 

# 低頻聽覺訊息對人工電子耳使用者在噪音環境之影響
# A Study of Low Frequency Acoustic Information for Cochlear Implant Users with Background Noise

黃柏誠[1] 李宜軒[*1] 蔡德明[2]
Bo-Cheng Huang[1], Yi-Hsuan Lee[*1], Charles T. M. Choi[2]

*摘要*

近年來許多研究發現人工電子耳搭配助聽器可以提升使用者的聽辨能力，特別是在噪音環境下的表現。這可能是因為助聽器所提供的低頻訊息可以幫助使用者從噪音中分離出語音訊息。本研究探討在噪音下低頻聽覺訊息對人工電子耳使用者語音辨識能力的影響。十位聽力正常的受試者分別在聆聽條件為只有單耳聽取模擬人工電子耳或雙耳分別聽取模擬人工電子耳與低通濾波語音下進行語音聽辨測試。結果顯示，當另一耳聽取低通濾波語音時能夠明顯的提升語音辨識率，在只有單耳聽取模擬人工電子耳且噪音來自另一側時，頭影效應明顯的提升聽辨能力。

*關鍵詞*：人工電子耳，低頻聽覺訊息，雙模式聽覺，聽覺實驗

*Abstract*

Recent studies have reported speech intelligibility from cochlear implant (CI) users fitting a contralateral hearing aid can be improved, especially in a noisy environment. This is partly true due to the contribution of the residual low frequency acoustic information provided from hearing aid. In this study, we evaluate the difference of speech recognition in Taiwanese Mandarin between CI users with and without low pass speech sound provided in noise. Ten normal hearing test-takers were measured speech recognition rate by listening to either unilateral simulated CI or bimodal (CI + low pass speech) hearing. The result shows the addition of low pass speech sound led to significant improvements in speech recognition under one–ear hearing condition and the head shadow effect was significant improved by performance in unilateral CI condition when noise coming from contralateral of CI.

*Keywords*: cochlear implant (CI), low frequency acoustic information, bimodal hearing, normal hearing experiment

## I. INTRODUCTION

Improvements in *Cochlear implant* (*CI*) technology and speech recognition outcomes have lead to some extension of the implantation criteria to include individuals with some residual hearing in one or both ears. Recent studies have reported that CI have successfully restored hearing performance in quiet in severely hearing impaired individuals. However, these individuals still remains significant difficulties in understanding speech in noise or the presence of competing talkers. This is partially because they do not have access to binaural hearing and the lack of pitch perception which can enable listener to segregated the target speech from noise.

For CI users with low frequency residual hearing may improve pitch perception ability by combined the electric and acoustic stimulation via using a *hearing aid* (*HA*) in the contralateral ear, this is known as bimodal hearing. Despite is unintelligible when presented alone, the low frequency sound is largely unavailable in the electric stimulation with CI. Previous studies have demonstrated that the addition of low frequency sound provides benefits in terms of speech recognition over the unilateral CI alone in noise condition [1-3]. Chang *et al.* [4] have suggested that listeners combine the relatively weak pitch information conveyed by the electric stimulation with the stronger pitch cue from the target in the low frequency acoustic to segregate target and background noise. Ching [5] showed the low frequency hearing from HA increases consonant voicing and manner of articulation information in bimodal hearing. Similarly,

[1] 國立臺中教育大學資訊工程學系 [2] 國立交通大學電機工程學系
[*]Corresponding author. E-mail: yslee@mail.ntcu.edu.tw
[1] Department of Computer Science, National Taichung University of Education, Taichung, Taiwan, R.O.C.
[2] Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C.
Manuscript received 19 November 2012; accepted 4 January 2013

Mok *et al.* [6] demonstrated that increased low frequency phonemes information (e.g., nasals, semivowels, diphthongs and the first formant frequency) were transmitted with bimodal hearing relative to the CI alone.

A tonal language is a language that uses tonal cues, which is called lexical tones, to identify word. For unilateral CI users who cannot percept lexical tones, we can expect their ability to recognize tonal language is poor. Ciocca *et al.* [7] demonstrated that CI users showed a deficit in recognizing these lexical tones for the Cantonese language, where pitch cues are not typically supplemented by temporal cues and only a few subjects able to score above chance levels. Taiwanese Mandarin is a tonal language variant of Mandarin and is derived from the official Standard Mandarin spoken in Taiwan. In this study, we were evaluate the benefits of speech recognition in Taiwanese Mandarin with simulated bimodal hearing.

## II. METHOD

### A. Subjects

10 normal-hearing subjects (6 females and 5 males) between the ages of 20 and 29 years old (mean 24.6 years) participated in this study. All subjects were native speakers of Taiwanese Mandarin.

### B. Stimuli

We use the material proposed in [8] which contains 300 Taiwanese Mandarin sentences for normal hearing experiment. Each sentence contains 7~10 words, and subjects are asked to recognize the last word. These sentences are evenly classified into two categories: *high predictability* (*HP*) and *low predictability* (*LP*). A sentence with HP means that it contains 2~3 cues to predict its last word. In contrast, a LP sentence doesn't contains any cues can predict its last word, which means the subject has to recognize the last word directly. We recruit a female talker to record this sentence material. Each sentences were recorded at a sampling rate of 44.1 kHz and a bit depth of 16 bits. After recording, each sentence would verify by five other subjects to ensure that all sentences could be fully recognized in quite condition. The masker was a multi-talker babble available from CD (AudiTec Ltd, St Louis).

### C. Procedure

All processing was done digitally via software routines in *MATLAB*. The directional hearing was simulated by using *head-related transfer functions* (*HRTFs*) [9]. The input signal which combined speech and noise was processed by filtering it with the corresponding HRTFs for each angle of incidence. The CI simulation was created using a vocoder to simulate the sound heard by CI users [10]. The low pass speech was derived from a 6th-order Butterworth filter with a cutoff at 500 Hz applied to the input signal which combined speech and noise to simulate residual low frequency hearing.

Fig. 1 shows how we use HRTFs to produce different source angle of incidence. Speech source is always fixedly originated in the front at 0 azimuth, and noise source is placed at an angle of incidence of 0° (S0N0), 90° (S0N90) or 270° (S0N270) azimuth, The distance between source and KEMAR was 1.4 meters. All sentences are mixed by multi-talker babble in signal-to-noise ratio (SNR) 5, 0 or -5 dB. For the listening mode, stimuli were presented either only vocoder in left ear to form unilateral CI-alone condition, or vocoder in left ear and low pass speech in right ear to form bimodal condition. The headphone we used is AKG K181DJ.

## III. RESULTS

Fig. 2 shows the experiment result conducted by normal hearing subject. Recognition % are plotted as the mean percent correct across subjects with the error bars indicating the standard error of the mean. A one way analysis of variance (ANOVA) was conducted to access for differences in recognition rates across SNR, listening mode or source angle of incidence. The S0N0, S0N90 and S0N270 indicated that the noise was presented at front, right and left, respectively. Significant differences were found between each SNR level except between SNR -5 and 0 dB at S0N270 for HP condition. At S0N270 and S0N0 conditions, bimodal hearing both showed significantly benefit compared with CI-alone at each SNR level. Especially at SNR 0 dB, the CI-alone condition yielded scores ranging from 11-46% and bimodal condition significantly improved to 54-80%. In CI-alone condition, significant differences were found between S0N270 and S0N90 at each SNR level, between S0N90 and S0N0 at SNR 5 dB for LP condition and SNR -5 dB and between S0N270 and S0N0 at SNR 5 dB for HP condition and SNR 0 dB. In bimodal condition, no significant differences were found between each source angle of incidence at SNR 0 and 5 dB, but significant differences were observed between S0N90 and S0N0 and between S0N270 and S0N90 at SNR -5 dB. Despite did not return performance to normal hearing levels, the addition of low frequency information was showed an improvement compared with CI-alone.

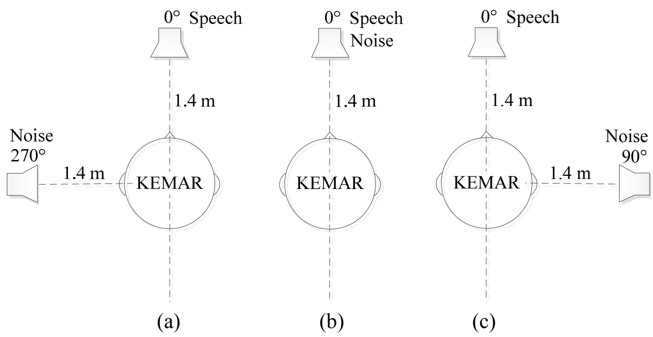


Fig. 1 The source angle of incidence in this experiment. Speech source is originated in the front at 0 azimuth. Noise source is placed at an angle of incidence of (a) 270°, (b) 0°, or (c) 90° azimuth

## IV. DISCUSSION

In this study, overall performance for the bimodal condition was greater than that for the CI-alone condition. Despite does not account for elevated thresholds of actual CI users, the low pass speech provides a theoretical basis for evaluating the usefulness of low frequency acoustic cues in our experiment. These results demonstrated that low frequency acoustic information can be an important cue under simulated bimodal hearing subject. Although the low pass speech was unintelligible, it was nevertheless improved the performance of speech recognition in Taiwanese Mandarin in noise.

The current findings, taken together with previous studies, such that when speech and noise were presented both from the front, Zhang *et al.* [11] showed adding low-frequency acoustic information to electrically stimulated information led to a significant improvement in sentence recognition in noise. They proposed that the robust representation of voicing allows access to low-frequency acoustic landmarks that mark syllable structure and word boundaries in noise. These landmarks can bootstrap word and sentence recognition. Whereas Berrettini *et al.* [12] evaluated the benefits from bimodal hearing when speech and noise are presented from different spatial directions and demonstrated improvements in speech perception from bimodal hearing, in comparison to the CI alone condition.

In a normal cochlea, pitch is coded by the temporal fine structure of the neural discharge or the place of excitation [13]. Almost current CI strategies extract only the temporal envelope of incoming signals and the fine structure information is discarded due to usage of fixed-rate pulsatile carrier. Although some methods have been proposed to enhance pitch representations in CI, e.g., using a varying pulse rate to restore the temporal cue [14] or using *virtual channels* to increase spectral resolution [15], all have shown limited functional benefit. The limited number of electrodes and the broad spread of electrical current also cause a very coarse place pitch cue. The relatively shallow insertion depth of current electrode arrays limit the transmitted of low frequency information. In bimodal condition, the addition of low pass speech provided low frequency acoustic information, such as F0, low-order harmonics, consonant voicing, lexical boundaries and contextual emphasis as well as manner. Above information are important to speech recognition in complex auditory scene [16], but in CI device, it couldn't transmitted.

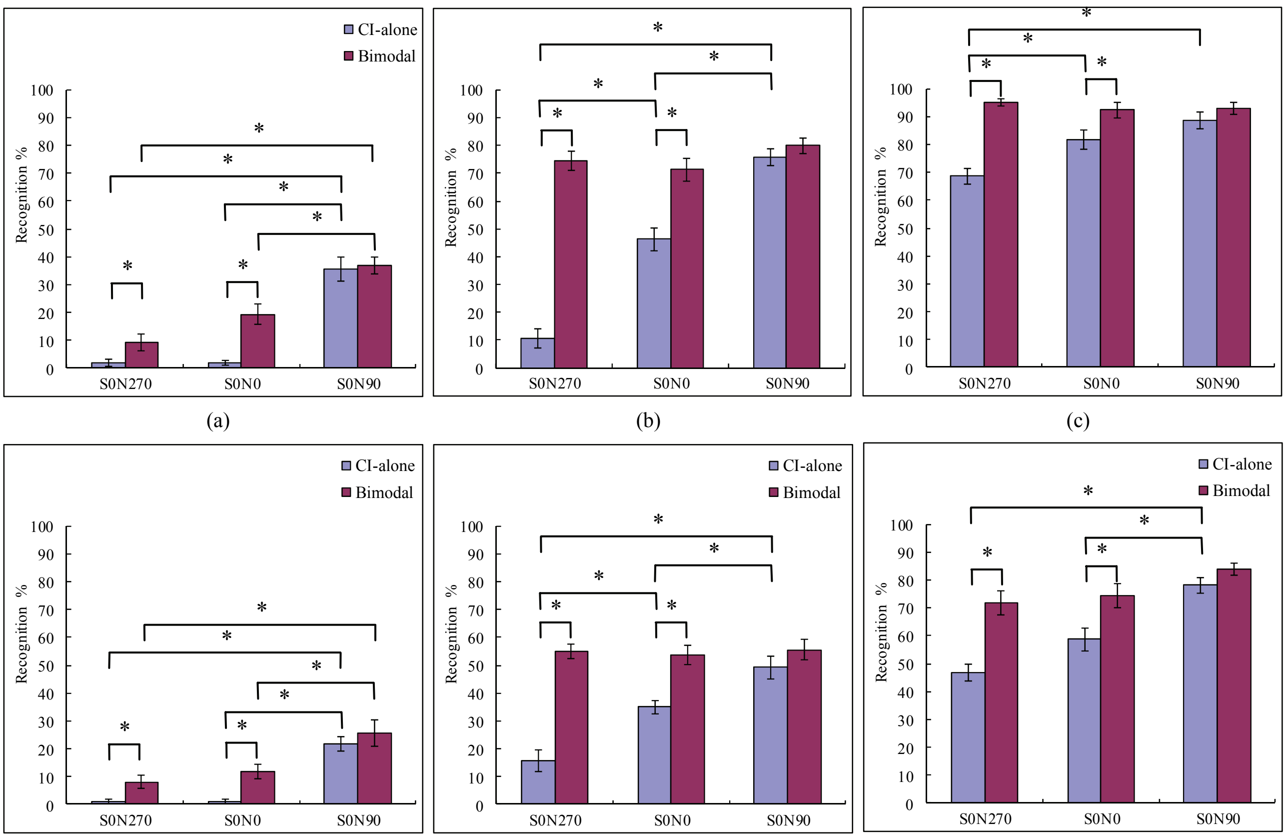


Fig. 2 The performance of speech recognition as a function of the source angle of incidence with comparing listening mode (CI-alone vs. Bimodal). Recognition rates are plotted as the mean percent correct across subjects with the error bars indicating the standard error of the mean in SNR (a) -5, (b) 0, and (c) 5 dB for HP (upper panel) and LP (lower panel) sentence. The symbol "*" indicates values that differed significantly between conditions at $P < 0.05$

With those low frequency cues, bimodal hearing significant enhanced performance range from 4-64%. Especially when noise was presented at the same side of CI and interference significantly, bimodal hearing improved by up to an average of 74%. We suggested that due to the lexical tone is critical to recognize Taiwanese Mandarin and it requires more accurate pitch perception ability which determines by F0.

In S0N270 condition, CI-alone yielded worst scores at each SNR level due to the significant interference by noise. A significant main effect between different noise source angle of incidences in CI-alone condition which referred to as *head shadow effect* was observed. The head shadow effect is a physical phenomenon due to the placement of the head. It refers to the benefit in speech recognition when the noise was moved from the ipsilateral side to the contralateral side of CI, so that the noise is shielded by the head. Bimodal hearing also provides binaural information to the central auditory nervous system enabling the utilization of binaural effect that assists in speech segregation in background noise. In S0N0 condition, the *binaural summation effect* led to bimodal condition was significantly benefit from CI-alone condition. The binaural summation effect describes the advantage of hearing with two ears results in signal being louder than with one ear. No significant difference was found between bimodal and CI-alone in S0N90 at each SNR level indicated that there was no *squelch effect* on simulated bimodal subject. The squelch effect describes if the speech and noise come from different spatial directions, adding an ear closer to the noise can significantly improve the speech recognition due to the ILD and ITD cues. However, because bimodal subjects use different strategies between ears which the ITD and ILD cues cannot be reliably coded, their speech recognition cannot be benefit by the squelch effect.

## V. CONCLUSIONS

In this study, we conduct the normal hearing experiment to evaluated the benefit of speech intelligibility by providing low frequency acoustic information to CI users. The results demonstrated that in general bimodal hearing enhanced the performance of speech recognition in noise in Taiwanese Mandarin. In our experiment, we found the head shadow and binaural summation effect, but the squelch effect is not exist. We suggested that speech recognition in Taiwanese Mandarin may improve by preserving residual low frequency hearing for CI users.

## ACKNOWLEDGEMENTS

This research was supported by the National Science Council of the Republic of China under contract numbers NSC 100-2221-E-142-001. We would like to thank the assistance from the following people. Prof. C. T. M. Choi from NCTU provided the vocoder. Prof. H. C. Chen from NKNU provided Taiwanese Mandarin materials.

## REFERENCES

[1] Y. K. Kong and R. P. Carlyon, "Improved speech recognition in noise in simulated binaurally combined acoustic and electric stimulation," *J. Acoust. Soc. Am.*, vol. 121, pp. 3717-3727, June 2007.

[2] C. A. Brown and S. P. Bacon, "Low-frequency speech cues and simulated electric-acoustic hearing," *J. Acoust. Soc. Am.*, vol. 125, pp. 1658-1665, Mar. 2009.

[3] Y. K. Kong, G. S. Stickney, and F. G. Zeng, "Speech and melody recognition in binaurally combined acoustic and electric hearing," *J. Acoust. Soc. Am.*, vol. 117, pp. 1351-1361, Mar. 2005.

[4] J. E. Chang, J. I. Bai, and F. G. Zeng, "Unintelligible low frequency sound enhances simulated cochlear-implant speech recognition in noise," *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 2598-2601, Dec. 2005.

[5] T. Y. C. Ching, "The evidence calls for making binaural-bimodal fittings routine," *Hear. J.*, vol. 58, pp. 32-41, Nov. 2005.

[6] M. Mok, D. Grayden, R. C. Dowell, and D. Lawrence, "Speech perception for adults who use hearing aids in conjunction with cochlear implants in opposite ears," *J. Speech Lang. Hear. Res.*, vol. 49, pp. 338-351, Apr. 2006.

[7] V. Ciocca, A. L. Frabcusm, R. Aisha, and L. Wong, "The perception of Cantonese lexical tones by early-deafened cochlear implantees," *J. Acoust. Soc. Am.*, vol. 111, pp. 2250-2256, May 2002.

[8] 蔡志浩和陳小娟，「噪音背景辨識語音測驗編製之研究」，*特殊教育研究學刊*，第 23 卷，第 121-140 頁，2002。

[9] B. Gardner and K. Martin, "HRTF measurements of a KEMAR," *J. Acoust. Soc. Am.*, vol. 97, pp. 3907-3908, June 1995.

[10] C. T. M. Choi, C. H. Hsu, W. Y. Tsai, and Y. H. Lee, "A vocoder for a novel cochlear implant stimulating strategy based on virtual channel technology," presented at Proc. of the 13th Int. Conf. on Biomed. Eng., Singapore, Dec. 6-10, 2008.

[11] T. Zhang, M. F. Dorman, and A. J. Spahr, "Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation," *Ear Hear*, vol. 31, pp. 63-69, Feb. 2010.

[12] S. Berrettini, S. Passetti, M. Giannarelli, and F. Forli, "Benefit from bimodal hearing in a group of prelingually deafened adult cochlear implant users," *Am. J. Otol. Head Neck Med. Surg.*, vol. 31, pp. 332-338, Sep. 2010.

[13] D. D. Greenwood, "A cochlear frequency-position function for several species–29 years later," *J. Acoust. Soc. Am.*, vol. 87, pp. 2592-2605, June 1990.

[14] H. J. McDermott and C. M. McKay, "Pitch ranking with nonsimultaneous dual - electrode electrical stimulation of the cochlea," *J. Acoust. Soc. Am.*, vol. 96, pp. 155-162, July 1994.

[15] P. Jones, H. McDermott, P. Seligman, and J. Millar, "Coding of voice source information in the Nucleus cochlear implant system," *Ann. Otol. Rhinol. Laryngol. Suppl.*, vol. 166, pp. 363-365, Sep. 1995.

[16] A. L. Giraud, C. Lorenzi, J. Ashburner, J. Wable, I. Johnsrude, R. Frackowiak, and A. Kleinschmidt, "Representation of the temporal envelope of sounds in the human brain," *J. Neurophysiol*, vol. 84, pp. 1588-1598, Sep. 2000.