern is different from that of Experiment 1. Although scores are still significantly correlated with ALPHAGE ($r = -0.38$, $n = 66$; $p \leqq 0.01$), YRSFORMAL ($r = 0.28$, $n = 66$; $p \leqq 0.05$), and FORMALAGE ($r = -0.37$, $n = 66$; $p \leqq 0.01$),

Table VIII. Experiment 2 Intercorrelations of the Variables ($n = 66$)

Variable	1	2	3	4	5	6	7
(1) Age	—						
(2) ALPHAGE	0.13						
(3) YRSFORMAL	-0.27^a	-0.53^b					
(4) FORMALAGE	0.11	0.76^b	-0.75^b				
(5) YRSIMM	-0.27^a	-0.40^b	0.48^b	-0.33^b			
(6) INVIMMAGE	-0.13	-0.45^b	0.56^b	-0.47^b	-0.74^b		
(7) SCORES	-0.03	-0.38^b	0.28^a	-0.37^b	0.16	0.13	

[a] Correlation is significant at the 0.05 level (two-tailed).
[b] Correlation is significant at the 0.01 level (two-tailed).

the correlations between correct responses and YRSIMM ($r = 0.16$, $n = 66$; $p > 0.05$) and INVIMMAGE ($r = 0.13$, $n = 66$; $p > 0.05$) are no longer statistically significant.

To better understand the relationship between ALPHAGE and SCORES and between FORMALAGE and SCORES, two partial correlation coefficients were calculated with YRSFORMAL controlled. Results show that both partial correlation coefficients are significant: correlation coefficient between SCORES and ALPHAGE with YRSFORMAL controlled is -0.29 ($p \leqq 0.021$, $n = 66$); correlation coefficient between SCORES and STRTFMRL with YRSFORMAL controlled is -0.26 ($p \leqq 0.04$; $n = 66$).

Results of variance partitioning via forced-entry regression converge with the findings of partial correlation. As shown in Table IX, two sets of regression analyses were separately performed. Both sets of variance partitioning were performed with the first four blocks [containing (1) sex and age, (2) native speaker instructors (NS), (3) native language, and (4) YRSIMM and INVIMMAGE] entered in the same order and with the order of the last two blocks switched [containing (1) FORMALAGE and ALPHAGE and (2) YRSFORMAL]. Results show that the first four blocks together contributed 7% of the response variation. ALPHAGE and FORMALAGE as a block, if entered last, independently contributed 10.0% of the variance [F(2,56) = 3.629, $p \leqq 0.05$]. This finding suggests that under noise, the two starting points together can significantly predict current performance of subjects on the auditory tasks. On the other hand, YRSFORMAL alone contributed 3% of the variation if entered last. However, this unique contribution is not statistically significant. In addition, NS alone contributed only 2.9% of the variance if entered

Table IX. Experiment 2 Forced-Entry by Block Regression Analyses

Predictors entered by block	R^2 change	F change	df1	df2	Significant F change
Model summary 1					
Age and sex	0.002	0.048	2	63	0.953
Native language	0.001	0.051	1	62	0.822
Native speaker instructors	0.061	3.949	1	61	0.051
YRSIMM, INVIMMAGE	0.007	0.229	2	59	0.796
FORMALAGE, ALPHAGE	0.155	5.696	2	57	0.006
YRSFORMAL	0.003	0.195	1	56	0.661
Model summary 2					
Age and sex	0.002	0.048	2	63	0.953
Native language	0.001	0.051	1	62	0.822
Native speaker instructors	0.061	3.949	1	61	0.051
YRSIMM, INVIMMAGE	0.007	0.229	2	59	0.796
YRSFORMAL	0.057	3.820	1	58	0.055
FORMALAGE, ALPHAGE	0.100	3.629	2	56	0.033

last (result not shown in tables), and it is not statistically significant. Besides, both correlation and forced-entry regression analyses indicate that immersion experiences (including both duration and starting point information) are not statistically valid predictors of performance under noise testing condition. Finally, a stepwise regression analysis was conducted as well. Only ALPHAGE was entered in the model with F(1,64) = 10.99 and $p \leqq 0.05$.

Subtypes of Stimuli on Performance

To test differential effects, if any, of various stimuli on performance, the 28 pairs of minimal pairs were divided into four subtypes in terms of characteristics of their contrastive phonemes. The four subtypes of stimuli are 1. vowel (V), 2. single consonant at initial position (CI), 3. single consonant at ending position (CE), and 4. second consonant of a cluster (C2). Mean correct responses are listed in Table X. Among them, mean correct performance of C2 is much lower than that of the others.

Although there are only two minimal pairs under C2, a series of forced-entry regression analyses were nevertheless performed on C2. It was found that the total percentage of variance accounted for is above 41%. It is much higher than those of the other three subtypes, which ranged from 14% to 25%. Variance partitioning shows that 22% of C2 score variance is uniquely accounted for by INVIMMAGE, 9.8% uniquely by FORMALAGE, and 5.7% uniquely by YRSIMM. YRS-FORMAL and ALPHAGE accounted for less than 1% of the total variance of C2 independently.

This pattern of results differs from that of the overall variance partitioning mainly in that immersion experience plays an essential role on distinguishing C2 minimal pairs, but not on other types of minimal pairs. In contrast, variance partitioning analyses of the other three subtypes did not reveal any independent effects by any predictors. Because there are only two minimal pairs under C2, further investigation is needed to examine the phenomenon. Nevertheless, this tentative result suggests that English consonant cluster contrasts are harder to discriminate than that of V, CI, and CE for adult native Chinese speakers of English.

Table X. Results of Four Subtypes ($n = 66$)

Subtypes	Mean[a]	SD
Vowel	4.58	0.40
CI	4.31	0.39
CE	4.37	0.56
C2	3.13	0.85

[a] Mean: mean-item correct out of six trials.

Discussion

The second experiment was an attempt to increase overall degree of difficulty of the previous task without altering the stimuli and also to verify findings of Mayo *et al*. (1997) regarding speech perception under noise. Our findings show that age effects can be detected if the task is made sensitive enough.

As discussed previously in Experiment 1, there was no statistically significant effect uniquely accounted for by age-related factors. The lack of age effects was suspected to be the result of a ceiling effect. And by increasing the difficulty level of the task, age effects did show up in this second experiment.

Overall performance under noise condition decreased and variation increased in comparison to no-noise condition. Descriptive analysis indicates that we again have a group of subjects whose starting points do not totally covary with learning duration. Results of subgroup *t* test analyses, matched-group analysis, and correlation analyses basically converge with one another. Age of exposure remains highly correlated with test scores as in Experiment 1. Both matched-group method and partial correlation results support this finding. Another starting point that is significantly correlated with performance level is the age that formal instruction of English starts. The results indicate that the earlier a learner receives formal instruction, the more likely he/she can distinguish minimal pairs under noise distraction.

According to the results of correlation analysis, although duration of formal instruction is significantly correlated with task performance, it was found that those who could recognize the English alphabet very early in life are more likely to outperform on the auditory task, even when duration of formal instruction was held constant. In addition, native-speakers-as-instructors seems to play an important role here according to correlation analysis, but not so according to variance partitioning analyses.

Variance partitioning results confirm the uniqueness of these two starting points as predictors of auditory competence in discriminating English minimal pairs. The results imply that, under noise-testing condition, relatively earlier English learning experience, such as age of exposure and formal instruction starting age, are better predictors than immersion learning experiences occurring relatively later.

Another important finding of our study concerns consonant clusters. Results of the four subtypes of stimuli indicate that vowel contrast is the easiest among them. Consonant contrasts in initial and final positions are equally hard to distinguish under noise, but not as hard as consonant cluster contrast. This finding resembles data collected from listeners who were native speakers of Japanese (cited in Strange, 1995). Consonants in clusters are the hardest to discriminate possibly due to the lack of such syllable structure in Mandarin. However, it could also be due to the liquid (/l/ vs. /r/) contrast contained in the two tested cluster minimal pairs, as the only two cluster minimal pairs in our stimuli are

both liquid contrasts. In regard to the performance on vowel contrast, Bohn (1995) pointed out that native Mandarin listeners relied almost exclusively on duration cue to solve a vowel contrast identifying task. We suspect that subjects in our study may have at least partially relied on duration cue when they tried to discriminate the contrast. Vowel duration is less distorted by the noise than other spectral cues. This might be the reason why vowel contrast was the easiest to discriminate in this study.

In addition, *t* test results show that subjects with different native language background did not perform differently. Many residents in Taiwan are multilinguals, though Mandarin is the official language. Our findings do not show significant difference between monolingual and multilingual groups. Given the facts that 1. more monolingual subjects in our study are early learners and more of the same group of subjects have been taught by native speakers, and 2. early starters and learners who have been taught by native speakers in our study performed better on the task, we would expect better performance from monolingual subjects. However, our results do not agree with this expectation.

Finally, we observe the fact that NS alone contributed a nonsignificant 2.9% of the variance. One possible explanation is that the effect of genuine input from native speakers is not independent of age effects. Subgroup *t* tests show that Taiwanese university students did perform better if they had been taught by native speakers of English, but this NS effect is no longer significant if other factors are held constant. However, we cannot rule out the possibility that some of our subjects, though not taught by native speakers of English, were exposed to native speaker input outside of the classroom.

GENERAL DISCUSSION

The most important result is that only under noise condition did starting age become an effective predictor of perceptual performance in our tests. This finding basically agrees with the Speech Perception In Noise (SPIN) experiment of Mayo and colleagues (Mayo *et al.*, 1997), although our data were based on language-learning settings that are distinct from those of Mayo. SPIN was conducted with native Spanish speakers of English who had lived in the United States for years. Our study was conducted in Taiwan with local subjects who were exposed to English mostly in the classroom. SPIN used a procedure in which noise level was adjusted until target words (which were the last words of each sentence and were partially masked with noise) were recognized around 50% of the time. Due to the information provided in the sentences, semantic and possibly grammatical knowledge may be involved in word identification. In fact, the SPIN study results showed that early starters benefited from contextual cues more than late starters did on high-predictable test items. In our task, the chance

for listeners to use word meaning to aid their performance was much reduced because all choices were high-frequency words. Furthermore, in the SPIN study, early learners were also younger than late learners when total years of learning were controlled. Thus, difference in performance could be due to cohort differ- ence or better general abilities, such as concentration, of younger subjects.

In terms of the long-term effects of teaching English to Chinese-speaking children before junior high school on auditory perception of spoken English, our findings seem to support the notion "the earlier, the better." Although only correlational data were available, starting points of exposure/formal instruction did independently account for more than 10% of the performance variances under noise background. One of the most likely reasons for failure to reveal sig- nificant age effects under no-noise condition is perhaps the ceiling effect of the tests. Another possible reason could be that, under noise condition, some more obvious acoustic distinctive features of the test stimuli were partially masked, and subjects had to rely on less obvious features to resolve the challenge. For example, rather than solely focusing on contrastive phonemes, they might have to rely on other acoustic cues in words besides the contrastive parts, such as vowels before or after the contrastive consonant.

Learning English in immersion settings and having native speakers of English as instructors were not effective predictor variables on performance variance, even under noise condition, according to all levels of analyses regard- ing NS. Given the high cost of these kinds of instruction arrangement, they do not seem worth the price tag.

According to the results of the subtypes of contrastive stimuli on performance in Experiment 2, the total variance accounted for is much higher in the case of con- sonant cluster. In addition, while the overall score variance partitioning pattern shows that starting age of formal instruction and age of exposure to the English alphabet together accounts for most of the total score variance (10%), starting age of English immersion experience independently accounts for 22% of the consonant cluster score variance. The results on consonant cluster also show that duration of formal instruction and age of exposure to the English alphabet are not predictors in effect if only the two pairs of consonant cluster are considered. This specific find- ing implies that mere exposure to the English alphabet and formal instruction expe- rience per se do not provide learners with sufficient and authentic consonant cluster input. It also implies that immersion experience at an early age can be beneficial for individuals' speech perception on consonant cluster contrast distinction. These unique findings on consonant cluster are thus worth paying more attention to.

As a final note for future research, more empirical studies conducted under settings in which English exposure primarily occurs in the classroom are needed. Given the limited hours of English-as-a-Foreign-Language instruction, learners can benefit in the long run if the curriculum is wisely designed. That is, instead of mere speculation or referring to conclusions drawn from studies

conducted in settings where English is used both inside and outside of the class-room, the contents of initial-level English-as-a-Foreign-Language instruction for younger children should be based on findings concerning age effects in EFL settings, where use of English is primarily restricted to the classroom.

ACKNOWLEDGMENTS

APPENDIX: MINIMAL PAIRS WITH AMPLITUDE

Contents	Amplitude (db)	
fat vs. vat—/f-/ vs. /v-/	70	71
light vs. right—/l-/ vs. /r-/	70	70
fly vs. fry—/l- / vs. /r-/	71	68
glass vs. grass—/l- / vs. /r-/	70	71
there vs. dare—/(ð)-/ vs. /d-/	70	73
think vs. sink—/θ-/ vs. /s-/	71	71
through vs. true—/θ-/ vs. /t-/	72	73
add vs. odd—/æ-/ vs. /a-/	73	71
safe vs. save—/-f/ vs. /-v/	77	75
not vs. nod—/-t/ vs. /-d/	72	72
then vs. them—/-n/ vs. /-m/	72	73
all vs. or—/-l/ vs. /-r/	74	74
back vs. bag—/-k/ vs. /-g/	76	75
bean vs. bin—/-i-/ vs. /-I-/	69	70
butter vs. better—/-^-/ vs. /-ɛ-/	72	72
man vs. men—/-æ-/ vs. /-ɛ-/	72	72
same vs. some—/-e-/ vs. /-^-/	73	71
down vs. town—/d-/ vs. /t-/	69	70
dry vs. try—/d-/ vs. /t-/	76	75
bear vs. pear—/b-/ vs. /p-/	72	72
came vs. game—/k-/ vs. /g-/	72	73
nail vs. mail—/n-/ vs. /m-/	72	73
hide vs. wide—/h-/ vs. /w-/	73	75
yet vs. wet—/j-/ vs. /w-/	75	72
mate vs. met—/-e-/ vs. /-ɛ-/	72	70
sin vs. sing—/-n/ vs. /-ŋ/	71	71
soap vs. soup—/-o-/ vs. /-u-/	71	70
boat vs. bought—/-o-/ vs. /-(ɔ)-/	72	73

REFERENCES

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345–360.

Bohn, O-S (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 279–304). Timonium, MD: York Press.

Bongaerts, T., Van Summeren, C., Planken, P., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, *19*, 447–465.

Brown, C. (2000). The interlanguage between speech perception and phonological acquisition from infants to adults. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory* (pp. 4–63). Malden, MA: Blackwell Publishers.

Dekeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*, 499–533.

Ervin-Tripp, S. M. (1974). Is second language learning like the first? *TESOL Quarterly*, *8*, 111–127.

Flege, J. E. (1992). Speech learning in a second language. In C. A. Ferguson, L. Menn, and C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications* (pp. 565–604). Timonium, MD: York Press.

Flege, J. E. (1995). Second language speech learning. Theory, findings, and problems, In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E., & Liu, S. (2001). The effects of experience on adults' acquisition of a second language. *Studies of Second Language Acquisition*, *23*, 527–552.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*, 78–104.

Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, *21*, 205–228.

Harley, B., & Wang, W. (1997). The critical period hypothesis: Where are we. In A.M.B. Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 19–51). Mahwah, NJ: Lawrence Erlbaum.

Huang, S (1994). *Yong-Yan, She-Hui Yong Zu-Qun-Yi-Shi* [*Language, Society and Ethnicity*]. Taipei, Taiwan: The Crane Publishing Co., Ltd.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study in a naturalistic environment. *Studies in Second Language Acquisition*, *16*, 73–98.

Johnson, J. S. (1992). Critical period effects in second language acquisition: The effects of written versus auditory materials on the assessment of grammatical competence. *Language Learning*, *42*, 217–248.

Johnson J. S., & Newport E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Johnson, J. S., Shenkman, K. D., Newport, E. L., & Medin, D. L. (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language*, *35*, 335–352.

Jusczyk, P. W. (1997). *The Discovery of Spoken Language*. Cambridge, MA: The MIT Press.

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

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l/: A first report. *Journal of the Acoustical Society of America*, *89*, 874–886.

Longman Dictionaries (1995). *Longman Dictionary of Contemporary English*. Harlow, England: Longman Group Ltd.

Marinova-Todd, S. H., Marshall, D. B., & Snow C. E. (2000). Three misconceptions about age and L2 learning. *TESOL Quarterly*, *34*, 9–34.

Mayo, L. H., Florentine, M., & Buss, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech & Hearing Research*, *40*, 686–693.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*, 11–28.

Oyama, S. C. (1976). A sensitive period for the acquisition of a primary language. *Journal of Psycholinguistic Research*, *5*, 261–283.

Patkowski, M. S. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics*, *11*, 73–89.

Slavoff, G. R., & Johnson, J. S. (1995). The effects of age on the rate of learning a second language. *Studies in Second Language Acquisition*, *17*, 1–16.

Snow C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, *49*, 1114–1128.

Strange, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 3–45). Timonium, MD: York Press.

Werker, J. F. (1995). Age-related changes in cross-language speech perception: Standing at the crossroads. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 155–169). Timonium, MD: York Press.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental changes. *Developmental Psychology*, *24*, 672–683.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, *37*, 278–286.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49–63.

Wode, H. (1994). Nature, nurture, and age in language acquisition: The case of speech perception. *Studies in Second Language Acquisition*, *16*, 325–345.

Wode, H. (1995). Speech perception, language acquisition, and some mutual implications. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 321–347). Timonium, MD: York Press.

Yamada, R. A. (1995). Age and acquisition of second language speech sounds perception of American English/r/and/1/by native speakers of Japanese. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 305–320). Timonium, MD: York Press.

Young-Scholten, M. (1995). The negative effects of 'positive' evidence on L2 phonology. In L. Eubank, L. Selinker, & M. Sharwood Smith (Eds.), *The Current State of Interlanguage: Studies in Honor of William E. Rutherford* (pp. 107–121). The Netherlands: John Benjamins Publishing Co.