# Speech Perception in Quiet and Noise Using the Hearing in Noise Test and the Japanese Hearing in Noise Test by Japanese Listeners

Kaori Nakamura and Sandra Gordon-Salant

**Objective:** The aim of this study was to evaluate the first-language (L1) and second-language (L2) perception abilities in quiet and noise of native Japanese listeners who acquired English late in life and lived in the United States. The study addressed two primary questions: (1) whether native Japanese listeners who developed some fluency in English showed poorer English speech perception ability in quiet and noise than native English listeners, and (2) whether native Japanese listeners living in an English-speaking environment demonstrated poorer speech perception ability in Japanese than native Japanese listeners who reside in Japan and rarely use a second language.

**Design:** Ten native Japanese adults who had excellent English word recognition ability in quiet were evaluated using the English and Japanese versions of the Hearing in Noise Test (HINT). In addition, 10 native English adults were evaluated using the English version of the HINT. All Japanese participants started learning English between the ages of 12 and 13 yrs while at school in Japan and had lived in the United States more than 4 yrs. An adaptive procedure was used to measure English and Japanese speech recognition thresholds (SRTs) in quiet and signal-to-noise ratios (SNRs) at which SRTs were obtained in noise conditions. In addition, a fixed-level procedure was used to obtain percent correct recognition scores at five different SNRs. Both HINT sentence stimuli and noise were presented at a simulated 0° azimuth under headphones.

**Results:** The Japanese group had significantly poorer English speech perception ability than the native English group in all the test conditions under both the adaptive and fixed-level procedures. The psychometric function for the Japanese listeners was shifted by approximately 3 to 4 dB SNR from that of the native English listeners. Japanese speech recognition ability by the Japanese group was not significantly different from that of native Japanese speakers who live in Japan.

**Conclusion:** The Japanese participants who had excellent English word recognition ability in quiet listening conditions failed to reach native-like English speech perception when presented with sentences in quiet conditions. In addition, all 10 Japanese listeners examined in this study had HINT SNRs (i.e., SNRs at which SRTs were obtained in noise conditions) outside the normal range. However, the possible deterioration of L1 speech perception after living in the L2 environment for an extended period of time was not observed among the Japanese, late L2 learners in this study. The implication of this study for clinical settings is that it is important to use L1 sentence materials in both quiet and noise to examine the speech recognition performance of L2 (English) users who had acquired L2 after puberty.

(Ear & Hearing 2011;32;121–131)

## INTRODUCTION

Non-native listeners have difficulty with the perception of English speech, especially in adverse conditions (e.g., noisy or reverberant conditions) (Gat & Keith 1978; Nabelek & Donahue 1984; Takata & Nabelek 1990; Mayo et al. 1997). Previous studies have indicated that non-native listeners across different native language (L1) groups have exhibited poorer second language (L2) word and sentence recognition ability in adverse conditions than native listeners (Gat & Keith 1978; Nabelek & Donahue 1984; Takata & Nabelek 1990; Mayo et al. 1997; van Wijngaarden et al. 2002; von Hapsburg et al. 2004; Rosenhouse et al. 2006; von Hapsburg & Bahng 2006; Bradlow & Alexander 2007; Weiss & Dempsey 2008). The difficulty that non-native listeners experience in adverse conditions seems to arise from various factors. For example, non-native listeners are less able to use contextual cues than native listeners (Mayo et al. 1997; van Wijngaarden et al. 2002). Non-native listeners are also disproportionately affected by informational maskers containing distracting factors such as speech (Garcia Lecumberri & Cooke 2006). It has also been reported that non-native listeners have poorer L2 phoneme and/or consonant identification ability than native listeners (Cutler et al. 2004, 2008; Garcia Lecumberri & Cooke 2006).

The aim of the present study was to examine the L1 (Japanese) and L2 (English) sentence perception abilities of native Japanese listeners who had excellent English word recognition performance. Sentence-length English speech perception by such listeners apparently has not been examined, although a few studies have tested the word-length English speech recognition ability of native Japanese listeners (Takata & Nabelek 1990; Shimizu et al. 1998, 2002). These studies have generally shown that the Japanese listeners have trouble perceiving English words in adverse conditions.

Native Japanese listeners were chosen for this study because the Japanese language has different syntactic and rhythmic structures from many Indo-European languages. For example, Japanese word order is classified as Subject (S) Object (O) Verb (V), whereas many European languages' word order is classified as SVO. Japanese is considered a mora-timed language, as opposed to stress-timed languages such as English and syllable-timed languages such as Spanish and French (Warner & Arai 2001). Mora-timed languages allocate each mora a similar duration, in the same way as syllable-timed languages allocate each syllable approximately the same duration. A mora can consist of a consonant-vowel, consonant-glide-vowel string, or a vowel alone in Japanese. A long vowel is considered to be two morae. For example, the word "Tokyo" consists of four morae (to-o-kyo-o) (Cutler & Otake 2002).

Native Japanese listeners use the moraic structure to perceive speech sounds. Studies by Dupoux et al. (1999) and Dehaene-Lambertz et al. (2000) demonstrated that Japanese listeners had difficulty discriminating between "vowel-consonant-consonant-vowel (VCCV)" and "VCuCV" nonwords

Department of Hearing and Speech Sciences, University of Maryland, College Park, College Park, Maryland.

0196/0202/10/3201-0121/0 • Ear & Hearing • Copyright © 2010 by Lippincott Williams & Wilkins • Printed in the U.S.A.

(e.g., ebzo versus ebuzo). The Japanese listeners tended to perceive epenthetic [u] vowels within a consonant cluster because a vowel usually comes after a consonant or consonant-glide string, and consonant clusters are not allowed in the Japanese moraic structure. The epenthesis effect was observed in all the Japanese listeners who participated in the Dupoux study, regardless of their proficiency in English or French. By contrast, Japanese listeners were better able to discriminate nonwords that differed only in vowel length than French listeners because a long vowel is considered to be two morae in Japanese (Dupoux et al. 1999). Accordingly, listeners are more sensitive to the rhythmic category of their own language than categories used for other languages. It is possible that English sentence perception by native Japanese listeners may be different from that of native speakers of other languages, because of their initial experience with a language that has unique syntactic and rhythmic structures, although this assertion has not been established.

Takata and Nabelek (1990) examined word-length English speech recognition performance of native Japanese listeners. The Japanese participants, who had acquired English after puberty and had been exposed to English in English-speaking countries for an average of 6 yrs, achieved excellent English word recognition scores in quiet conditions but poorer English word recognition scores in noise (−3 dB signal-to-noise ratio [SNR]) and reverberation (1.2-sec reverberation time) conditions than native English listeners. A study by Shimizu et al. (1998) investigated English and Japanese word recognition performance in quiet by native Japanese listeners who were mostly elective bilinguals (e.g., opted to take English courses in school in Japan) and resided in Japan at the time of testing. The Japanese participants demonstrated poorer recognition of English words than Japanese words. It is unknown whether native Japanese listeners who have excellent word-length English speech recognition ability in quiet conditions would be able to achieve similarly excellent performances on sentence-length English speech recognition tasks.

In addition to investigating the nature of L2 speech perception, it is important to examine whether L1 speech perception ability of late L2 learners is altered or deteriorates after immersion in the L2 environment. If L1 speech perception of a late L2 learner becomes poorer after living in the L2 environment for a substantial period of time, it may be necessary to consider whether it is appropriate to examine the listeners with speech audiometry tests in their native language. Only a few studies have been conducted on this issue, and their results indicated that L1 speech perception declined as L2 experience or L2 proficiency increased (Weiss & Dempsey 2008; von Hapsburg & Bahng 2009). Specifically, Weiss and Dempsey (2008) reported that as L2 experience increased, Spanish (L1) speech perception in both quiet and noise by Spanish–English bilinguals decreased. von Hapsburg and Bahng (2009) reported that Korean–English listeners who had acquired English late in their life and had moderate English proficiency performed poorer on Korean (L1) speech comprehension tasks in difficult listening conditions than their peers with low English proficiency.

Further investigations on the influence of L2 exposure on L1 speech perception are necessary to understand the nature of L1 speech perception while living in an L2 environment. No study apparently has yet compared L1 speech perception of late L2 learners with their counterparts who are monolinguals or use L1 almost solely in their native country. This need is especially pressing given that approximately 18.4% of the U.S. population aged 5 and older (i.e., 47 million people) speak a language other than English at home, including approximately 400,000 people who speak Japanese at home in the United States (Shin & Bruno 2003). Approximately 45% of those people (i.e., 21 million people) speak English less than "very well" (Shin & Bruno 2003).

Certain aspects of L1 seem to change subsequent to extensive L2 exposure. The impact of L2 exposure seems to vary for different aspects of L1. For example, L1 lexicon and semantics seem to be more susceptible to changes than morphosyntax (Pavlenko 2003). In addition, the degree of change might be different depending on the age of L2 acquisition and length of residence in the L2 environment (Silva-Corvalan 1991; Yeni-Komshian et al. 2000; Laufer 2003). For instance, those immersed in the L2 environment for a long period of time might have greater difficulty in lexical retrieval than those immersed for a shorter period of time (Laufer 2003). Those who acquired L2 in their early childhood tend to have poorer L1 pronunciation proficiency and/or verb systems than those who acquired L2 in their late childhood (Silva-Corvalan 1991; Yeni-Komshian et al. 2000).

This study addressed two primary questions. The first objective was to examine whether native Japanese listeners who had excellent word-length English speech recognition ability showed poorer sentence-length English speech recognition performance in quiet and noise than native English listeners. This question was addressed because the Japanese language has different rhythmic and syntactic structures from many other languages and hence, native Japanese listeners might perform differently from individuals with different native language backgrounds on these types of speech recognition tasks. No previous reports have examined English sentence recognition performance by native Japanese listeners. The second objective was to explore whether native Japanese listeners who were surrounded by an English-speaking environment demonstrated poorer speech perception ability in Japanese than native Japanese listeners who resided in Japan and rarely used a second language. This second question was addressed because immigrants may have fewer occasions to speak and listen to their L1 than before coming to the United States. As a result, it is possible that they fail to maintain the same L1 speech perception ability as they achieved before immigrating to the United States. If L1 speech perception of a listener becomes poorer after immersion in the L2 environment, it may be necessary to consider whether it is appropriate to examine the listeners with speech audiometry tests in their native language.

Native Japanese listeners were expected to exhibit English sentence recognition performance in quiet, which was similar to the performance of native English listeners. This hypothesis was based on two factors: (1) the speech materials used in this study consisted of simple, first-grade reading level sentences, and (2) one of the criteria for selection of native Japanese participants was to have excellent English word recognition ability in quiet. Native Japanese listeners, however, were hypothesized to exhibit poorer performance on English sentence recognition tasks presented in noise than native English listeners. This was because earlier findings on speech percep-

**TABLE 1. Demographic data of Japanese listeners**

| Japanese Participants | Age (yrs) | Age of Arrival in the U.S. | No. of Years Residing in the U.S. | Use of English/ Japanese per Day (%)* | Nationality of Spouse | Listen to Radio/ Watching TV in English (hrs/Day) | Reading in English (hrs/Day) | Writing in English (hrs/Day) | Speaking English (hrs/Day) | Overall Self-Rated Scores† |
|---|---|---|---|---|---|---|---|---|---|---|
| J1 | 35 | 26 | 9 | 25/75 | American | 1 | 0 | 0 | 3 | 4.5 |
| J2 | 36 | 30 | 6 | 25/75 | American | 1 | 0 | 0 | 3 | 4.0 |
| J3 | 33 | 27 | 6 | 50/50 | American | 1 | 1 | 0 | 6 | 5.0 |
| J4 | 37 | 28 | 9 | 50/50 | American | 1 | 2 | 1 | 2 | 4.5 |
| J5 | 31 | 25 | 6 | 25/75 | Japanese | 1 | 3 | 2 | 5 | 5.0 |
| J6 | 36 | 32 | 4 | 25/75 | Japanese | 0 | 2 | 1 | 2 | 6.0 |
| J7 | 39 | 34 | 4 | 50/50 | American | 3 | 1 | 0 | 3 | 5.0 |
| J8 | 37 | 15 | 22 | 25/75 | Japanese | 2 | 1 | 2 | 4 | 5.5 |
| J9 | 32 | 28 | 4 | 25/75 | Japanese | 4 | 2 | 1 | 1 | 5.0 |
| J10 | 37 | 30 | 7 | 50/50 | American | 1 | 4 | 4 | 4 | 6.3 |
| Average | 35.3 (2.4) | 27.5 (4.9) | 7.7 (5.1) | 35/65 (12.5/12.5) | | 1.5 (1.1) | 1.6 (1.2) | 1.1 (1.2) | 3.3 (2.0) | 5.1 (0.7) |

*Values in parentheses represent 1SD.*
** Frequency of English and Japanese use in all daily activities combined.*
*† Seven-point scales were used ranging from one (very poor) to seven (native-like).*

tion by non-native listeners have reported that non-native listeners had poorer speech perception in adverse conditions than native listeners. In addition, native Japanese listeners who had acquired English after puberty were expected to perform similarly to Japanese listeners residing in Japan on Japanese sentence recognition tasks presented in quiet and noise conditions. This finding would suggest that immersion in the L2 environment did not affect L1 speech perception performance. The last assumption, however, was uncertain because few studies have been conducted on this matter.

## MATERIALS AND METHODS

### Participants

The study involved 20 adult participants with normal hearing. Participants were divided into two groups: (1) 10 Japanese listeners (9 females and 1 male) who had lived in the United States for more than 4 yrs (mean: 7.7 yrs) and had excellent English word recognition ability in quiet, and (2) 10 native English speakers (3 females and 7 males). A minimum length of residency in the United States was set for the Japanese listeners to ensure that they had been exposed to English in the United States for a sufficient period of time. A study by Flege et al. (1997) indicated that non-native listeners who had arrived United States after the age of 20 yrs and resided in the United States for 7.3 years on average were better able to produce and perceive English vowels than their counterparts who had resided in the United States for 0.7 yrs. Accordingly, an extended period of residence in the United States may improve the English perception of non-native listeners. The 10 native English speakers were primarily monolinguals, except for one participant who had acquired Japanese late in his life. Most of the native English speakers had taken L2 courses in college but reported that they had retained little knowledge about L2 and no daily use of L2.

The group of native Japanese listeners ranged in age from 32 to 39 yrs (mean: 35.3 yrs), and the group of native English listeners ranged in age from 19 to 38 yrs (mean: 31 yrs). The demographic data for the native Japanese listeners are shown in Table 1. The Japanese participants were selected so that they would represent a typical group of native Japanese listeners residing in the United States. All Japanese participants had started learning English between the ages of 12 and 13 yrs at school in Japan and received 6 to 10 yrs of formal English education in Japan. Approximately half of the Japanese participants had enrolled in English programs or taken courses in college in the United States since their arrival. All Japanese participants were circumstantial bilinguals at the time of testing. Circumstantial bilinguals are those who must learn L2 to function in society, work, or school, which is different from elective bilinguals, who choose to learn L2 out of personal interest or as school work (von Hapsburg & Pena 2002). English proficiency of the Japanese participants was obtained through the L2 Language History Questionnaire developed by Li et al. (2006). The participants were asked to self-evaluate their proficiency in four aspects of English (reading, writing, speaking, and listening) using seven-point scales ranging from one (poor) to seven (native-like). The average self-evaluation score by the Japanese participants was 5.1, with a score of 5.0 being "good." The results of the language questionnaire are shown on Table 1.

All participants had normal middle ear function, as evidenced by a normal tympanogram and acoustic reflex thresholds elicited at levels within the 90th percentile for individuals with normal hearing (Gelfand et al. 1990) at 500, 1000, and 2000 Hz at the time of testing. They also had normal hearing sensitivity, with air conduction pure tone thresholds equal to or better than 20 dB HL at the octave frequencies of 250 to 8000 Hz in each ear (ANSI 2004). In addition, all participants had monosyllabic word recognition scores of 90% or greater at 60 dB HL using the Northwestern University Test No. 6 (NU6) (Tillman & Carhart 1966).

### Stimuli

The Hearing in Noise Test (HINT), an open-set format speech recognition test, was used to measure English-sentence speech recognition thresholds (SRTs) in quiet and SNRs at which SRTs were obtained in noise conditions (Nilsson et al. 1994). The HINT sentences and noise were taken from the HINT for Windows 6.2 Audiometric System compact disc (developed by House Ear Institute and distributed by Maico Diagnostics). The test consists of 12 phonetically balanced, equivalent lists of 20 sentences each. Each sentence varies from six to seven syllables in length and is graded at a first-grade reading level. The HINT sentences were recorded by a male native English speaker. The steady state noise was matched to the average long-term spectrum of the sentences.

The Japanese HINT (J-HINT) was used to measure Japanese-sentence SRTs in quiet and SNRs at which SRTs were obtained in noise conditions (Shiroma et al. 2008). The J-HINT sentences and noise were taken from the J-HINT for Windows 6.2 Audiometric System compact disc (developed by House Ear Institute and distributed by Maico Diagnostics). The test consists of 12 phonetically balanced lists of 20 sentences each. Each sentence is constructed using 15 to 18 morae. The sentences used third- to fourth-grade vocabulary and were recorded by a male native Japanese speaker. The steady state noise was matched to the average long-term spectrum of the J-HINT sentences.

Both English and Japanese versions of the HINT, along with other foreign-language versions of the HINT, were developed under the same methodological assumptions and procedures to have comparable measures of speech perception ability (Soli & Wong 2008). Test materials were selected from a large number of simple, everyday sentences (336 sentences for the English version, 1000 sentences for the Japanese version), which were rated according to the naturalness of the sentences by native speakers of the target language. Sentences with low ratings were discarded. The difficulty of the sentences within a test was equalized using the same procedures for the different language versions of the HINT. Sentences that were difficult to perceive were discarded. All speech materials, including English and foreign language materials, were recorded and processed at the same institute under the same procedures (Shiroma et al. 2008; Soli & Wong 2008).

The means and SDs from the normative sample of both English and Japanese versions of the HINT are available. The norms for the English version of the HINT for the 20-sentence lists were obtained from three test centers with a total of 67 native speakers of American English (Vermiglio 2008). Those for the Japanese version of the HINT were obtained from two sites with a total of 85 native Japanese speakers for sound field norms and 65 native Japanese speakers for headphone norms (Shiroma et al. 2008). All the native Japanese speakers who participated in the Shiroma study had been raised and resided in Japan at the time of testing. It is, however, unknown whether all the participants were monolingual.

### Preliminary Procedures

Pure tone thresholds at 250, 500, 1000, 2000, 4000, and 8000 Hz and ipsilateral and contralateral acoustic reflex thresholds at 500, 1000, and 2000 Hz were measured, and a tympanogram was obtained from all participants. Word recognition scores were also measured from all participants using the NU6. Items from the NU6 were presented at 60 dB HL to the participants. All participants were asked about their audiological/otologic history orally. The Japanese participants completed the L2 Language History Questionnaire (Li et al. 2006). The questionnaire includes questions related to (1) language history, such as age of L2 acquisition; years of residence; and proficiency in L2 reading, writing, speaking, and listening, (2) language use and environment, such as frequency of daily use in L1 and L2 and hours per day spent watching TV, reading, writing, and speaking in L1 and L2, and (3) language preference, such as languages used at home, work, and school. Consent forms were obtained from all participants before testing.

### Experimental Procedures

Japanese participants were tested using both the HINT and the J-HINT. The HINT and J-HINT sentences and noise were presented via a laptop computer. The participants listened to the sentences and competing speech-spectrum noise binaurally through TDH-39 earphones in Maico cushions. The output of the headphone was calibrated so that the level of the speech-spectrum noise was 65 dBA, as measured in a 6-$cm^3$ coupler. Verbal instructions were given to the Japanese participants in Japanese.

HINT and J-HINT SRTs, referring to the SRTs obtained using the HINT or J-HINT in quiet conditions, were obtained using the standard adaptive procedure to measure a 50% correct response. HINT and J-HINT SNRs, reflecting the SNRs at which the listeners achieve a 50% correct score in noise conditions, were additionally measured using the standard adaptive procedure. One list of 20 sentences each was used to determine HINT and J-HINT SRTs and SNRs. For quiet conditions, the first sentence in each list was presented at 20 dBA. When it was not repeated correctly, the same sentence was presented repeatedly, and the level was increased in 4-dB increments until the listener repeated the sentence correctly. The subsequent sentences in the list were presented once each. For the first four sentences, the speech level was decreased by 4 dB after a correct response and increased by 4 dB after an incorrect response. After the presentation of the fourth sentence, a threshold was estimated by taking the average of (1) the final presentation level of the first sentence, (2) the presentation levels of the second through fourth sentences, and (3) the level at which the fifth sentence would be presented (i.e., either 4 dB higher or lower than the presentation level of the fourth sentence). The fifth sentence was then presented at the level of the estimated threshold. For the 5th through 20th sentences, the speech level varied in 2-dB steps according to the participant's response. The final SRT was calculated from the average presentation level of the 5th through 20th sentences and the level at which the 21st sentence would be presented.

For measurement of HINT and J-HINT SNRs, the noise was presented at a fixed level of 65 dBA, and the first sentence in the list was presented at 65 dBA. The speech level varied according to the listener's response. The same adaptive protocol was used in the noise conditions as described previously for the quiet conditions to determine the stimulus level at which the listener achieved 50% correct (i.e., the SRT); this value was converted to the SNR.

The participants were asked to repeat the entire sentences that they heard. To familiarize the listeners with the task, one list for each test was presented first to each listener for practice. Some word variations (e.g., "a/the" and "is/was" for the HINT and "ga/wa" for the J-HINT) were accepted and were scored as correct. Both sentence stimuli and noise were presented at a simulated 0° azimuth.

Speech recognition performance was additionally evaluated using fixed SNRs to construct performance-intensity functions. The SNR in dB for the estimated 50% correct point and a slope of each individual's psychometric function at the 50% correct response point were calculated. This is because previous studies have demonstrated that thresholds obtained using adaptive procedures were different from those obtained using the constant stimuli method, even when the two methods had the same target probabilities for correct responses (Taylor et al. 1983; Kollmeier et al. 1988; Stillman 1989). In addition, a study by Gordon-Salant (1987) reported that different test paradigms (i.e., an adaptive procedure versus a fixed-level paradigm) resulted in different sensitivity in identifying age effects on speech recognition in noise among older listeners. Given these differences in results drawn from different test paradigms, both the adaptive procedure and the fixed-level paradigm were used in this study to investigate non-native listeners' L1 and L2 speech perception ability.

The HINT sentences were presented with the noise at five SNRs of −6, −4, −2, 0, and +2 dB. To achieve these SNRs, the noise level was constant at 65 dBA, and the sentence level was fixed at 59, 61, 63, 65, and 67 dBA. The J-HINT sentences were presented with the noise at five SNRs of −9, −7, −5, −3, and −1 dB. For these sentences, the noise level was constant at 65 dBA, and the sentence level was fixed at 56, 58, 60, 62, and 64 dBA. These SNRs were chosen to encompass percent correct scores ranging from approximately 10 to 90%. Normative data for listeners with normal hearing indicated that the average 50% correct response occurred at −2.6 dB SNR for the HINT and −5.3 dB SNR for the J-HINT (Soli & Wong 2008). Slopes (% increment/dB) of performance-intensity functions were reported to be 10.6%/dB for the HINT and 10.2%/dB for the J-HINT (Soli & Wong 2008). In other words, a 1-dB change in SNR resulted in approximately 10% change in speech intelligibility for the HINT and J-HINT. In this study, one list of 20 sentences was presented at each SNR. The sentence stimuli and noise were presented at a simulated 0° azimuth. Percent correct scores were determined by dividing the total number of correctly identified key words by the total number of key words in the list of 20 sentences.

Native English listeners were tested with the HINT only. Verbal instructions were given to native English participants in English. Three types of data were obtained from these listeners: (1) HINT SRTs; (2) HINT SNRs with the same configuration (i.e., signal and noise at a simulated 0° azimuth) and procedure as used for Japanese listeners; and (3) percent correct scores at the same SNRs for the English HINT as used for Japanese participants.

The lists of sentences were randomly presented to each listener, including both Japanese and native English listeners, and no lists were presented more than once for each participant. In addition, the sentence lists were randomly assigned across listening conditions. Because the interlist equivalency of the J-HINT was unknown, this random list assignment was assumed to counterbalance possible differences among the lists. To avoid an order effect, half of the Japanese participants first performed tasks using the HINT and then tasks using the J-HINT. The order was reversed for the other half of the Japanese participants. The HINT and J-HINT each involved seven tasks: SRT, SNR at which SRT was obtained in noise, and percent correct recognition at five fixed SNRs. The order of these tasks was also randomized. Native English participants performed the same seven tasks as Japanese participants with the HINT only. The order of these seven tasks was also randomized for native English listeners.

All participants were tested individually in a sound-treated booth in a single session. The procedure, including preliminary tests, lasted approximately 2 hrs for Japanese participants and 1 hr for native English participants. This protocol was approved by the University of Maryland Institutional Review Board for Research Involving Human Subjects.

## RESULTS

Word recognition scores (NU6) obtained from the native English and Japanese groups were similar: the mean scores for the native English and Japanese groups were 97.7% (SD = 1.6) and 96.3% (SD = 0.7), respectively. The scores were arcsine transformed to attempt to fit the assumption of normal distribution, and then Levene's test was conducted to examine the assumption of homogeneity of variance between the two groups. The result indicated that the assumption of homogeneity of variance was violated ($p < 0.01$) because the variance of word recognition scores for the native English group was greater than the variance for the Japanese group. A possible factor underlying this observation is that the variance of pure-tone thresholds for the native English listeners was significantly greater than that for the Japanese listeners. In addition, the Japanese listeners tended to make the same errors. For example, 9 of 10 Japanese listeners misheard the word "check" as "Jack," and some of these listeners misheard the word "chain" as "Jane." There was no systematic error observed among the native English listeners.

Welch's $t$ test (i.e., $t$ test for unequal variances) was conducted, and the results indicated that the mean word recognition scores for the two groups were statistically different ($t[9.9] = 2.36$, $p < 0.05$). A 1.4% difference in scores is, however, usually considered to be clinically insignificant. This is because a difference in the number of correctly recognized words between the two groups was less than one word.

### Speech Recognition Performance in English

HINT SRTs for the two groups and the normative data from the HINT provided by the test developers (Soli & Wong 2008) are shown in Figure 1. Individual data are listed in Table 2. Levene's test revealed that the variance among the native English listeners was not equal to the variance among the Japanese listeners ($p < 0.05$): the native English listeners had greater variance than the Japanese listeners. This might again be due to a difference in the variances of pure-tone thresholds between the two groups. In addition, the Japanese listeners seemed to make similar errors. For example, they tended to omit possessive adjectives, such as "her" and "his," if these appeared in the middle of sentences. In addition, some words in the speech materials were not familiar to the Japanese listeners: none of the Japanese listeners correctly repeated "cuckoo clock" and "handstand."

The Japanese listeners as a group performed more poorly than the native English listeners. Welch's $t$ test confirmed that the mean HINT SRTs were significantly different between the two groups ($t[13.2] = -4.67$, $p < 0.01$). A comparison of the normative data with the performance data of the two groups was conducted subsequently. The mean HINT SRT obtained from the native English group conformed to these normative data. However, the mean HINT SRT for the Japanese listeners was better than the fifth percentile of the normative data (i.e.,

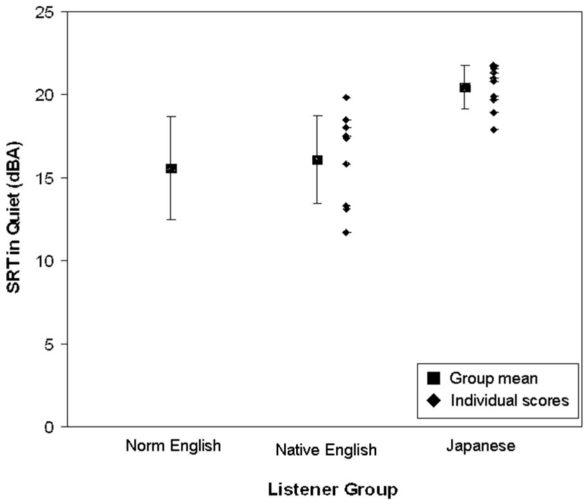


Fig. 1. Individual HINT SRTs, group mean HINT SRTs, and 1SD for the native English and Japanese listeners. The mean HINT SRT and 1SD from the normative sample of native English speakers are shown for comparison (Soli & Wong 2008).

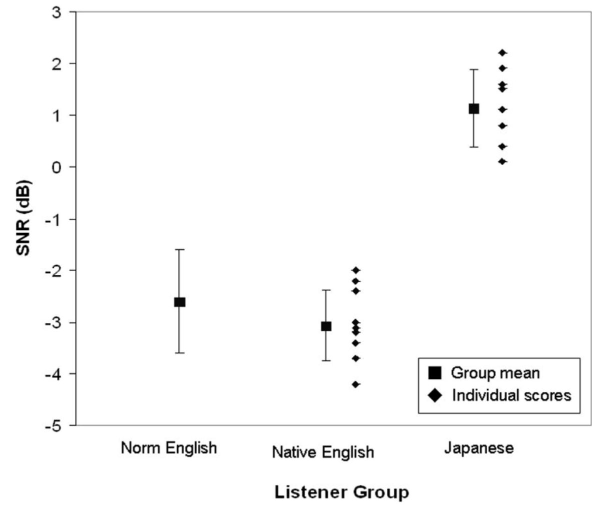


Fig. 2. Individual HINT SNRs, group mean HINT SNRs, and 1SD with the adaptive procedure for the native English and Japanese listeners. The mean HINT SNR and 1SD from the normative sample of native English speakers are shown for comparison (Soli & Wong 2008).

20.7 dBA) (Vermiglio 2008). Six of 10 Japanese participants examined in this study had HINT SRTs that were below the fifth percentile of the normative data.

Figure 2 presents HINT SNRs from the two groups and the normative data from the HINT provided by the test developers (Soli & Wong 2008). Levene's test indicated that the assumption of homogeneity of variance was retained ($p > 0.05$). The Japanese group performed more poorly in noise conditions than the native English group. The independent sample $t$ test confirmed that the mean HINT SNRs were significantly different between the two groups ($t[18] = -13.00$, $p < 0.01$). Although the mean HINT SNR measured from the native English group conformed to the normative data, all the individual HINT SNRs obtained from the Japanese participants were below the fifth percentile of the normative data (i.e., $-1.0$ dB SNR) (Vermiglio 2008).

It is possible that the statistical difference in HINT SNRs between the two groups was due to the significant group difference in HINT SRTs, rather than to the difference in language history between the two groups. Thus, analysis of covariance was conducted to examine the group effect on the dependent variable (i.e., HINT SNRs) after adjusting for the covariate or the initial differences (i.e., HINT SRTs) between the two groups. The results indicated that the assumption of homogeneity of variances was retained ($F[1, 18] = 1.36$, $p > 0.05$), and there was a significant group effect on HINT SNRs after controlling for the effect of HINT SRTs ($F[1, 17] = 96.46$, $p < 0.01$, partial $\eta^2 = 0.85$).

Correlation analyses were additionally conducted to evaluate whether the Japanese listeners' age at the time of arrival to United States or their length of U.S. residency correlated to the HINT SRTs or SNRs. The results indicated that HINT SRTs and SNRs did not correlate with their age at the time of arrival to United States and/or length of U.S. residency ($r = 0.08$ for age versus SRT; $r = 0.02$ for length versus SRT; $r = -0.17$ for age versus SNR; $r = -0.03$ for length versus SNR; $p > 0.05$, for all correlations).

Percent correct recognition scores in the five fixed SNR conditions for the two listener groups are shown in Figure 3. The Japanese group had lower mean percent correct scores at all SNRs tested in this study than the native English group. In addition, only one Japanese listener attained >90% correct score at +2 dB SNR, and the average percent correct score for the Japanese group was 79.2% at +2 dB SNR. By contrast, all the native English listeners attained >97% correct recognition scores at +2 dB SNR. The scores were arcsine transformed to attempt to fit the assumption of normal distribution, and then mixed model repeated-measures analysis of variance was conducted to analyze the data. The within-subject factor was SNR condition (five levels), and the between-subject factor was listening group (native English and Japanese). Mauchly's sphericity test indicated that the sphericity assumption was not violated ($p > 0.05$). The results indicated that there were significant main effects of SNR ($F[4, 72] = 264.62$, $p < 0.01$, partial $\eta^2 = 0.94$) and listening group ($F[1, 18] = 190.37$, $p < 0.01$, partial $\eta^2 = 0.91$) and a significant interaction between the two variables ($F[4, 72] = 3.89$, $p < 0.01$, partial $\eta^2 = 0.18$).

**TABLE 2. Individual data of HINT SRTs and SNRs for the Japanese listeners**

| Japanese Participants | HINT SRT | HINT SNR (Adaptive)* | HINT SNR (Fixed Level)† |
|---|---|---|---|
| J1 | 21.3 | 1.9 | −1.5 |
| J2 | 18.9 | 2.2 | −0.3 |
| J3 | 20.8 | 1.6 | −1.3 |
| J4 | 17.9 | 0.1 | −1.6 |
| J5 | 19.7 | 1.6 | −1.8 |
| J6 | 21.7 | 1.5 | −2.0 |
| J7 | 21.6 | 0.4 | −0.1 |
| J8 | 21.0 | 1.1 | −2.1 |
| J9 | 19.9 | 0.8 | −2.0 |
| J10 | 21.8 | 0.1 | −1.2 |
| Average | 20.5 (1.25) | 1.1 (0.71) | −1.4 (0.66) |

*Values in parentheses represent 1SD.*
** HINT SNR obtained using the adaptive procedure.*
*† HINT SNR obtained using the fixed-level procedure.*

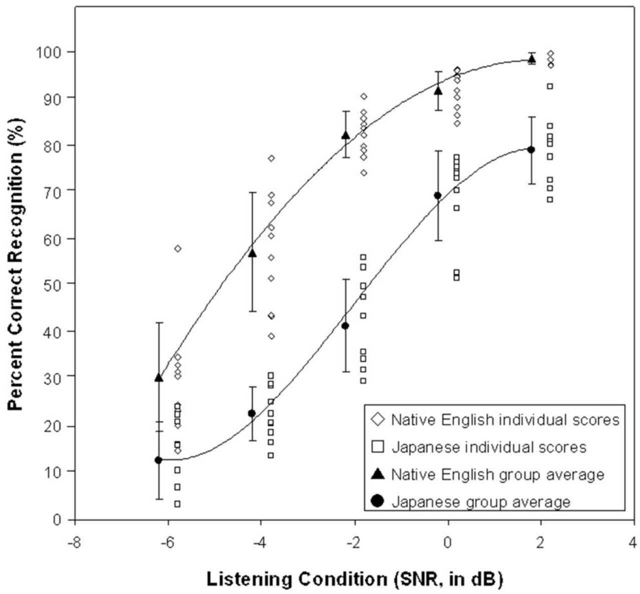


Fig. 3. Psychometric functions using the mean group percent correct scores for the HINT and individual scores obtained from the native English and Japanese listeners. Functions were fitted with third-degree polynomials. Error bars indicate 1SD.

The effect of listening group at each SNR was investigated further to examine the interaction. Levene's test revealed that homogeneity of variance was not retained when SNRs were at −4 and −2 dB ($p < 0.05$ at both SNRs). One possible source of the different group variances was that the Japanese listeners performed uniformly poorly on these conditions, whereas the native English listeners did not. Welch's $t$ tests using a Bonferroni correction for −4 and −2 dB SNR conditions and independent sample $t$ tests using a Bonferroni correction for −6, 0, and +2 dB SNR conditions were conducted. The results indicated that the mean percent correct scores at each SNR were significantly different between the two groups ($t[18] = 4.03$, $p < 0.01$ for −6 dB SNR; $t[13.6] = 7.87$, $p < 0.01$ for −4 dB SNR; $t[15.5] = 11.69$, $p < 0.01$ for −2 dB SNR; $t[18] = 7.61$, $p < 0.01$ for 0 dB SNR; $t[18] = 10.25$, $p < 0.01$ for +2 dB SNR).

In addition, multiple comparison tests using the Bonferroni method were conducted on the data for each group separately to examine whether the mean percent correct scores obtained at the five SNRs were significantly different from each other. For the native English group, all comparisons were statistically significant, except for the mean difference between −2 and 0 dB SNR. For the Japanese group, all comparisons reached a significant level, except for the mean differences between −6 and −4 dB SNR and between 0 and +2 dB SNR. These results indicate that the Japanese group performed significantly poorer on all the SNR conditions than the native English group. For the native English group, a robust improvement in scores was observed when SNR improved from −6 to −2 dB SNR, whereas a robust improvement took place between −4 and 0 dB SNR for the Japanese group.

The estimated 50% correct point was calculated using percent correct recognition scores in the five fixed SNR conditions using the Psignifit software (Wichmann & Hill 2001a,b) implemented with Matlab version 7.6.0 (The MathWorks, Inc.). The mean 50% correct points using the fixed-level procedure for the native English and Japanese groups were −4.5 dB SNR (SD = 0.6) and −1.4 dB SNR (SD = 0.7), respectively. These mean 50% correct points were better than those obtained using adaptive procedures (mean = −3.1 dB SNR, SD = 0.7 for the native English group; mean = 1.1 dB SNR, SD = 0.8 for the Japanese group). Regression analysis was conducted to examine the correlation between the HINT SNRs obtained using the adaptive procedure and the derived SNRs (corresponding to the 50% correct points) obtained using the fixed-level procedure. The results indicated that the correlation coefficient was statistically significant ($r = 0.92$, $p < 0.01$). In addition, the independent sample $t$ test indicated that the mean 50% correct points in dB SNR obtained using the fixed-level procedure were statistically different between the two groups ($t[18[ = -10.70$, $p < 0.01$), as was the case for the adaptive procedure.

Psychometric functions for each group were generated by fitting the mean scores with third-degree polynomials, as shown in Figure 3. The fitted psychometric function for the Japanese listeners was shifted by approximately +3 to 4 dB SNR from that of the native English listeners throughout the range of percent correct scores. This means that in order for the Japanese listeners to perform equally as well as the native English listeners, the signal presentation levels needed to be 3 to 4 dB higher than those for the native English group. The slope of each individual's psychometric function at the 50% correct response point was estimated using the Psignifit software (Wichmann & Hill 2001a,b) implemented with Matlab version 7.6.0 (The MathWorks, Inc.). The mean slope of the psychometric function for the HINT was shallower for the Japanese listeners than for the native English listeners. The independent sample $t$ test confirmed that the average slopes of the functions between the two groups were significantly different ($t[18] = 3.25$, $p < 0.01$). In addition, an analysis of covariance was conducted to examine the group effect on slopes of the functions after controlling for the initial difference in HINT SRTs between the groups. The results indicated that the assumption of homogeneity of variances was retained ($F[1, 18] = 0.003$, $p > 0.05$), and there was a significant group effect on slopes of the function after controlling for the effect of HINT SRTs ($F[1, 17] = 5.49$, $p < 0.05$, partial $\eta^2 = 0.24$). These results indicated that the Japanese group had more gradual improvement in speech recognition performance as SNR improved, compared with the native English group.

In sum, the Japanese listeners had poorer English speech recognition ability than the native English listeners in both quiet and noise conditions. In addition, two different test procedures (an adaptive procedure and a fixed-SNR paradigm) resulted in the same outcome: English sentence recognition performance by the Japanese group was worse than that of the native English group, although both groups exhibited higher SNRs (poorer scores) with the adaptive procedure. Moreover, the Japanese group had a shallower psychometric function slope than the native English group, and the psychometric function for the Japanese group was shifted by +3 to 4 dB from that for the native English group.

**Speech Recognition Performance in Japanese**

J-HINT SRTs and SNRs were additionally compared with the normative data from native Japanese listeners on the

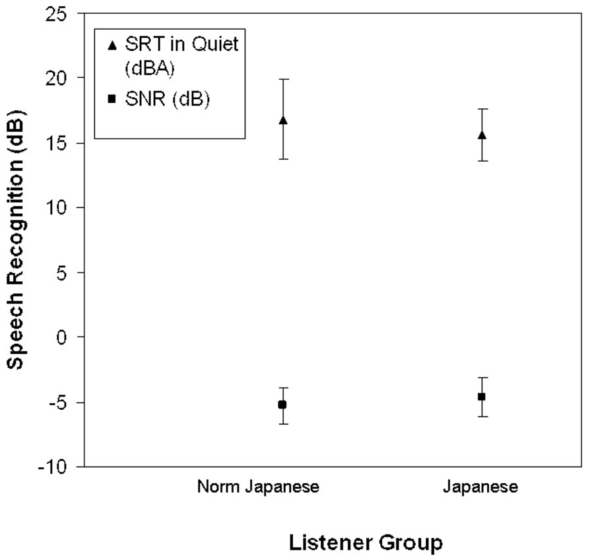


Fig. 4. Mean J-HINT SRTs and SNRs for the Japanese listeners and reference normative comparison group (Shiroma et al. 2008). Error bars reflect 1SD.

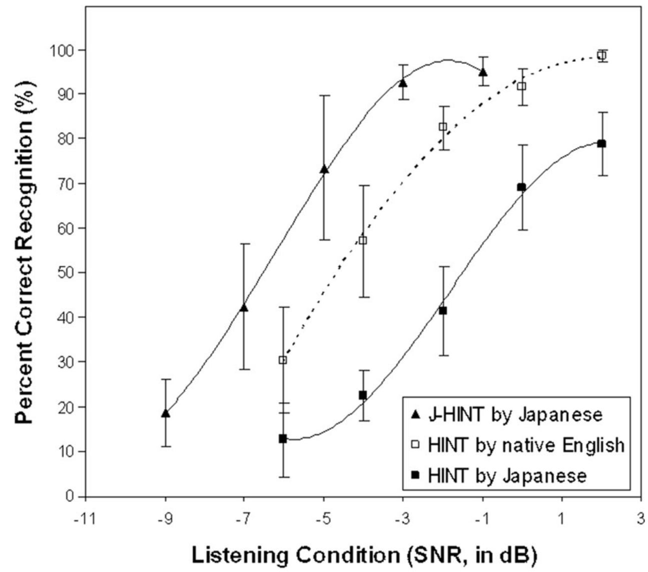


Fig. 5. Psychometric functions of the Japanese listeners measured for the J-HINT and HINT. Also shown is the mean psychometric function of the native English listeners for the HINT. Functions were fitted with third-degree polynomials. Error bars indicate 1SD.

J-HINT provided by Shiroma et al. (2008). These comparisons were conducted to examine whether the Japanese listeners who were late L2 learners living in the United States for an extended period of time had poorer L1 (Japanese) speech perception than native Japanese listeners who had resided in Japan. These data are shown in Figure 4. One-sample $z$ tests indicated that the mean J-HINT SRTs and SNRs obtained from the Japanese group were not significantly different from the normative data ($z = -1.19$, $p > 0.05$ for SRT; $z = 1.49$, $p > 0.05$ for SNR). Accordingly, Japanese speech recognition ability by the Japanese group was not significantly different from the native Japanese speakers who had lived in Japan.

Percent correct recognition scores in the five fixed SNR conditions obtained from the Japanese group using the J-HINT are shown in Figure 5, which also includes the scores obtained using the HINT from the two groups as a reference. The estimated 50% correct points in dB SNR were further calculated using percent correct recognition scores in the five fixed SNR conditions using the Psignifit software (Wichmann & Hill 2001a,b) implemented with Matlab version 7.6.0 (The MathWorks, Inc.). The mean 50% correct point using the fixed-level procedure was $-6.6$ dB SNR (SD $= 0.79$), which was statistically different from that obtained using the adaptive procedure (mean $= -4.5$ dB SNR, SD $= 1.5$) ($t = -5.01$, $p < 0.01$). Regression analysis revealed that the estimated 50% correct point in dB SNR obtained using the fixed-level procedure was not significantly correlated to that obtained using the adaptive procedure ($r = 0.59$, $p > 0.05$). It is unknown why these two SNRs did not correlate to each other. As was the case for the HINT SNRs, the mean J-HINT SNR obtained from the fixed-level procedure was better than that from the adaptive procedures. In general, the speech perception ability of the Japanese listeners seemed to be unchanged even after they had been immersed in English for an extended period of time.

The slope of the psychometric function for the J-HINT obtained from the Japanese group was 20.4% per dB, which was steeper than the slope listed in the normative data obtained from the native Japanese listeners residing in Japan (10.2%/dB) (Soli & Wong 2008). This might be due to a difference in the method used to calculate slopes between the studies.

## DISCUSSION

### Speech Recognition Performance in English

The first question addressed in this study was whether English sentence recognition ability by the Japanese listeners who had excellent word recognition scores in quiet was different from that of the native English listeners. The results of this study showed that the Japanese participants as a group performed more poorly on English sentence recognition tasks in both quiet and noise conditions than the participants who were native English listeners. Poor performance in quiet conditions by the Japanese group was surprising, considering that the HINT consisted of simple, first-grade reading level sentences (Soli & Wong 2008).

Two different test procedures (i.e., adaptive procedure and fixed-level conditions) were used to measure their English sentence perception ability in noise. HINT SNRs obtained using the adaptive and fixed-level procedures strongly correlated to each other. Under both procedures, the performance by the Japanese group was worse than that of the native English group. Even when the effect of group difference in HINT SRTs was controlled statistically, the Japanese listeners had poorer HINT SNRs than the native English listeners. All the 10 Japanese listeners examined in this study had HINT SNRs outside the normal range. In addition, the average percent correct score for the Japanese group was 79.2% at +2 dB SNR, and only one Japanese listener attained greater than a 90% correct score at +2 dB SNR. By contrast, all the native English listeners attained >97% correct recognition scores at +2 dB SNR.

The slopes of the psychometric functions were additionally compared, and the Japanese group had a shallower average slope than the native English group even after controlling for the effect of HINT SRTs. These results were partially expected: it was hypothesized that the Japanese group would have poorer HINT SNRs, lower percent correct scores, and a shallower slope of the psychometric function than the native English group. These expectations were based on results of previous studies indicating that non-native English listeners who had acquired English late in life had poorer English speech recognition ability in adverse conditions than native English listeners (Gat & Keith 1978; Nabelek & Donahue 1984; Takata & Nabelek 1990; Mayo et al. 1997). In addition, a study by Mayo et al. (1997) reported that the slopes of psychometric functions were steeper for native English listeners than for native Spanish listeners who had acquired English in both early and late childhood.

The Japanese group had unexpectedly poorer SRTs in quiet than the native English group. It was hypothesized that the Japanese listeners would demonstrate HINT SRTs that were similar to those obtained from the native English listeners. This was expected because the HINT speech materials comprised simple sentences with a first-grade reading level, and one of the inclusion criteria for the Japanese participants was to have excellent English word recognition ability in quiet (i.e., word recognition scores of ≥90% at 60 dB HL using the NU6). In fact, all the Japanese participants had word recognition scores of ≥94%. However, 6 of 10 Japanese participants had HINT SRTs that were outside the normal range (i.e., below the fifth percentile of the normative data). Three of these participants had resided in the United States for at least 7 yrs at the time of testing.

There have been only a few studies examining English speech recognition performance by native Japanese listeners. A study by Shimizu et al. (1998) investigated English and Japanese speech recognition performance in quiet by native Japanese listeners, who were mostly elective bilinguals and resided in Japan at the time of testing. The Japanese participants had English word recognition scores ranging from 78 to 100%, whereas all participants had Japanese word recognition scores of 100%. The results obtained from the Shimizu study were different from this study in which excellent English word recognition scores (i.e., >94%) were obtained from all the Japanese participants. The better performance by the Japanese listeners in this study may be associated with the more extensive exposure to English as circumstantial bilinguals who reside in the United States.

Takata and Nabelek (1990) used monosyllabic words from the Modified Rhyme Test to examine English speech perception ability by native Japanese listeners who had acquired English after puberty and were circumstantial bilinguals. The Japanese participants in this earlier study achieved excellent English word recognition scores in quiet conditions, similar to findings observed in this study. However, these results also indicated that Japanese participants who had excellent English word recognition ability in quiet listening conditions failed to reach native-like English speech perception when presented with sentences in quiet conditions. This observation has not been reported previously.

One of the reasons for differences in performance between word and sentence materials by the native Japanese listeners is that sentence recognition tasks require use of short-term memory to a greater degree than word recognition tasks. A previous study reported that there was a strong correlation between the familiarity with a language and verbal short-term memory capacity (Thorn & Gathercole 1999). It is possible that the native Japanese listeners had poorer short-term memories for English words because they are less familiar with the language. It also seems that the Japanese listeners had limited knowledge of some of the vocabulary items used in the HINT test, because all of them missed specific words. Thus, use of non-native sentence materials, even in quiet, may place a greater cognitive demand on listeners, resulting in abnormal performance.

Two previous studies investigating English speech perception ability by late English learners using the HINT reported that the late English learners had SRTs in the normal range in quiet listening conditions (von Hapsburg et al. 2004; Weiss & Dempsey 2008). Specifically, von Hapsburg et al. observed that the average HINT SRT by native Spanish listeners who had acquired English after the age of 10 yrs was similar to that of native English listeners. Weiss and Dempsey found that those who acquired English at the age of 11 yrs and older had significantly better HINT SRTs than those who acquired English at the age of 7 yrs and younger.

Differences in linguistic profile and/or language history of participants might have contributed to the differences in English sentence speech perception ability in quiet between the native Japanese listeners examined in this study and native Spanish listeners reported in previous studies (von Hapsburg et al. 2004; Weiss & Dempsey 2008). For example, the average ages at the time of arrival to United States were 13.8, 15.8, and 27.5 yrs for the von Hapsburg et al. study, the Weiss and Dempsey study, and this study, respectively. Accordingly, the native Spanish listeners in the two studies were immersed in English at an earlier age than the native Japanese listeners in this study. There was one Japanese participant (J8) in this study who had a similar age at the time of arrival to United States and the length of U.S. residency as the participants of the von Hapsburg et al. (2004) and Weiss and Dempsey (2008) studies. J8 had migrated to the United States at the age of 15 yrs and had been in the United States for 22 yrs, whereas the remainder of the Japanese participants had arrived United States at the age of 25 yrs and older and had been in the United States between 4 and 9 yrs. The HINT SRT for J8 was 21.0 dBA, which was below the fifth percentiles of the normative data (i.e., 20.7 dBA), and thus outside the normal limits (Soli & Wong 2008). J8 had been in school in the United States since the age of 15 yrs, had been in the process of obtaining a Ph.D. degree at a University, and worked full time in the United States at the time of testing. Despite the extended exposure to English, J8 demonstrated poor English sentence perception ability in quiet.

It is possible that the poor performance by J8 might be attributed to the frequent use of L1. This participant reported using Japanese in 25% of the time on a daily basis. A series of studies by Flege and coworkers indicated that frequency of L1 use affected L2 learners' English vowel perception and production ability: early L2 learners who used L1 frequently (i.e., 25 to 80% of time) had poorer English vowel perception and production performance than their counterparts who used L1 less frequently (i.e., 1 to 15% of time) (Piske et al. 2002, Flege et al. 2003, Flege & MacKay 2004).

L1 differences among the participants of the previous studies (von Hapsburg et al. 2004; Weiss & Dempsey 2008) and this study might also have contributed to the different outcomes. The previous studies examined native Spanish listeners, whereas the present study examined native Japanese listeners. Lexicon, morphology, phonology, and syntax of the two languages, Spanish and Japanese, are different. For example, Japanese word order is classified as Subject (S) Object (O) Verb (V), whereas Spanish word order is classified as SVO, similar to many other Indo-European languages including English. The majority of Japanese vocabulary (i.e., >80% of total vocabulary) comprised original Japanese words and Chinese-influenced words (Kindaichi et al. 2001). In contrast, Spanish vocabulary is derived from Latin, and it is estimated that approximately 28.3% of English vocabulary is influenced by Latin (Finkenstaedt & Wolff 1973). In terms of morphology, a plural marker for both English and Spanish is "s," whereas no plural makers are used in Japanese (Nitta 1997). Hence, Spanish and English have a greater degree of similarity in lexicon, syntax, and morphology than Japanese and English, which may lead to different patterns of English speech perception performance between native Spanish and Japanese listeners.

### Speech Recognition Performance in Japanese

The second question addressed in this study was whether the Japanese listeners who were late L2 learners living in the United States for an extended period of time had poorer L1 (Japanese) speech perception ability than native Japanese listeners who had resided in Japan. All Japanese listeners examined in this study had J-HINT SRTs and SNRs in the normal range (i.e., above the fifth percentile of the normative data), and their J-HINT SRTs and SNRs did not differ significantly from normative statistics obtained from those who had resided in Japan (Shiroma et al. 2008). Thus, the possible decline in L1 speech perception as a result of immersion in an L2 environment for an extended period of time was not observed among the Japanese, late L2 learners in this study. One possible reason for the failure to find a decline in L1 speech perception among these Japanese participants is that they used their native language at least 50% of the time in daily situations, as shown in Table 1. In addition, the Japanese participants reported anecdotally that they maintained their Japanese culture.

By contrast, a study by Weiss and Dempsey (2008) found that L2 experience influenced L1 speech perception. The Weiss and Dempsey study demonstrated that L1 (Spanish) speech perception ability in both quiet and noise tended to decrease as bilingual age (defined as the ratio between the number of years being bilingual and chronological age) increased. In addition, a study by von Hapsburg and Bahng (2009) reported that late English learners with moderate English proficiency had poorer L1 speech comprehension performance in noise conditions than their peers with low English proficiency. These two studies demonstrated some decline in L1 speech perception as L2 experience or L2 proficiency increased. In the Weiss and Dempsey study, however, early English learners were included in the data analyses, and therefore it is possible that a decline in L1 speech perception occurs only in early English learners but not to late English learners. In the von Hapsburg and Bahng study, the Korean Speech Perception in Noise test was used to assess the participants' L1 speech perception ability. This test involved speech comprehension tasks, rather than speech recognition tasks, which were used in this study. Specifically, the participants in the von Hapsburg and Bahng study were asked to answer specific questions about the sentences that they heard in the presence of noise, rather than repeating sentences they heard. Because there were differences in the participants' language histories and the task participants were asked to perform among the studies, it is difficult to compare the results directly. In a future study, late L2 learners who use L1 less frequently than L2 and/or have lived in an L2 environment longer than the Japanese participants examined in this study should be examined to confirm the current findings (i.e., no decline in L1 speech perception ability).

Taken together with the results from previous findings, the implication of this study for clinical settings is that it is important to use L1 speech materials to examine the speech recognition performance of L2 (English) users who had acquired L2 after puberty. This might be especially critical for native Japanese listeners because all 10 Japanese listeners examined in this study had HINT SNRs outside the normal range, and six had HINT SRTs below the fifth percentile of the normative data when test materials were English sentences. Even those who had been in the United States for 9 yrs or greater and were immersed in English in daily life exhibited poor English speech perception in both quiet and noise. By contrast, there was no indication of decline in Japanese speech perception ability among the Japanese participants after they had been immersed in an English language environment for a substantial period of time. Accordingly, use of Japanese speech materials with Japanese–English bilinguals who are immersed in English late in life is strongly recommended.

## ACKNOWLEDGMENTS

The authors thank Rochelle Newman, Monita Chatterjee, and Tracy Fitzgerald for their helpful and valuable comments during the study.

Address for correspondence: Kaori Nakamura, BA, Department of Hearing and Speech Sciences, University of Maryland, College Park, 0100 Lefrak Hall, College Park, MD 20742. E-mail: knaamura@hesp.umd.edu.

Received September 15, 2009; accepted June 2, 2010.

## REFERENCES

American National Standards Institute. (2004). *Specification for audiometers (ANSI S3.6-2004)*. New York, NY: ANSI.

Bradlow, A., & Alexander, J. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *J Acoust Soc Am*, *121*, 2339–2349.

Cutler, A., Garcia Lecumberri, M. L., Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: Effects of local context. *J Acoust Soc Am*, *124*, 1264–1268.

Cutler, A., & Otake, T. (2002). Rhythmic categories in spoken-word recognition. *J Mem Lang*, *46*, 296–322.

Culter, A., Weber, A., Smits., R., et al. (2004) Patterns of English phoneme confusion by native and non-native listeners. *J Acoust Soc Am*, *116*, 3668–3678.

Dehaene-Lambertz, G., Dupoux, E., Gout, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. *J Cogn Neurosci*, *12*, 635–647.

Dupoux, E., Kakehi, K., Hirose, Y., et al. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *J Exp Psychol Hum Percept Perform*, *25*, 1568–1578.

Finkenstaedt, T., & Wolff. D. (1973). *Ordered Profusion. Studies in Dictionaries and the English Lexicon.* Heidelberg, Germany: Carl Winter Universitatsverlag.

Flege, J. E., Bohn, O. S., Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *J Phon*, *25*, 437–470.

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Stud Sec Lang Acquis*, *26*, 1–34.
Flege, J. E., Schirru, C., MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Commun*, *40*, 467–491.
Garcia Lecumberri, M. L., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *J Acoust Soc Am*, *119*, 2445–2454.
Gat, I. B., & Keith, R. W. (1978). An effect of linguistic experience. Auditory word discrimination by native and non-native speakers of English. *Audiology*, *17*, 339–345.
Gelfand, S. A., Schwander, T., Silman, S. (1990). Acoustic reflex thresholds in normal and cochlear-impaired ears: Effects of no-response rates on 90th percentiles in a large sample. *J Speech Hear Disord*, *55*, 198–205.
Gordon-Salant, S. (1987). Age-related differences in speech recognition performance as a function of test format and paradigm. *Ear Hear*, *8*, 277–282.
House Ear Institute (n.d.). *Operation Instructions HINT for Windows Audiometric System.* Eden Prairie, MN: Maico Diagnostics.
House Ear Institute (n.d.). User Guide. Japanese Hearing in Noise Test (HINT) Speech Module Version 1.0. Eden Prairie, MN: Maico Diagnostics.
Kindaichi, K., Saeki, U., Oishi, H., et al. (Eds). (2001). *新選国語辞典 (第 8 版)* [Newly Edited Japanese Dictionary (8th ed.)]. Tokyo, Japan: Shogakukan.
Kollmeier, B., Gilkey, R. H., Sieben, U. K. (1988). Adaptive staircase techniques in psychoacoustics: A comparison of human data and a mathematical model. *J Acoust Soc Am*, *83*, 1852–1862.
Laufer, B. (2003). The influence of L2 on L1 collocational knowledge and on L1 lexical diversity in free written expression. In V. Cook (Ed). *Effects of the Second Language on the First* (pp, 19–31). Clevedon, England: Multilingual Matters.
Li, P., Sepanski, S., Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behav Res Methods*, *38*, 202–210.
Mayo, L. H., Florentine, M., Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *J Speech Lang Hear Res*, *40*, 686–693.
Nabelek, A. K., & Donahue, A. M. (1984). Perception of consonants in reverberation by native and non-native listeners. *J Acoust Soc Am*, *75*, 632–634.
Nilsson, M., Soli, S. D., Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am*, *95*, 1085–1099.
Nitta, Y. (1997). *Nihongo bunpo kenkyu jyosetu: Nihongo no kijyutsu o mezashite* [日本語文法研究序説 – 日本語の記述を目指して] Tokyo, Japan: Kuroshio Shuppan.
Pavlenko, A. (2003). 'I feel clumsy speaking Russian': L2 influence on L1 in narratives of Russian L2 users of English. In V. Cook (Ed). *Effects of the Second Language on the First* (pp, 32–61). Clevedon, England: Multilingual Matters.
Piske, T., Flege, J. E., MacKay, I. R. A. (2002). The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica*, *59*, 49–71.
Rosenhouse, J., Haik, L., Kishon-Rabin, L. (2006). Speech perception in adverse listening conditions in Arabic-Hebrew bilinguals. *Int J Billing*, *10*, 119–135.
Shimizu, T., Makishima, K., Yoshida, M., et al. (1998). 日本人に対する英語語音聴力検査の試み [Speech audiometry using the American word lists for Japanese subjects with normal hearing]. *Nippon Jibiinkoka Gakkai Kaiho*, *101*, 879–883.
Shimizu, T., Makishima, K., Yoshida, M., et al. (2002). Effect of background noise on perception of English speech for Japanese listeners. *Auris Nasus Larynx*, *29*, 121–125.
Shin, H. B., & Bruno, R. (2003). *Language Use and English Speaking Ability: 2000. Census 2000 Brief (C2KBR-29).* Washington, DC: U.S. Census Bureau. Retrieved July 29, 2005, from http://www.census.gov/prod/2003pubs/c2kbr-29.pdf.
Shiroma, M., Iwaki, T., Kubo, T., et al. (2008). The Japanese hearing in noise test. *Int J Audiol*, 47:381–382.
Silva-Corvalan, C. (1991). Spanish language attrition in a contact situation with English. In H.W. Seliger & R.M. Vago (Eds). *First Language Attrition* (pp, 151–171). Cambridge, England: Cambridge University Press.
Soli, S. D., & Wong, L. L. N. (2008). Assessment of speech intelligibility in noise with the Hearing in Noise Test. *Int J Audiol*, *47*, 356–361.
Stillman, J. A. (1989). A comparison of three adaptive psychophysical procedures using inexperienced listeners. *Percept Psychophys*, *46*, 345–50.
Takata, Y., & Nabelek, A. K. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *J Acoust Soc Am*, *88*, 663–666.
Taylor, M. M., Forbes, S. M., Creelman, C. D. (1983). PEST reduces bias in forced choice psychophysics. *J Acoust Soc Am*, *74*, 1367–1374.
Thorn, A. S. C., & Gathercole, S. E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *Q J Exp Psychol A*, *52*, 303–324.
Tillman, T. W., & Carhart, T. (1966). An Expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words: Northwestern University Auditory Test No. 6 (Tech. Rep. No. SAM-TR-66-55). Brooks Air Force Base, TX: United States Air Force School of Aerospace Medicine.
van Wijngaarden, S. J., Steeneken, H. J., Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *J Acoust Soc Am*, *111*, 1906–1916.
Vermiglio, A. (2008). The American English Hearing in Noise. *Int J Audiol*, *47*, 386–387.
von Hapsburg, D., & Bahng, J. (2006). Acceptance of background noise levels in bilingual (Korean-English) listeners. *J Am Acad Audiol*, *17*, 649–658.
von Hapsburg, D., & Bahng, J. (2009). Effects of noise on bilingual listeners'first language (L1) speech perception. *Perspect Hear Hear Disord: Res Diagn*, *13*, 21–26.
von Hapsburg, D., Champlin, C. A., Shetty, S. R. (2004). Reception thresholds for sentences in bilingual (Spanish/English) and monolingual (English) listeners. *J Am Acad Audiol*, *15*, 88–98.
von Hapsburg, D., & Pena, E. D. (2002). Understanding bilingualism and its impact on speech audiometry. *J Speech Lang Hear Res*, *45*, 202–213.
Warner, N., & Arai, T. (2001). The role of the mora in the timing of spontaneous Japanese speech. *J Acoust Soc Am*, *109*, 1144–1156.
Weiss, D., & Dempsey, J. J. (2008). Performance of bilingual speakers on the English and Spanish versions of the Hearing in Noise Test (HINT). *J Am Acad Audiol*, *19*, 5–17.
Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys*, *63*, 1293–1313.
Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Percept Psychophys*, *63*, 1314–1329.
Yeni-Komshian, G. H., Flege, J. E., Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Biling (Camb Engl)*, *3*, 131–149.