Telephone Speech Perception by Mandarin-Speaking Cochlear Implantees

Mei-Ji Horng, Hsiao-Chuan Chen, Chuan-Jen Hsu, and Qian-Jie Fu

Objective: **To evaluate Mandarin-speaking cochlear implant patients' understanding of telephone speech.**

Design: **Telephone speech was simulated by band-limiting broadband speech stimuli (300–3200 Hz) and adding Gaussian noise (35 dB signal-to-noise ratio). Recognition of multitalker vowels, consonants, voice gender, and Chinese tones was measured for both simulated telephone speech and broadband speech in fifteen Mandarin-speaking cochlear implant patients.**

Results: **Results showed no significant difference in Chinese tone recognition scores between broadband and telephone speech. However, mean recognition scores for vowels, consonants and voice gender were significantly lower with telephone speech. The effect of the limited telephone bandwidth on speech recognition was highly variable among subjects. Some subjects were more sensitive to high-frequency speech cues, resulting in a significant drop in performance with band-limited telephone speech, while other subjects were less sensitive to high-frequency speech cues, resulting in similar performance between broadband and band-limited telephone speech.**

Conclusions: **These results suggest that the limited bandwidth negatively affects cochlear implant patients' understanding of telephone speech. Because the effect of band-limited speech was highly variable among subjects, the results also suggest that the contribution of high frequency information to speech recognition may vary significantly among cochlear implant patients. For patients who receive little benefit from high-frequency speech cues, speech processor adjustments may be necessary to access the additional cues provided in broadband speech.**

(Ear & Hearing 2007;28;66S–69S)

Many cochlear implant (CI) patients are capable of good speech understanding in quiet listening conditions. For CI patients who are able to understand natural (i.e., broadband) speech without difficulty, the channel noise and limited bandwidth associated with telephone speech may cause speech understanding to decline. While many CI patients are capable of some degree of telephone communication (Cohen, Waltzman, & Shapiro, 1989), speech understanding was significantly worse with telephone speech than with broadband speech (Ito, Nakatake, & Fujita, 1999; Milchard & Cullington, 2004). Milchard & Cullington (2004) investigated the effect of band-limited speech on consonant recognition by 10 adult CI users and 10 normal-hearing (NH) listeners. Consonant recognition scores were measured for broadband speech and for three conditions of band-limited speech (300–4500 Hz, 300–3400 Hz, 300–2500 Hz). For CI subjects, there was no significant difference in performance with the broadband and the 300–4500 Hz band-limited speech; however, performance was significantly worse with the 300–3400 Hz and 300–2500 Hz band-limited speech. For NH subjects, there was no significant difference between broadband speech and the three band-limited speech conditions. These previous studies suggest that, while a significant amount of speech information can be conveyed to CI users with band-limited speech, the telephone speech bandwidth may not preserve important high-frequency speech cues. While the effect of band-limited speech has been measured for English phonemes and sentences, it is unclear how band-limited speech may affect the reception of suprasegmental speech cues, e.g., tones in a tonal language.

The reception of tonal information is important for Mandarin-speaking CI users' speech recognition (e.g., Fu, Zeng, Shannon, et al., 1998; Luo & Fu, 2004). There are four tonal patterns in Mandarin Chinese, which can be characterized by the variation in speech patterns' fundamental frequency (F0) during voiced speech. In general, Tone 1 has a flat and high F0 pattern, Tone 2 has a rising F0 pattern, Tone 3 has a falling and then rising F0 pattern and Tone 4 has a falling F0 pattern. For example, the monosyllable /ma/, pronounced using Tone 1, 2, 3, or 4 can mean "mother," "hemp," "horse," or "scold," respectively. Currently, most CI speech processing strategies do not provide sufficient spectral or temporal cues to support tone recognition; the fundamental frequency (F0) and its harmonics are not explicitly encoded, and temporal pitch cues in electric hearing are generally weak. With the latest CI technology, native Mandarin-speaking CI patients

Graduate Institute of Audiology and Speech Therapy, National Kaohsiung Normal University, Kaohsiung, Taiwan (M.-J.H., H.-C.C.); Department of Otolaryngology, National Taiwan University College of Medicine, Taipei, Taiwan (M.-J.H., C.-J.H.); and the Department of Auditory Implants and Perception, House Ear Institute, Los Angeles, California (Q.-J.F.).

0196/0202/07/282 Supplement-0066S/0 • Ear & Hearing • Copyright © 2007 by Lippincott Williams & Wilkins • Printed in the U.S.A.

recognize only 50–70% of Chinese tones, performance comparable to that of NH subjects listening to a 4-channel acoustic CI simulation (e.g., Fu, Hsu, & Horng, 2004; Wei, Cao, & Zeng, 2004). The goal of the present study is to evaluate CI patients' understanding telephone speech; vowel, consonant, voice gender and Chinese tone recognition was measured in 15 CI subjects, for both broadband and simulated telephone speech.

METHODS

Subjects

Fifteen Mandarin-speaking children (7 males and 8 females) implanted with the Nucleus-24 CI device participated in the current experiment. All subjects used the ACE strategy in their clinically assigned processor; all subjects were assigned frequency allocation Table 7 (188–7938 Hz). Subject ages ranged from 8 to 16 years, with a mean age of 10.9 years; all children were prelingually deafened. The duration of implant use ranged from 3 to 7 years, with a mean of 5.5 years. Relevant demographic details are shown in Table 1.

Signal Processing

Telephone speech was simulated by band-pass filtering (300–3200 Hz; 96 dB/octave filter slopes) broadband speech; since all speech stimuli were sampled at 16 kHz, the bandwidth of broadband speech was 0–8000 Hz. Because telephone speech quality can also affected by different kinds of noise during transmission (e.g., circuit noise, line noise), Gaussian noise was added to the band-limited speech signal (35 dB signal-to-noise ratio, or SNR).

TABLE 1. Cochlear implant subject demographics for the present experiment

Subject	Age	Gender	Duration of use (yr)
1	12	F	7
2	13	F	7
3	9	M	3
4	8	F	6
5	11	M	7
6	8	M	6
7	9	F	5
8	11	M	5
9	16	F	5
10	12	M	3
11	12	F	3
12	14	M	6
13	11	F	7
14	9	M	6
15	9	F	6

Speech Materials and Procedures

The speech materials for vowel, consonant, voice gender, and Chinese tone recognition tests were drawn from the *Chinese Standard Database* (Wang, 1993). All speech stimuli were sampled at a 16-kHz sampling rate, without high-frequency pre-emphasis. One male and one female speaker each produced 4 tones for 6 Mandarin Chinese single-vowel syllables (/a/, /o/, /e/, /i/, /u/, /ü/), resulting in a stimulus set of 48 speech tokens. Vowel recognition was measured using a 6-alternative identification paradigm. Chinese tone recognition was measured with the vowel stimulus set, using a 4-alternative identification paradigm. Voice gender discrimination was also measured with the vowel stimulus set, using a 2-alternative identification paradigm. For consonant test materials, one male and one female speaker each produced Tone 1 (flat tone) for 19 Mandarin Chinese syllables (/a, ba, pa, ma, fa, da, ta, la, ga, ka, ha, za, ca, sa, zha, cha, sha, ya, wa/), resulting in a stimulus set of 38 speech tokens (utterances). Chinese consonant recognition was measured using a 19-alternative identification paradigm.

Chinese vowel, consonant, voice gender, and tone recognition were measured for both broadband and band-limited telephone speech. The speech tests and conditions were randomized and counterbalanced across subjects. During testing, a stimulus was presented to the subject in free field at a comfortable listening level. The subject responded by clicking on one of the response choices (6 choices for the vowel recognition test, 19 for the consonant recognition test, 2 for the voice gender recognition test and 4 for the Chinese tone recognition test), after which a new stimulus was presented. No training or feedback was provided.

RESULTS

Figure 1 shows individual subject and group mean data for vowel (panel A), consonant (panel B), voice gender (panel C) and Chinese tone recognition (panel D), for both broadband and simulated telephone speech. Note that 3 CI subjects (1, 7, and 9) did not complete the voice gender recognition test due to the time constraint. Consonant recognition performance was most affected by the reduced bandwidth of telephone speech. Paired *t*-tests showed that performance was significantly worse with telephone speech for vowel ($p = 0.016$), consonant ($p < 0.001$) and voice gender recognition ($p = 0.043$). There was no significant difference in Chinese tone recognition scores between broadband and telephone speech (paired *t*-test, $p = 0.164$).

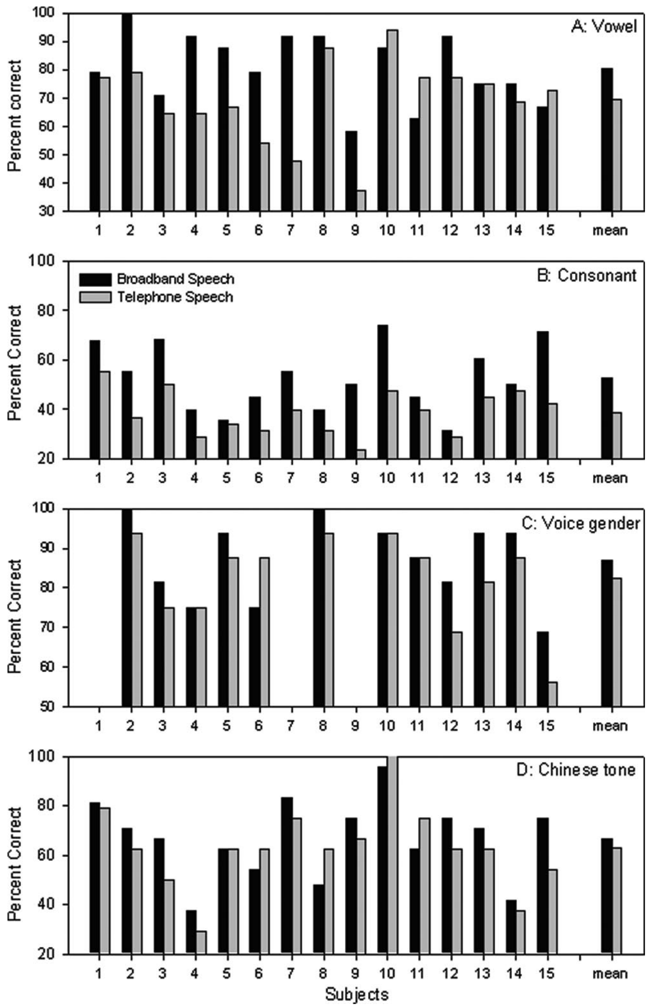

Fig. 1. Individual and group mean data with broadband and simulated telephone speech.

Discussion

In the present study, mean Chinese tone recognition performance with broadband speech (67% correct) was similar to that observed in previous studies with adult Mandarin-speaking CI users (e.g., Fu et al., 2004; Wei et al., 2004). Mean tone recognition performance did not significantly change when high-frequency speech cues were removed by the telephone speech simulation. Previous studies have shown that the amplitude contour and periodicity information are primary cues for Chinese tone recognition by CI users and NH subjects listening to acoustic CI simulations (e.g., Fu et al., 1998, 2004; Fu & Zeng, 2000). The data from the present study suggest that these cues are largely preserved with telephone speech, resulting in comparable tone recognition performance between broadband and telephone speech.

Mean voice gender recognition performance with broadband speech (87% correct) was comparable to that observed in previous studies with adult English-speaking CI users (e.g., Fu, Chinchilla, Nogaki, et al., 2005). A small but significant deficit in mean voice gender recognition was observed with simulated telephone speech. These results suggest that, although voice gender recognition strongly depends on periodicity cues (Fu et al., 2005), which were preserved with the telephone speech simulation, high-frequency speech cues may also contribute to voice gender recognition. Mean vowel recognition dropped by 11 percentage points with telephone speech. This result is somewhat surprising, as vowel recognition is thought to strongly depend on first and second formant cues, which would have been well preserved by the telephone speech bandwidth; the present results suggest that the high-frequency information (e.g., third formant cues), may also significantly contribute to Chinese vowel recognition. The largest drop in mean performance with telephone speech was observed for Chinese consonant recognition (14 percentage points), similar to the performance deficit (17 percentage points) observed with adult English-speaking CI users in the study by Milchard & Cullington (2004).

There was also significant intersubject variability in terms of the effect of band-limited speech on performance, suggesting significant differences in subjects' access to or use of high-frequency speech cues. For subjects who were sensitive to high-frequency speech cues, the telephone speech bandwidth caused performance to significantly decline. For subjects who were less sensitive to high-frequency speech cues, there was no significant difference between broadband and telephone speech. For example, the deficit in consonant recognition performance with telephone speech was significantly correlated with subjects' broadband consonant recognition ($r = 0.75$; $p = 0.001$). Speech processor adjustments and/or auditory training may be necessary to improve CI users' access to high-frequency speech cues, thereby improving broadband speech recognition. Auditory training may also improve CI users' recognition of band-limited telephone speech.

Conclusions

The limited bandwidth associated with telephone speech significantly reduced CI subjects' Chinese speech recognition performance, relative to that with broadband speech. Chinese tone recognition was not significantly affected by the reduced bandwidth. The contribution of high-frequency information to speech understanding may vary significantly among CI patients.

ACKNOWLEDGMENTS

We would like to thank all subjects who participated in the present study. We thank John J. Galvin III for editorial assistance. All testing was conducted using the "Melody Auditory Training Software," provided by Melody Medical Instruments Corp., Taiwan and House Ear Institute.

Address for correspondence: Dr. Qian-Jie Fu, Department of Auditory Implants and Perception, House Ear Institute, 2100 West 3rd Street, Los Angeles, CA 90057. E-mail: qfu@hei.org.

Received December 15, 2005; accepted May 17, 2006.

REFERENCES

Cohen, N. L., Waltzman, S. B., & Shapiro, W. H. (1989). Telephone speech comprehension with use of the nucleus cochlear implant. *Annals of Otology, Rhinology & Laryngology Supplement*, *142*, 8–11.

FuQ.-J., Chinchilla, S., Nogaki, G., & Galvin, J.J. (2005). Voice gender identification by cochlear implant users: the role of spectral and temporal resolution. *Journal of the Acoustical Society of America*, *118*, 1711–1718.

Fu, Q.-J., Hsu, C.-J., Horng, M.-J. (2004). Effects of speech processing strategy on Chinese tone recognition by Nucleus-24 cochlear implant patients. *Ear & Hearing*, *25*, 501–508.

Fu, Q.-J., Zeng, F.-G., Shannon, R. V., & Soli, S. D. (1998). Importance of tonal envelope cues in Chinese speech recognition. *Journal of the Acoustical Society of America*, *104*, 505–510.

Fu, Q.-J., Zeng, F.-G,. (2000). Effects of envelope cues on Mandarin Chinese tone recognition. *Asia-Pacific Journal of Speech, Language and Hearing*, *5*, 45–57.

Ito, J., Nakatake, M., & Fujita, S. (1999). Hearing ability by telephone of patients with cochlear implants. *Otolaryngology—Head and Neck Surgery*, *121*, 802–804.

Luo, X., Fu Q.-J,. (2004). Enhancing Chinese tone recognition by manipulating amplitude envelope: Implications for cochlear implants. *Journal of the Acoustical Society of America, 116*, 3659–3667.

Milchard, A. J., & Cullington, H. E. (2004). An investigation into the effect of limiting the frequency bandwidth of speech on speech recognition in adult cochlear implant users. *International Journal of Audiology*, *43*, 356–362.

Wang, R.-H. (1993). The standard Chinese database. University of Science and Technology of China, internal materials.

Wei, C.-G., Cao, K.-L., Zeng, F.-G. (2004). Mandarin tone recognition in cochlear-implant subjects. *Hearing Research, 197*, 87–95.