Applied Psycholinguistics **23** (2002), 599–621
Printed in the United States of America
DOI: 10.1017.S0142716402004058

Long-term language attainment of bilingual immigrants: Predictive variables and language group differences

GISELA JIA
Lehman College, City University of New York

DORIS AARONSON
New York University

YANHONG WU
Peking University

ADDRESS FOR CORRESPONDENCE
Gisela Jia, Psychology Department, Lehman College/CUNY, 250 Bedford Park Blvd., West Bronx, NY 10468. E-mail: giselaj@lehman.cuny.edu

ABSTRACT
This study examined the variables related to US immigrants' long-term attainment in English, their second language (L2), and their native language (L1). For 44 Mandarin–English bilinguals, with increasing age of arrival (AOA) in the United States, their accuracy in L2 grammaticality judgment tasks decreased and accuracy in an L1 grammaticality judgment task increased. Moreover, both AOA in the United States and mothers' English proficiency uniquely predicted a significant proportion of the variance for bilinguals' L2 proficiency. Finally, as a group, 72 speakers of three Asian languages showed lower levels of L2 proficiency and stronger AOA effects on the task performance than 32 speakers of six European languages. These differences in language proficiency were associated with differences in language use, language learning motivation, and cultural identification between the two groups. These findings suggest that L2 acquisition in the immigration setting is a complicated process involving the dynamic interactions of multiple variables.

Most immigrants face the task of learning the language of their host country as a second language (L2). Understanding the factors influencing the speed at which they acquire their L2 and the level of L2 proficiency they obtain carries both practical and theoretical significance. One major factor that has been the focus of research is the impact of age of arrival (AOA) in the L2-speaking country on L2 acquisition. Although adults typically have been found to be faster than younger children in the initial stage of learning (e.g., Olson & Sam-

uels, 1973; Snow & Hoefnagel-Höhle, 1977, 1978), in the long run, younger children attain a higher level of L2 proficiency (e.g., Asher & Garcia, 1969; Johnson & Newport, 1989; Oyama, 1976, 1978; Patkowski, 1982). The latter finding, called long-term L2 attainment decline, has received much attention because it is regarded as a genuine age effect (attainment after sufficient input) that has implications for brain plasticity in relation to language acquisition (e.g., Patkowski, 1990). However, the causes of the long-term L2 attainment decline have been a topic of much debate (see Birdsong, 1998, Harley & Wang, 1997, and Marinova-Todd, Marshall, & Snow, 2000, for reviews). While some deem long-term L2 decline as evidence for a neurobiologically based and domain-specific critical period for L2 acquisition (e.g., Johnson, 1992; Johnson & Newport, 1989; Patkowski, 1990), others propose cognitive, social, and environmental factors as major causes of this phenomenon (e.g., Bialystok & Hakuta, 1994, 1998; Snow, 1983; Snow & Hoefnagel-Höhle, 1987).

Recently, researchers have made several empirical advances by increasing the scope and depth of the investigation of the long-term L2 attainment by immigrants. Some investigators have examined long-term attainment in both the first language (L1) and the L2 in relation to AOA, considering that the dynamic interactions between L1 and L2 proficiency can help us better understand long-term L2 attainment. Studies adopting this approach have consistently found that long-term L2 proficiency decreases but long-term L1 proficiency increases with increasing AOA. Younger arrivals are more proficient in the L2 than the L1, whereas older arrivals show the opposite pattern. Such trends have been found with the accuracy of pronunciations of Korean–English bilinguals (Yeni-Komshian, Flege, & Liu, 2000) and Italian–English bilinguals (Flege, Munro, & Mackay, 1995), as well as with the speed and accuracy of lexical retrieval of Russian–English bilinguals (McElree, Jia, & Litvak, 2000) and Spanish–English bilinguals (Kohnert, Bates, & Hernandez, 1999). These findings suggest that younger and older arrivals in general may go through different processes of L2 acquisition: younger arrivals switch their dominant language from the L1 to the L2, whereas older arrivals maintain the L2 as their dominant language.[1]

Another line of inquiry has examined variables independent of AOA that are predictive of language proficiency. In a long-term attainment study of 240 Korean–English bilinguals' L2 grammatical proficiency and accent, Flege, Yeni-Komshian, and Liu (1999) identified many factors that predicted a significant amount of L2 proficiency variance. More media input in L2 (e.g., movies, videos, TV, and radio), as well as stronger integrative motivation (e.g., to gain American friends) and instrumental motivation (e.g., to get a good job), were significantly associated with higher L2 proficiency. In addition, when the number of years of education in the United States was controlled, AOA effects disappeared. Advances along this line of inquiry situate L2 proficiency in its acquisition context and allow for an assessment of the relative contributions of various factors and their interactions with AOA.

The pairing of L1–L2 is another predictor of long-term L2 proficiency that has emerged from recent studies with multiple bilingual groups. When other background variables of L2 acquisition were matched, Spanish–English bilinguals showed a higher level of long-term L2 attainment and a weaker AOA

effect than Chinese–English[2] or Vietnamese–English bilinguals (Bialystok & Miller, 1999; Birdsong & Molis, 2001; McDonald, 2000). These findings imply that variables associated with different bilingual groups, such as L1–L2 linguistic distance, as well as social and cultural elements related to language learning and use, can all be independent predictors of L2 proficiency.

Taken together, findings from these recent studies indicate that AOA effects in long-term L2 attainment by immigrants are considerably more complicated than those indicated by previous research. A variety of issues raised in these recent studies must be explored further in order to more fully understand the complexity of L2 acquisition, including the age effect, as indicated by long-term L2 attainment decline. Given the findings of long-term L2 attainment decline and long-term L1 attainment increase with increasing AOA, to which aspects of language proficiency and to which bilingual groups can these findings be generalized? What are the mechanisms leading to the dominant language switch or maintenance processes among the younger or older arrivals? Given the findings that some environmental and motivational variables are predictive of L2 attainment independently of AOA, are other environmental or affective variables involved? Given the findings of different long-term L2 attainment and AOA effects between Chinese–English and Spanish–English bilinguals, can such differences be found in other language groups? In addition to the linguistic distance between L1 and L2, are there social and cultural attributes related to language learning that differentiate among bilingual groups? The current study addressed these questions by extending four aspects of previous research.

First, long-term attainment studies conducted to date that simultaneously examine L1 and L2 have yet to focus on grammar, a critical part of language. Until simultaneous long-term attainment studies on L1 and L2 grammatical proficiency are done, we cannot draw a firm conclusion about the existence of the dominant language switch or maintenance processes for younger and older arrivals. Therefore, in the current study, the L1 and L2 grammatical proficiency of 44 native Mandarin speakers was measured using two grammaticality judgment tasks to examine the relation between long-term L1 and L2 grammatical attainment.

Second, the richness and complexity of language environment indicated by previous research (e.g., Flege et al., 1999) requires us to continue to examine a broader array of social, environmental, and affective variables. Therefore, in the current study, a comprehensive set of such variables with relevance to language acquisition was examined using a language background questionnaire to see how they are related to long-term L1 and L2 attainment.

Third, research on L1 acquisition has demonstrated L1 proficiency growth into adolescence and adulthood (e.g., Braine, Brooks, Cowan, Samuels, & Tamis-LeMonda, 1993; Scott, 1984). Consequently, it seems reasonable to hypothesize that immigrant children may come to a new country with different levels of L1 proficiency, which may, in turn, influence both their L1 and L2 development after their arrival. As little information is available regarding the developmental trends in the acquisition of grammar in a monolingual Mandarin population, this study collected data about the grammatical proficiency of 122 Mandarin monolinguals aged 9 to 16 years and living in China. The goal of this aspect of

the study was to begin to gather normative data regarding the acquisition of grammar in the monolingual Mandarin population in the hope of laying the groundwork for a better understanding of how L1 proficiency associated with different developmental levels affects the acquisition of L2.

Fourth, to examine the generalization of the group differences in L2 proficiency between Chinese–English and Spanish–English bilinguals found in previous studies (Bialystok & Miller, 1999; Birdsong & Molis, 2001), the long-term L2 attainments of the subgroups of 72 Korean–, Mandarin–, and Cantonese–English bilinguals and 32 European English bilinguals were compared. Additionally, social and cultural variables related to language learning for these two groups were also contrasted to explore the possible causes for group differences in long-term L2 attainment.

METHOD

Participants

Bilinguals. The bilingual participants were 112 adult (62 females, 50 males) nonnative English speakers (referred to as bilinguals), with no learning or hearing disabilities based on self-report. Participants were native in 11 different languages, including Mandarin (n = 44), Russian (n =19), Cantonese (n = 17), Korean (n = 11), Spanish (n = 6), Japanese (n = 4), Haitian Creole (n = 3), Urdu (n = 2), Polish (n = 2), Swahili (n = 2), German (n = 1), Punjabi (n = 1), and Italian (n = 1). Coming to the United States between ages 1 and 38, their AOA ranged from infancy to full adulthood, covering the entire range of possible maturational sensitive periods discussed in previous literature. They also had a varied length of residence in the United States (M = 10.7 years; range = 5–32 years; SD = 5.0), years of education in the United States (M = 8.6 years; range = 1–19 years; SD = 3.6), and age of learning English (M = 9.4 years; range = 3–17 years; SD = 3.3). Most of the participants were undergraduates at New York University (NYU), and the others were recruited by personal contacts and by advertisements in Chinese newspapers. All participants were foreign-born and none had gone to English-speaking international schools or received professional English training prior to their arrival in the United States. To ensure sufficient exposure to L2, all participants had been studying in an English-speaking school or working in an English-speaking environment after their arrival and had lived in the United States for at least 5 consecutive years previous to the study. Most younger arrivals started to learn English after they came to the United States, whereas most older arrivals began to do so in junior high school in their home country.

Adult monolinguals. Adult English and Mandarin monolinguals were included in two control groups for the English and Mandarin tasks. The 26 English monolinguals and 20 Mandarin monolinguals were undergraduates from NYU and Peking (Beijing) University, respectively.

Child monolinguals. Five groups of monolingual Mandarin-speaking children (N = 122), with mean ages of 9;3 (n = 25), 10;2 (n = 21), 11;2 (n = 25), 12;3 (n = 25), and 16;4 (n = 25), were recruited from elementary, junior, and senior high schools in Beijing, China.

Measures

English Grammaticality Judgment Tasks. The L2 proficiency of the 112 bilinguals was measured with a listening grammaticality judgment task and a reading grammaticality judgment task. As the listening and reading tasks involve different performance demands, presenting stimuli in both modalities provides a more comprehensive picture of English grammatical proficiency. Several previous studies have adopted this method of presenting sentence stimuli (e.g., Bialystok & Miller, 1999; Goto-Butler, 2000).

The listening and reading tasks had 257 and 256 sentences, respectively. The English sentences resembled those used in previous studies (Flege et al., 1999; Johnson & Newport, 1989). Such similarities facilitate a comparison of our findings about L2 attainment to those of previous studies, and thus yield a baseline for interpretations of our findings about L1 attainment and environment. The sentences tested 12 grammatical structures representing a wide variety of the most basic standard English structures. Four of them dealt with English morphology: past tense, plurals, third person singular present, and present/past progressive form. The other eight structures targeted English syntax: articles, pronominalization, particle movement, predicate structure,[3] auxiliaries, yes/no questions, *wh*- questions, and word order (see the Appendix for sample sentences for each rule type). Each structure was tested by from 12 to 32 pairs of sentences. Only high-frequency words were used, and most words had only one or two syllables. Sentence length (range = 6–11 words; M = 6.8 words) was approximately balanced across the different sentence structures. For each task, half of the sentences were grammatical and the other half were their matched ungrammatical counterparts. The two members of these matched pairs were presented in separate halves of the test to minimize possible memory effects.

Mandarin Grammaticality Judgment Task. The L1 proficiency of the 44 native Mandarin speakers was measured with a Mandarin listening grammaticality judgment task. No reading version of this task was used because many younger arrivals could not read in Chinese.

The Mandarin grammaticality judgment task consisted of 94 sentences: 47 grammatical sentences and 47 matched ungrammatical counterparts. These sentences were constructed based on some standard tests used on US campuses and some other references (e.g., Li & Thompson, 1981; Lu, 1994). To maximally match the difficulty level of the Mandarin sentences to that of the English sentences, only high-frequency words were used and sentence lengths (range = 8–13 Chinese characters; M = 8.3 Chinese characters) were close to those of the English ones. As discussed later, the mean accuracy for monolingual Mandarin and English control groups was almost identical for the listening grammati-

cality judgment tasks in those languages. Because Mandarin grammar differs markedly from that of English, the Mandarin sentences were designed to test some language-specific properties by focusing on three rule types: word order, inappropriate insertion of words, and inappropriate omission of words.

WORD ORDER.

Noun modifier. In Mandarin, a noun modifier (adjective phrase) generally precedes the noun it modifies (1a). An ungrammatical sentence places the noun in front of the noun modifier (1b).

1a. *Ta1 diu1-diao4 le gang1 mai3 de na4 ben3 shul.*
He lose PFV just now buy NOM that CL book.

*1b. *Ta1 diu1-diao4 le na4 ben3 shu1 gang1 mai3 de.*
He lost PFV that CL book just now buy NOM.

He lost the book that he just bought.

Verb modifier. In Mandarin, a verb modifier (adverbial phrase) generally precedes the verb it modifies (2a). An ungrammatical sentence places the verb in front of the verb modifier (2b).

2a. *Ta1 zai4 Bei3jing1 xue2-xi2 han4-yu3.*
He in Beijing learn Chinese.

*2b. *Ta1 xue2-xi2 han4-yu3 zai4 Bei3jing1.*
He learn Chinese in Beijing.

He is learning Chinese in Beijing.

Noun phrase. An ungrammatical sentence in this category contains a fixed noun phrase in an incorrect order (3b).

3a. *Ta1 qu4-nian2 qi1-yue4 bi4-ye4.*
He last year July graduate.

*3b. *Ta1 qi1-yue4 qu4-nian2 bi4-ye4.*
He July last year graduate.

He graduated last July.

Ba structure. *Ba* inverts the order of a verb and an object by placing the direct object immediately after *ba* and before the verb (4a). An ungrammatical sentence in this category fails to reverse the order of the verb and the object (4b).

4a. *Ta1 ba3 zhe4-zhi1 bei1-zi da3-sui4 le.*
He *ba* this glass break PFV.

*4b. *Ta1 da3-sui4 le ba3 zhe4-zhi1 bei1-zi.*
He break PFV *ba* this glass.

He broke the glass.

INAPPROPRIATE INSERTION.

Insert *le*. *Le* functions both as a perfective aspect verbal suffix and a sentence-final particle (5a). An ungrammatical sentence inserts *le* when the sentence is not in the perfective state (5b).

5a. *Qing3 jiang3 yi1-ge gu4-shi gei3 wo3 ting1.*
Please tell one story to I listen.
*5b. *Qing3 jiang3 le yi1-ge gu4-shi gei3 wo3 ting1.*
Please tell one story to I listen.
Please tell me a story.

Insert *ba*. *Ba* inverts the order of a verb and an object by placing the direct object immediately after *ba* and before the verb (6a). However, the use of *ba* is sensitive to whether the noun phrase after *ba* is definite or indefinite and sensitive to the nature of the action. An ungrammatical sentence in this category inserts *ba* when it is unnecessary (6b).

6a. *Wo3 ying1-wen2 xue2 de hen3 nu3-li1.*
I English learn COM very hard.
*6b. *Wo3 ba3 ying1-wen2 xue3 de hen3 nu3-li4.*
I *ba* English learn COM very hard.
I've been working hard on English.

Insert *de*. *De* links a noun modifier and a noun in some situations but not in others (7a). An ungrammatical sentence inserts *de* when it is not needed for the modifying relation (7b).

7a. *Wo3 mai3 le hen3 duo1 ping2 pi2jiu3.*
I buy PFV very many CL beer.
7b. *Wo3 mai3 le hen3 duo1 ping2 de pi2jiu3.*
I buy PFV very many CL *de* beer.
I have bought many bottles of beer.

OMISSION OF REQUIRED WORDS.

Omit *de*. One of the functions of *de* in Mandarin is to link a verb with a stative clause or adverbial phrase to indicate the manner of the event. An ungrammatical sentence in this category inappropriately omits *de* (8b).

8a. *Ta1 dian4-nao3 wan2 de hen3 hao3.*
He computer play *de* very good.
*8b. *Ta1 dian4-nao3 wan2 hen3 hao3.*
He computer play very good.
He is very good at using computers.

Language Background Questionnaire. A 32-item questionnaire was designed to examine age, time, environmental, and affective variables that may predict

language attainment. Some of the items were taken from related studies (Gardner & Clement, 1990; Sanchez & Fernandez, 1993; Schumann, 1978), and the others were constructed after consulting with social psychologists and conducting pilot interviews with bilinguals.

AGE/TIME VARIABLES. Participants reported the age at which they immigrated to the United States, the age at which they started to learn English in the United States or in their native country, the number of years that they had lived in the United States, and the number of years of education that they had received in the United States.

ENVIRONMENTAL VARIABLES. The eight environmental variables examined included the number of L2 speakers at home, their frequency of speaking L2 at home (4-point scale: 1 = *never,* 4 = *always*), the number of L1 speakers at home, their frequency of speaking L1 at home (4-point scale), their frequency of speaking either L1 or L2 at their workplace (5-point scale: 1 = *always L1,* 5 = *always L2*), and their fathers', mothers', and siblings' L2 proficiency for speaking, reading, and writing (4-point scale: 1 = *not at all,* 4 = *fluently*). For participants who had more than one sibling, the average rating over all siblings was used. Among the 112 bilinguals, 93 (83.0%) lived with various family members during their residence in the United States.

AFFECTIVE VARIABLES.

Self-consciousness. Four variables measured participants' self-consciousness in learning and using English. Participants rated (4-point scale: 1 = *never,* 4 = *always*) whether they prepared English sentence grammar before they spoke, whether they cared about grammar when they spoke, whether they avoided opportunities to speak English because they felt that their English was not good enough, and whether they felt embarrassed or neutral when others did not understand their English.

Cultural preference and identity. Two items about culture preference and identity were used to obtain information on which culture the participants most liked or identified with on a 3-point scale (1 for *native culture,* 2 for *both cultures* and 3 for *American culture*).

Motivation. Five motivation variables were measured. Participants rated their motivation to learn English on five aspects using 7-point scales (1 = *not feel this way at all,* 7 = *feel this way very strongly*). The five aspects were motivation to learn English to get a good job, to make more American friends, to become more like Americans, to understand American culture better, and because of the beauty of English as a language.

SELF-EVALUATED L1 AND L2 PROFICIENCY. All of the bilinguals self-evaluated their speaking, reading, and writing abilities in both L1 and L2 as a supplementary measure of their language proficiency.

Procedures

English sentences in the listening version were recorded on tape by a female native English speaker with a standard American accent. Each sentence was read twice with a 1–2 s pause between the repetitions and a 3–4 s delay between different sentences. All sentences were read clearly with normal intonation at an average rate of 4.2 words/s. Sentences in the English reading version were clearly typed on 11 double-spaced pages. A female native Mandarin speaker with standard Mandarin accent recorded the Mandarin sentences with a normal intonation at an average rate of 3.8 Chinese characters/s. Each sentence was read twice, with a 1–2 s pause separating the repetitions and a 3–4 s delay between different sentences. All bilinguals filled out the language background questionnaire (LBQ). They were also verbally interviewed to provide additional information about their language learning experiences, and the experimenters took notes during the interviews.

Each bilingual participated individually in two 1–1.5 hr sessions at NYU. One session was for the English listening task and the other was for the English reading task, the LBQ, the Mandarin listening task (only for the Mandarin–English bilinguals), and the interview. The order of the listening and reading tasks was counterbalanced across participants, and the two sessions were at least 3 days apart. All of the monolingual Mandarin speakers participated in the study in groups in a classroom setting in Beijing.

RESULTS

Relation between long-term L1 and L2 attainment

Consistent with previous findings of AOA effects, for the 112 bilinguals, a younger AOA was associated with higher accuracy on the English listening task (M = 86.7, range = 60.4–98.3, SD = 8.6; r = –.69, $p < .001$) and on the reading task (M = 88.7, range = 68.9–98.0, SD = 5.6; r = –.46, $p < .001$. The latter correlation was significantly smaller than the former one, t (112) = 8.8, $p < .01$. The 26 monolingual English speakers scored a mean of 94.6% on the listening task (range = 86.7–97.6, SD = 2.6) and a mean of 92.8% on the reading task (range = 87.0–97.6, SD = 2.5).

For the subgroup of 44 Mandarin–English bilinguals, a younger AOA was associated with lower accuracy on the L1 (Mandarin) listening task (range = 66.7–97.8, M = 88.8, SD = 7.4, r = .55, $p < .001$). The 20 monolingual Mandarin speakers had a mean accuracy of 92.9% (range = 89.4–96.8, SD = 2.1).

L1 and L2 listening performance also had a significant negative correlation (r = –.33, $p < .05$), indicating that the better the participants did on the L2 task, the poorer they did on the L1 task. Figure 1 provides a direct visual comparison of the participants' listening performance on the two tasks plotted against the AOA.

The participants' 5-point self-evaluated proficiency in speaking, reading, and writing for each language was summed to get a composite score for their general proficiency in each language. Similar to performance on the listening task, a younger AOA was associated with higher L2 ratings (r = –.44, $p < .001$) but

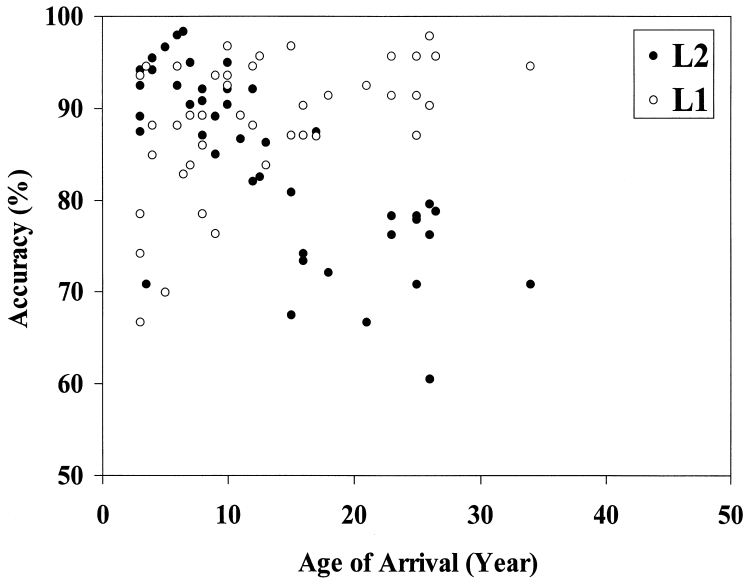

Figure 1. A scatterplot of the accuracy versus the arrival age for the (●) English listening task and (○) Mandarin listening task.

lower L1 ratings ($r = .62$, $p < .001$). Also analogous to participants' performance on the two language tasks, a higher self-rating on one language was associated with a lower self-rating on the other language ($r = -.30$, $p < .001$).

Participants' task performance and their self-ratings were also significantly correlated with each other. A higher composite L2 proficiency rating was associated with better performance on the English listening task for the 112 bilinguals ($r = .53$, $p < .001$), as well as for the 44 Mandarin–English bilinguals ($r = .48$, $p < .001$). A higher composite L1 proficiency rating was associated with better performance on the Mandarin listening task for the 44 Mandarin–English bilinguals ($r = .74$, $p < .001$). Such a positive relation between self-evaluated and behaviorally measured language proficiency was also demonstrated in previous studies (e.g., Hakuta & D'Andrea, 1992).

Variables predicting L2 and L1 performance

This part of the analysis is aimed at locating, among the large number of variables measured, those that were significantly associated with the performance variance for the grammaticality tasks. The participants' performance scores on the English listening and reading tasks and the Mandarin listening task were viewed as dependent variables (DV). Three other sets of variables were viewed as independent variables (IV), which were related to time, environment, and affective variables. The analysis consisted of three steps. First, for each independent variable set, simple bivariate correlations with the dependent variables were

obtained. Second, all variables yielding significant bivariate correlations with each given performance variable were selected for a backward-elimination regression analysis, with each of the grammaticality performance variables separately taken as the dependent variable in these analyses.[4] Third, variables that remained significant in each set were selected for a simultaneous multiple regression analysis with each performance variable as a DV. This further determined the predictive power of the variables remaining while their interrelations were taken into account (also see Flege et al., 1999, for alternative analysis strategies).

Simple bivariate correlations by set and backward elimination analyses

AGE/TIME VARIABLES. As shown in Table 1, better performance on the English listening task or the English reading task was associated with younger AOA, younger age of learning English, and more years of education in the United States, but not with length of residence in the United States (Table 1). In contrast, better performance on the Mandarin task was associated with older AOA, older age of learning English, and fewer years of education in the United States, but not with length of residence in the United States.

However, as shown in Table 2, these three time variables were highly correlated with each other, suggesting that they may be responsible for roughly the same portion of the performance variance. Therefore, a backward-elimination regression was run to locate the significant predictors when the collinearity was taken into account. As shown in Table 3, AOA was the only significant predictor for each of the performance variables and was therefore selected for the general regression analysis.

ENVIRONMENTAL VARIABLES. As shown in Table 1, of the eight environmental variables analyzed, better performance on the L2 (English) listening task was associated with more people speaking the L2 at home, a higher frequency of speaking the L2 at home, higher L2 proficiency of mothers, and higher L2 proficiency of siblings. Better performance on the L2 reading task was associated with higher L2 proficiency of fathers and higher L2 proficiency of mothers. Consistent with this trend, better performance on the L1 task was associated with fewer people speaking the L2 at home, lower frequency of L2 speaking at home, and higher frequency of L1 speaking at the workplace.

Based on the statistically significant results of the backward elimination analysis (Table 3), among the environmental variables, mother's English proficiency and frequency of home L2 speaking were selected for the general English listening performance analysis. The mother's English proficiency was selected for the general English reading performance analysis, and the frequency and number of people at home speaking the L2 were selected for the general Mandarin performance analysis.

AFFECTIVE VARIABLES. As shown in Table 1, of the four self-consciousness variables measured, better performance on the English listening and reading tasks predicted less preparation before speaking English and less avoidance of

Table 1. *Bivariate correlations between language performance and participant variables*

	Task performance		
Participant variables	Listening	Reading	Mandarin
Age/time variables			
Age of arrival	−.69***	−.42***	.55***
Age of learning	−.54***	−.33***	.58***
Length of stay	.04	−.03	−.19
Education in US	.56***	.36***	−.56***
Environmental variables			
No. of L2 speakers (at home)	.20*	.14	−.37*
L2 speaking frequency	.27**	.18	−.32*
No. of L1 speakers (at home)	.18	.13	−.27
L1 speaking frequency	.09	.04	−.05
Work language	.10	.12	.42**
Father's L2 proficiency	.20	.23*	−.25
Mother's L2 proficiency	.31**	.30**	−.28
Sibling's L2 proficiency	.27*	.08	−.29
Self consciousness			
Prepare before speaking	−.49***	−.25**	.28
Care about grammar	.09	.17	.29
Avoid opportunity	−.32***	−.22*	.20
Sensitive to nonunderstanding	−.15	−.07	.00
Cultural preference and identity			
Culture identity	.13	.08	−.21
Culture preference	.09	.05	−.08
Motivation			
Know more about the culture	−.17	.03	.14
Get a good job	−.03	−.05	.07
Like English as a language	−.09	−.08	−.02
Make more friends	−.14	−.07	.02
Become more like Americans	−.03	−.11	−.06

*$p < .05$; **$p < .01$; ***$p < .001$.

Table 2. *Correlation matrix of time variables*

	Arrival age	Learning age	US residence (years)	Education in US (years)
Arrival age	1.00	.76***	−.25**	−.77***
Age of learning	.76***	1.00	−.46***	−.83***
US residence (years)	−.25***	−.46***	1.00	.76***

**$p < .01$; ***$p < .001$.

Table 3. *Summary results of backward-elimination analyses for variables predicting language task performance*

	Task performance								
	English listening			English reading			Mandarin listening		
Participant variables	B	SE B	β	B	SE B	β	B	SE B	β
Time variables									
Age of arrival	–2.00	.21	–.68***	–.72	.20	–.34	.48	.11	.61***
Environmental variables									
L2 number			.00			–2.77		.95	–.46***
L2 frequency	2.27	1.26	.21*			.17			–.21
Mother's L2	6.70	1.84	.42***	4.77	1.62	.33***			
Psychological variables									
Preparation[a]	–7.76	1.42	–.48***	–2.85	1.13	–.25**			

[a]Whether participants prepare before speaking English.
*$p < .05$; **$p < .01$; ***$p < .001$.

opportunities to speak English. There was no relation between language performance and either cultural identification or cultural preference. However, AOA had a significant correlation with cultural identity ($r = -.23$, $p < .05$), indicating that younger arrivals identified more with American culture and older arrivals identified more with their native culture. Of the five motivation variables measured, none of the bivariate correlations approached statistical significance, so the whole subset was dropped from further analysis.

For the Mandarin task, none of the correlations reached significance, perhaps due to the smaller sample of Mandarin–English bilinguals. However, there was a trend for those who performed better on the Mandarin task to prepare more before they spoke English and to care more about their English grammar.

Based on the results of the backward elimination analyses in Table 2, preparation before speaking was selected from among the affective variables for the general English listening and reading analyses.

General analysis. The general simultaneous regression analysis with each of the performance variables (Table 4) obtained three sets of important relations. First, AOA and number of L2 speakers at home both remained significant predictors of the variance in Mandarin listening task performance. Second, AOA and mother's English proficiency both remained significant predictors of the English listening task as well as the reading task performance. Third, when taken together, L1 and L2 performance variance was significantly predicted by AOA and two environmental variables (mother's English proficiency and number of English speakers at home).

Table 4. *Summary results of simultaneous regression analyses for variables predicting language task performance*

Participant variables	Task performance								
	English listening			English reading			Mandarin		
	B	SE B	β	B	SE B	β	B	SE B	β
Age of arrival	−1.63	.33	−.47***	−.85	.34	−.27*	.40	.11	.50***
Mother's L2	7.5	1.36	.46***	5.88	1.57	.39***			
L2 number[a]							−1.70	.83	−.29*
L2 frequency[b]	.21	.86	.02						
Preparation[c]	−1.33	1.08	−.11	.78	1.23	.07			

[a]The number of people who spoke L2 at home.
[b]The frequency of speaking L2 at home.
[c]Whether participants prepare in advance before speaking L2.
*$p < .05$; ***$p < .001$.

Table 5. *Performance on the Mandarin listening task by monolingual Mandarin speakers*

Mean age (year;month)	*N*	*M* (% correct)	*SD*	Range
9;3	25	77	6.1	67–89
10;3	21	81	7.1	65–93
11;4	25	84	5.1	75–91
12;3	25	85	4.0	78–91
16;1	25	93	4.8	80–100
19;6	25	92	3.7	85–98

Mandarin monolinguals

Table 5 presents group means for performance on the Mandarin listening task by the five groups of monolingual Mandarin children and the adult monolingual Mandarin control group. A one-way analysis of variance (ANOVA) revealed a significant age effect for the task performance, $F(5, 140) = 35.29$, $p < .001$. Modified LSD (Bonferroni) tests with $p = .05$ showed significant differences between the younger and older groups: 9 versus 11, 12, 16, 19; 10 versus 16, 19; and 11, 12 versus 9, 16, 19. The group means show developing L1 syntactic proficiency with adult-level performance at about age 16. Thus, Mandarin-speaking children immigrating to an L2-speaking country would, on average, arrive with less than adult level L1 proficiency, and some aspects of their L1 abilities are positively correlated with their AOA.

Table 6. *Participant variable information for the Asian and European groups*

Participant variables	Range	*M*	*SD*
Age of arrival			
Asian	3–38	12.3	7.7
European	3–22	11.6	4.9
Age of learning English			
Asian	4–15	9.5	3.3
European	3–17	9.6	3.4
Length of residence			
Asian	5–28	11.4	4.9
European	5–32	10.2	5.3
Education in US			
Asian	2–19	8.8	3.6
European	2–14	8.0	3.6

Note: Data are in number of years.

Language groups and L2 sentence structures

Combined analyses were conducted to compare L2 performance on each sentence structure for the European and Asian groups. The *European group* consisted of 32 participants with L1s of Russian ($n = 19$), Spanish ($n = 6$), Haitian Creole ($n = 3$), Polish ($n = 2$), German ($n = 1$), and Italian ($n = 1$). The *Asian group* had 72 participants, with L1s of Mandarin ($n = 44$), Cantonese ($n = 17$), and Korean ($n = 11$). The two groups were similar on AOA, age of English learning, length of US residence and number of years of US education (see Table 6). Independent sample *t* tests showed no significant differences between these two groups on any of these variables.

The Asian and European groups differed in some environmental and social affective variables. European mothers were reported to speak better English, $t(62) = -3.6$, $p < .001$, than Asian mothers. In comparison to Asian language speakers, European language speakers reported stronger motivation to learn English because of the beauty of the language, $t(102) = -2.4$, $p < .05$, they did less preparation of grammar before speaking English, $t(81) = 2.5$, $p < .05$, and they had a stronger identity with American culture, $t(98) = 2.6$, $p < .05$.

The percentage of errors for each rule type was computed only for ungrammatical sentences because judging a correct sentence does not localize grammatical knowledge to a particular part of the sentence. As shown in Table 7, in general, the Asian group had stronger AOA effects and significantly lower accuracy in the listening task, $t(97) = -4.8$, $p < .001$, and the reading task, $t(98) = -2.7$, $p < .01$, than the European group. There were significant correlations between overall accuracy and AOA for both the listening and reading tasks for the Asian group, whereas these correlations were not significant for the European group. For the Asian group, accuracy and AOA had significant correlations for all of the rule types for the listening task and eight of the rule types for the

Table 7. *L2 task performance (accuracy and AOA correlations) by modality, rule type, and language group*

	Listening				Reading			
	Asian		European		Asian		European	
Rule type	%	Corr.	%	Corr.	%	Corr.	%	Corr.
Total accuracy	85	−.76**	92	−.27	88	−.53**	91	−.24
Articles	72	−.69**	81	−.42*	79	−.37**	82	−.31
Predicate structure	74	−.68**	93	−.27	79	−.50**	88	−.42*
Third person singular	78	−.33*	92	.00	90	.05	86	−.04
Plurals	79	−.61**	94	.05	80	−.22	88	−.24
Wh- question	81	−.63**	83	−.26	90	−.25*	90	−.15
Auxiliaries	85	−.63**	97	.07	92	−.19	93	−.08
Pronominalization	85	−.55**	91	.03	93	−.29*	90	−.20
Present progressive	87	−.52**	94	−.23	89	−.13	89	−.25
Particles	90	−.56**	97	.08	92	−.33**	93	.00
Past tense	90	−.49**	98	.08	90	−.29*	90	−.13
Word order	93	−.29*	96	.14	97	−.12	98	−.13
Yes or no question	96	−.56**	99	.13	90	.39**	88	−.04

Note: The rule types are listed according to the accuracy levels (low to high) of the listening task by the Asian group.
*$p < .05$; **$p < .01$.

reading task. The European group showed a significant age effect on only one rule type for the listening task (article) and one for the reading task (predicate structure).

One may wonder whether the much weaker AOA effect for the European group is due to the group's ceiling performance on the tasks. However, significant AOA effects could occur even when the accuracy score was high (e.g., performance on particles by the Asian group for both tasks), and similar accuracy scores were related to very different AOA effects (e.g., performance on predicate structure and plurals in the reading task by the European group). Therefore, the current language group differences in AOA effects were not completely due to a ceiling effect of the European group. However, this does not exclude the possibility that, if a more difficult language task had been given, the European group would still show minimal age effects, as they did with the current measure.

Another possible explanation of the minimal AOA effect for the European group is its much smaller sample size in comparison to the Asian group (32 vs. 72). In addition, the European group had a smaller range of AOA (3–22 years) than the Asian group (3–38 years). To address these issues, 32 Asian language speakers from the current study were randomly selected from the total sample of 72, with the constraints that their AOAs, age at which they started to learn English, length of residence in the United States, and number of years of education in the United States matched those of the 32 European language speakers.

Still, the correlation between AOA and accuracy on the listening task was −.78 ($p < .001$), and that between AOA and accuracy on the reading task was −.53 ($p < .001$). Both correlations were identical to those obtained for the whole group of 72 participants. Further, this subgroup of Asian language speakers also showed significant AOA effects on 10 or 5 sentence structures for the listening or reading task, respectively. These results are consistent with previous findings in the literature that strong AOA effects existed with moderate sample sizes of 33 (Bialystok & Miller, 1999) and 24 (Goto-Butler, 2000) native Chinese speakers. Therefore, ceiling effects and the smaller sample size could not completely explain the minimal AOA effects for the European group.

There was also an interaction between task modality and language group. The Asian group obtained higher total accuracy and weaker AOA effects on the L2 reading task than on the L2 listening task. The group obtained higher accuracy on the reading task than on the listening task for 11 out of 12 rule types and lower AOA–accuracy correlations for all 12 rule types. However, the European group showed little difference in accuracy level and AOA effects between the two tasks.

DISCUSSION

Findings from the current study add to the emerging body of literature concerning the complexity of the L2 acquisition process in the immigration setting. The findings indicate that long-term L2 attainment is associated with multiple variables, including long-term L1 attainment, language environment, and L1–L2 pairing, in addition to AOA.

The study made the first efforts to examine the relation between long-term L1 and L2 grammatical proficiency. For the 44 Mandarin–English bilinguals, accuracy in the L2 (English) task decreased and accuracy in the L1 (Mandarin) task increased with increasing AOA, and higher scores on one task predicted lower scores on the other. Analogously, with increasing AOA, the 112 bilinguals' composite self-ratings of their L2 proficiency decreased and self-ratings of their L1 proficiency increased, and their self-ratings of L1 and L2 proficiency were negatively correlated. Performance on the language tasks and self-rated proficiency were also highly consistent with each other.

These results are in accord with findings from previous studies using comparable tasks to simultaneously measure long-term L1 and L2 attainment in phonology (Flege et al., 1995; Yeni-Komshian et al., 2000) and lexicon (Kohnert et al., 1999; McElree et al., 2000). Such findings indicate that the relative levels of L1 and L2 proficiency have opposite patterns among younger and older arrivals. The opposite patterns of relative L1 and L2 proficiency suggest that, as a group, younger arrivals tend to switch their dominant language from L1 to L2, whereas older arrivals tend to maintain L2 as their dominant language.[5] Future research is necessary to examine these processes in greater depth by addressing the following questions. What are the ongoing attributes of the dominant language switch or maintenance processes? Why do dominant language switch or maintenance processes occur among younger or older arrivals, respectively? What does the existence of these processes indicate about the causes of the long-term L2 attainment decline? A longitudinal study of 11 Chinese immigrant

children and adolescents living in the United States has been conducted to explore these issues (Jia & Aaronson, in press).

A host of environmental variables was found to predict L1 and L2 proficiency in the current study. Bilinguals whose mothers had higher L2 proficiency performed significantly better on the L2 listening and reading tasks, and participants who were surrounded by more L2 speakers at home performed significantly worse on the L1 listening task. Many other environmental variables either approached significance or showed similar trends. In the multiple regression analysis, mothers' English proficiency uniquely explained the same amount of L2 performance variance as did AOA. Although these data are correlational, the supplementary interviews with bilinguals revealed that, for the most part, language environment influenced language proficiency, rather than vice versa. For example, many bilinguals reported that their language choice when interacting with their parents was based on their parents' English proficiency. These findings add to the set of environmental variables located by Flege et al. (1999). Future research should further expand the scope of language environment variables being investigated and, more importantly, explore causal relations between the environmental variables and language proficiency.

Analyses of the 122 monolingual Mandarin speakers' overall performance on the Mandarin grammaticality judgment task showed a significant growth between ages 9 and 16. This result is consistent with a large body of literature that has documented developing L1 proficiency until adolescence and beyond. For example, a developing trend exists from early childhood to adulthood in understanding grammatical categories (Braine et al., 1993) and universal quantifiers (Brooks, Jia, Braine & Dias, 1998), in the accuracy and speed for lexical retrieval (Wiegel-Crump & Dennis, 1986), and the ability to express discourse connectivity and logical relations (Scott, 1984). These findings suggest that, although immigrants of different ages all start L2 immersion with L1 being dominant, due to different developmental levels, they have different levels of L1 proficiency that may influence their L2 acquisition. Future research should examine whether, and how, the different levels of L1 proficiency can lead to differences in L2 acquisition. Findings from the longitudinal study (Jia & Aaronson, in press) provide some initial evidence for the role of different L1 proficiency in L2 acquisition.

Group differences in long-term L2 attainment were found between 72 Asian language speakers and 32 European language speakers. The Asian group showed lower levels of L2 proficiency and stronger AOA effects than the European group. This is congruent with previous findings contrasting Spanish–English bilinguals with Chinese–English and Vietnamese–English bilinguals (Bialystok & Miller, 1999; Birdsong & Molis, 2001; McDonald, 2000). In light of these findings, one may ask why long-term L2 attainment varies among different language groups. The language groups are associated with different L1 linguistic attributes as well as different social and cultural behaviors. Regarding linguistic attributes, European languages usually are linguistically closer to English whereas Asian languages are linguistically further from English. As a nonlinguistic, but approximate indicator of linguistic distance, Odlin (1997) used the different lengths of language courses given during 1985 at the Foreign Service

Institute of the U.S. State Department designed to help L1 English speakers with similar language learning aptitudes to reach comparable levels of proficiency in different languages. The average length of Asian language classes was substantially longer than that of the European language classes. Consistent with this phenomenon, the Vietnamese speakers (McDonald, 2000) and Chinese speakers (Bialystok & Miller, 1999) had particular trouble with English structures that are absent in their native languages such as plurals, articles, third person singular present, and subject–verb agreement.

Socially and culturally, in comparison to Asian language speakers, European language speakers tended to use English more frequently (Birdsong & Molis, 2001; McDonald, 2000). In the current study, in comparison to the Asian mothers, European mothers were reported to speak better English. In addition, European language speakers reported stronger motivation to learn English because of the beauty of the language, did less preparation of grammar before speaking English, and had a stronger identity with American culture. However, reasons for these social and cultural differences among various L1 groups and their causal links to language proficiency are not fully understood and merit further research.

The superior performance of the Asian group on the reading task to that on the listening task for all the sentence structures tested by the Asian group has also been found by two other studies with Chinese–English bilinguals (Bialystok & Miller, 1999; Goto-Butler, 2000). This finding suggests that Chinese or Asian language speakers, especially older arrivals, demonstrate higher levels of grammatical knowledge during visual sentence processing when they can control the reading speed and when the stimuli are not transient. In other words, AOA effects on long-term grammatical attainment are highly sensitive to the modality through which the grammatical knowledge is tapped.

Combined together, these several aspects of findings indicate that L2 acquisition is a complicated process involving the dynamic interactions of multiple variables. The well-documented phenomenon of AOA effects on long-term L2 attainment is precisely embedded in this intricate picture. Until we can gain a deep and comprehensive understanding of how these variables relate to one another and to long-term L2 attainment, there is not sufficient empirical ground to equate the negative correlation between long-term L2 attainment and AOA with a neurobiologically based, genetically preprogrammed, and domain-specific critical period for L2 acquisition. Future long-term attainment studies may simultaneously compare multiple language groups with larger sample sizes to disentangle the various variables at work. In addition, future research can also use a longitudinal design to identify variables that covary with AOA and at the same time affect L2 acquisition (e.g., Jia & Aaronson, in press).

APPENDIX

SAMPLE ENGLISH SENTENCES

Plurals
1a. Two brothers made the first airplane.
*1b. Two brother made the first airplane.
Present progressive form
2a. More and more snow was falling.
*2b. More and more snow was fall.
Third person singular present
3a. The flower smells very sweet.
*3b. The flower smell very sweet.
Past tense
4a. I watched a new movie last night.
*4b. I watch a new movie last night.
Articles
5a. The first airplane was like a big funny bird.
*5b. First airplane was like a big funny bird.
Pronominalization
6a. Everybody thinks that he is nice.
*6b. Everybody thinks that him is nice.
Particle movement
7a. He showed off his toy to the girl.
*7b. He showed his toy off to the girl.
Predicate structure
8a. The man allows his son to watch TV.
*8b. The man allows his son watch TV.
Auxiliaries
9a. Kathy should brush her teeth.
*9b. Kathy should brushes her teeth.
Yes or no question
10a. Have you been waiting for me?
*10b. Have been you waiting for me?
Wh- question
11a. Why did you wake up so early?
*11b. Why you woke up so early?
Word order
12a. The professor gives his students hard tests.
*12b. The professor hard tests gives his students.

ACKNOWLEDGMENTS

We are grateful to many individuals who offered their generous help with the research reported in this article. Martin Braine and Patricia Brooks provided suggestions for the design of the current research. Zuohong Xue, Peihong Zhou, and Hua Zhang helped collect the data from Mandarin monolingual speakers. Gary Marcus, Grace Yeni-Komshian, Kevin Sailor, Brian McElree, and Susan Carey read through drafts of this manu-

script and offered feedback on data analyses, theoretical conceptualization, and writing. Jin Fan, Ed Colet, Tom Donnelly, Jennifer Thomas, John Daws, Michael Palij, Pat Shrout, and Niall Bolger offered suggestions for statistical analyses. Yasuhiro Shirai and several other anonymous readers of versions of this manuscript provided their constructive and thoughtful comments. Richard Bock made valuable contributions concerning matters of style. Christine Lau, Shing-Hou Chen, Ruth Williams, and Michael Young assisted with data entry, stimulus design, and proofreading related to the research and this article. This research was partially supported by a Dean's Dissertation Fellowship, a PSC-CUNY Research Grant (69683-0029), and a Shuster Fellowship to the first author, and NIMH MRISP Grant MH49747 to Lehman College, and a SCORE/NICHD grant (413531117) to the first author. Portions of this research were Gisela Jia's dissertation and were presented at the meetings of the 10th American Psychological Society Annual Convention in June 1998, the Boston University Conference on Language Development in October 1998, and the Second Conference of the Japanese Society for Language Sciences in August 2000. Portions of the data were published in the conference proceedings for the last two conferences.

NOTES

1. Language dominance refers to a bilingual's relative proficiency in L1 and L2 as indicated by quantitative proficiency measurements in the two languages (Nicholadis, 1994). Such measurements involve a number of linguistic attributes, such as phonology, vocabulary, and syntactic knowledge. Dominance in one attribute does not necessarily imply dominance in the others. However, some researchers suggest that adult bilinguals' dominance in one proficiency domain may provide reliable evidence for a more global picture of dominance patterns (Lambert, Havelka, & Gardner, 1959). An important distinction is between dominance in language use (using one language more often than the other) and dominance in language proficiency (being more proficient in one language than the other). In this article, dominant language only refers to dominance in language proficiency.
2. Mandarin and Cantonese are two speech dialects of Chinese that share basically the same writing system. In this article, Chinese will be used to refer to the written language. In addition, when a cited study lumped Mandarin and Cantonese speakers together in a sample or did not make a distinction between these two, the term Chinese speakers is used when findings from such studies are discussed.
3. In Johnson and Newport (1989), this rule type was called subcategorization.
4. Because there was high multicollinearity among the variables in most sets of independent variables, the next step was to locate the variables that had significant unique predictive power for the performance variance when the correlations among the independent variables were taken into account. Among the possible strategies for multiple regression analysis, simultaneous regression is not appropriate for this case. When the data exhibit high multicollinearity, it is possible that none of the variables uniquely explains a significant amount of performance variance, although together they are explanatory (Cohen & Cohen, 1983). Further, there is no strong theoretical basis for prediction of a causal or priority relation among these independent variables to justify the use of hierarchical regression. Therefore, a backward elimination procedure was adopted to drop out all variables that did not add significant predictive power to the combination of the other variables (Draper & Smith, 1998).

5. The dominant language switch or maintenance processes refer to group trends only and do not imply that individual bilinguals (or multilinguals) can have only one language dominant. Some individuals can maintain high levels of proficiency in both the L1 and L2 due to exceptional language learning abilities or unique environmental demands.

REFERENCES

Asher, J., & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal, 53,* 334–341.

Bialystok, E., & Hakuta, K. (1994). *In other words.* New York: Basic Books.

Bialystok, E., & Hakuta, K. (1998). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161–181). Mahwah, NJ: Erlbaum.

Bialystok, E., & Miller, B. (1999). The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition, 2,* 127–145.

Birdsong, D. (Ed.). (1998). *Second language acquisition and the critical period hypothesis.* Mahwah, NJ: Erlbaum.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints on second language acquisition. *Journal of Memory and Language, 44,* 235–249.

Braine, M. D. S., Brooks, P. J., Cowan, N., Samuels, M., & Tamis-LeMonda (1993). The development of categories at the semantics/syntax interface. *Cognitive Development, 8,* 465–494.

Brooks, P. J., Jia, G., Braine, M. D. S., & Dias, M. G. (1998). A cross-linguistic study of children's comprehension of universal quantifiers: A comparison of Mandarin Chinese, Portuguese, and English. *First Language, 18,* 33–79.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis.* New York: Wiley.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97,* 3125–3134.

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

Gardner, R. C., & Clement, R. (1990). Social psychological perspectives on second language acquisition. In H. Giles & W. P. Robinson (Eds.), *Handbook of language and social psychology* (pp. 495–517). West Sussex, UK: Wiley.

Goto-Butler, Y. (2000). The age effect in second language acquisition: Is it too late to acquire native-level competence in a second language after the age of seven? In Y. Oshima-Takane, Y. Shirai, & H. Sirai (Eds.), *Studies in language sciences* (Vol. 1). Nagoya, Japan: Chyukyo University.

Hakuta, K., & D'Andrea, D. (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics, 13,* 72–99.

Harley, B., & Wang, W. (1997). The critical period hypothesis: Where are we now? In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives.* Hillsdale, NJ: Erlbaum.

Jia, G., & Aaronson, D. (in press). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics.*

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

Johnson, J., & Newport, E. (1989). Critical period effect in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60–69.

Kohnert, K. J., Bates, E., & Hernandez, A. E. (1999). Balancing bilinguals: Lexical–semantic production and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research, 42,* 1400–1413.

Lambert, W. E., Havelka, J., & Gardner, R. C. (1959). Linguistic manifestations of bilingualism. *American Journal of Psychology, 72,* 77–82.

Li, C., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar.* Los Angeles: University of California Press.

Lu, J. J. (1994). Some common grammatical mistakes made by foreigners who learn Chinese. *Language Teaching and Learning, 1,* 63–74.

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

McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics, 21,* 395–423.

McElree, B., Jia, G., & Litvak, A. (2000). The time course of conceptual processing in three bilingual populations. *Journal of Memory and Language, 42,* 229–254.

Nicholadis, E. (1994). *Code-mixing in young bilingual children.* Unpublished doctoral dissertation, McGill University, Canada.

Odlin, T. (1997). *Language transfer: Cross-linguistic influence in language learning.* New York: Cambridge University Press.

Olson, L., & Samuels, S. J. (1973). The relationship between age and accuracy of foreign language pronunciation. *Journal of Educational Research, 66,* 263–267.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5,* 261–283.

Oyama, S. (1978). The sensitive period and comprehension of speech. *Working Papers on Bilingualism, 16,* 1–17.

Patkowski, M. S. (1982). The sensitive period for the acquisition of syntax in a second language. In S. D. Krashen, R. C. Scarcella, & M. H. Long (Eds.), *Child–adult differences in second language acquisition* (pp. 52–63). Rowley, MA: Newbury House.

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

Sanchez, J. I., & Fernandez, D. M. (1993). Acculturative stress among Hispanics: A bidimensional model of ethnic identification. *Journal of Applied Social Psychology, 23,* 654–668.

Schumann, J. H. (1978). Affective factors and the problem of age in second language acquisition. *Language Learning, 25,* 209–235.

Scott, C. M. (1984). Adverbial connectivity in conversations of children 6 to 12. *Journal of Child Language, II,* 423–452.

Snow, C. E. (1983). Age differences in second language acquisition: Research findings and folk psychology. In K. M. Bailey, M. H. Long, & S. Peck (Eds.), *Second language acquisition studies* (pp. 141–150). Rowley, MA: Newbury House.

Snow, C. E., & Hoefnagel-Höhle, M. (1977). Age differences and the pronunciation of foreign sounds. *Language and Speech, 20,* 357–365.

Snow, C. E., & Hoefnagel-Höhle, M. (1978). Age differences in second language acquisition. In E. M. Hatch (Ed.), *Second language acquisition* (pp. 333–344). Rowley, MA: Newbury House.

Snow, C. E., & Hoefnagel-Höhle, M. (1987). The critical period for language acquisition: Evidence from second language learning. *Child Development, 49,* 1115–1128.

Wiegel-Crump, C. A., & Dennis, M. (1986). Development of word-finding. *Brain and Language, 27,* 1–23.

Yeni-Komshian, G., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition, 3,* 131–149.