

RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 19 (1993-1994)
Indiana University

**New Directions for Assessing Speech Perception
in Persons with Sensory Aids¹**

**Karen Iler Kirk,² David B. Pisoni, Mitchell S. Sommers,³
Marilyn Young,⁴ and Christopher Evanson,⁴**

*Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405*

¹This research was supported by NIDCD Research Grants DC-0064 and DC-00111 and NIDCD Training Grant DC-00012 to Indiana University. The authors gratefully acknowledge the assistance of Luis Hernandez and John Karl of the Speech Research Laboratory, Indiana University, Bloomington, in stimulus preparation. We also thank Susan Todd, Allyson Riley, David Ertmer, and Terri Kerr for their help with data collection and analysis.

²Also at Department of Otolaryngology-Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, IN.

³Now at Department of Psychology, Washington University, St. Louis, MO.

⁴Also at Indiana University School of Medicine, Indianapolis, IN.

Abstract

This study examined the influence of stimulus variability and lexical difficulty on the speech perception performance of adults who used either multichannel cochlear implants or conventional hearing aids. The effects of stimulus variability were examined by comparing word identification in single-talker vs. multiple-talker conditions. Lexical effects were assessed by comparing recognition of "easy" words (i.e., words that occur frequently and have few phonemically similar words, or neighbors) with "hard" words, (i.e., words with the opposite lexical characteristics). Word recognition performance was assessed using either closed- or open-set response formats. The results demonstrated that both stimulus variability and lexical difficulty influenced word recognition performance. Identification scores were poorer in the multiple-talker than in the single-talker conditions. Also, scores for lexically "easy" items were better than for "hard" items. The effects of stimulus variability were not evident when a closed-set response format was employed.

New Directions for Assessing Speech Perception in Persons with Sensory Aids

Introduction

Human listeners typically display a great deal of "perceptual robustness." They are able to perceive and understand spoken language very accurately over a wide range of conditions that modify the acoustic-phonetic properties of the signal. Among these conditions are differences due to talkers, speaking rates, dialects, speaking styles and alternative pronunciations, as well as several kinds of signal degradations. Previous research with listeners with normal hearing suggests that the nervous system preserves a great deal of the fine acoustic-phonetic detail present in the speech signal. Indexical properties of speech, including talker-specific attributes, are not lost or filtered out during perceptual analysis. Instead, these attributes are encoded by listeners with normal hearing, and become part of a very rich and elaborate representational system in long-term memory.¹

In listeners with normal hearing (NH), introducing stimulus variability in the acoustic signal (e.g. by varying the number of talkers), causes word identification accuracy to *decrease* and response latency to *increase*.² Introducing variability by varying the lexical properties of words, such as their frequency of occurrence in the language, or their phonetic similarity to other words in memory, influences spoken word recognition in a similar fashion. One measure of lexical similarity is the number of "lexical neighbors" or words that differ by one phoneme from the target word. "Easy" words (high frequency words from sparse neighborhoods) are recognized *faster* and with *greater accuracy* than "hard" words (low frequency words from dense neighborhoods)³. At present very little is known about the effect of stimulus variability on spoken word recognition by listeners with hearing impairment, in part because their perceptual abilities are often assessed with tests that were specifically designed to minimize variability, such as the NU-6⁴. These tests consist of *phonetically-balanced monosyllabic words* spoken by a single talker at a single speaking rate; stimulus items are not balanced for lexical difficulty. Including different sources of variability in measures of word recognition might provide insight into the underlying speech perception processes employed by listeners with hearing impairment.

Experiment I

In the first experiment, we examined speech perception under conditions of stimulus and lexical variability in adults with a mild-to-moderate hearing loss. We also compared their performance on these conditions with their performance on a more traditional test of word recognition.

Subjects

The subjects were 12 adults between the ages of 18-65 years who had a mild-to-moderate bilateral sensorineural hearing loss, with speech discrimination scores of at least 50% words correct on a recorded version of the NU-6 lists. All subjects were current hearing aid (HA) users, and had a minimum of three months experience with their sensory aids. Six of the subjects used bilateral amplification, and six wore a hearing aid in only one ear. The subjects' mean age was 55.3 years.

Stimulus Materials

Both recorded monosyllabic word and sentence materials were employed. The monosyllabic word lists contained words selected from the Modified Rhyme Test (MRT).⁵ From a digital database containing

6000 recorded monosyllabic words (300 MRT words recorded by 20 talkers), a subset of 100 words containing 50 "easy" and 50 "hard" words was selected using lexical statistics obtained from a computerized version of Webster's Pocket Dictionary.² The single-talker condition contained 50 tokens produced by the same male talker, and the multiple-talker condition contained 50 tokens selected from five male and five female talkers. Half of the tokens in each condition were lexically "easy" and half were "hard." The sentence materials were selected from a computerized database containing 100 Harvard sentences produced by 10 male and 10 female talkers.⁶ Each sentence contained five key words to be identified. A male talker was selected for the single-talker condition. Both the single- and multiple-talker conditions contained 48 sentences.

Procedure

With the subjects seated inside a sound-treated room, the stimuli were presented via the free field at approximately 70 dB SPL, a level approximating that of loud conversational speech. Subjects wore their personal amplification during testing. Subjects repeated the word they heard and the experimenter transcribed the response. To be counted as correct, a response had to be identical to the target word. For example, the plural form of a singular target was counted as incorrect. The sentence stimuli were scored by the percent of key words correctly identified.

Results

Figure 1a presents the mean percent of "easy" and "hard" monosyllabic words correctly identified in the single-talker and multiple-talker test conditions, along with their average performance on the NU-6. Performance was better on the NU-6 than on the single- and multiple-talker conditions containing lexical variability. In addition, the HA subjects identified words with significantly greater accuracy in the single-talker condition than in the multiple-talker condition ($p < .0001$), and "easy" words were identified with significantly greater accuracy than "hard" words ($p < .003$). Figure 1b presents the results for the sentence materials. As expected, the subjects' word recognition scores improved when sentences were presented. Figure 1b illustrates that word recognition declined when multiple talkers were introduced, despite the presence of syntactic and semantic cues to word identity that were available in the sentence stimuli. Thus the effects of stimulus variability appear to be relatively robust for both isolated words and sentences.

Insert Figure 1a about here.

Insert Figure 1b about here.

Discussion

Using traditional clinical criteria based on NU-6 scores, these adult listeners with mild-to-moderate hearing loss would be considered to have "good" spoken word recognition skills. However, when word recognition performance was assessed under more demanding perceptual conditions containing several sources of stimulus variability, their performance declined. These listeners not only had lower word recognition scores in the multiple-talker condition compared to the single-talker condition, but they also had more difficulty making fine acoustic-phonetic discriminations among *lexically difficult words* in the

HEARING AID USERS

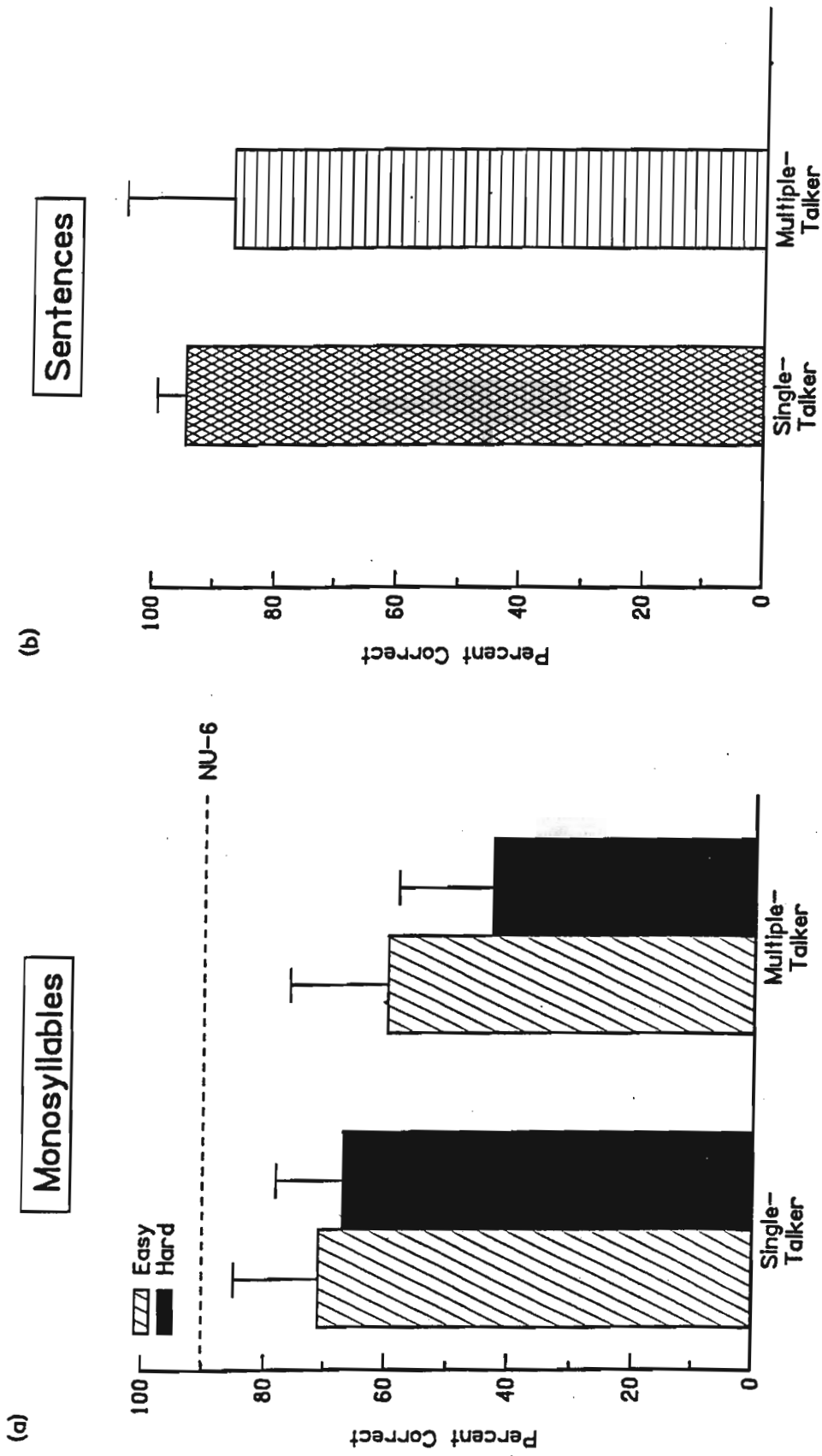


Figure 1. HA aid subjects' word recognition performance in two talker conditions for word (1a) and sentence (1b) stimuli. The percent of monosyllabic words correctly identified as a function of lexical difficulty also is presented (1a).

multiple-talker conditions. That is, they identified "easy" words with greater accuracy than "hard" words in the multiple-talker condition. This finding suggests that although listeners with mild-to-moderate hearing impairment may have reduced frequency selectivity and audibility and have difficulty making fine phonetic distinctions, they do recognize words "relationally." That is, they perceive words in the context of other phonetically similar words, just as normal listeners do.

Experiment II

In this second experiment we investigated the effects of stimulus variability on spoken word recognition by adult multichannel cochlear implant (CI) users, using both isolated words and sentence stimuli. In addition, the effects of stimulus variability on monosyllabic word recognition were examined utilizing both an open- and closed-set response format.

Subjects

All subjects were 18 years or older, with a bilateral profound hearing loss. These subjects received a CI because they derived no benefit from conventional HAs. Subject selection criteria included a minimum of six months CI use, and NU-6 scores $\geq 10\%$ words correct. Nineteen CI users were initially tested as part of this investigation. Eleven CI subjects met the inclusion criteria; 10 used a Nucleus CI and one used a Clarion CI. Their mean age was approximately 49 years, and their mean length of device use was 4.3 years.

Procedure

The stimulus materials and procedures used to assess open-set speech recognition were the same as those in Experiment I, except that the CI subjects used their CI rather than a HA during testing. In addition, the CI subjects were administered a closed-set test of word identification, the Word Intelligibility by Picture Identification (WIPI).⁷ The WIPI consists of four different lists of 25 monosyllabic words. Subjects respond by selecting the item they heard from among a six-picture matrix. For the present study, the subjects were presented with two different recorded lists: one produced by a male talker, and one in which the items were produced by five male and five female talkers. Only eight of the CI subjects were administered the Harvard sentences and the WIPI due to time constraints.

Results and Discussion

Figure 2a presents the mean percent of words correctly identified for the open-set single- and multiple-talker conditions, along with the average performance on the NU-6. These CI subjects performed better on the NU-6 than on new stimulus lists containing talker or lexical variability. Furthermore, their performance was better marginally on the single-talker than on the multiple-talker conditions ($p < .1$), and significantly better on the lexically "easy" than the "hard" words ($p < .03$). Figure 2b presents the results for the sentence stimuli. The CI subjects' word recognition performance improved when sentence stimuli were presented. For the CI subjects, a larger difference between the single- and multiple-talker conditions was observed for the sentence stimuli than for isolated words, perhaps because the baseline performance in the single-talker condition was higher for sentences. The CI subjects' performance declined marginally in the multiple-talker condition compared to the single-talker condition ($p < .1$) when sentence stimuli were employed.

Insert Figure 2a about here.

Insert Figure 2a about here.

Table 1 presents a comparison of the CI subjects' recognition of isolated words as a function of talker condition and response format. As expected, these subjects' word recognition performance was better when a closed-set response format is employed. However, Table 1 shows that talker variability did not influence word recognition performance when the set of potential responses is greatly constrained through the use of a closed-set task.

Insert Table 1 about here.

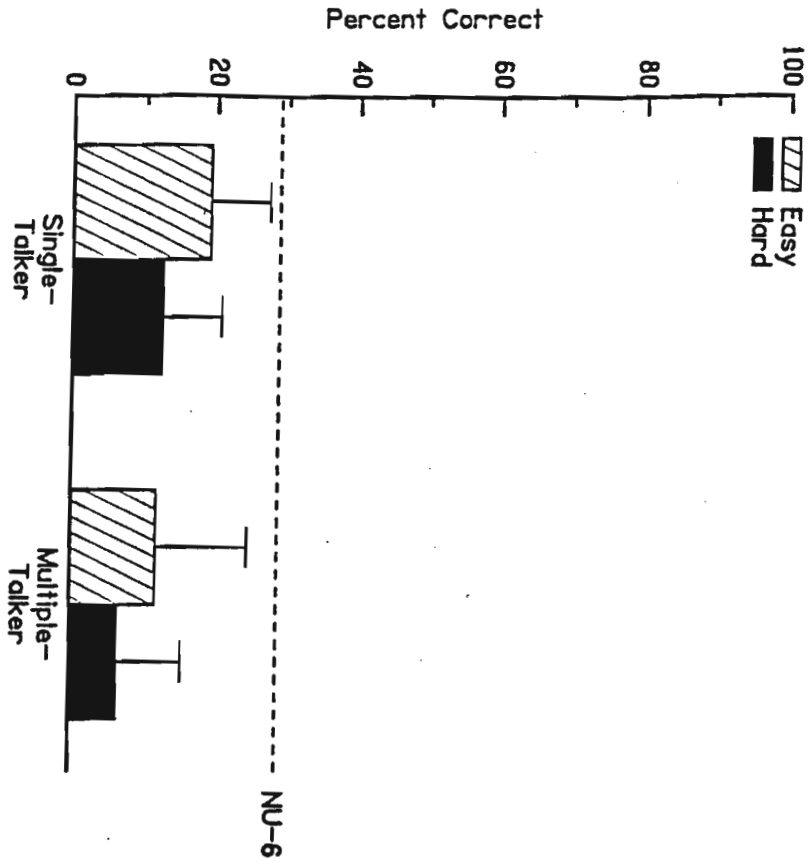
GENERAL DISCUSSION

The results suggest that listeners with sensory aids (either hearing aids or cochlear implants) employ underlying perceptual processes that are similar to those of listeners with normal hearing. When spoken word recognition was assessed under more demanding perceptual conditions containing stimulus variability, recognition performance declined. That is, performance was poorer in the multiple- than in the single-talker condition for both the monosyllabic and sentence stimuli. These results suggest that the effect is relatively robust when open-set stimuli are employed. However, when other structural cues to word recognition were present, as in the closed-set response mode, the effect of talker variability was eliminated. Lastly, the finding that lexically "easy" words were identified with greater accuracy than lexically "hard" words suggests that listeners with hearing impairment perceive words in the context of other phonemically similar words, just as listeners with normal hearing do. In summary, the present results suggest that stimulus variability is important in speech perception by persons with hearing impairment who demonstrate at least some open-set speech perception. It appears that these listeners with sensory aids (either hearing aids or cochlear implants) encode instance-specific information that is present in the speech signal. Finally, the present results suggest that low-variability tests of speech recognition may not adequately assess listeners' underlying perceptual capabilities when they have to recognize speech in the presence of stimulus variability, such as might be encountered in natural communication situations, when there is noise, speech input from many different talkers, or other sources of signal degradation.

CI USERS

Monosyllables

(a)



Sentences

(b)

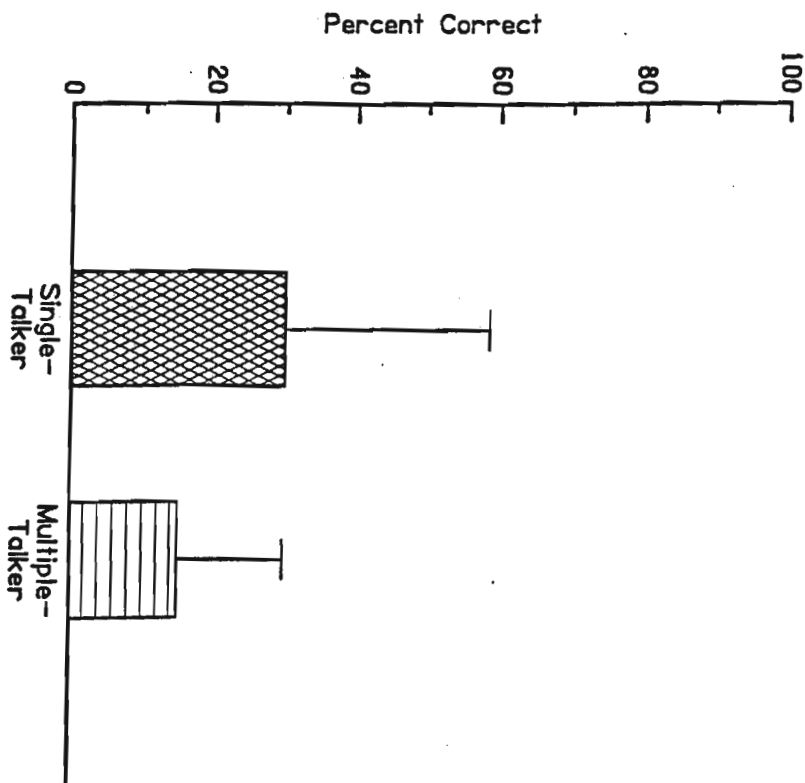


Figure 2. CI aid subjects' word recognition performance in two talker conditions for word (2a) and sentence (2b) stimuli. The percent of monosyllabic words correctly identified as a function of lexical difficulty also is presented (2a).

Table 1.

The percent of words correctly identified by the CI subjects as a function of talker condition and response mode.

	Open-Set Response (N=11)		Closed-Set Response (N=8)	
	Single-talker	Multiple-talker	Single-talker	Multiple-talker
Mean	15.9%	9.14%	43.7%	43.6%
SD	8.2%	10.9%	27.9%	29.2%

References

1. Pisoni, D.B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate, and perceptual learning. *Speech Communication*, 13, 109-125.
2. Pisoni, D.B. (1992). Some comments on invariance, variability and perceptual normalization in speech perception. *Proceedings 1992 International Conference on Spoken Language Processing, Banff, Canada*, 587-590.
3. Luce, P.A. (1986). Neighborhoods of words in the mental lexicon. *Research on Speech Perception Technical Report No 6*. Bloomington, IN: Indiana University.
4. Tillman, T.W. & Carhart, R. (1966). An expanded test for speech discrimination utilizing CNC monosyllabic words. *Northwestern University Auditory Test No. 6 Technical Report No. SAM-TR-66-55*. Brooks Air Force Base, TX: USAF School of Aerospace Medicine.
5. House, A.S., Williams, C.E., Hecker, M.H.L. & Kryter, K.D. (1965). Articulation-testing methods: Consonantal differentiation with a closed-response set. *Journal of the Acoustical Society of America*, 35, 158-166.
6. Karl, J.R. & Pisoni, D.B. (1994). The role of talker-specific information in memory for spoken sentences. *Journal of the Acoustical Society of America*, 95, 2873.
7. Ross, M. & Lerman, J. (1970). A picture identification test for hearing impaired children. *Journal of Speech & Hearing Research*, 13, 44-53.