The interlanguage speech intelligibility benefit[a)]

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

(Received 20 December 2002; revised 6 June 2003; accepted 16 June 2003)

This study investigated how native language background influences the intelligibility of speech by non-native talkers for non-native listeners from either the same or a different native language background as the talker. Native talkers of Chinese ($n=2$), Korean ($n=2$), and English ($n=1$) were recorded reading simple English sentences. Native listeners of English ($n=21$), Chinese ($n=21$), Korean ($n=10$), and a mixed group from various native language backgrounds ($n=12$) then performed a sentence recognition task with the recordings from the five talkers. Results showed that for native English listeners, the native English talker was most intelligible. However, for non-native listeners, speech from a relatively high proficiency non-native talker from the same native language background was as intelligible as speech from a native talker, giving rise to the ''matched interlanguage speech intelligibility benefit.'' Furthermore, this interlanguage intelligibility benefit extended to the situation where the non-native talker and listeners came from different language backgrounds, giving rise to the ''mismatched interlanguage speech intelligibility benefit.'' These findings shed light on the nature of the talker–listener interaction during speech communication.
 [DOI: 10.1121/1.1603234]

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

I. INTRODUCTION

It has long been recognized that adult second-language learners from a given native language background progress in a relatively consistent way from a monolingual to a bilingual system. Conversely, adult second-language learners from different native language backgrounds exhibit different deviations from the target language norms. For example, while native English talkers consistently produce and perceive the French high front rounded vowel [y] as [u] or [u]-like, native Portuguese talkers consistently produce and perceive French [y] as [i] or [i]-like (Rochet, 1995). This observation has formed the basis for a wealth of research and theorizing in phonetics and phonology on the perception and production of individual speech sound categories in second languages (e.g., see contributions to Strange, 1995). In this paper, we explore the phenomenon of native and target language interaction during the acquisition of second language sound structure at the level of overall sentence intelligibility. We investigate this issue by testing sentence intelligibility amongst adult non-native talkers and listeners who share a native language and amongst adult non-native talkers and listeners who do not share a native language. Since individuals from the same native language background who are in the process of acquiring a given target language all share an ''interlanguage,''[1] we predicted that target language intelligibility between non-native talkers and listeners from the same native language background would be enhanced relative to intelligibility between a native talker and a non-native listener. Indeed, second-language learners often report that the speech of a fellow non-native talker is easier to understand than the speech of a native talker. The present study aimed to investigate this claim under controlled laboratory conditions. Furthermore, we included talkers and listeners from several native language backgrounds in order to examine whether any ''interlanguage speech intelligibility benefit'' that we may observe (i.e., the benefit afforded by a shared interlanguage between a non-native talker and listener) is entirely dependent on the talker and listener sharing the native language or if non-native speech is generally more intelligible than native speech for all non-native listeners regardless of native language background.

Findings in support of these predictions would demonstrate that any measure of overall speech intelligibility must take into account both talker- and listener-related factors. Speech that is of relatively low intelligibility for one group of listeners (e.g., non-native accented English for native listeners of English) may be highly intelligible for another group of listeners (e.g., non-native listeners of English). Therefore, rather than considering variability in intelligibility as a function of talker characteristics, it is more accurate to measure intelligibility in terms of the talker–listener relationship. In this study, we approach this issue with respect to intelligibility between talkers and listeners from either the same or different native language backgrounds.

Previous research has demonstrated that native listeners generally find native talkers more intelligible than non-native talkers, particularly in noisy conditions (Munro, 1998; Munro and Derwing, 1995). However, the findings from studies investigating whether native or non-native talkers are more intelligible for non-native listeners have been less clear. For example, based only on casual observation rather than on controlled laboratory study, Nash (1969) claims that non-native speakers are more intelligible to other non-natives than to native speakers: ''A [non-native] speaker who cannot make himself understood when speaking English to a native English speaker will have no difficulty conversing in English

[a)]An earlier version of this study was presented at the 143rd meeting of the Acoustical Society of America, Pittsburgh, PA, June 2002.
[b)]Electronic mail: t-bent@northwestern.edu

with another [non-native] speaker'' (p. 4). Similarly, Weinreich (1953) states, ''When the other interlocutor is also bilingual, the requirements of intelligibility...are drastically reduced. Under such circumstances, there is hardly any limit to interference; forms can be transferred freely from one language to the other and often used in unadapted shape'' (p. 81). Smith and Rafiqzad (1979) provide some experimental evidence in support of this claim in a study in which foreign-accented talkers of English from multiple native language backgrounds and one native English talker recorded English passages of their own choosing. Then, listeners from various language backgrounds, including native languages that matched and mismatched the native languages of the talkers, completed a cloze procedure test in which the listeners filled in words that had been removed from transcripts of the recorded passages. Results showed that the non-native listeners found the native talker equally or less intelligible than the non-native talkers. Nevertheless, a talker–listener match with respect to native language did not necessarily produce the highest intelligibility scores. Unfortunately, the passages in this study differed significantly in difficulty and a strong positive correlation between passage difficulty and intelligibility was reported. This correlation suggests that the observed intelligibility differences may have been due to the difficulty of the passage rather than to the language backgrounds of the talkers and listeners.

More recently, van Wijngaarden (2001) and van Wijngaarden *et al.* (2002) provided strong evidence in support of the claim that under certain conditions non-native listeners find sentences produced by non-native talkers at least as intelligible as sentences produced by native talkers. Specifically, van Wijngaarden (2001) found that for native listeners, all four of the non-native talkers included in the study were less intelligible than each of the four native talkers included in the study. However, the two higher intelligibility non-native talkers (based on their ranking for the native listeners) were more intelligible for non-native listeners than each of the four native talkers. van Wijngaarden *et al.* (2002) measured intelligibility of Dutch, English, and German sentences produced by native and non-native talkers of these three languages for trilingual listeners (native language = Dutch, second language = English, third language = German). The listeners were all more proficient in English than in German as evidenced by the fact that they learned English earlier, used English more frequently, and gave higher self-reported scores on overall proficiency in English than in German. When listening to English (their higher proficiency foreign language) the listeners found the native English talkers more intelligible than the non-native talkers. However, when listening to German (their lower proficiency foreign language) they found the non-native talkers more intelligible than the native talkers. The findings of these studies demonstrate an intelligibility advantage between non-native talker–listener pairs when the listener is at a relatively early stage of target language acquisition and the talker is at a relatively advanced stage of target language acquisition. Similarly, Imai *et al.* (2003) found an advantage of shared language background for word recognition. They investigated the ability of native and non-native listeners with a first language of Spanish to recognize English words spoken by one native talker and one non-native talker also with Spanish as a first language. Native listeners performed better than the non-native listeners (i.e., recognized more words) for the words produced by the native talker, whereas the non-native listeners performed better than the native listeners for words produced by the non-native talker. An important finding from this study was that the non-native listeners outperformed the native listeners only for the words produced by the non-native talker that came from dense lexical neighborhoods (i.e., words that have many similar sounding ''neighbors'' with which they can easily be confused). This finding is consistent with the finding of Bradlow and Pisoni (1999) that word recognition by non-native listeners is worse for ''hard'' words (low-frequency words in high-density neighborhoods) than for ''easy'' words (high-frequency words in sparse neighborhoods). Since lexical neighborhood characteristics are defined in terms of segments shared across words, these findings suggest a connection between lower-level processing at the segmental level and higher-level processing at the word level.

Other work on non-native speech perception has shown that the success of acquiring the second language system depends on many factors, including (but not limited to) age of acquisition, duration of exposure, amount of continued first language usage, and quantity and quality of second language input (e.g., Flege, 2002). Although studies have shown that adults are capable of learning novel phonetic contrasts (Logan *et al.*, 1991; Lively *et al.*, 1993, 1994; Bradlow *et al.*, 1997), the impact of the native language has been shown to influence speech perception even for very early bilinguals. For example, Sebastian-Galles and Soto-Faraco (1999) showed that bilinguals who learned their second language as early as 4 years of age were less efficient at processing phonemic contrasts on-line than native listeners. The perception and production of non-native contrasts in late bilinguals and naive listeners is even more strongly influenced by the first language. However, an important generalization to emerge from studies of non-native phoneme perception is that not all non-native contrasts are equally difficult for second language learners. The relative difficulty with which non-native listeners perceive non-native sounds is related to the relationship between the status of sounds in the overall systems of phonemic contrasts in the first and second languages (e.g., Flege *et al.*, 1999; Best *et al.*, 2001). These difficulties in phoneme perception have been shown to affect lexical representation for early bilinguals (Pallier *et al.*, 2001) and to impair word recognition for early and late bilinguals (Meador *et al.*, 2000). Experiments with non-native listeners and native talkers have shown that non-native listeners can perform with native-like proficiency in the quiet, but even early bilinguals are impaired in noise compared to native listeners (e.g., Mayo *et al.*, 1997; McAllister, 1990). The cause for this decline is not fully understood, but a contributing factor may be the fact that non-native listeners are less able to take advantage of contextual cues for word identification (Mayo *et al.*, 1997), and so when in the presence of noise they have less signal-independent information to rely on as an aid to sentence recognition.

Two important issues regarding speech intelligibility between native and non-native talkers and listeners remain unresolved in controlled laboratory studies. First, all of the studies where both the talker and listener were non-native have involved a shared native language between talker and listener. The only exception is Smith and Rafiqzad (1979), where the uncontrolled variable of passage difficulty confounded the results. Therefore, we still do not know how the intelligibility of non-native talkers for non-native listeners varies depending on whether there is a match or mismatch in native language background between talker and listener. Data addressing this issue will allow us to assess the extent to which the possible non-native talker–listener intelligibility benefit is due to a particular shared interlanguage or to certain more general features of non-native language production and perception that are independent of the particular native and target languages in question. Second, the findings of van Wijngaarden (2001) and van Wijngaarden *et al.* (2002) suggested that talker and listener proficiency in the target language is important for determining the relative intelligibility of native and non-native talkers, but this factor deserves more attention. By including non-native talkers who are known to vary in target language production proficiency, the present study aimed to gain further insight into the factors that contribute to an interlanguage speech intelligibility benefit.

Using a database of sentence recordings by non-native talkers from various native language backgrounds, the present study investigated how the talker–listener match or mismatch with respect to language background and variation in second language proficiency affected non-native talker intelligibility. We made two specific predictions. First, we predicted, based on the findings of van Wijngaarden (2001) and van Wijngaarden *et al.* (2002), that a non-native talker with relatively high proficiency in English speech production would be at least as intelligible as a native talker for non-native listeners from the same native language background. This prediction supposes that non-native language learners from a given native language have a broad base of shared linguistic and phonetic knowledge that facilitates speech communication in the non-native language. We expected this interlanguage intelligibility benefit to be attenuated in the case of less proficient non-native talkers due to the fact that their speech production may stray so far from the target language norm that important lexical contrasts may not be adequately conveyed, resulting in word and sentence intelligibility that is low enough to produce very low overall intelligibility for both native and non-native listeners. Second, we predicted that a relatively high proficiency non-native talker will also be at least as intelligible as a native talker for non-native listeners from a different native language background. This prediction supposes that, regardless of native language background, certain features of non-native speech will make non-native talkers more intelligible to all non-native listeners. For example, non-native talkers may be less likely than native talkers to apply certain reduction phenomena that characterize native accented running speech in English such as alveolar flapping and failure to release final stops. While these features of non-native English speech may contribute to the overall impression of a foreign accent for native listeners, they may in fact be beneficial for speech perception by all non-native listeners of English regardless of native language background.

II. THE NORTHWESTERN UNIVERSITY FOREIGN ACCENTED ENGLISH SPEECH DATABASE

The Northwestern University Foreign Accented English Speech Database (NUFAESD) contains recordings of 64 sentences produced by 32 non-native talkers for a total of 2048 recorded sentences. Along with these recordings, the database includes demographic information about each talker and an overall intelligibility score for each talker as measured by a perception test with native English listeners.

The sentences in this database were taken from the sentence lists included in the revised Bamford–Kowal–Bench Standard Sentence Test (BKB-R) developed by the Cochlear Corporation for use with American children. The original Bamford–Kowal–Bench Standard Sentence Test was developed for use with British children (Bamford and Wilson, 1979; Bench and Bamford, 1979). These sentences were chosen for this database because they include words that are highly familiar to non-natives and are syntactically simple. Each list consists of 16 simple, declarative sentences with 3 or 4 keywords for a total of 50 keywords per list. From the 21 lists included in the BKB-R test, four lists (lists 7, 8, 9, and 10) were selected for the NUFAESD. These lists were selected based on their equivalent intelligibility scores for normal-hearing children as reported in Bamford and Wilson (1979). Ratings of age of acquisition, written and verbal frequency, imagability, concreteness, and familiarity for the 155 keywords in these four lists were gathered from the MRC Psycholinguistic Database: Machine Usable Dictionary Version 2.00 and are shown in Table I. Because some of the keywords are repeated within and across lists, there are not 200 unique keywords.

The talkers recorded for this database were recruited from the Northwestern University International Summer Institute and English as a Second Language (ESL) program over the course of the summer of 2000 and the 2000–2001 academic year. The International Summer Institute provides 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. However, these students all had limited experience with spoken English communication. 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. The talkers came from a variety of language backgrounds: Chinese[2] ($n=20$), Korean ($n=5$), Bengali ($n=1$), Hindi ($n=1$), Japanese ($n=1$), Romanian ($n=1$), Slovakian ($n=1$), Spanish ($n=1$), and Thai ($n=1$). The talkers ranged in age from 22 to 32 years with an average age of 25.5 years. They had spent on average 2.7 months in the United States

TABLE I. Frequency and psychological ratings for keywords.

	Verbal frequency[a]	Age of acquisition rating[b]	Written frequency[c]	Familiarity rating[d]	Imagability rating[d]	Concreteness rating[d]
Average	120	231	588	587	524	503
Standard deviation	395	58	2266	35	104	122
Number of items	134	37	134	106	104	97

[a]Verbal frequency: frequency of occurrence in verbal language from Brown (1984).
[b]Age of acquisition: age multiplied by 100 to produce a range from 100 to 700.
[c]Written frequency: the norms of Kucera–Francis (1967).
[d]Concreteness, familiarity, and imageability values range 100 to 700.

with a range from 0.25 months to 24 months at the time of recording. The average age at which the participants began to study English in school was 12.0 years and the average length of English learning was 9.8 years. All talkers had no known speech or hearing impairments. All talkers were paid for their participation.

The talkers were recorded in a sound-attenuated booth. The recording were made on an Ariel Proport A/D sound board 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 with a 16-kHz sampling rate and transferred to a PC-based computer. The digital speech files were then segmented into sentence length files and the root-mean-square amplitude was equated across all files.

For assessing the intelligibility of the talkers, the digitized recordings were embedded in white noise, yielding a speech-plus-noise file with a +5-dB signal-to-noise ratio. Each of the stimulus files consisted of a 400-millisecond silent leader, followed by 500 ms of noise, followed by the speech-plus-noise file, and ending with 500 ms of noise only.

In order to keep the number of native English listeners manageable while still evaluating the intelligibility of all the non-native talkers, eight test conditions with five listeners each were constructed. Each native English listener evaluated 64 sentences from the four BKB-R lists and all 32 non-native talkers described above. Each of the 32 non-native talkers supplied two different sentence recordings to each of eight test conditions (for a total of 64 sentences per condition). Therefore, each talker's intelligibility was evaluated on the basis of 16 sentences (8 conditions$\times$2 sentences per condition) from one full BKB-R list in a multiple talker presentation format (i.e., in each condition, listeners evaluated sentences by all 32 talkers). The native English listeners were presented with each sentence once and their task was to transcribe the sentence in standard English orthography. Intelligibility scores for each talker were based on the number of keywords correctly transcribed across all of the 16 sentences submitted to perceptual evaluation by the native English listeners. The raw percent-correct scores were converted to rationalized arcsine transform units (rau). This transformation places the scores on a linear and additive scale, thus facilitating meaningful statistical comparisons across the entire range of the scale (Studebaker, 1985). The rau scale ranges from -23 to 123. The middle of the scale corresponds closely to percent correct, but at the edges of the scale, below 15% and above 85%, the rau scale diverges from percent correct. For example, 80% corresponds to 79.9 rau but 95% corresponds to 102 rau. The non-native subjects' sentence production scores (i.e., the average percentage of keywords correctly transcribed by the native English listeners) ranged from 43 to 93 rau.

III. TEST OF FOREIGN-ACCENTED ENGLISH PERCEPTION

Using materials from the NUFAESD, we designed an experiment to investigate the perception of foreign-accented English by listeners from various language backgrounds. The overall design involved presenting English sentences produced by five talkers to four listener groups. The talkers were native talkers of Chinese ($n=2$), Korean ($n=2$), and English ($n=1$). The listeners were native talkers of Chinese ($n=21$), Korean ($n=10$), English ($n=21$), and a mixed group of non-native talkers of English from various native language backgrounds ($n=12$). None of the non-native listeners had served as talkers in the collection of the NUFAESD materials and none of the native English listeners had evaluated the sentences in the NUFAESD. Of particular interest for this study was the effect of a listener–talker match or mismatch in native language background, and the interaction of this listener–talker match or mismatch with talker proficiency in the target language (i.e., English).

A. Method

1. Talkers

The talkers selected for this experiment were from three language backgrounds: monolingual English ($n=1$), non-native (NN) talkers of English with a first language of Chinese ($n=2$), and NN talkers of English with a first language of Korean ($n=2$). The four non-native talkers' productions were taken from the NUFAESD. These talkers were selected based on their first language (either Korean or Chinese), gender (female), and production intelligibility scores obtained from the perceptual evaluation test conducted at the time of the database compilation. For the Chinese and Korean talkers, one from each language background was of higher proficiency and one of lower proficiency as defined by their production intelligibility scores. The Chinese high- and low-proficiency talkers had intelligibility scores (average percent of keywords correctly identified converted to the rau scale) of 80 and 43 rau, respectively. The Korean high- and low-

TABLE II. General information about non-native listener groups. Mean and standard deviations (in parentheses) are shown. The Northwestern University Graduate School minimum for TOEFL is 560 (paper) or 220 (computer).

	NN-Chinese ($n=21$)	NN-Korean ($n=10$)	NN-Mixed ($N=12$)	Significant differences
Age (years)	23.60 (1.60)	28.00 (4.72)	25.18 (2.35)	Korean>Chinese[a] Mixed>Chinese[b]
TOEFL[c]—computer	n/a	($n=1$) 253	($n=8$) 269.63 (22.06)	n/a
TOEFL—paper	($n=17$) 641.82 (19.63)	($n=7$) 602.43 (18.80)	($n=4$) 629.25 (37.25)	Chinese>Korean[a]
Time in US (months)	1.12 (1.92)	8.25 (9.48)	5.23 (10.34)	Korean>Chinese[a]
AOA[d] (years)	10.95 (2.22)	12.25 (1.39)	10.00 (4.22)	none
LOA[e] (years)	10.11 (2.26)	7.00 (1.41)	8.17 (4.37)	Chinese>Korean[a]

[a]$=p<0.01$.
[b]$=p<0.05$.
[c]TOEFL=Test of English as a Foreign Language.
[d]AOA=age of acquisition.
[e]LOA=length of acquisition (i.e., English study).

proficiency talkers had scores of 90 and 55 rau, respectively. An effort was made to match the intelligibility scores for the high-proficiency Chinese and Korean and the low-proficiency Chinese and Korean. However, no talkers in our sample with the selected demographic variables had identical production intelligibility scores. The talkers were not matched for any other features of their speech and we did not control for speaking rate in either the recording of the talkers or the selection of the talkers. The same procedure and equipment was used to record the monolingual English talker as was used in the compilation of the NUFAESD. No analogous intelligibility score is available for the native English talker since she was not included in any previous intelligibility testing.

2. Listeners

A total of 65 adults with normal speech and hearing participated in the experiment. The listeners came from four different language backgrounds: monolingual English ($n=21$), non-native speakers of English with a first language of Chinese (NN-Chinese, $n=21$), non-native speakers of English with a first language of Korean (NN-Korean, $n=10$), and non-native speakers of English with native language backgrounds other than Chinese or Korean (NN-mixed, $n=12$). The NN-mixed group included individuals from the following native language backgrounds: Bulgarian ($n=1$), Dutch ($n=1$), French/Douala ($n=1$), German ($n=1$), Greek ($n=2$), Hindi ($n=1$), Japanese ($n=2$), Serbian ($n=1$), Spanish ($n=1$), and Tamil ($n=1$). The non-native listeners were recruited from the Northwestern University International Summer Institute and ESL program over the course of the summer of 2001 and the 2001–2002 academic year. All listeners were paid for their participation in the study. Additional data for the three groups of non-native listeners are shown above in Table II.

The monolingual English listeners were all undergraduates at Northwestern University and ranged in age from 17 to 22 years with an average age of 19.1 years. The native English listeners were significantly younger than the non-native listeners [$t(59)=5.78$, $p<0.0001$]. They were recruited from the Linguistics Department subject pool and received course credit for their participation in the study. None of the listeners reported any speech or hearing problems.

3. Stimuli and procedure

Sentences from the four BKB-R lists (a total of 60 sentences) were divided into five lists of 12 sentences covering 37 keywords each (11 sentences with 3 keywords and 1 sentence with 4 keywords). See the Appendix for the sentence lists with the keywords underlined.

Subjects were seated in front of a computer monitor in a sound-attenuated booth. 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 HD 580). 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 heard all five talkers and all five sentence lists in a blocked format. The monolingual English talker was always in the third position. The high-proficiency Chinese and Korean talkers were either first or second and the low-proficiency Chinese and Korean talkers were either fourth or fifth. The NN-Chinese and monolingual English listeners always heard the talkers in the following order: Chinese high proficiency, Korean high proficiency, native English, Chinese low proficiency, Korean low proficiency. For the NN-Korean listeners the order of the Chinese and Korean talkers was reversed. The rationale behind this ordering of the talkers was to ensure consistency across the NN-Chinese and NN-Korean groups with respect to the ordering of the listener–talker native language match and mismatch. Furthermore, the high-proficiency non-native talkers were ordered before the native talker so that superior performance

on the non-native talkers could not be attributed to a practice effect. Half of the NN-mixed listener group heard the lists in the order for the NN-Chinese group while the other half heard the lists in the order for the NN-Korean group. The particular sentence list read by each talker was counterbalanced across listeners.

In a separate session after the perception test, a word familiarity test was administered. For this test, each of the 144 unique keywords in the complete set of sentences used in this study 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." A set of 75 filler items was also presented as part of this test. These filler items 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 scale would be represented by the items in this test. An additional 128 words were also included in this test for the purpose of a separate study not reported here. 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 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.

4. Data analysis

Sentence-in-noise perception scores were determined by a strict keyword-correct count. For each set of sentences heard by each listener, the talker could receive a score from 0 to 37 keywords correct. This score was obtained by counting the number of keywords transcribed perfectly. Words with added or deleted morphemes were counted as incorrect. However, obvious spelling errors were not counted as incorrect. Raw intelligibility scores were converted to percent correct and then to rationalized arcsine units (rau).

B. Results

1. Word familiarity

Data from the word-familiarity rating task showed that the vast majority of words was highly familiar to the vast majority of the non-native listeners. Familiarity data from two of the 44 listeners were missing: one did not return for the second data collection session when the word-familiarity test was administered and the other's data had to be discarded due to a computer error. Of the remaining 42 listeners, all gave high ratings (5 or greater) to at least 94% of the words. Thirty-two listeners gave scores of less than 5 to no more than two words; six listeners gave scores of less than 5 to three to five words; four listeners gave scores of less than 5 to six to nine words, and no listeners gave scores of less than 5 to more than nine words. Furthermore, there was no correlation between the listeners' average familiarity rating score and their average score on the sentence-in-noise perception test (rho = 0.120, $p = 0.47$). A vast majority of the words, 79%, was given a rating of 7 by all subjects. Additionally, only five words had average scores lower than 6: broom (5.8), buckets (5.8), janitor (5.6), jug (5.1), and saucepan (5.9). Of the 144 target items, only 19 words were given any scores lower than 5. Nine words were given a score under 5 by one listener; three words were given scores under 5 by two listeners; and seven words were given scores lower than 5 by more than two listeners (from 5 to 18 listeners). Last, the non-natives' scores on the low-, mid-, and high-familiarity filler items were similar to scores given by native listeners in Bradlow and Pisoni (1999). The low-, mid-, and high-familiarity filler items were given scores of 1.83, 3.88, and 6.93, respectively, by the native listeners in Bradlow and Pisoni's study and were given scores of 2.85, 3.79, and 6.53, respectively, by the non-native listeners in the current study. Therefore, we performed all analyses of the 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 independently of word familiarity.

2. Foreign-accented English perception

Table III summarizes the five talkers' intelligibility scores for each of the four listener groups.

A repeated measures ANOVA with listener (native English, NN-Chinese, NN-Korean, NN-mixed) as the between-subjects factor and talker (high-proficiency Chinese, low-proficiency Chinese, high-proficiency Korean, low-proficiency Korean, native English) as the within-subjects factor showed highly significant main effects of listener [$F(3,240) = 39.34$, $p < 0.0001$] and talker [$F(4,240) = 194.15$, $p < 0.0001$]. The interaction of talker and listener was highly significant [$F(12,240) = 11.37$, $p < 0.0001$]. For each listener group the low-proficiency Chinese and Korean talkers were less intelligible than the native talker and the high-proficiency Chinese and Korean talkers. However, the rankings for the two high-proficiency non-natives and the native talker depended on the language background of the listener group.

Post hoc pairwise comparisons (Bonferroni/Dunn tests) of talker intelligibility within each listener group were conducted. Due to the large number of paired comparisons (ten for each listener group), the p value must be less than 0.005 to be significant. Pairwise comparisons within a talker across listener groups were not conducted as those comparisons were not of primary interest for this study. For the NN-Chinese listeners, the high-proficiency Chinese talker was not significantly different from the high-proficiency Korean talker or the native English talker. However, the high-proficiency Korean talker was significantly more intelligible than the native English talker ($p = 0.001$). All three of these talkers were significantly more intelligible for these listeners than the low-proficiency Chinese and Korean talkers ($p < 0.001$). For the NN-Korean listeners, there were no significant differences between the high-proficiency Chinese, high-proficiency Korean, and native English talkers. Addi-

TABLE III. Percent keywords correctly transcribed for talkers and listener groups in rau. Standard deviations are shown in parentheses. Scores in bold are significantly higher than the other scores in the row and are not significantly different from one another except for the underlined scores. Specifically, for the NN-Chinese listeners, the high-proficiency Korean talker is significantly more intelligible than the native English, and for the NN-Korean listeners the high-proficiency Korean talker is significantly more intelligible than the low proficiency Korean talker.

	Talker					
Listener group	Chinese high proficiency	Korean high proficiency	Native English	Chinese low proficiency	Korean low proficiency	All talkers
NN-Chinese ($n=21$)	**64** (**10.8**)	**66** (**11.7**)	**56** (**10.4**)	30 (12.1)	41 (9.8)	51 (17.4)
NN-Korean ($n=10$)	**60** (**15.5**)	**74** (**15.8**)	**60** (**11.7**)	22 (11.4)	**53** (**12.0**)	54 (21.7)
NN-Mixed ($n=12$)	**62** (**11.0**)	**70** (**7.7**)	**67** (**15.8**)	19 (20.9)	41 (14.8)	52 (24.1)
Native English ($n=21$)	77 (12.2)	91 (8.4)	**109** (**14.7**)	38 (13.8)	60 (12.1)	75 (27.7)
All listeners	67 (13.9)	76 (15.2)	76 (26.8)	29 (15.9)	49 (14.5)	

tionally, the high-proficiency Chinese and native English talkers were not significantly different from the low-proficiency Korean talker. However, the high-proficiency Korean talker was significantly more intelligible than the low-proficiency Korean talker ($p<0.001$). For the NN-mixed listener group, the high-proficiency Chinese, high-proficiency Korean, and the native English talkers were not significantly different from one another, and were all significantly more intelligible than the low-proficiency Chinese and low-proficiency Korean talkers ($p<0.0001$). For the native English listeners, the native English talker was significantly more intelligible than all the other talkers ($p<0.001$).

In summary, the native listeners found the native talker more intelligible than any of the other talkers, and all listeners generally found the lower-proficiency non-native talkers less intelligible than either their high-proficiency counterparts or the native talker. Of particular interest for this study was the finding that the non-native listeners found the high-proficiency non-native talkers with whom they shared native language as intelligible as the native English talker. This finding also extended, in one case, to a low-proficiency non-native talker in that the low-proficiency Korean talker was as intelligible as the native English talker for the NN-Korean listener group. This finding demonstrates a *matched interlanguage speech intelligibility benefit*, such that a native language match between a non-native talker and a non-native listener facilitates speech intelligibility. Furthermore, when the non-native listeners and high-proficiency non-native talkers did not share a native language, the non-native listeners found the non-native talkers equally as or more intelligible than the native talker. This general finding suggests that the interlanguage benefit can extend to the situation of a talker–listener native language mismatch, demonstrating a *mismatched interlanguage speech intelligibility benefit*.

To investigate the possible contribution of speaking rate to the intelligibility results, we measured sentence duration for each of the five speakers. All pairwise comparisons amongst the five talkers were significant ($p<0.005$) except for the high-proficiency Korean and the low-proficiency Chinese talker, whose average sentence durations did not differ significantly. The average sentence durations for the five talkers are shown in Table IV.

For each of the non-native listener groups, the average sentence durations for the five talkers did not significantly correlate with their intelligibility scores (Spearman rank correlation, NN-Chinese listeners: rho$=0.300$, $p=0.55$; NN-Korean listeners: rho$=0.103$, $p=0.84$; NN-Mixed listeners: rho$=-0.100$, $p=0.84$; Native English listeners: rho $=-0.300$, $p=0.55$). For example, the low-proficiency Korean talker had longer sentence durations than the high-proficiency Korean, yet almost all the non-native listeners (except for three NN-Korean listeners) found the high-proficiency Korean more intelligible than the low-proficiency Korean. This analysis suggests that the observed interlanguage intelligibility benefit is not simply due to variability in speaking rate across the talkers.

IV. SUMMARY AND DISCUSSION

Four groups of listeners (monolingual English, NN-Chinese, NN-Korean, and NN-mixed) transcribed sentences produced by a native English talker, two Chinese non-native talkers of English, and two Korean non-native talkers of English. The major finding of this study was that the relative intelligibility of each talker depended on the language background of the listener such that

TABLE IV. Average sentence durations for the five talkers in order of increasing duration. Means in milliseconds and standard deviations (in parentheses) are shown.

	Average sentence duration
Native English	1223 ms (161)
Low-proficiency Chinese	1512 ms (244)
High-proficiency Korean	1587 ms (210)
Low-proficiency Korean	1680 ms (280)
High-proficiency Chinese	1717 ms (196)

(1) For native listeners, intelligibility of the native talker was greater than the intelligibility of any of the non-native talkers.
(2) For non-native listeners, intelligibility of a high-proficiency non-native talker (and in one case a low-proficiency talker) from the same native language background was equal to the intelligibility of the native talker. This is the ''matched interlanguage speech intelligibility benefit.''
(3) For non-native listeners, intelligibility of a high-proficiency non-native talker from a different native language background was greater than or equal to the intelligibility of the native talker. This is the ''mismatched interlanguage speech intelligibility benefit.''

The matched interlanguage speech intelligibility benefit can be explained by the fact that non-native speech production and perception are both systematically linked to native language sound structure (for a wealth of supporting research and theory see Strange, 1995; Best, 1994, 1995; Flege, 1992, 1995; Kuhl and Iverson, 1995 and many others). Thus, the speech of a non-native talker is more intelligible to non-native listeners with whom they share a native language than for native listeners due to the fact that the overall shared phonetic and phonological knowledge between the non-native talker and non-native listeners from the same language background is likely to be more extensive than a native/non-native pair. For the non-natives who share a native language, their linguistic knowledge covers aspects of both the native and target languages, whereas for the non-native/native pair the shared knowledge base includes only their knowledge of the target language insofar as it is developed in the non-native talker. This shared knowledge base includes the system of consonant and vowel categories, phonotactics, stress patterns, and intonation as well as other features of the sound system. Thus, a non-native listener is well equipped to interpret certain acoustic–phonetic features of the speech of a native-language-matched non-native talker as the talker intended them to be interpreted, even though they may deviate markedly from the target language norm, whereas native listeners are better equipped to interpret the speech of a native talker. For example, even foreign-accented talkers who have gained control over producing the full inventory of vowel contrasts of the target language may produce the vowels of the target language in the region of the vowel space (i.e., with a base of articulation) that is typical of the native language rather than of the target language. While this may result in misinterpretation of a particular vowel for native listeners or non-native listeners from a different native language background, non-native listeners from the same native language background as the talker will be more likely to access the correct vowel category, thereby contributing to the matched interlanguage speech intelligibility benefit that we observed in this study.

A possible explanation for the mismatched interlanguage speech intelligibility benefit is that it results from the talker's and listener's shared knowledge of the structure of the target language in conjunction with the influence of general strategies that listeners and talkers apply when learning to produce and perceive a foreign language. For example, in native-accented American English word-final stop consonants are frequently unreleased. Native listeners know to listen for cues to the presence and identity of word-final consonants in other parts of the signal and to interpret the lack of a salient word-final stop release as reflecting structural and/or contextual influences at the phrase and discourse levels, but non-native listeners may miss these cues, thereby compromising their overall comprehension of native-accented English. However, in their own English speech, non-native talkers who have not yet mastered all the details of American English allophony may produce particularly salient word-final stop consonant releases, thereby facilitating the intelligibility of their speech for other non-native listeners from a wide range of native language backgrounds (though not necessarily for native listeners). For example, Smith, Bradlow, and Bent (2003) demonstrated that non-native listeners from various native language backgrounds are better at identifying words in minimal pairs that differ only in the voicing of the final stop (e.g., cap vs cab, pick vs pig) produced by non-native talkers than produced by native talkers. In this case, even though the native listeners generally performed better than the non-native listeners, the performance of the non-native listeners surpassed that of the native listeners for words produced by one non-native talker, indicating that the non-native listeners must be listening for certain cues that native listeners are not attending to.

Alternatively, the mismatched interlanguage speech intelligibility benefit that we observed in this study may stem from similarities in the sound structure of the two languages that we investigated, in which case it is in fact just another manifestation of the matched interlanguage benefit rather than a separate phenomenon. For example, both Chinese and Korean have a much more constrained syllable structure than English, including a constraint against final consonant clusters. The transfer from a language which does not allow final consonant clusters such as Chinese or Korean to a language that does allow coda clusters such as English may result in similar features of Chinese-accented and Korean-accented English which serve to facilitate perception of English speech between native Chinese and native Korean listeners. However, evidence against this account based on structural similarities between Chinese and Korean comes from the results of the test with the NN-mixed listener group which included very small number of listeners (only 1 or 2) from each of a wide range of language backgrounds. Seven of these 12 listeners found the high-proficiency Chinese talker equally or more intelligible than the English talker, and five of the 12 listeners from this highly heterogeneous listener group found the high-proficiency Korean talker more intelligible than the native English talker. These listeners (i.e., those from the NN-mixed group who showed a mismatched interlanguage speech intelligibility benefit) came from a diverse group of native language backgrounds including Bulgarian, Dutch, French/Douala (bilingual), Greek, Hindi, Japanese, and Spanish. This finding suggests that the observed mismatched interlanguage speech intelligibility benefit is unlikely to be the result of structural similarities between the native languages of the talkers and listeners, but

rather due to certain tendencies in foreign-accented English regardless of native language background. Nevertheless, data from studies with additional carefully selected languages are needed in order to rule out this alternative account.

From the current study, we cannot determine at which stage of spoken language processing the interlanguage benefit arises since perceiving words in sentences requires processing on many different levels and our task was an off-line measure of comprehension that represents the accumulation of processing at multiple levels. Because we controlled for the linguistic content of the sentences, we can assume that the advantage that many of the non-native listeners received when listening to non-native speech came from the differences in the acoustic signal and not from differences in lexical choices or syntactic structures. Therefore, it is very likely that the source of the observed interlanguage effect was at a relatively early, phonetic stage of processing. Nevertheless, it is possible that the interlanguage benefit also operates at higher levels of sentence comprehension where other aspects of linguistic structure and of extralinguistic factors come into play. Additional tests that specifically tap into various levels of spoken language processing will be required to determine whether the early processes are most important for the interlanguage benefit or if the localization of the phenomenon occurs later in the processing of spoken language as well.

Large individual differences were found in the magnitude of the interlanguage benefit. The difference in intelligibility between the high-proficiency non-native talker and the native talker ranged from −23 to 52 rau for the matched interlanguage benefit and from −24 to 39 rau for the mismatched interlanguage benefit. The basis of these large individual differences remains unknown. The subjects in this study were rather homogeneous in terms of demographic variables (e.g., age of English study onset, length of residence, etc.) and in terms of English proficiency [e.g., TOEFL scores ranged only from 573–677 (paper) and 220–290 (computer)]. Future research will be needed to determine the source of these differences. For example, independent tests of language proficiency, particularly with respect to target language speech perception, will help to determine how listener proficiency in the non-native language influences the preference for non-native over native speech.

The findings from the present study are consistent with the findings of van Wijngaarden (2001) and van Wijngaarden *et al.* (2002), which demonstrated a non-native talker speech intelligibility advantage for non-native listeners at a relatively early stage of target language acquisition. Two important methodological differences between the present study and the studies by van Wijngaarden and colleagues are the target language (English vs Dutch) and the intelligibility measurement technique: van Wijngaarden and colleagues measured the signal-to-noise ratio required for a 50%-correct response (the speech reception threshold, or SRT), whereas the present study measured intelligibility in terms of a keyword recognition accuracy score for sentences presented at a fixed signal-to-noise ratio. Taken together, the present study and those of van Wijngaarden and colleagues provide converging evidence for an interlanguage speech intelligibility benefit and demonstrate that any measure of speech intelligibility must take into account both talker- and listener-related factors.

We conclude by noting an implication of the present findings for language variation and change. The demonstration of the interlanguage speech intelligibility benefit suggests a mechanism that may underlie the establishment of new pronunciation norms across a community of non-native talkers. Due to various social, political, and historical circumstances, non-native talkers who share a native language may sometimes communicate in a shared foreign language. This situation can arise in settings where the shared foreign language dominates in the broader context. For example, in many university research laboratories where the director and experimenters share a native language, the typical language of the laboratory may be English due to the fact that the broader scientific community is English dominated. This situation can occur on an even larger scale such as in a country like India, where English is widely spoken as a second language and, in certain settings, even talkers from the same native language background will communicate in English. Under such circumstances, a characteristic variety of the target language (e.g., Indian English) may develop as a result of the interlanguage speech intelligibility benefit which will cause certain interlanguage features to become firmly entrenched in the speech of the community. This intelligibility-based factor will likely operate in conjunction with other sociolinguistic factors to reinforce the establishment of a new and lasting variety of the target language.

ACKNOWLEDGMENTS

We are grateful to Martin Armstrong, Lyla Miller, Mengting Shieh, and Saundra Wright for data collection assistance. This research was supported by NIH-NIDCD Grant No. DC 03762.

APPENDIX: FOREIGN-ACCENTED ENGLISH PERCEPTION TEST MATERIALS

Based on BKB-R materials (Bamford and Wilson, 1979; Bench and Bamford, 1979).

List 1

(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) The ball is bouncing very high.
(11) Father forgot the bread.
(12) The girl has a picture book.

List 2

(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.

List 3

(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) The shoes were very dirty.
(9) They went on a vacation.
(10) The baby broke his cup.
(11) The lady packed her bag.
(12) The dinner plate is hot.

List 4

(1) A dish towel is by the sink.
(2) She looked in her mirror.
(3) The good boy is helping.
(4) They followed the path.
(5) The kitchen clock was wrong.
(6) Someone is crossing the road.
(7) The mailman brought a letter.
(8) They are riding their bicycles.
(9) He broke his leg.
(10) The milk was by the front door.
(11) The shirts are hanging in the closet.
(12) The chicken laid some eggs.

List 5

(1) The orange was very sweet.
(2) He is holding his nose.
(3) The new road is on the map.
(4) She writes to her brother.
(5) The football player lost a shoe.
(6) The three girls are listening.
(7) The coat is on a chair.
(8) The train is moving fast.
(9) The child drank some milk.
(10) The janitor used a broom.
(11) The ground was very hard.
(12) The buckets hold water.

[1]Selinker (1972) defined an interlanguage as ''a separate linguistic system based on the observable output which results from a learner's attempted production of a target language norm'' (p. 214).
[2]All Chinese subjects were native speakers of Mandarin Chinese in the sense that all schooling including University had been in Mandarin. However, many of the Chinese subjects spoke another dialect of Chinese at home.

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 Linguistic Experience: Issues in Cross-language Research*, edited by W. Strange (York, Baltimore), pp. 171–204.

Best, C., McRoberts, G., and Goodell, E. (**2001**). ''Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system,'' J. Acoust. Soc. Am. **109**, 775–794.

Bradlow, A. R., Pisoni, D. B., Yamada, R. A., and Tohkura, Y. (**1997**). ''Training Japanese listeners to identify English /r/ and /l/. IV. Some effects of perceptual learning on speech production,'' J. Acoust. Soc. Am. **101**, 2299–2310.

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.

Brown, G. D. A. (**1984**). ''A frequency count of 190,000 words in the London-Lund corpus of English conversation,'' Behav. Res. Methods Instrum. **16**(6), 502–532.

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. Steol-Gammon (York, Timonium, MD), pp. 565–604.

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

Flege, J. E. (**2002**). ''Interactions between the native and second-language phonetics systems,'' in *An Integrated View of Language Development: Papers in Honor of Henning Wode*, edited by P. Burmeister, T. Piske, and A. Rohde (Trier, Wissenschaftlicker Trier).

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.

Imai, S., Flege, J. E., and Walley, A. (**2003**). ''Spoken word recognition of accented and unaccented speech: Lexical factors affecting native and non-native listeners,'' in Proceedings of the International Congress on Phonetic Science, Barcelona, Spain.

Kucera, H., and Francis, W. N. (**1967**). *Computational Analysis of Present-Day American English* (Brown University Press, Providence).

Kuhl, P. K., and Iverson, P. (**1995**). ''Linguistic experience and the perceptual magnet effect,'' in *Speech Perception and Linguistic 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.

Lively, S. E., Logan, J. S., and Pisoni, D. B. (**1993**). ''Training Japanese listeners to identify English /r/ and /l/. II. The role of phonetic environment and talker variability in learning new perceptual categories,'' J. Acoust. Soc. Am. **94**, 1242–1255.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., and Yamada, T. (**1994**). ''Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories,'' J. Acoust. Soc. Am. **96**, 2076–2087.

Logan, J. S., Lively, S. E., and Pisoni, D. B. (**1991**). ''Training Japanese listeners to identify English /r/ and /l/: A first report,'' J. Acoust. Soc. Am. **89**, 874–886.

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

McAllister, R. (**1990**). ''Perceptual foreign accent: L2 user's comprehension ability,'' in Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-language Speech, pp. 216–228.

Meador, D., Flege, J. E., and MacKay, I. R. A. (**2000**). ''Factors affecting the

recognition of words in a second language,'' Bilingualism: Lang. Cogni. **3**, 55–67.
Munro, M. J., and Derwing, T. M. (**1995**). ''Foreign accent, comprehensibility, and intelligibility in the speech of second language learners,'' Lang. Learn. **45**, 73–97.
Munro, M. J. (**1998**). ''The effects of noise on the intelligibility of foreign-accented speech,'' Stud. Second Lang. Acquisit. **20**, 139–153.
Nash, R. (**1969**). ''Intonational interference in the speech of Puerto Rican bilinguals,'' J. English **4**, 1–42.
Pallier, C., Colome, A., and Sebastian-Galles, N. (**2001**). ''The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries,'' Psychol. Sci. **12**, 445–449.
Rochet, B. (**1995**). ''Perception and production of second-language speech sounds by adults,'' in *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, edited by W. Strange (York, Baltimore), pp. 379–410.
Sebastian-Galles, N., and Soto-Faraco, S. (**1999**). ''Online processing of native and non-native phonemic contrasts in early bilinguals,'' Cognition **72**, 111–123.
Selinker, L. (**1972**). ''Interlanguage.'' IRAL **10**, 209–231.
Smith, B., Bradlow, A. R., and Bent, T. (**2003**). ''Production and perception of temporal contrasts in foreign-accented English,'' in Proceedings of the International Congress on Phonetic Science, Barcelona, Spain.
Smith, L. E., and Rafiqzad, K. (**1979**). ''English for cross-cultural communication: The question of intelligibility,'' TESOL Quarterly **13**, 371–380.
Strange, W., Editor (**1995**). *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (York, Baltimore).
Studebaker, G. A. (**1985**). ''A rationalized arcsine transform,'' J. Speech Hear. Res. **28**, 455–462.
van Wijngaarden, S. J. (**2001**). ''Intelligibility of native and non-native Dutch speech.'' Speech Commun. **35**, 103–113.
van Wijngaarden, S. J., Steeneken, H. J. M., and Houtgast, T. (**2002**). ''Quantifying the intelligibility of speech in noise for non-native listeners.'' J. Acoust. Soc. Am. **111**, 1906–1916.
Weinreich, U. (**1953**). *Languages in Contact: Findings and Problems* (Mouton, The Hague).