The clear speech effect for non-native listeners[a)]

Ann R. Bradlow[b)] and Tessa Bent
Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, Illinois 60208

(Received 24 July 2001; accepted for publication 24 April 2002)

Previous work has established that naturally produced clear speech is more intelligible than conversational speech for adult hearing-impaired listeners and normal-hearing listeners under degraded listening conditions. The major goal of the present study was to investigate the extent to which naturally produced clear speech is an effective intelligibility enhancement strategy for non-native listeners. Thirty-two non-native and 32 native listeners were presented with naturally produced English sentences. Factors that varied were speaking style (conversational versus clear), signal-to-noise ratio (-4 versus -8 dB) and talker (one male versus one female). Results showed that while native listeners derived a substantial benefit from naturally produced clear speech (an improvement of about 16 rau units on a keyword-correct count), non-native listeners exhibited only a small clear speech effect (an improvement of only 5 rau units). This relatively small clear speech effect for non-native listeners is interpreted as a consequence of the fact that clear speech is essentially native-listener oriented, and therefore is only beneficial to listeners with extensive experience with the sound structure of the target language. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1487837]

PACS numbers: 43.71.Gv, 43.71.Es, 43.71.Hw [KRK]

I. INTRODUCTION

Previous work has shown that naturally produced clear speech is more intelligible than conversational speech for a variety of listener populations under various listening conditions. In a seminal paper, Picheny, Durlach, and Braida (1985) reported a 17–20-percentage point clear speech intelligibility advantage for hearing-impaired adults. This clear speech effect was robust across several talkers, listeners, presentation levels, and frequency-gain characteristics. This result has since been extended to both normal and hearing-impaired listeners in noise and reverberation (Payton *et al.*, 1994; Uchanski *et al.*, 1996), and to elderly hearing-impaired listeners (Schum, 1996). While there is some variance across talkers in the magnitude of the clear speech intelligibility advantage they can provoke, the available data indicate that the clear speech effect (i.e., the intelligibility difference between conversational and clear speech) is generally robust across a wide range of talkers, listeners, and signal degradation conditions.

The non-native listener population is in many respects the opposite of the listener population that has typically been the focus of clear speech research. Most previous studies of clear speech perception have focused on communicative situations in which speech communication is challenged by the introduction of some factor that impedes the listener's access to the signal either by signal degradation (such as the addition of background noise or reverberation) or by the listener's own hearing loss. In contrast to the problems of signal access experienced by normal listeners in noise and reverberation or by hearing-impaired listeners, the underlying source of the speech perception problems experienced by non-native listeners is their limited experience with the sound system (i.e., the system of phonological contrasts and their phonetic realization) and higher levels of linguistic structure (i.e., vocabulary, syntax, semantics, and pragmatics) of the target language. In other words, non-native listener speech perception deficits arise primarily from problems of access to the language-specific, linguistic code, rather than from problems of access to the speech signal.

The major goal of the present study was to investigate the extent to which naturally produced clear speech is an effective intelligibility enhancement strategy for non-native listeners. If clear speech enhances speech perception in the face of either signal or code access problems (that is, clear speech is effective for both hearing-impaired and non-native listeners), then we may conclude that the intelligibility advantage of clear speech is not limited to listeners who already have extensive experience with the target language. However, if naturally produced clear speech is not as effective a speech perception enhancement strategy for non-native listeners as for native listeners with signal access problems, then we may conclude that clear speech production is essentially native-listener oriented, and its perceptual benefits are only accessible to listeners who already have extensive experience with the target language. Either of these two results would be of considerable interest from both theoretical and practical perspectives since it would provide insight into the mechanism that underlies the intelligibility advantage of clear speech and suggest directions to follow for the design of speech perception improvement strategies that can be optimized for specific listener populations.

Previous research on clear speech production has identified a wide range of acoustic phonetic features that characterize the conversational-to-clear speech transformation. These include modifications that serve to enhance the overall acoustic salience of the signal such that it is more resistant to the adverse effects of background noise or a listener-related perceptual deficit. These modifications include a decreased

---

[a)] An earlier version of this study was presented at the 141st meeting of the Acoustical Society of America, Chicago, IL, June 2001.

[b)] Electronic mail: abradlow@northwestern.edu speaking rate (including the insertion of longer and more frequent interword pauses), a wider dynamic pitch range, greater sound-pressure levels, more salient stop consonant releases, and greater obstruent rms intensity (Picheney *et al.*, 1986).

Additionally, clear speech production involves modifications that serve to enhance the acoustic distance between contrasting speech categories. For example, clear speech production involves less vowel reduction, thereby enhancing the acoustic distance between individual vowels (Picheney *et al.*, 1986; Moon and Lindblom, 1994; Johnson, Flemming, and Wright, 1993; Bradlow, in press; Ferguson and Kewley-Port, in press). Furthermore, Uchanski (1988, 1992) showed that clear speech production involves nonuniform segment duration increases such that durational contrasts are enhanced; for example, while both long/tense and short/lax vowels increased in duration from conversational to clear speech, the lengthening was far greater for the long/tense vowels than for the short/lax vowels, thereby enhancing the length distinction between these contrasting categories. Similarly, Cutler and Butterfield (1990) showed that English syllables before word boundaries were lengthened to a greater extent in clear speech than in conversational speech, thereby enhancing the duration difference between preboundary and non-preboundary syllables. Thus, the available data indicate that clear speech production is guided by both signal enhancement strategies (i.e., those that make the signal generally more acoustically salient) and code enhancement strategies (i.e., those that exaggerate the acoustic distance between contrasting categories).

In addition to incorporating these signal and code enhancements, clear speech production also involves the maintenance of certain language-specific pronunciation norms. For example, Bradlow (in press) showed that clear speech production in both English and Spanish involved maintenance (rather than reduction) of CV coarticulation throughout the duration of long, clear speech syllables where passive constraints on articulator movements could have been overcome (see also Matthies *et al.*, 2001). Furthermore, Ohala (1995) showed that temporal differences that distinguish English stop consonant categories, such as voice onset time and overall word duration, were not exaggerated in clear speech; rather, talkers exhibited a strict adherence to pronunciation norms of the language vis à vis these parameters during clear speech production. In combination with the findings about the enhancement of the acoustic distance between contrasting categories, these findings suggest that clear speech production is guided, at least in part, by principles that directly reflect the phonology and phonetics of the target language.

Previous research on non-native speech perception has identified two major sources of non-native listener difficulty. The first of these sources is specific to the particular native-target language (L1–L2) pairing in question; the second is apparently independent of the language backgrounds of the listener and talker. The L1–L2 specific source of non-native listener difficulty has been well documented (for example, see the contributions to Strange, 1995), and results from the fact that listeners perceive an incoming L2 speech sound as belonging to the L1 category whose members are most similar to it (Best, 1994, 1995; Flege, 1992, 1995; Kuhl and Iverson, 1995). That is, non-native speech sounds are perceived with respect to the sound system, or code, of the native language. In addition to this code mismatch between L1 and L2 which interferes with L2 speech perception, non-native listeners also appear to be disproportionately challenged by degraded signals relative to native listeners. For example, several studies have shown that non-native listener performance declines more sharply than native-listener performance on speech perception tasks with increasing levels of signal distortion either through the addition of background noise or reverberation (Nábělek and Donahue, 1984; Mayo, Florentine, and Buus, 1997; Meador, Flege, and McKay, 2000). Thus, the difficulties imposed by the non-native listeners' lack of experience with the phonological structure and the patterns of phonetic implementation of the target language (i.e., their code access problems) are aggravated by the presence of a signal distorting influence.

Based on this information about clear speech production and about patterns of non-native listener speech perception, we predicted that non-native listeners who have had very limited experience with the spoken form of the target language would show only a small intelligibility benefit for clear speech relative to conversational speech. That is, we expected to find that clear speech would not be as effective a speech intelligibility enhancement tool for non-native listeners as it is for hearing-impaired listeners or for normal listeners with degraded speech. The rationale behind this prediction was that, while the signal enhancements of clear speech production should provide a perceptual benefit to all listeners, the code enhancements (including those that enhance the acoustic distance between contrasting categories and those that involve adherence to pronunciation norms) are likely to be beneficial only for those listeners who are already experienced with the sound structure of the target language. In particular, the slower speaking rate, the wider dynamic pitch range, greater sound pressure levels, more salient stop consonant releases, and greater obstruent rms intensity (i.e., the signal enhancements of clear speech) should promote intelligibility even for listeners who are relatively unskilled in decoding connected speech in the target language (Hazan and Simpson, 2000). However, only a listener who is already sensitive to the important dimensions of contrast that define the sound system of the target language or who is familiar with the detailed pronunciation patterns of native speakers is likely to benefit from the code enhancements of clear speech. For example, only a listener who is already sensitive to the consistent length difference between lax/short and tense/long vowels in English will benefit from the enhanced duration difference between these vowel categories in clear speech. Therefore, we predicted that overall non-native listeners would derive a significant, but relatively small, clear speech benefit.

In order to test this prediction we designed a test of sentence-in-noise perception in which we manipulated speaking style (conversational versus clear) and signal-to-noise ratio (−4 vs −8 dB). Our specific research questions were: (1) How does the magnitude of the clear speech benefit compare across native and non-native listeners? (2) How do

TABLE I. General information about the non-native subjects.

	Number of subjects reporting	Mean	Range
Age	32	25.5 yrs.	22–32 yrs.
Length of English study	31	9.8 yrs.	6–17 yrs.
Age of acquisition of English	31	12.0 yrs.	5–18 yrs.
Length of time in an English speaking community	32	0.26 yrs.	0.02–2 yrs.
TOEFL score[a]	25	626	597–667
SPEAK score[b]	27	37.5	27–52

[a] TOEFL = Test of English as a Foreign Language (mostly written).

[b] SPEAK = Speaking Proficiency English Assessment Kit (a test of spoken English).

increasing levels of noise affect sentence perception by native and non-native listeners? In addition to these primary research questions, we were also interested in investigating some of the factors that may be related to sentence-in-noise perception abilities across individual non-native listeners. Specifically, we wanted to see whether non-native listener sentence-in-noise perception ability correlated with speech production ability and with various demographic variables, including age of acquisition, length of acquisition, and written language ability (as measured by the standardized Test of English as a Foreign Language, or TOEFL).

II. METHOD

The overall design of the study involved testing a group of adult non-native listeners who had very limited prior exposure to spoken English on a sentence-in-noise recognition task. Performance on this task was measured as a function of speaking style and level of background noise. A control group of normal-hearing native American English listeners was also tested on this task. As part of this experiment, we also collected data on the non-native subjects' familiarity with the words in the test sentences, and subsequently recorded them producing the set of test sentences. These recordings were then submitted to intelligibility testing by native American English listeners, thereby allowing us to examine the correlation between sentence perception and production for the group of non-natives. This overall design yielded two primary sets of data: (1) performance on the sentence-in-noise perception task by groups of non-native and native listeners, and (2) sentence production accuracy by the same non-native subjects who performed the perception test as judged by an independent group of native listeners (i.e., not the same native listeners who participated in the perception test).

A. Subjects

A total of 64 normal-hearing adults participated in this study. The test group included 32 non-native speakers of English, recruited from the Northwestern University International Summer Institute and English as a Second Language Program. The International Summer Institute is designed to provide incoming international graduate students from across the university with intensive English language training as well as a general introduction to academic life in the USA during the month before they begin their graduate studies at Northwestern University. All of the participants in this program had already been admitted to a doctoral program and had therefore demonstrated a high level of proficiency with written English communication (as measured by a minimum score of 560 on the paper-and-pencil TOEFL examination or 220 on the computer-based version of the test). However, these students had been identified by their admitting department as likely to have some difficulty with spoken English communication (based on their department's subjective experience of the spoken English skills of previous students from each individual student's home country). Similarly, the subjects recruited from the ESL program all came to the program due to their own (or their department's) recognition of their need to improve their oral and aural English skills. Thus, the population in this study does not represent a random sampling of non-native speakers in the Northwestern University community, nor were the subjects selected based on their native language background; rather, these individuals were selected based on their limited experience in an English-speaking environment. The breakdown according to native language background was as follows: Bengali ($n=1$); Chinese ($n=20$); Hindi ($n=1$); Japanese ($n=1$); Korean ($n=5$); Romanian ($n=1$); Slovakian ($n=1$); Spanish ($n=1$); Thai ($n=1$). Additional information about the group of non-native subjects is provided in Table I. (As indicated in the table, some subjects were not able or did not wish to report some of the information we requested.) The non-native subjects were all paid for their participation in this study.

The native speakers in the control group were all currently enrolled undergraduates at Northwestern University. They ranged in age from 18–30 years, and were all monolingual speakers of American English. They were recruited from the Linguistics Department subject pool and received course credit for their participation in this study. All subjects (from both the non-native test group and the native control group) reported normal speech and hearing at the time of testing.

B. Stimuli

The stimuli for this study consisted of slightly modified versions of the sentence lists included in the Revised Bamford–Kowal–Bench Standard Sentence Test, which was originally developed for use with British children (Bench and Bamford, 1979). The revised set of sentence lists was developed by the Cochlear Corporation for use with American children. Each list consists of 16 simple English sentences with either three or four keywords for a total of 50 keywords per list. Of the original 21 lists, four lists (lists 7, 8, 9, and 10) were selected based on their equivalent intelligibility scores for normal hearing children as reported in Bamford and Wilson (1979). The rationale behind the selection of these sentences for this study was that, due to their limited vocabulary, they would be appropriate for use with a variety of listener populations, including non-native listeners and children (both of whom are likely to have limited receptive and productive vocabularies). The sentence lists used in this study are provided in the Appendix.

Two adult native speakers of Standard American English with no known speech or hearing impairment (one male, age 33 years, and one female, age 40 years) were recorded producing these sentences in a sound-treated booth in the Phonetics Laboratory in the Department of Linguistics at Northwestern University. They read the sentences from a printed list, speaking into a microphone that fed directly into the sound card (SoundBlaster Live) of a desktop computer. Recording was done on a single channel at a sampling rate of 16 kHz using the PRAAT speech analysis software package (developed at The Institute of Phonetic Sciences at the University of Amsterdam, copyright by Paul Boersma and Paul Weenink). The input level was continuously monitored and adjusted so as to ensure maximum gain without exceeding the dynamic range of the recording system.

The talkers produced the sentences under two conditions. First they read the sentences in a conversational speaking style, then they repeated the sentences in a clear speaking style. For the conversational speaking style, they were instructed to read at their normal pace without any particular attention to clarity, as if addressing someone highly familiar with their voice and speech patterns. For the clear speaking style, they were instructed to read the sentences as if speaking to a listener with a hearing loss or from a different language background.

Following the recording sessions, the digital speech files were segmented into sentence-length files. The root-mean-square amplitude of each of the digital speech files was then rescaled to 65 dB SPL. Each file was then digitally mixed with white noise (generated in PRAAT) to yield two speech-plus-noise files, with -4 and -8 dB as the two signal-to-noise ratios. These signal-to-noise ratios were selected on the basis of pilot testing which indicated that normal-hearing, native listeners performed in the mid-to-high range on a sentence intelligibility test with these sentences at these signal-to-noise ratios. Each of the final stimulus files consisted of a 400-millisecond silent leader, followed by 500 milliseconds of noise, followed by the speech-plus-noise file, and ending with a 500-millisecond noise-only tail. The noise in the 500-millisecond, noise-only header and tail was always at the same level as the noise in the speech-plus-noise portion of the stimulus file, that is at 69 and 73 dB for the -4- and -8-dB signal-to-noise ratio conditions, respectively.

C. Procedure

1. Non-native listeners

Each subject participated in two data collection sessions: a perception test session followed by a production recording session. For the first session, subjects were tested in groups of one to three. The data collection session began with a language background questionnaire which probed the subjects' language learning experiences (both native and foreign languages) as well as their self-reported performance on standardized tests of English language proficiency. Next, subjects were seated in front of a computer monitor in a sound-attenuated booth and the sentence-in-noise perception test was administered. Stimulus presentation was controlled by special-purpose experiment running software (SUPERLAB PRO 2.01). The audio files were played out through the computer sound card (SoundBlaster Live) over headphones (Sennheiser HD580) at a comfortable listening level, which was set by the experimenter before the start of the experiment. The subject's task was to listen to the sentence stimulus and to write down whatever she or he heard on specially prepared answer sheets. After each trial, the subject pressed a button on a response box (supplied as part of the SUPERLAB PRO 2.01 package) to elicit the next trial. Each trial was presented only once, but subjects could take as long as they needed to record their responses.

Each subject responded to stimuli from only one talker (either the male or the female talker). Signal-to-noise ratio (-4 versus -8 dB), and speaking style (conversational versus clear) were factors that varied within subjects. Each subject responded to a total of 64 sentences (4 sets of 16 sentences), which were distributed evenly across the two signal-to-noise and speaking style conditions. For all subjects, the first two sets of sentences were presented in the -4-dB signal-to-noise ratio condition, and the second two sets were presented in the -8-dB signal-to-noise ratio condition. This order of presentation of the two signal-to-noise ratio conditions was done to ensure that any practice effect would be counteracted by the decline in signal-to-noise ratio (which was presumed to be more powerful than any practice effect). Within each of the two signal-to-noise ratio blocks, one set of sentences was presented in conversational speaking style and the other was presented in the clear speaking style. The order of presentation of the four sentence lists was counterbalanced across subjects. Over the course of the entire session, each subject heard each of the 64 sentences only once.

After all subjects had finished the sentence-in-noise perception test, a word familiarity rating test was administered. For this test, each of the 154 unique keywords in the complete set of sentences used in the sentence-in-noise perception test was presented to the subjects for a familiarity rating on a scale of 1 to 7 where 1="I don't know this word," 4="I recognize this as an English word but I don't know its meaning," and 7="I know this word." An additional set of 75 filler items was presented as part of this test. These additional words were selected from lists of words that were given low, medium, and high familiarity ratings by native listeners in Lewellen *et al.* (1993) and that were used in previous tests with both native and non-native listeners (Bradlow and Pisoni, 1999). The inclusion of these words ensured that the full range of the familiarity rating scale would be represented by the items in this test. On each trial, the target word was presented in standard American English orthography on the computer screen (using SUPERLAB PRO 2.01 software), and the subject entered his or her familiarity rating by pressing the appropriate button on the keyboard. The current item remained on the screen until a response was recorded, which then triggered the start of the next trial. The order of presentation of the items was randomized. Both ratings and response times were recorded. Subjects were instructed to work as fast as they could without sacrificing accuracy. The entire first data collection session (including the language background questionnaire, the sentence-in-noise perception test, and the word familiarity rating test) lasted approximately 1 hour.

Subjects returned to the Phonetics Laboratory approximately 1–2 weeks after the first data collection session for a recording session. In this session, subjects were recorded individually reading the four sets of sentences that were used in the sentence-in-noise perception test. They were instructed to read at their natural pace without any particular attention to clarity and were given no indication as to the nature of the intended listener. These recordings were made on an Ariel Proport with a Shure SM81 microphone. All subjects read the four sets of sentences in the same order. After the recording, the sound files were converted to the WAV format and transferred to a PC-based computer. The digital speech files were then segmented into sentence length files, and each file's root-mean-square amplitude was rescaled to 65 dB SPL. Each file was then digitally mixed with white noise (generated in PRAAT) to yield a speech-plus-noise file with a +5-dB signal-to-noise ratio. The final files had the same composition as the sentence-in-noise test files: a 400-millisecond silent leader, followed by 500-milliseconds of noise, followed by the speech-plus-noise file, and ending with a 500-millisecond noise-only tail.

2. Native listeners

Two independent groups of native listeners participated in this study. The first group of 32 native listeners served as a control group for the sentence-in-noise perception test that was administered to the non-native listeners. Data collection on this test with the native listener controls followed exactly the same procedure as for the non-native listeners.

The second group of native listeners ($n=40$) served as judges for the non-native sentence productions that were recorded in the second session with the non-native subjects. In order to keep the number of required native listener judges manageable, each non-native talker's overall intelligibility was assessed on the basis of native listener transcriptions of only one of the four sentence lists. Furthermore, in order to avoid possible adaptation on the part of the native listener to the non-native talker's voice and speech patterns, we compiled 8 sentence presentation conditions each of which included 2 sentence productions from each of the 32 non-native talkers for a total of 64 sentences per condition. According to this scheme, each talker contributed 2 sentences to each of the 8 conditions resulting in each talker's overall intelligibility score being based on transcriptions of one set of 16 sentences (2 sentences in each of the 8 condition) from a total of 40 independent listeners (5 listeners in each of the 8 conditions).

D. Data analysis

Accuracy scores on the sentence-in-noise perception test from both the non-native and native listeners were obtained by counting the number of keywords correctly transcribed. We adopted a strict scoring criterion by which a word was counted as correctly transcribed if it included all and only the appropriate affixes (e.g., if a plural marking ‘‘s’’ or past tense marking ‘‘ed’’ was omitted or added the word was counted as incorrect.) Listeners were not penalized for obvious spelling errors. This scoring method resulted in a keyword-correct score out of a possible 50 for each of the four sentence lists for each listener: −4-dB signal-to-noise ratio conversational style, −4-dB signal-to-noise ratio clear style, −8-dB signal-to-noise ratio conversational style, −8-dB signal-to-noise ratio clear style. The scores were converted to percent-correct scores, and then converted to rationalized arcsine transform units (rau) (Studebaker, 1985). This transformation places the scores on a linear and additive scale, thus facilitating meaningful statistical comparisons across the entire range of the scale. An identical scoring system was applied to the native listener transcriptions of the non-native productions.

It is important to note that both the use of meaningful sentences and the strict scoring criterion (in which all and only the appropriate affixes had to be present for a word to be counted as correct) may have penalized the non-native listeners based on their more limited knowledge of English morphology and syntax relative to the native listener controls. Therefore, while the data from this experiment do not allow us to view the effect of language background (native versus non-native) on speech perception in the absence of higher-level linguistic influences on sentence recognition accuracy, they provide a benchmark for assessing sentence recognition under conditions that, at least to some extent, resemble real-world communicative situations. It remains for further research to compare performance under the sentence and data analysis conditions of the present experiment with performance under conditions that isolate speech perception ability from other higher-level linguistic influences.

III. RESULTS

Data from the word familiarity rating task with the non-native listeners showed that the vast majority of the words was highly familiar to the vast majority of the listeners. Familiarity rating data from 2 of the 32 listeners had to be discarded due to an equipment failure in one case and a misunderstanding of the task on the subject's part in the other case. Of the remaining 30 subjects, 6 gave high ratings (greater than 5 on the 7-point scale) to all of the 154 unique keywords in the sentence-in-noise perception test. Of the remaining 24 listeners, only 2 gave low ratings (less than 5) to more than 4% of the words: 1 gave low ratings to 7 words (4.5%) and 1 gave low rating to 11 words (7%). The rest of the listeners gave high ratings (greater than 5) to at least 97% of the words. We found no relationship between the number of low-familiarity ratings given by an individual subject and his or her overall performance on the sentence-in-noise perception test. Therefore, we performed all analyses of the

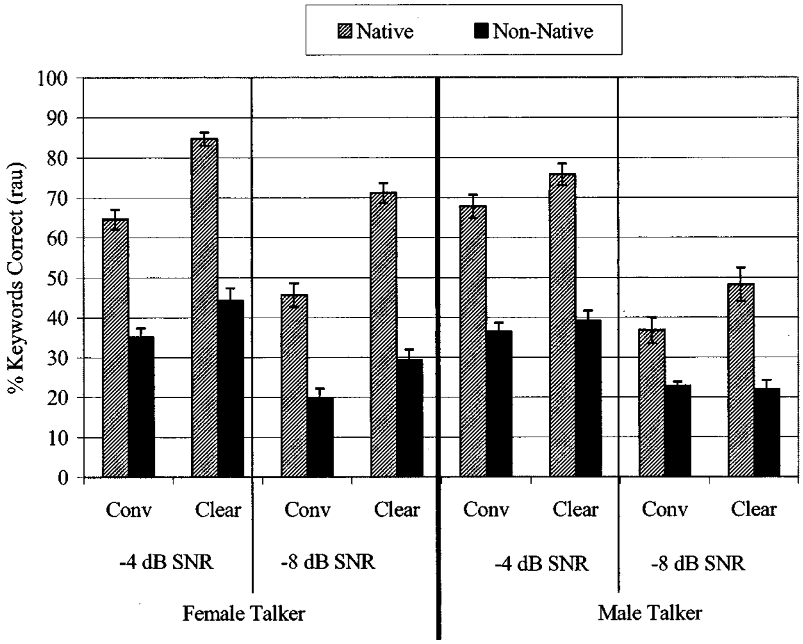

FIG. 1. Overall sentence-in-noise perception test scores (mean keywords correct on the rau scale) for native and non-native listeners. Error bars are the standard error of the mean.

sentence-in-noise perception test with the assumption that the non-native listeners were all sufficiently familiar with the keywords to ensure that this test provided a valid measure of their ability to perceive sentences in noise under the various speaking styles and signal-to-noise ratio conditions incorporated into this test.

Figure 1 shows the overall results of the sentence-in-noise perception test for both non-native and native listeners. Data for subjects that responded to the stimuli from the female and male talkers are shown on the left and right, respectively. As expected, the non-native listeners as a group performed at a much lower level on this test than the native listeners. Furthermore, the two groups of subjects showed different patterns of responses across the various conditions. The data in absolute rau units were submitted to a four-factor repeated measures ANOVA with talker (male versus female) and listener (native versus non-native) as between-subjects factors, and signal-to-noise ratio (-4 vs -8 dB), and speaking style (conversational versus clear) as within-subjects factors. All four main effects were significant [talker: $F(1,60)=9.23$, $p<0.005$; listener: $F(1,60)=272.64$, $p<0.0001$; signal-to-noise ratio: $F(1,60)=248.75$, $p<0.0001$; speaking style: $F(1,60)=93.32$, $p<0.0001$]. Listener interacted significantly with speaking style [$F(1,60)=25.08$, $p<0.0001$] due to the fact that the native listeners exhibited a much greater clear speech effect than the non-native listeners [mean difference$=$11.11 rau; $t(62)=4.28$, $p<0.0001$]. The speaking style by talker interaction was also significant [$F(1,60)=23.66$, $p<0.0001$] due to the fact that the female talker elicited a much greater clear speech effect than the male talker [mean difference$=$10.79 rau; $t(62)=4.125$, $p<0.0001$]. Signal-to-noise ratio entered into significant three-way interaction with talker and listener [$F(1,60)=7.49$, $p<0.01$] due to the fact that, for the native listeners, the decline in performance for the -8-dB signal-to-noise ratio condition relative to the -4-dB signal-to-noise ratio condition was greater for the stimuli produced by the male talker than for the stimuli produced by the female talker [mean difference$=$13 rau; $t(30)=3.47$, $p<0.001$], whereas for the non-native listeners there was no significant difference in the signal-to-noise ratio effect across the two talkers [mean difference$<$1 rau; $t(30)=0.06$, $p=0.96$]. None of the other interactions was significant.

Table II shows the proportional changes that correspond to the absolute scores shown in Fig. 1. For the improvement from conversational to clear speech perception, these proportions were calculated as (clear speech score–conversational speech score)/conversational speech score, and therefore represent the clear speech benefit as a proportion of the conversational speech score. For the decline in performance from the -4-dB signal-to-noise ratio condition to the -8-dB signal-to-noise ratio condition, these proportions were calculated as (-4 dB)–(-8 dB)/(-4 dB), and therefore represent the decline in speech perception performance as a proportion of performance in the -4-dB signal-to-noise ratio condition.

Figure 2 shows the magnitude of the clear speech effect (top panels) and of the signal-to-noise ratio effect (bottom panels) for individual subjects in the male (left panels) and female (right panels) talker conditions. The clear speech effect sizes shown in this figure were calculated by subtracting

TABLE II. Proportional changes in sentence-in-noise performance.

		Native	Non-native
Proportional increase (Clear-Conv)/Conv:			
Female	-4 dB	31%	26%
	-8 dB	56%	49%
Male	-4 dB	12%	7%
	-8 dB	31%	-4%
Proportional decrease (-4 dB)$-$(-8 dB)/(-4 dB):			
Female	Clear	16%	34%
	Conv.	29%	44%
Male	Clear	36%	44%
	Conv.	46%	37%

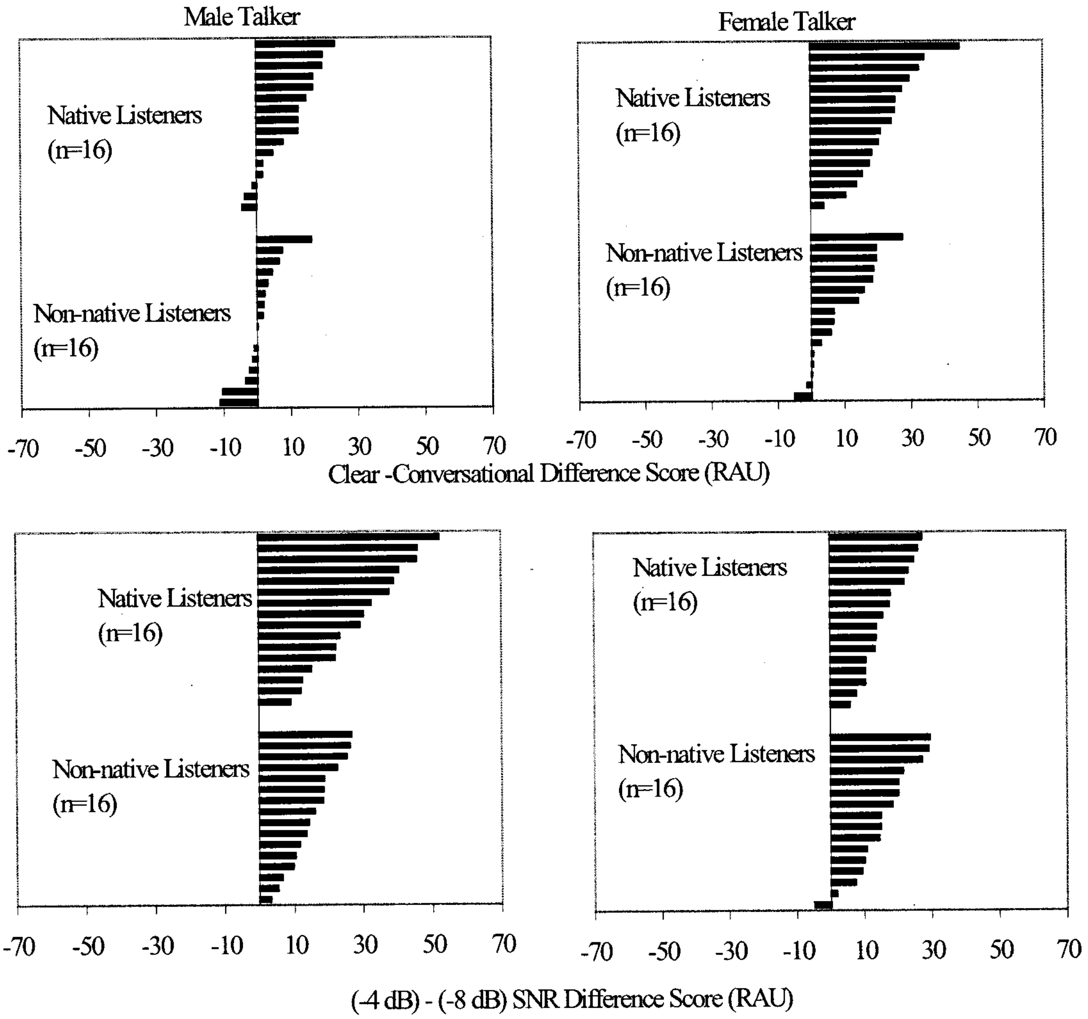

FIG. 2. Magnitude of the clear speech effect (top panels) and of the signal-to-noise ratio effect (bottom panels) for individual subjects in the male (left panels) and female (right panels) talker conditions.

each listener's average sentence-in-noise perception score (across the two signal-to-noise ratio conditions) for the conversational speech stimuli from the corresponding score for the clear speech stimuli. Similarly, the signal-to-noise ratio effect sizes were calculated by subtracting each listener's average sentence-in-noise perception score (across the two speaking style conditions) for the −8-dB signal-to-noise ratio from the corresponding score for the −4-dB signal-to-noise ratio. The plots show that the group mean effects, discussed above and shown in Fig. 1, are reflected quite well when we look at the individual subject data.

In order to gain some insight into the factors that may have contributed to some of the individual variability observed in the size of the clear speech effect across individual non-native listeners, we looked at the relationship between the sentence-in-noise perception scores and the various demographic variables that were collected in the language background questionnaire. This analysis was performed using various sentence-in-noise perception scores: only the conversational speech stimuli, only the clear speech stimuli, and the clear-conversational speech difference score. The most notable result of this correlational analysis (given in Table III) was a negative correlation ($r = -0.358$, $p < 0.05$) between conversational speech perception score and the magnitude of the clear speech effect, indicating that those non-native listeners with poorer conversational speech perception benefit more from clear speech than their fellow non-native listeners who begin with better conversational speech perception. There was also a significant positive correlation between clear speech perception scores and the magnitude of the clear speech effect ($r = +0.738$, $p < 0.05$), probably because relatively good performance with clear speech typically arises from a relatively large clear speech effect. Also of note in this correlational analysis was the positive correlation between sentence production (as assessed by the native listeners' transcriptions of the non-native talkers' productions) and clear speech perception ($r = +0.449$, $p < 0.01$), indicating that those non-natives with relatively good perception abilities also had relatively good production abilities. This perception–production correlation was weaker and just failed to reach significance when perception ability was indexed by conversational speech perception ability ($r = +0.34$, $p < 0.0564$), possibly due to the fact that the conversational speech perception scores showed less variance than the clear speech perception scores (standard deviation for conversational and clear speech perception were 6.67 and 9.23 rau, respectively).

TABLE III. Correlations between sentence-in-noise perception scores and various other demographic variables for the non-native listeners.

	Conv.	Clear	Clear-Conv.	TSE	TOEFL	AOA	LOA	LOI
Conversational ($n=32$, 5–42 rau)								
Clear ($n=32$, 15–53 rau)	0.366[a]							
Clear-conversational ($n=32$, −11–28 rau)	−0.358[a]	0.738[b]						
SPEAK[c] ($n=27$, 27–52)	0.205	0.234	0.081					
TOEFL[d] ($n=25$, 597–667)	−0.085	−0.250	−0.187	0.554				
AOA[e] ($n=31$, 5–18 years)	−0.063	−0.177	−0.132	−0.643	0.270			
LOA[f] ($n=31$, 6–17 years)	0.003	−0.139	−0.142	−0.071	−0.170	−0.436[a]		
LOI[g] ($n=32$, 0.02–2 years)	0.245	0.280	0.104	−0.421	0.094	0.008	−0.259	
Sentence production ($n=32$, 43–93 rau)	0.340[h]	0.449[b]	0.204	0.232	−0.229	−0.211	0.059	−0.061

[a] $p<0.05$.
[b] $p<0.01$.
[c] SPEAK = Speaking Proficiency English Assessment Kit (a test of spoken English).
[d] TOEFL = Test of English as a Foreign Language (mostly written).
[e] AOA = Age of Acquisition.
[f] LOA = Length of Acquisition.
[g] LOI = Length of immersion in an English speaking environment.
[h] $p<0.06$.

Finally, Fig. 3 shows individual subject data for the non-native listeners by native language background. The data are shown separately for those who responded to the female (top panel) and male (bottom panel) talkers. Within each plot, the data are presented in order of performance on the conversational speech materials (striped bars are in ascending order from left to right). The filled bars represent the amount of change in performance for the clear speech materials relative to the conversational speech materials. As shown in this figure, performance varied widely across listeners, even for listeners from the same native language background who responded to stimuli from the same talker. For example, performance of native speakers of Chinese covered a broad range in both baseline performance with conversational materials and in the magnitude of the clear speech benefit regardless of the talker, suggesting that native language background alone did not determine performance on the present test of sentence-in-noise perception.

IV. DISCUSSION

This study was designed to address three main issues: the magnitude of the clear speech effect across native and non-native listeners, the effect of increasing levels of noise on sentence perception by native and non-native listeners, and the factors that are related to sentence-in-noise perception abilities across individual non-native listeners. The findings of this study demonstrated that

(1) The group of non-native listeners exhibited a comparatively small clear speech benefit;
(2) The non-native listeners were not more adversely affected by increasing levels of noise than the native listeners;
(3) None of the demographic variables or test scores that we investigated was related to sentence-in-noise perception ability by the non-native listeners in this study. However, we found a significant negative correlation between the magnitude of the clear speech effect and conversational speech perception, and a significant positive correlation between sentence production ability and clear speech perception ability.

We discuss these finding in reverse order, ending with a discussion of the comparatively small clear speech benefit for the non-native listeners.

The lack of correlation between the various demographic variables and sentence-in-noise perception is somewhat surprising, given the results of earlier work that found significant relationships between some of these variables (e.g., age of acquisition and length of L2 study/exposure) and L2 speech perception performance (e.g., Flege, MacKay, and Meador, 1999). However, it is likely that the lack of a relationship between the sentence-in-noise perception test and any of the other factors that we investigated is due to the sample's uniformly limited amount of exposure to spoken English. Despite the fact that the subjects varied considerably in terms of age of English acquisition (5–18 years) and length of English study (6–17 years), none of these subjects had spent a significant amount of time in an English-speaking environment (none had spent more than 2 years, and all but two had spent under 1 year immersed in an English-speaking environment). In contrast, the non-native population studied by Flege and colleagues had long-term exposure to spoken English (several decades worth). Thus, while the non-native listeners in the present study showed a

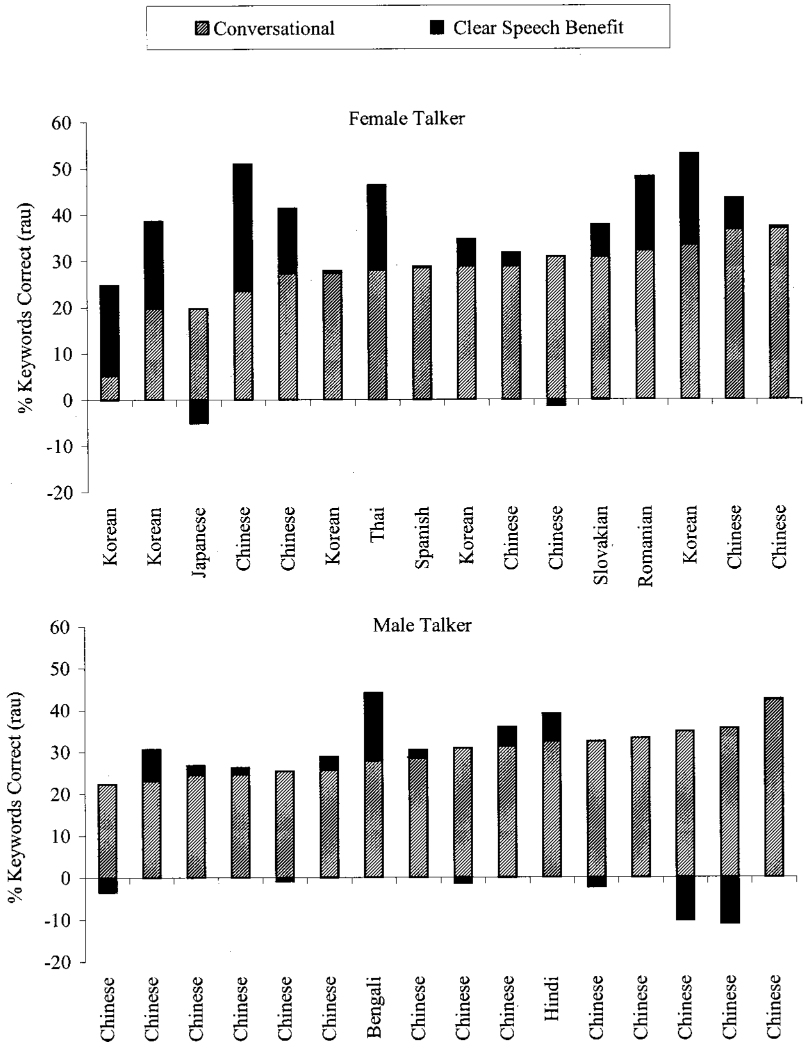

FIG. 3. Individual non-native subject's performance on the sentence-in-noise perception test broken down by native language background. Striped bars show performance with conversational speech stimuli; dark bars show the change in performance level from conversational to clear speech. Data for subjects who listened to the female and male talkers' productions are shown on the top and bottom panels, respectively.

considerable range of overall performance on the sentence-in-noise perception test (range$=$39 rau; minimum$=$15 rau; maximum$=$54 rau), this variability could not be explained by any of the factors that we investigated. This result suggests that non-native listener spoken language processing in the absence of immersion in the target language depends on other factors, which still remain to be identified.

In contrast to the lack of correlation between demographic variables and the sentence-in-noise perception test, we found a significant negative correlation between the magnitude of the clear speech effect (i.e., the ability to take advantage of the talker-related clear speech modifications) and conversational speech perception ($r = -0.358$, $p < 0.05$). This relationship suggests that those non-native listeners with the most room for improvement are generally the ones who derive the most benefit from clear speech, and conversely that as sentence perception accuracy increases, the clear speech benefit decreases. This pattern of results seems to contradict the finding that the group of non-native listeners had a much smaller clear speech benefit than the native listeners, since the group of native listeners certainly had far higher overall speech perception accuracy than the group of non-native listeners. A possible resolution to this apparent paradox is that, whereas the native listeners derived most of their clear speech benefit from the language-specific, code enhancements of clear speech production (i.e., those that exaggerate the acoustic distance between contrasting categories and those that involve adherence to pronunciation norms), the non-native listeners derived most of their clear speech benefit from the signal enhancements of clear speech production (i.e., those that enhance the overall acoustic salience of the signal). Under this view, the negative correlation between conversational speech perception and the magnitude of the clear speech benefit for the non-native listeners is accounted for by supposing that the more proficient non-native listeners (as measured by the conversational speech perception accuracy) have shifted to a more native-like strategy of attempting to take advantage of the code enhancements of clear speech. While this strategy is ultimately efficient and effective, it requires extensive experience with the target language.

Initially, we hypothesized that the beneficial effects of the signal and code enhancements of clear speech production would be additive, that is, we hypothesized that all listeners would benefit from the signal enhancements and that listeners with experience with the spoken form of the target language would derive an additional benefit from the code enhancements. However, the pattern of results for the non-native listeners suggests that perhaps for normal-hearing non-native listeners, the beneficial effects of the signal enhancements are attenuated once the listener reaches a level of sufficient sentence perception proficiency where she or he begins to pay attention to the code enhancements of clear speech. This account for the apparently contradictory findings of (a) a negative conversational correlation between conversational speech perception and the magnitude of the clear speech benefit for the non-native listeners, and (b) a much smaller clear speech benefit for the group of non-native listeners than for the group of native listeners is highly speculative at this point and further investigation is required before any conclusive explanation can be offered.

The present study also found that the group of non-native listeners was not more adversely affected by noise than the group of native listeners; rather, the native listeners showed a greater difference between signal-to-noise ratio conditions than the non-native listeners for the male talker's productions. This result was surprising in view of previous findings to the contrary (Nábèlek and Donahue, 1984; Mayo, Florentine, and Buus, 1997; Meador, Flege, and McKay, 2000); however, it is possible that the non-native listeners in the present study exhibited a floor effect for the −8-dB signal-to-noise ratio condition. Future testing with inexperienced subjects such as those in the present study should involve more favorable signal-to-noise ratios so that overall performance can be elevated, thereby avoiding the possibility of a floor effect.

Furthermore, given that the native listeners in the present study showed different signal-to-noise ratio condition differences across the two talkers, it appears that some talkers may produce speech that is more resistant to noise than others. This talker-related factor is also seen when we compare the effect of speaking style across the two talkers (averaged across both signal-to-noise ratio conditions). For both the native and the non-native listeners, the two talkers had equivalent overall conversational speech intelligibility scores [native listeners: mean male versus female diff. = 2.8 rau, $t(30)=0.86$, $p=0.40$; non-native listeners: mean male versus female diff. = 2.3 rau, $t(30)=0.97$, $p=0.34$]. However, the female talker's clear speech intelligibility was significantly greater than the male's for the native listeners [mean male versus female diff. = 15.6 rau, $t(30)=4.70$, $p<0.0001$]. This difference approached significance for the non-native listeners [mean male versus female diff. = 6.21 rau, $t(30)=1.99$, $p=0.06$]. Furthermore, for the native listeners, the decline in intelligibility for the female's speech in the −8-dB signal-to-noise ratio condition (in the conversational speech condition) was smaller than the male's (male = −31 rau; female = −19 rau). Thus, although both talkers had comparable baseline intelligibility scores, the female talker was more effective at producing clear speech and her conversational speech was more resistant to increasing levels of noise.

In a companion study to the present study (Bradlow, Kraus, and Hayes, submitted), we report the results of a series of acoustic–phonetic measurements of the male's and female's productions that may be responsible for these differences. This analysis must be taken as preliminary because it was based on only two individual talkers; nevertheless, it revealed some notable differences between the clear speech productions of these two individual talkers. In particular, the acoustic measurements showed that for both talkers, conversational-to-clear speech modifications included a decreased speaking rate, longer and more frequent pauses, less alveolar stop flapping, more final stop releasing, greater consonant-to-vowel intensity ratios, a higher mean pitch, a wider pitch range, and an expanded vowel space. However, there were substantial differences between the degree to which each talker's clear speech exhibited these characteristics. In particular, the female decreased overall speaking rate and increased the frequency and duration of pauses far more than the male. The female also increased $F0$ mean and expanded her overall vowel space more than the male. However, the male avoided reduction processes, such as alveolar flapping and unreleased final stop consonants, and increased his pitch range more than the female. Given that the female's clear speech was significantly more intelligible than the male's (despite equivalent conversational speech intelligibility), this pattern of intertalker differences in clear speech production suggests that modifications to temporal characteristics, as well as increased articulatory effort and precision, may be particularly effective for enhancing intelligibility. In contrast, increased pitch range and elimination of ''reduction'' processes (such as alveolar flapping and final consonant releasing) may be less effective for enhancing intelligibility. It remains for further research with more talkers to determine whether these preliminary generalizations are valid.

The major purpose of the present study was to investigate whether non-native listeners would show a speech intelligibility benefit for naturally produced clear speech over conversational speech. This purpose was inspired by the finding of a robust clear speech effect for a wide range of listeners including hearing-impaired adults, elderly adults, and normal-hearing adults in degraded listening conditions. We were particularly interested in investigating the clear speech effect for non-native listeners, because the nature of their speech perception deficit is quite different from that of the other populations that have been the subjects of previous clear speech research. Whereas previous clear speech studies have focused on listeners whose knowledge of the target language is intact but their access to the speech signal is impeded, the present study focused on non-native listeners for whom the speech perception deficit results primarily from a lack of access to the underlying linguistic code. We reasoned that since clear speech production is essentially ''native-listener oriented,'' only those listeners who have experience with the phonology and phonetics of the target language would be able to take full advantage of the conversational-to-clear speech modifications. Nevertheless, we predicted that since clear speech production involves some features that are likely to increase intelligibility for all listeners, such as the slower speaking rate, wider dynamic pitch range, greater sound-pressure levels, more salient stop consonant releases, and greater obstruent rms intensity, non-native listeners would derive a significant, but relatively small, clear speech effect.

TABLE IV. Comparison of the clear speech effect size across various listener populations as reported in the literature (representative, but non-exhaustive).

Study	Materials	Listeners	Presentation conditions	Average clear speech effect
Payton *et al.*, 1994	nonsense sentences	2 HI[a] adults	OMCL freq-gain char. self-adjusted levels noise and reverb. as for NH 1 talker	26 rau
		10 NH[b] adults	3 reverberation conditions 3 noise conditions 3 reverb/noise combinations 1 talker	20 rau
Schum, 1996	meaningful sentences	60 HI adults	multitalker babble 10 young, 10 elderly talkers	20 rau
Picheny *et al.*, 1985	nonsense sentences	5 HI adults	2 freq.-gain characteristics 3 levels 3 talkers	17 rau
Helfer, 1997	nonsense sentences	39 NH adults	multitalker babble, 3 SNRs audio, visual, audiovisual	17 rau
Present study	meaningful sentences	32 NH adults	2 noise levels 2 talkers	16 rau
Bradlow *et al.* (submitted)	meaningful sentences	38 NL[c] children	2 noise levels 2 talkers	9 rau
		77 LI[d]children	2 noise levels 2 talkers	9 rau
Present study	meaningful sentences	32 NN[e]adults	2 noise levels 2 talkers	5 rau

[a] HI = hearing-impaired.
[b] NH = normal hearing.
[c] NL = normal hearing and not learning-impaired.
[d] LI = learning-impaired.
[e] NN = non-native.

The findings of the present study confirmed this prediction by demonstrating a far smaller clear speech effect for the non-native listener group than for the native listener group. Indeed, the average non-native listener clear speech effect of 5 rau was far smaller than the average native listener clear speech effect of 16 rau on this sentence-in-noise perception test. This native listener clear speech effect is comparable in magnitude to previously reported clear speech effects for both hearing-impaired listeners and normal listeners under degraded listening conditions, confirming that our test provided data that could be used for a valid comparison of the clear speech effect across several populations. Table IV presents such a comparison, where the data have been listed in descending order of the magnitude of the clear speech effect. Included in this table are the data from a companion study to the present study which used the same stimuli and experimental design as the present study to investigate the clear speech effect in normal and learning-impaired children (Bradlow, Kraus, and Hayes, submitted). While the data included in this table do not represent an exhaustive list of all clear speech perception studies, they are representative of the general range of the clear speech effect across various listener populations.

It is evident from the data in Table IV that for hearing-impaired listeners and normal listeners with degraded signals the clear speech effect is remarkably consistent across studies. Excluding the data from one study that found an unusually large effect of 26 rau for two hearing-impaired listeners (Payton *et al.*, 1994), the range of the effect is just 4 rau (16–20 rau). [If we include the data from the Payton *et al.* (1994) study, the range is 10 rau.] This consistency indicates that for these listeners the clear speech effect is robust and stable in magnitude across a variety of talker and listening conditions.

Of greater interest for the present study is the finding that the data in Table IV appear to fall into three groups: a group with a large clear speech effect (>15 rau), a group with an intermediate clear speech effect (9 rau), and a group with a small clear speech effect (5 rau). We propose that an important factor that distinguishes the three groups of listeners represented by these levels of the clear speech effect is the stage of target language development, and that this factor plays a major role in determining the likelihood that a listener will benefit from naturally produced clear speech. The listeners in the first group get the large >15-rau clear speech benefit because of their extensive experience with the target language. Thus, despite the fact that their access to the speech signal is impeded, they are able to take advantage of most, if not all, of the clear speech enhancements. The children in the second group (mean age = 10.6 and 10.4 years, s.d. = 1.8 and 2.2 years for the normal and learning-impaired children, respectively) represent a listener population in which the process of language acquisition is not complete. Thus, according to this interpretation, these children were able to derive some benefit from the clear speech enhancements, but this benefit was smaller than for the adults in the first group. (For more discussion of these data see Bradlow *et al.*, submitted.) The third group of subjects, the non-native listeners (who had the least amount of experience with spoken English), showed only a very small clear speech effect. A prediction from this pattern of data is that the magnitude of the clear speech effect develops with age and exposure to the spoken language, and therefore that the individuals in the child and non-native listener groups will eventually attain the large clear speech benefit exhibited by the native listener adults.

A noteworthy limitation of the present data was the use of meaningful sentence materials in the sentence-in-noise perception test. Even though the non-native subjects in this study were quite proficient in English grammar (they all performed well enough on a standardized test of English reading and writing proficiency to be admitted into an American graduate school), it is very likely that they were less able than the native listeners to take advantage of the contextual information offered by the meaningful sentence materials. This important difference may have been particularly salient under the more difficult listening conditions, that is for the −8-dB signal-to-noise ratio and for the conversational speech materials. It is therefore possible that the difference in the magnitude of the clear speech benefit across the two groups of subjects in this study reflected a difference in their ability to take advantage of contextual information rather than a difference in their ability to take advantage of the acoustic–phonetic modifications of English clear speech. It remains for future research to address this important limitation.

We conclude by emphasizing the importance of broadening the range of listener populations that are tested on their ability to take advantage of naturally produced clear speech. It is very likely that different listener populations will find different enhancement features of clear speech more or less beneficial. Without a better understanding of how talker- and listener-related factors interact to influence overall speech intelligibility, we will remain limited in our ability to enhance speech intelligibility through the use of either naturally or synthetically produced clear speech over a wide range of communicative situations.

ACKNOWLEDGMENTS

We are grateful to Rebecca Schwarzlose and Juri Matsubara for data collection and processing, and to Xueqing Xu for programming assistance. We also thank two anonymous reviewers and Keith Kluender for extremely helpful comments on an earlier version of the manuscript. This research was supported by NIH-NIDCD Grant No. DC 03762.

APPENDIX: SENTENCE-IN-NOISE PERCEPTION TEST MATERIALS

List 1 (BKB-R List 7, keywords underlined)

(1) The children dropped the bag.
(2) The dog came back.
(3) The floor looked clean.
(4) She found her purse.
(5) The fruit is on the ground.
(6) Mother got a saucepan.
(7) They washed in cold water.
(8) The young people are dancing.
(9) The bus left early.
(10) They had two empty bottles.
(11) The ball is bouncing very high.
(12) Father forgot the bread.
(13) The girl has a picture book.
(14) The orange was very sweet.
(15) He is holding his nose.
(16) The new road is on the map.

List 2 (BKB-R List 8, keywords underlined)

(1) The boy forgot his book.
(2) A friend came for lunch.
(3) The match boxes are empty.
(4) He climbed his ladder.
(5) The family bought a house.
(6) The jug is on the shelf.
(7) The ball broke the window.
(8) They are shopping for cheese.
(9) The pond water is dirty.
(10) They heard a funny noise.
(11) The police are clearing the road.
(12) The bus stopped suddenly.
(13) She writes to her brother.
(14) The football player lost a shoe.
(15) The three girls are listening.
(16) The coat is on a chair.

List 3 (BKB-R List 9, keywords underlined)

(1) The book tells a story.
(2) The young boy left home.
(3) They are climbing the tree.
(4) She stood near her window.
(5) The table has three legs.
(6) A letter fell on the floor.
(7) The five men are working.
(8) He listened to his father.
(9) The shoes were very dirty.
(10) They went on a vacation.
(11) The baby broke his cup.
(12) The lady packed her bag.
(13) The dinner plate is hot.
(14) The train is moving fast.
(15) The child drank some milk.
(16) The car hit a wall.

List 4 (BKB-R List 10, keywords underlined)

(1) A dish towel is by the sink.
(2) The janitor used a broom.
(3) She looked in her mirror.
(4) The good boy is helping.
(5) They followed the path.
(6) The kitchen clock was wrong.
(7) The dog jumped on the chair.
(8) Someone is crossing the road.
(9) The mailman brought a letter.
(10) They are riding their bicycles.
(11) He broke his leg.
(12) The milk was by the front door.
(13) The shirts are hanging in the closet.
(14) The ground was very hard.
(15) The buckets hold water.
(16) The chicken laid some eggs.

Bamford, J., and Wilson, I. (**1979**). ''Methodological considerations and practical aspects of the BKB sentence lists,'' in *Speech-hearing Tests and the Spoken Language of Hearing-impaired Children*, edited by J. Bench and J. Bamford (Academic, London), pp. 148–187.

Bench, J., and Bamford, J., Editors (**1979**). *Speech-hearing Tests and the Spoken Language of Hearing-impaired Children* (Academic, London).

Best, C. T. (**1994**). ''The emergence of native-language phonological influences in infants: A perceptual assimilation model,'' in *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, edited by J. Goodman and H. C. Nusbaum (MIT Press, Cambridge, MA), pp. 167–224.

Best, C. T. (**1995**). ''A direct realist view of cross-language speech perception,'' in *Speech Perception and Linguistics Experience: Issues in Cross-language Research*, edited by W. Strange (York, Baltimore), pp. 171–204.

Bradlow, A. R., and Pisoni, D. B. (**1999**). ''Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors,'' J. Acoust. Soc. Am. **106**, 2074–2085.

Bradlow, A. R. (**in press**). ''Confluent talker- and listener-related forces in clear speech production,'' in *Papers in Laboratory Phonology VII*, edited by C. Gussenhoven and N. Warner (Mouton de Gruyter, Berlin).

Bradlow, A. R., Kraus, N., and Hayes, E. (submitted). ''Speaking clearly for learning-impaired children: Sentence in noise perception,'' J. Speech Lang. Hear. Res. .

Cutler, A., and Butterfield, S. (**1990**). ''Durational cues to word boundaries in clear speech,'' Speech Commun. **9**, 485–495.

Ferguson, S. H., and Kewley-Port, D. (**in press**). ''Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am.

Flege, J. E. (**1992**). ''Speech learning in a second language,'' in *Phonological development: Models, Research, and Applications*, edited by C. Ferguson, L. Menn, and C. Stoel-Gammon (York, Timonium, MD), pp. 565–604.

Flege, J. E. (**1995**). ''Second language speech learning: Theory, findings, and problems,'' in *Speech Perception and Linguistics Experience: Issues in Cross-language Research*, edited by W. Strange (York, Baltimore), pp. 233–277.

Flege, J. E., MacKay, I. R. A., and Meador, D. (**1999**). ''Native Italian speakers' perception and production of English vowels,'' J. Acoust. Soc. Am. **106**, 2973–2987.

Hazan, V., and Simpson, A. (**2000**). ''The effect of cue-enhancement on consonant intelligibility in noise: Speaker and listener effects,'' Lang Speech **43**, 273–294.

Helfer, K. S. (**1997**). ''Auditory and auditory-visual perception of clear and conversational speech,'' J. Speech Lang. Hear. Res. **40**, 432–443.

Johnson, K., Flemming, E., and Wright, R. (**1993**). ''The hyperspace effect: Phonetic targets are hyperarticulated,'' Language **69**, 505–528.

Kuhl, P. K., and Iverson P. (**1995**). ''Linguistics experience and the perceptual magnet effect,'' in *Speech Perception and Linguistics Experience: Issues in Cross-language Research*, edited by W. Strange (York, Baltimore), pp. 121–154.

Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., and Greene, B. G. (**1993**). ''Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision and semantic categorization,'' J. Exp. Psychol. Gen. **122**, 316–330.

Matthies, M., Perrier, P., Perkell, J. S., and Zandipour, M. (**2001**). ''Variation in anticipatory coarticulation with changes in clarity and rate,'' J. Speech Lang. Hear. Res. **44**, 340–353.

Mayo, L. H., Florentine, M., and Buus, S. (**1997**). ''Age of second-language acquisition and perception of speech in noise,'' J. Speech Hear. Res. **40**, 686–693.

Meador, D., Flege, J. E., and MacKay, I. R. (**2000**). ''Factors affecting the recognition of words in a second language,'' Bilingualism **3**, 55–67.

Moon, S.-J., and Lindblom, B. (**1994**). ''Interaction between duration, context, and speaking style in English stressed vowels,'' J. Acoust. Soc. Am. **96**, 40–55.

Nabelek, A. K., and Donahue, A. M. (**1984**). ''Perception of consonants in reverberation by native and non-native listeners,'' J. Acoust. Soc. Am. **75**, 632–634.

Ohala, J. J. (**1995**). ''Clear speech does not exaggerate phonemic contrast,'' in Proceedings of the 4th European Conference on Speech Communication and Technology, pp. 1323–1325.

Payton, K. L., Uchanski, R. M., and Braida, L. D. (**1994**). ''Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing,'' J. Acoust. Soc. Am. **95**, 1581–1592.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1985**). ''Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech,'' J. Speech Hear. Res. **28**, 96–103.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1986**). ''Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech,'' J. Speech Hear. Res. **29**, 434–446.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1989**). ''Speaking clearly for the hard of hearing. III. An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech,'' J. Speech Hear. Res. **32**, 600–603.

Schum, D. J. (**1996**). ''Intelligibility of clear and conversational speech of young and elderly talkers,'' J. Am. Acad. Audiol **7**, 212–218.

Strange, W., Editor (**1995**). *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (York, Baltimore).

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., and Nishi, K. (**2001**). ''Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners,'' J. Acoust. Soc. Am. **109**, 1691–1704.

Studebaker, G. A. (**1985**). ''A rationalized arcsine transform,'' J. Speech Hear. Res. **28**, 455–462.

Uchanski, R. M., Choi, S. S., Braida, L. D., Reed, C. M., and Durlach, N. I. (**1996**). ''Speaking clearly for the hard of hearing IV. Further studies of the role of speaking rate,'' J. Speech Hear. Res. **39**, 494–509.

Uchanski, R. M. (**1988**) ''Spectral and temporal contributions to Speech Clarity for hearing impared listeners,'' Doctoral dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (unublished).

Uchanski, R. M. (**1992**) ''Segment durations in conversational and clear speech'' (unpublished).