JSLHR, Volume 40, 686–693, June 1997

Age of Second-Language Acquisition and Perception of Speech in Noise

Lynn Hansberry Mayo
Mary Florentine
Søren Buus
Northeastern University
Boston, MA

To determine how age of acquisition influences perception of second-language speech, the Speech Perception in Noise (SPIN) test was administered to native Mexican-Spanish-speaking listeners who learned fluent English before age 6 (early bilinguals) or after age 14 (late bilinguals) and monolingual American-English speakers (monolinguals). Results show that the levels of noise at which the speech was intelligible were significantly higher and the benefit from context was significantly greater for monolinguals and early bilinguals than for late bilinguals. These findings indicate that learning a second language at an early age is important for the acquisition of efficient high-level processing of it, at least in the presence of noise.

KEY WORDS: second language, speech in noise, age of acquisition, age of learning, non-native speech perception

Non-native speakers may demonstrate native-like speech recognition in quiet, but not in the presence of background noise (Bergman, 1980; Buus, Florentine, Scharf, & Canévet, 1986; Florentine, 1985a, 1985b; Florentine, Buus, Scharf, & Canévet, 1984; McAllister, 1990; Roussohatzaki, 1990; Spolsky, Sigurd, Masahito, Walker, & Arterburn, 1968; Takata & Nábelek, 1990). The reason for this difficulty is unclear, but appears to be related to such factors as the duration of foreign-language study (Florentine, 1985a, 1985b; Roussohatzaki, 1990; Yamada, 1995), the age of the listener (Bergman, 1980), and the listening conditions (Takata & Nábelek, 1990).

Florentine (1985b) addressed some of these issues by comparing native and non-native speakers' perception of American-English in noise. Although her non-native listeners' ability to understand speech in noise improved as their exposure to American-English increased, they did not perform as well as natives, even after extensive exposure. Her data also indicate that non-native listeners did not benefit as much from contextual cues as did native listeners. Only two non-native listeners performed with native-like proficiency, and they had been exposed to American-English since infancy. Florentine (1985b) interpreted these initial data as indicative of a possible sensitive period for second-language acquisition that influences a non-native listener's speech perception in the presence of noise.

The purpose of this experiment was to determine if the listeners identified in Florentine's study are typical of non-native listeners and if age of second-language acquisition influences speech perception in

 1092-4388/97/4003-0686

noise. Native Mexican-Spanish-speakers who learned American-English before age 6 or after age 14 were compared with a group of native American-English speakers. As in our earlier studies (Florentine, 1985a, 1985b; Roussohatzaki, 1990), the Speech Perception in Noise Test (SPIN; Kalikow, Stevens, & Elliott, 1977; Bilger, Nuetzel, Rabinowitz, & Rzeczkowski, 1984) was used.

Method

Listeners

The listeners were divided into four groups on the basis of their language background. The monolingual group (MON) consisted of 9 listeners who learned English from birth and ranged in age from 20 to 29 years old. The bilingual-since-infancy group (BSI) included 3 listeners who could not determine which language they learned first, because they spoke Spanish with one parent and English with the other. They were between 21 and 35 years old. The bilingual-since-toddler group (BST) included 9 listeners who learned English as a second language before the age of 6. They were between 22 and 37 years old and spoke English for an average of 24 years (range: 17–32 years). The bilingual-post-puberty (BPP) group consisted of 9 listeners who learned English after the age of 14. They were between 21 and 49 years old and spoke English for an average of 9.4 years (range: 3–26 years). All non-native listeners spoke fluent Spanish and English, as determined from their interactions with faculty at the University of Texas at El Paso and/or an interview with the first author, who is bilingual in English and Spanish. The non-native listeners (or at least one of their parents) were born in Chihuahua, Mexico, and they communicated with at least one parent in Spanish. At the time of the study, listeners were being taught in English at the University of Texas at El Paso and had been using English for daily communication for at least 3 years.

All listeners had normal hearing as determined by a screening with 20 dB HL pure tones at 250, 500, 1000, 2000, and 4000 Hz and a negative history for hearing impairments.

Test Materials and Apparatus

All listeners were tested with the SPIN test, which consists of sentences with controlled predictability presented in a competing babble-type noise (Bilger et al., 1984; Kalikow et al., 1977). The test consists of eight lists of 50 short sentences. The target word, a monosyllabic noun, falls at the end of every sentence. The predictability of the target word is controlled so that a listener can rely on context to help determine the final word in half of the sentences (e.g., The boat sailed across the *bay*); no contextual cues are provided in the other half of the sentences (e.g., John was thinking about the *bay*).

Analog recordings of the SPIN sentences were dubbed to a digital audiotape (DAT) and played back on a digital audiotape player (Sony TCD-D8). Signals in the two channels were attenuated separately (Hewlett Packard 350B and 350D), summed and amplified (Scott 406A), and presented diotically through earphones (Sony MDR-V6). All testing was carried out in a quiet room.

Procedure

The experiment consisted of three parts, which were conducted over one session lasting 2 hours. The listeners were given both written and verbal instructions in English. In all three parts, the sentences were presented at 70 dB SPL. First, the practice list of sentences was presented without noise. After each sentence, the listener wrote the target word and repeated it to the experimenter, who also wrote it. If the first 25 responses were correct, the test was terminated; otherwise, a complete list of 50 sentences was presented. Only listeners who attained 96% correct responses in quiet were tested in the second and third parts of this experiment. No listener failed to reach this criterion.

In the second part, sentences were presented together with a competing babble noise. One sentence was presented on each trial. The level of the noise was varied according to an adaptive psychophysical procedure to determine the level at which the listener could repeat the target word correctly 50% of the time. Because high- and low-predictability sentences are mixed within each list, the performance level of 50% correct is an average between the two types of sentences. The noise level was initially set to 55 dB SPL and was increased by 5 dB for each trial until the first error. After an error, the noise level was reduced by 2-dB steps until the listener responded correctly, at which time the noise level was raised by 2-dB steps. This procedure was continued until eight reversals in accuracy occurred. The noise levels at the final six reversals were averaged to obtain an estimate of the level at which the listener could repeat the target words with 50% accuracy.

In the third part of the experiment, each listener was tested at five to seven signal-to-noise ratios chosen to encompass the range of 15–85% accuracy. Noise levels were presented in mixed order and were chosen for each listener around the listener's midpoint, yielding 50% accuracy as established in part two. Listeners in the MON, BSI, and BST groups were tested in 2-dB steps. Listeners in the BPP group were tested in 3-dB steps, which was necessary to obtain results that encompassed the range of 15–85% accuracy.

Data Analysis

The data from the third part of the experiment were analyzed to determine how the age of second-language acquisition affected the listener's ability to discriminate speech in the presence of background noise. The percentages of correct responses were calculated separately for each listener, level of predictability, and noise level. To reduce the effects of the floor and ceiling of 0% and 100% correct responses, the percentages of correct responses were then transformed into *z* scores. This transformation eliminates scores of 0% and 100% correct response because they yield infinite values of *z*. Finally, lines were fitted by the method of least squares to the *z* scores plotted as a function of noise level. Separate lines were fitted for each listener and level of predictability. The Noise Tolerance Level (NTL), defined as the point at which listeners discriminated the target words with 50% accuracy, was determined as the noise level at which this line intercepted a *z* score of 0.

To investigate effects in the raw data, a three-way analysis of covariance (ANCOVA; Data Desk 4.2, 1994) was performed on the *z* scores with independent variables of grouping by linguistic experience, noise level as a continuous variable, and predictability as a discrete factor. To investigate the differences among the groups of listeners in more detail, separate two-way (linguistic background x predictability) analyses of variance (ANOVAs) were performed on the NTLs and slopes. Scheffé tests (level of significance set at $p < 0.05$) were used as appropriate to determine the sources of significant effects indicated by the ANOVA.

Results

To illustrate the data analysis, Figure 1 shows raw data from one typical listener in each of the four groups. In each panel, the percentage of correctly identified target words is plotted as a function of noise level. These psychometric functions are much shallower for the BPP listener than for the MON, BSI, and BST listeners. In addition, the BPP listener shows almost no difference between the slopes for the low- and high-predictability sentences, whereas there is a clear difference for the other listeners.

The percentages of correct responses shown in Figure 1 are plotted as *z* scores in Figure 2. The difference in slopes and tolerable levels of noise are maintained, but the data now can be fit reasonably well by simple lines. Averaged across all listeners, the lines account for 90% of the variance in the high-predictability data and 85% of the variance in the low-predictability data. The slopes of these lines and the noise levels at which they yield *z* scores of zero [i.e., the Noise Tolerance Level

Figure 1. Raw data from four representative listeners (MON23 = monolingual, BSI19 = bilingual since infancy, BST1 = bilingual since toddler, and BPP 11 = bilingual post-puberty). Open circles represent scores obtained for sentences with high predictability; filled circles represent scores obtained for sentences with low predictability.

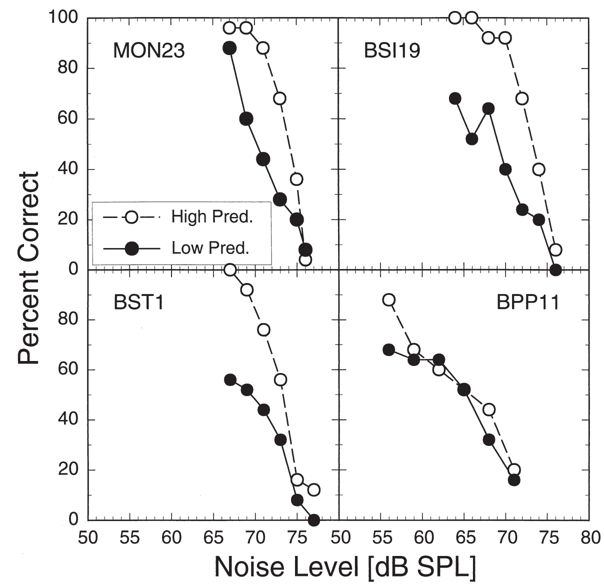
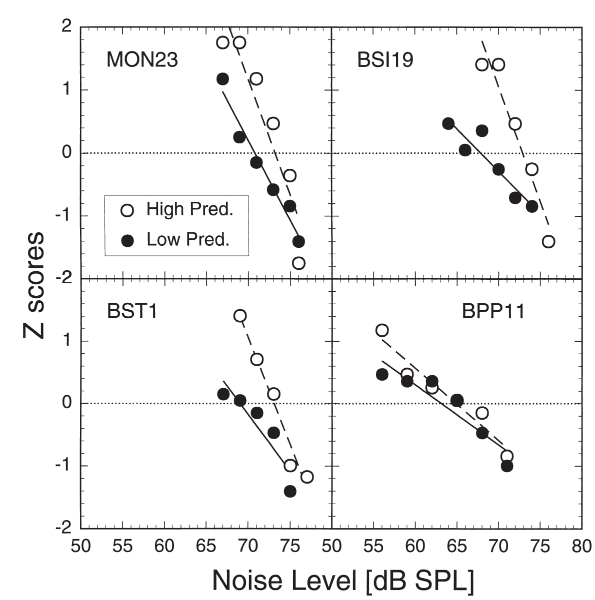


Figure 2. Corresponding *z* scores for the raw data presented for the four listeners in Figure 1. Open circles represent *z* scores transformed from scores obtained for sentences with high predictability; filled circles represent scores obtained for sentences with low predictability. The lines are fit by the method of least squares.



(NTL)] will be used to summarize the listeners' performance. The MON, BSI, and BST listeners show roughly similar NTLs, whereas the BPP listener shows lower NTLs. The lines are steeper for the MON, BSI, and BST listeners than for the BPP listeners, especially for the high-predictability sentences.

Figure 3 shows the average NTLs for the four groups. Clearly, the MON, BSI, and BST listeners can tolerate greater amounts of background noise than the BPP listeners for both the low- and high-predictability sentences. For the high-predictability sentences the difference varies between 5 (BST-BPP) and 6 dB (BSI-BPP and MON-BPP); for the low-predictability sentences it varies between 4 (BST-BPP and BSI-BPP) and 6 dB (MON-BPP). The standard deviations are larger for the BPP listeners than for the other groups. This reflects the fact that this group encompasses individuals with a wide range of exposure to English.

Figure 4 shows the group averages of the slopes of the psychometric functions for the high- and low-predictability sentences. For the MON, BSI, and BST listeners the slopes are much steeper for the high- than for the low-predictability sentences. For the BPP listeners the high- and low-predictability slopes are roughly equal, and both are similar to those of the low-predictability functions for the BSI and BST listeners. As discussed later, these observations are supported by statistical analysis.

Figure 5 shows the average psychometric functions for each group and level of predictability. The lines show z scores as a function of noise level. These lines have slopes equal to the average slope for each group and level of predictability and intersect z equal to zero at the average NTL for each group and level of predictability. As shown

Figure 3. The average Noise Tolerance Level (see text) plotted for high- and low-predictability sentences for each group of listeners: MON (n = 9), BSI (n = 3), BST (n = 9), and BPP (n = 9). Error bars show ±1 standard deviation.

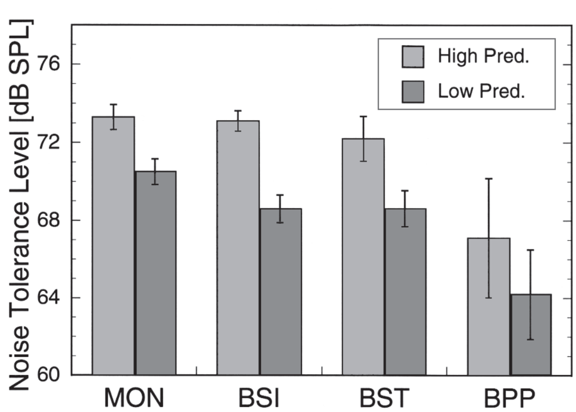


Figure 4. The average slopes of the psychometric functions, taken from lines fit to the z scores transformed from the raw data, plotted for each group of listeners: MON (n = 9), BSI (n = 3), BST (n = 9), and BPP (n = 9). Error bars show ±1 standard deviation. For clarity, the sign of the slope has been reversed to make all values positive.

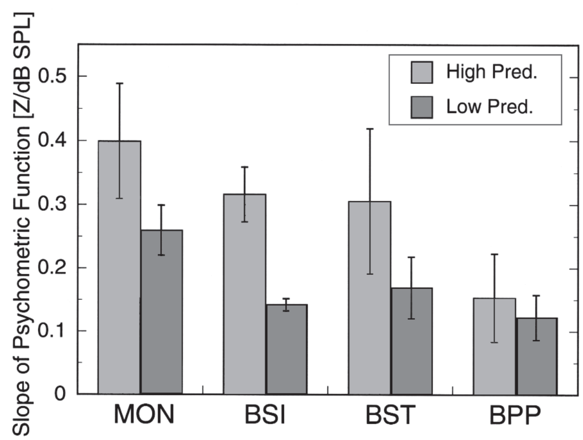
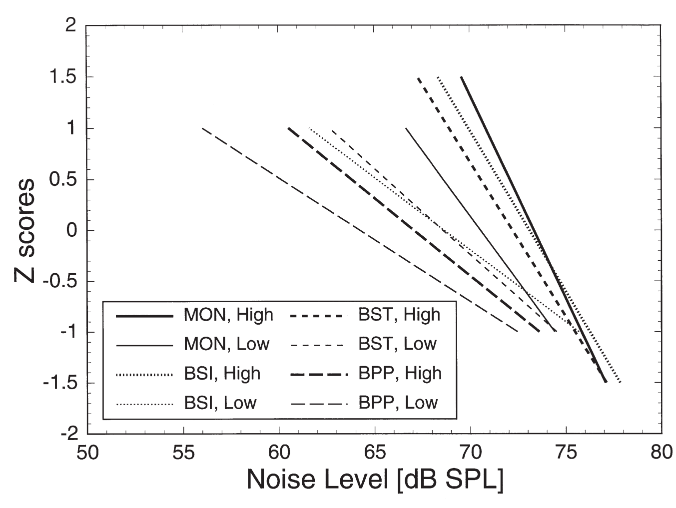


Figure 5. The average psychometric functions for each group: MON (n = 9), BSI (n = 3), BST (n = 9), and BPP (n = 9). The lines are drawn to represent the average slopes and NTLs obtained for each group and level of predictability. Thick lines are for high-predictability sentences; thin lines are for low-predictability. Separate lines are shown for the MON group (solid line), BSI group (dotted lines), BST group (short dashes), and BPP group (long dashes).



in Figures 3 and 4, it is evident that the NTLs and slopes decrease with increasing age of onset of learning. Moreover, Figure 5 shows that the differences between NTLs depend critically on the level of performance used as a criterion for the NTLs. In Figure 3, the NTL was chosen as the noise level at which the listener could just maintain a performance of 50% correct responses. This choice

JSLHR, Volume 40, 686–693, June 1997

of criterion was arbitrary. One could easily argue that normal conversation requires that more than half the words be understood. Therefore, it is of interest to examine the noise levels at which other criterion performances can just be maintained. Generally, the differences in NTLs increase as the criterion increases. For example, the NTLs for the high-predictability sentences would be 10.3 dB lower for the BPP listeners than for the MON listeners if the criterion were $z = 1$ (84% correct) instead of the 6.2-dB difference obtained with the criterion of $z = 0$ (50% correct). On the other hand, if the criterion were $z = -1$ (16% correct) the difference between the MON and BPP listeners would be only 2.2 dB.

The trends described above are supported by the statistical analyses, which were performed with the BSI and BST groups combined into a single group of Early Bilinguals (EB). The BSI and the BST groups were combined because the BSI group consisted of only 3 listeners. One might question whether the BSI group should be combined with the MON group or with the BST group. However, Figures 3 and 4 clearly show that the BSI group is more similar to the BST group than to the MON group. The BSI group shows virtually the same performance on all four dependent measures as the BST group and lower performance than the MON group, except for the high-predictability NTLs. Therefore, the BSI and the BST groups were combined into the EB group. With this grouping, the ANCOVA of the z scores indicates highly significant main effects of noise level, age at onset of learning, and predictability, as shown in Table 1. It also shows that all interactions among these variables are highly significant. These findings indicate that the effect of noise level depends both on when the second language was acquired and on the predictability of the sentences.

To explore these complex interactions, separate analyses were performed on the slopes and the NTLs derived from the individual listeners' psychometric functions. For the NTLs, the ANOVA shows significant effects of group [$F(2, 54) = 74.2$; $p \le 0.0001$] and predictability [$F(1, 54) = 54.3$; $p \le 0.0001$], but no significant interaction [$F(2, 54) = 0.717$; $p = 0.493$]. Scheffé tests for contrast showed that the NTLs were significantly higher for the high- than for the low-predictability sentences. They also showed that the MON group had significantly higher NTLs than the EB group, which in turn had significantly higher NTLs than the BPP group.

For the slopes, the ANOVA also shows significant effects of group [$F(2, 54) = 35.4$; $p \le 0.0001$] and predictability [$F(1, 54) = 35.4$; $p \le 0.0001$] as well as a significant interaction between them [$F(2, 54) = 4.25$; $p \le 0.02$]. The Scheffé test for contrast showed that the MON group had significantly steeper slopes than the EB group for both high- and low-predictability psychometric functions. The EB group in turn had significantly steeper slopes than the BPP group for high-predictability sentences, but not low-predictability sentences. The MON and EB groups had significantly steeper slopes for the high- than the low-predictability sentences, but the BPP group did not.

Table 1. Results of the three-way analysis of covariance of the z scores.

Source	*df*	Sums of squares	Mean square	*F* ratio	*P*
Const	1	.0124	.0124	.0708	0.7903
NL	1	145.5	145.5	834.3	0.0001
LT	2	32.39	16.20	92.84	0.0001
NL*LT	2	27.97	13.98	80.15	0.0001
Prd	1	6.270	6.270	35.94	0.0001
NL*Prd	1	4.936	4.936	28.30	0.0001
LT*Prd	2	4.564	2.282	13.08	0.0001
NL*LT*Prd	2	3.931	1.965	11.27	0.0001
Error	309	53.91	.1745		
Total	320	228.5			

Note. Noise level (NL) was a continuous *factor*, whereas Learn Time [LT; 3 levels: Monolingual (MON), Early Bilingual (EB), and Bilingual Post Puberty (BPP)] and Predictability (Prd; 2 levels: High and Low) were fixed factors.

Discussion

The present study clearly shows that learning a second language at an early age is important for the ability to understand it in noisy situations. The EB listeners tolerated more background noise than BPP listeners, as evidenced by their higher NTLs. This difference is unlikely to be due to differences in education or use of English. All listeners were university students and reported using English for daily communication. On the other hand, it is likely that part of the difference between the EB and BPP listeners may be ascribed to differences in the duration of exposure to English. Yamada (1995) reported that the duration of exposure to a foreign language significantly influences the speech-perception skills of that language, but the age of acquisition and duration of exposure were confounded in her study.

In the present study, duration of exposure also varied with age of learning, because the listeners in the EB and BPP groups were of similar ages. However, duration of exposure does not account for all the differences between the EB and BPP listeners. Regression analyses on the data for the BPP group show no significant correlation between the exposure duration and any of our dependent variables, although the duration of exposure ranged from 3 to 26 years. The correlation coefficients were 0.59 [$F(1, 8) = 3.78$, $p = 0.093$] for the high-predictability slopes, 0.61 [$F(1, 8) = 4.23$, $p = 0.079$] for the high-predictability NTLs, 0.16 [$F(1, 8) = 0.196$, $p = 0.67$] for the low-predictability slopes, and 0.39 [$F(1, 8) = 1.24$, $p = 0.30$] for the low-predictability NTLs.

That exposure is less important than age of onset is further supported by a comparison between the three BPP listeners with the longest exposure and the three EB listeners with the shortest exposure. As shown in Table 2, long-exposure BPP listeners performed worse than the short-exposure EB listeners, despite the equal exposure durations. Not only are the averages different, the ranges do not even overlap for three of the four dependent variables. This indicates that the superior performance of the EB group as a whole is unlikely to be explained by their longer average duration of exposure. Rather, it seems that the age of acquisition is the important variable, and it appears that even extensive exposure will not result in native-like speech-perception-in-noise skills if the second language is not learned in early childhood.

The slopes of the psychometric functions may provide some insight into the reasons for the difficulty experienced by BPP listeners. The present data show a lack of difference between the high- and low-predictability slopes for BPP listeners, in agreement with Florentine's (1985b) data. This finding indicates that late learning of a second language hampers BPP listeners' use of across-word contextual information to improve guessing of the target when its phonetic content is partially obscured by noise. In contrast, the steep high-predictability slopes in the MON and EB listeners indicate that these listeners derive considerable advantage from hearing parts of words in sentences in which the target word can be deduced from the carrier phrase. These listeners appear better able to take advantage of across-word contextual information than the BPP listeners.

To appreciate this interpretation of the slope differences, consider the performance on high-predictability sentences when the average listener in each group is just able to recognize the individual words in the sentence with 50% accuracy. This should be the case when the noise level is set equal to the NTL for the low-predictability sentences because they provide no clues about the target word. To the extent that the listeners' performance on the target word in high-predictability sentences exceeds 50% correct, the listener must have been able to guess some of the target words from the incomplete context provided by 50% of the words of the carrier sentence. (Because the noise level under consideration yields 50% correct with the low-predictability sentences, one must assume that only half the words in the carrier sentence were recognized as isolated words.) In fact, the guessing accuracy should be 2*(Pc–50%), where Pc is the overall percentage of correct responses for the high-predictability sentences. This follows because the improvement in performance comes from guessing the target word from the incomplete context in that 50% of the sentences for which the target word itself was not recognized. For example, assume that the listener gets 65% of the high-predictability sentences correct. Of this score, 50% comes from recognizing (as an isolated word) the target word in half the sentences. The remaining 15% comes from guessing the target word from the incomplete context in 30% of the other half of the sentences, in which the target word itself was not recognized. On the other hand, if a listener gets 85% of the high-predictability sentences correct, the guessing accuracy must be 2*(85%–50%) = 70% for that half of the sentences in which the target word was not recognized. Clearly, the latter listener makes much more efficient use of the context than the former listener.

According to these arguments, the efficiency with which each group of listeners uses across-word contextual information can be assessed by their guessing accuracy for the high-predictability sentences when the noise level is set to the group's NTL for the low-predictability sentences. The psychometric functions in Figure 5 indicate that the MON listeners obtained 86.7% correct ($z = 1.11$) for the high-predictability sentences when

Table 2 . Comparison between the three BPP listeners with the longest exposures to American-English and the three EB listeners with the shortest exposures. For each subgroup, the numbers on top are the averages across the three listeners and the numbers in parentheses are the ranges.

		High-predictability sentences		Low-predictability sentences	
Listeners	Exposure duration	NTL (dB SPL)	Slope	NTL (dB SPL)	Slope
Short-exposure EB	18.3	72.6	–0.226	68.7	–0.178
	(17.0 to 19.6)	(71.7 to 74.4)	(–0.158 to –0.285)	(68.6 to 68.9)	(–0.163 to –0.196)
Long-exposure BPP	18.4	68.7	–0.175	64.7	–0.117
	(13.4 to 25.7)	(67.2 to 71.0)	(–0.115 to –0.282)	(62.7 to 66.3)	(–0.089 to –0.148)

JSLHR, Volume 40, 686–693, June 1997

the noise level was equal to their NTL for low-predictability sentences (70.5 dB SPL). This indicates that context allowed them to guess the target words in 2*(86.7%–50%) = 73.4% of that half of sentences in which they did not recognize the target word itself. The EB listeners obtained 88% correct ($z = 1.18$) on the high-predictability sentences when the noise level was equal to their NTL for low-predictability sentences (68.6 dB SPL), which indicates a guessing accuracy of 2*(88%–50%) = 76%. On the other hand, the BPP listeners obtained only 66.6% correct ($z = 0.43$) on the high-predictability sentences when the noise level was equal to their NTL for low-predictability sentences (64.2 dB SPL). Thus, their guessing accuracy was only 2*(66.6%–50%) = 33.2%, or less than half that for the MON and EB listeners. This indicates that the BPP listeners are much less able to take advantage of across-word context when listening to speech in noise. Thus, it appears that learning a second language at an early age is important for the acquisition of efficient high-level processing of the language, at least in the presence of noise. This is a significant consideration because many situations involve communication in noisy environments (e.g., classrooms, restaurants, social gatherings, trains, airports, etc.).

The efficiency of high-level processing may not be the only factor that depends on early learning. Recall that the EB listeners had shallower psychometric functions for both high- and low-predictability sentences than did the MON listeners. In addition, the NTLs for the low-predictability stimuli are lower for the EB than for the MON listeners. One possible explanation for these findings is that speech perception in noise is degraded by the interaction at the phonetic level of two different language systems. Because performance from the low-predictability sentences primarily depends on low-level phonetic processes, it seems likely that phonetic interference from their Spanish could have influenced their NTLs for American-English. This follows from Grosjean's (1989) suggestion that both language systems are constantly engaged in bilinguals.

In agreement with earlier findings on phonetic discrimination of non-native speech (see Flege, 1995; Yamada, 1995), informal analysis of phonetic errors made by non-native listeners in the present experiment suggests that the listeners' phonetic repertoire was influenced by their linguistic experience. One example was the common perception by late learners of /b/ for /v/, even for target words with high predictability presented at high signal-to-noise ratios. Contrastive sound rules in Spanish are predictive of this error. The /b/ in Spanish is not clearly plosive as is the American /b/; thus it possesses some characteristics of the English /v/.

Similar findings have been reported by Williams (1979). She found that the boundaries for categorical perception of voice-onset time (VOT) in English speech by native Spanish-speaking youngsters gradually shifted as they spent more time in the United States. However, even after 3 years of English exposure, their VOT boundaries still differed from those of American-English monolinguals. Taken together, these studies clearly indicate that listeners who learn a second language after early childhood are likely to have difficulty discriminating some phonemes of their second language, especially when the auditory system is taxed by noise.

Whereas some theories of second-language acquisition suggest that age of acquisition has a gradual effect on second-language proficiency (e.g., Flege, 1995), other theories suggest that a sensitive period for second-language acquisition may exist (Lenneberg, 1967; for further discussion, see Johnson & Newport, 1991). Specialization to native speech sounds is evident as early as 6 months (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). Beyond this age, a language-specific brain emerges and the ability to discriminate non-native speech sounds is reduced (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993).

Marler (1970) examined the existence of critical periods for song birds' acquisition of song by controlling their learning environments. White-crowned sparrows exposed to their native song only during the first 10 days of their lives and those introduced to the song after they were 100 days old showed no evidence that they could acquire the song (Marler, 1970). Marler (1970) suggested that the influence of the critical period for song exposure can be attributed to sensory, rather than motor, constraints and postulated an auditory template. On this basis, it seems that a sensitive period might exist for humans during which exposure to languages profoundly affects the perception of speech. Unfortunately, the bimodal distribution of age of learning and the considerable variability among listeners make it difficult to determine if the effect of age is gradual or if a critical period exists. However, the present data do show beyond a reasonable doubt that early is better for learning a second language.

Conclusions

This study indicates that even when non-native listeners have developed a high level of fluency, their ability to discriminate second-language speech in the presence of noise is influenced by the age during which they acquired the language. Results show:

1. Highly proficient non-native listeners who learned English after puberty perform worse in noise than listeners who learned the language as infants or toddlers, although all listeners performed with native-like proficiency (better than 96% correct) in the quiet.

2. The slopes of the psychometric functions for early bilinguals indicate that they can take better advantage of context than listeners who acquired their second language after puberty.

3. Although early bilinguals demonstrate significantly better speech-discrimination-in-noise skills than listeners who learned the language after puberty, they may not function at the same level as monolinguals, who demonstrate higher NTLs and steeper psychometric functions. Although early second-language acquisition is advantageous for listening in noise, interference from a first language may reduce even an early-bilingual listener's ability to perceive second-language speech in noise.

Acknowledgments

Part of this work was carried out at the Department of Speech-Language Pathology at the University of Texas at El Paso. Professors Joan Manley and Belinda Reyes at the University of Texas at El Paso kindly helped locate listeners. Professors Linda Ferrier, Joanne Miller, and Stephen Sadow from Northeastern University provided helpful advice. The reviewers provided useful comments.

References

Bergman, M. (Ed.). (1980). *Aging and the perception of speech*. Baltimore: University Park Press.

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research, 27*, 32–38.

Buus, S., Florentine, M., Scharf, B., & Canévet, G. (1986). Native French listeners' perception of American-English in noise. *Proceedings of Inter-noise '86*, 895–898.

Data Desk 4.2 [Computer software]. (1994). Ithaca, NY: Data Description Inc.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233–277). Baltimore: York Press.

Florentine, M. (1985a). Non-native listeners' perception of American-English in noise. *Proceedings of Inter-noise '85*, 1021–1024.

Florentine, M. (1985b). Speech perception in noise by fluent, non-native listeners. *Proceedings of the Acoustical Society of Japan*, H-85-16.

Florentine, M., Buus, S., Scharf, B., & Canévet, G. (1984). Speech perception thresholds in noise for native and non-native listeners. *Journal of the Acoustical Society of America, 75*, Suppl. 1, s84.

Grosjean, F. (1989). Neurolinguists beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*, 3–15.

Johnson, J. S., & Newport, E. N. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition, 39*, 215–258.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language, 32*, 402–420.

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America, 61*, 1337–1351.

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

Lenneberg, E. H. (1967). Language in the context of growth and maturation. In E. H. Lenneberg (Ed.), *Biological foundations of language* (pp. 169–181). New York: John Wiley & Sons.

Marler, P. (1970). A comparative approach to vocal development: Song learning in the white-crowned sparrow. *Journal of Comparative Physiology and Psychology, 71*, 1–25.

McAllister, R. (1990). Perceptual foreign accent: L2 users' comprehension ability. In J. Leather & A. James (Eds.), *New Sounds 90* (pp. 216–228). Amsterdam: University of Amsterdam.

Roussohatzaki, M. (1990). *Perception of American-English in noise by Greek listeners*. Unpublished master's thesis, Northeastern University, Boston.

Spolsky, B., Sigurd, B., Masahito, S., Walker, E., & Arterburn, C. (1968). Preliminary studies in the development of techniques for testing overall second language proficiency. *Language Learning, 18*, Special Issue No. 3, 79–101.

Takata, Y., & Nábelek, A. K. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *Journal of the Acoustical Society of America, 88*, 663–666.

Williams, L. (1979). The modification of speech perception and production in second-language learning. *Perception and Psychophysics, 26*, 95–104.

Yamada, R. A. (1995). Age and acquisition of second language speech sounds: Perception of American-English /ɤ / and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 305–320). Baltimore: York Press.

Received August 2, 1996

Accepted January 11, 1997

Contact author: Mary Florentine, PhD, Department of Speech-Language Pathology and Audiology, Northeastern University, 133 Forsyth Building, Boston, MA 02115