ISCA Archive
http://www.isca-speech.org/archive

AVSP 2010 - International Conference
on Audio-Visual Speech Processing
Hakone, Kanagawa, Japan
September 30-October 3, 2010

# Long-term Cochlear Implant Users Have Resistance to Noise, but Short-Term Users Don't

*Yori Kanekama*[1], *Satoko Hisanaga*[2], *Kaoru Sekiyama*[1],
*Narihiro Kodama*[3], *Yasuhiro Samejima*[3], *Takao Yamada*[3], *and Eiji Yumoto*[4]

[1]*Faculty of Letters, Kumamoto University, Japan*
[2]*Graduate School of Social and Cultural Sciences, Kumamoto University, Japan*
[3]*Department of Otolaryngology-Head and Neck Surgery, Kumamoto University*
[4]*Department of Otolaryngology-Head and Neck Surgery, Graduate School of Medicine, Kumamoto University*

yxkanekama@yahoo.co.jp, Sekiyama@kumamoto-u.ac.jp

## Abstract

Hearing impaired people who received cochlear implants (CI) typically need some adaptation period over several years to maximize the benefit of CI. During the adaptation period, visual speech should play an important role. Although such a role of visual speech has been largely investigated for child CI users, it is not clear for adults. Here we investigated audiovisual speech perception in people who received CI as adults. Thirteen Japanese adult CI users were tested with Japanese meaningful word lists presented in auditory-visual (AV), auditory-only (AO), and visual-only (VO) conditions. The AV and AO conditions consisted of both quiet and noisy conditions. Percent correct responses showed that the resistance to noise is developed in 4 years of CI use for AV speech perception, and in about 8 years for AO speech perception. Results also showed that AO and AV performances were generally poor when the length of hearing loss was longer than 5 years. The performance in the VO condition was not related to that in AV condition, indicating that the AV benefit is a combining effect.

**Index Terms**: cochlear implant, speech perception in noise, adult

## 1. Introduction

### 1.1. Speech perception in CI users

The cochlear implant (CI) is an electric device that provides hearing sensation to patients with profound hearing loss. Nowadays, CI is increasingly becoming a treatment option for patients who cannot obtain benefits from hearing aids. As the CI technology has developed, CI users' speech comprehension performance has improved. With the most advanced CI technology and speech processing technology, many of the CI users receive great benefit such as conversing with friends and family over the telephone. There is, however, considerable variability remains in individual CI users.

One of the greatest difficulties for CI users is in listening with background noise. The difficulty may be due to the combinations of various factors such as the limited spectral information conveyed by the CI [1], the CI users' reduced frequency selectivity [2], their loss of binaural integration [3], their inability to distinguish fundamental frequency and noise [4]. To clarify the nature of speech comprehension in CI users, it is fair to evaluate both in quiet and noisy situations.

It is also known that the CI users' speech comprehension is improved after surgery not instantaneously, but gradually over several years. Apparently, the long learning process is very important to reorganize the CI users' brain system for the new sensory inputs. In prelingually deaf children with CI, a five-year longitudinal study has shown continuing improvements in the children's speech comprehension over five years [5]. In this study, it was also found that the age at implantation greatly affects the manner of adaptation, i.e., purely auditory improvement was limited in prelingually deaf children who received CI after 4.5 years of age. Therefore, the age at implantation and the duration of CI use are critical factors to predict children's speech comprehension.

Concerning adult CI recipients, the scene is quite different from children's because these recipients were postlingually deaf. Supported by their language competence and previous auditory experience, most adult CI recipients improve auditory speech comprehension irrespective of the age at implantation. However, two factors have been also noted to affect, that is, the duration of CI use and the length of hearing loss.

For adults, a large-scale longitudinal study is not available, but cross-sectional studies have shown that sentence comprehension performance in noise is related to the duration of CI use [3], [6] and the length of hearing loss [3], [7]. Generally, auditory speech comprehension performance is positively correlated with the duration of CI use and negatively correlated with the length of hearing loss.

To date, there is no study on CI users' speech comprehension in noise in which the contribution of visual speech is considered. The present study examined this aspect, by using not only auditory-only but also auditory-visual speech stimuli. The primary goal of the present study was to look at how the roles of auditory and visual cues change during years of adaptation process that adult CI users undergo. We also looked at the effect of the length of hearing loss.

### 1.2. Auditory-visual integration & CI users

Recently, there have been increasing interests in using the CI to study auditory-visual speech perception and integration. When the visual cue is available, CI users rely more than normal listeners on the visual cue, forcing them to become not only better speechreader but also better multisensory integrators [8].

For instance, adult CI users show significant AV benefit for syllable identification when the auditory and visual cues are congruent [9]. Child CI recipients' word comprehension increases over several years, and it is better in AV presentation than in AO or VO presentations [5]. By applying the McGurk effect paradigm to children, Schorr, Fox, van Wassenhove, and Knudsen (2005) reported that auditory-visual integration declined with age at implantation beyond 2.5 years: When the auditory and visual cues were incongruent, the late-implanted children's perception was dominated by visual cues whereas early-implanted children showed auditory-visual integration [10].

Moreover, an event-related potential study demonstrated a profound cortical reorganization in CI users, with good users being able to recruit a larger cortical area, even the visual cortex, than poor users to perform an auditory task [11]. The implication of this finding is a strong relationship between auditory and visual processing for CI users, perhaps as a result of adapting to their sensory characteristics.

In the present study, the AV benefit in noise was examined, especially in terms of its relationship with the duration of CI use and the length of hearing loss.

# 2. Method

## 2.1. Participants

Thirteen CI users (mean age of 61.7 years, ranging from 39 to 78 years) participated in this study. All CI participants were postlingually deafened and monolingual Japanese speakers with no known neurological, cognitive, or learning deficits, and used oral communication only. The mean time of deafness was 3.7 years ranging from .3 to 15 years and mean duration of CI use was 4.5 years ranging from .08 to 10 years at the time of the assessment. All CI users were monaurally implanted, and none used a hearing aid on the nonimplanted ear. Due to the limited available participant pool, factors typically related to CI use were not controlled (Table 1). For instance, it is important to note that the processing schemes used by each implant varied across listeners. Thus, the participant pool represents a heterogeneous group, in terms of hearing histories and processing schemes used.

## 2.2. Apparatus

Participants were seated 1 m from a loudspeaker (Sony SMS-1P) located in a quiet room for speech therapy. All speech stimuli were played by a DVD player (Sony HDMI DVP-NS700H) and background noise was produced by a CD player (Panasonic SL-S370). These were routed through a two-channel DJ mixer (Pioneer DJM-400) to the loudspeaker located in the sound-treated room. Visual portion of the speech stimuli was routed to a 20 inch flat TV monitor (Sony KLV-17HR1) from a DVD player. Speech and background noise stimuli were presented from $0^{o}$ azimuth for all testing. The output levels of the speech stimuli and background noise were calibrated at the position to be occupied by the listener and were checked periodically throughout the experiment (Onosokki LA 230, A-weighting, slow averaging).

### *2.2.1. Speech stimuli*

Speech stimuli were audio-visually recorded while a male speaker (a licensed Speech-Language Pathologist) was articulating words from CI 2004 list (Technical Committee on Cochlear Implant in Japan, 2004). The list consisted of 8 lists of 25 Japanese words. The speech stimuli was edited for two conditions [auditory-only (AO) and auditory-visual (AV)]. AO condition consisted of auditory speech and a visual fixation point. This visual fixation point was presented while each word was presented. In other words, the visual fixation point was the visual cue for CI users to notify when the word started and finished. The AV condition consisted of auditory and visual speech. Video was digitized by 29.97 frames per second in 720 × 480-pixel. Sound was digitized by 16-bit 44,100Hz resolution and stored in stereo. The speech stimuli were presented at 70 dB SPL.

### *2.2.2. Background noise*

CI users commonly report difficulty of their performance in noise. Thus, they need to be assessed their speech comprehension in competitive noise conditions, as in the reality. In the current study, 2-talker speech babble was applied to resemble real world environments. Moreover, a time-reversed speech babble was introduced in this study. Usually, when the speech sounds are time-reversed, the

Table 1. Demographics of the cochlear implant listeners. Including age, implant type, speech coding strategy, speech processor, length of hearing loss, and duration of implant use.

| Participant | Age (years) | Implant | Speech coding strategy | Speech processor | Length of hearing loss (year) | Duration of implant use (year) |
|---|---|---|---|---|---|---|
| 1 | 64 | Freedom Implant-Necleus24 Contour | ACE | Freedom | .3 | .08 |
| 2 | 64 | CI 24 R(CS)-Necleus 24 Contour | ACE | Freedom | 15 | .5 |
| 3 | 75 | CI 24 R(CS)-Necleus 24 Contour | ACE | Freedom | 2.6 | 1.16 |
| 4 | 50 | CI 24 R(CS)-Necleus 24 Contour | ACE | Freedom | 3.42 | 1.16 |
| 5 | 74 | CI 24 R(CS)-Necleus 24 Contour | ACE | Freedom | 5 | 1.33 |
| 6 | 78 | CI 24 R(CS)-Necleus 24 Contour | ACE | SPrint | 3 | 3.58 |
| 7 | 60 | CI 24 R(CS)-Necleus 24 Contour | ACE | SPrint | 2 | 3.91 |
| 8 | 79 | CI 24 R(CS)-Necleus 24 Contour | ACE | ESPrit-3G | 12 | 3.91 |
| 9 | 55 | CI24M-Necleus 24 | ACE | SPrint | .58 | 6.5 |
| 10 | 47 | CI24M-Necleus 24 | ACE | SPrint | .41 | 6.5 |
| 11 | 73 | CI24M-Necleus 24 | ACE | ESPrit | 3 | 7.66 |
| 12 | 39 | CI24M-Necleus 24 | ACE | ESPrit | .75 | 8.08 |
| 13 | 46 | CI24M-Necleus 24 | ACE | ESPrit | .5 | 9.91 |

spectral contents of speech and fundamental frequency are in essence untouched [12]. For normal hearing listeners, time-reversed speech masker would result in a slight improvement in performance instead of normal speech as a masker [2], [13]. In other words, the effects of interference on time-reversed speech masker are lower than speech masker. Therefore, in this study, the performance in time-reversed speech masker was expected to be better than that in normal speech masker.

A recording of two males' simultaneous speech served as background noise (2-talker babble). One read an article from the newspaper and one read from a story for children, "the north wind and the sun." Both sounds were digitized by 16-bit 44,100 Hz resolution and stored in stereo and combined by using audio editing software. A time-reversed speech babble was created by reversing the above 2-talker babble by using an audio editing software. The background noise was presented at 60 dB SPL (+10 dB SNR). The presentation levels for both speech stimulus and background noise were determined based on our preliminary experiments. In the preliminary study, ceiling effect was occurred for normal hearing listeners with +10 dB SNR, however, to avoid floor effect, +10 dB SNR was applied for CI users.

### *2.2.3. Condition*

Audio-only (AO), auditory-visual (AV), and visual-only (VO) conditions were administered in this study.

### *2.2.4. Procedures*

Each participant was instructed to listen carefully to the speech stimuli. When they heard the word, they were instructed to repeat aloud. If they were not sure what they heard, they were encouraged to guess. AO and AV conditions were administered with 2 types of background noise (2-talker babble & time-reversed babble) and quiet condition. VO condition was only administered in quiet condition (Table 2).
Half participants were tested in AV-AO-VO sequence, and another half was in AV-VO-AO sequence. CI 2004 list and the condition within each AO, AV, and VO conditions were counterbalanced across the participants.

Table 2. Experimental Conditions

| Modality | Noise |
|---|---|
| AO | 2-talker Babble (AO2) |
| | Reversed 2-talker Babble (AOR-2) |
| | Quiet (AO Q) |
| AV | 2-talker Babble (AV2) |
| | Reversed 2-talker Babble (AVR-2) |
| | Quiet (AV Q) |
| VO | Quiet |

## 3. Results

### 3.1. Duration of CI use and speech perception in noise

Figure 1 shows mean percent correct responses in AO and AV conditions for 3 groups of participants varying in the duration of CI use (short-CI: 1 to 3 years, medium-CI: 4 to 7 years, and long-CI: 8 to 10 years).

The performance of the long-CI group was the highest across the noise conditions compared to the other 2 groups in both conditions. The medium-CI group performed generally better than the short-CI group except AO condition in quiet condition. Generally, performance in AV condition was better than AO condition in all groups. First, a three-way ANOVA was conducted with factor being duration of CI use, AO and AV conditions, and noise, and the dependent variable being with percent correct. The results indicated that performance in AV condition was significantly higher than in AO condition. However, we would like to analyze the data within AO and AV conditions in detail. Therefore, a two-way ANOVA was applied in this study.

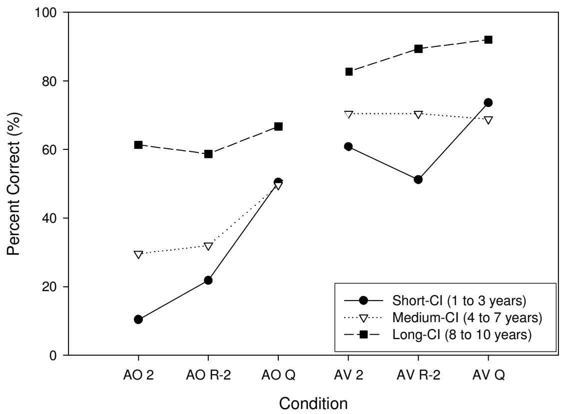


Figure 1. Duration of CI use (short-CI: 1 to 3 years, medium-CI: 4 to 7 years, and long-CI: 8 to 10 years) and percent correct responses for AO conditions (AO 2, AO R-2, & AO Q) and AV conditions (AV 2, AV R-2, & AV Q).

A two-way ANOVA was conducted with factor being duration of CI use and noise in AO condition and the dependent variable being with percent correct. The results for ANOVA indicated a significant duration of CI use effect [$F$ (2, 20) =5.08, $p$<.05] and noise condition [$F$ (2, 20) = 13.96, $p$<.001]. No significant interaction was measured between duration of CI use and noise [$F$ (4, 20) =2.63, $p$=.07].

Paired-samples t tests were conducted to evaluate the differences among 3 CI groups, using the Holm's sequential Bonferroni method to control for type I error. The results indicated that performance of long-CI group was significantly higher than other 2 group. There was a significant difference between performance in quiet and 2-talker babble, and quiet and reversed 2-talker babble. In other words, performance of CI users was negatively influenced by noise.

A two-way ANOVA was conducted with factor being duration of CI use and noise in AV condition and the dependent variable being with the percent correct. The results for ANOVA indicated a significant duration of CI use effect [$F$ (2, 10) =4.97, $p$<.05 and noise effect [$F$ (2, 20) =4.60, $p$<.05]. There was a significant interaction between duration of CI use and noise [$F$ (4, 20) =4.05, $p$<.05]. Paired-samples t tests were conducted to evaluate the differences among types of noise, using the Holm's sequential Bonferroni method to control for type I error. The results indicated that performance in quiet was significantly higher than 2-talker babble and reversed 2 talker babble. Furthermore, paired-samples t tests were conducted to evaluate the differences between duration of CI use and noise. Significant differences were measured in reversed 2-talker babble between long-CI group and short-CI group, and medium-CI group and short-CI group. Moreover, a significant difference was measured in quiet condition between long-CI group and medium-CI group. Performance of short-CI group was significantly higher in quiet condition than 2-talker babble and reversed 2-talker babble. However, in long-CI and medium-CI groups performance was not significantly different in among noise conditions. From the results, long-term CI users performed better not only in quiet condition but also in noise condition indicating the longer CI users tend to have resistance to noise.

### 3.2. Length of hearing loss and speech perception in noise

Figure 2 shows mean percent correct responses in AO and AV conditions for 3 groups of participants varying in the length of hearing loss [short-hearing loss (HL): 0 to 1 year, medium-HL: 2 to 4 years, and long-HL: 5 to 15 years].

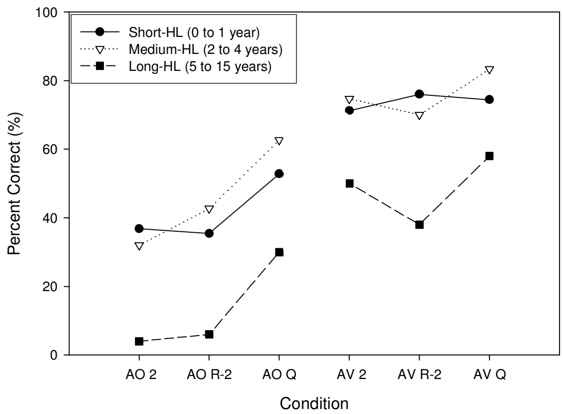


Figure 2. Length of hearing loss [short-hearing loss (HL): 0 to 1 year, medium-HL: 2 to 4 years, and long-HL: 5 to 15 years] and percent correct responses for AO 2, AO R-2, AO Q, AV 2, AV R-2, & AV Q).

The short-HL (0 to 1 year) and medium-HL (2 to 4 years) groups were very similar in their performances under all conditions. On the other hand, the long-HL group (5 to 15 years) was dramatically poorer than the other 2 groups across the conditions. Obviously, the performance in AV condition was better than AO condition in all groups. First, a three-way ANOVA was conducted with factor being length of CI use, AO and AV conditions, and noise, and the dependent variable being with percent correct. The results indicated that performance in AV condition was significantly higher than in AO condition. Again, we would like to analyze the data within AO and AV conditions in detail. Therefore, a two-way ANOVA was applied in this study.

A two-way ANOVA was conducted with factor being length of hearing loss and noise in AO condition and the dependent variable being with the percent correct. The results for ANOVA indicated a significant noise effect [$F$ (2, 20) =10.88, $p$<.001. There was no significant interaction between length of hearing loss and noise [$F$ (4, 20) = .38, $p$ =.37]. Paired-samples t tests were conducted to evaluate the differences among noise, using the Holm's sequential Bonferroni method to control for type I error. A significant difference was measured between quiet and 2-talker babble, and quiet and reversed-2 talker babble. Performance of CI users was negatively influenced by noise.

A Two-way ANOVA was conducted with factor being length of hearing loss and noise in AV condition and the dependent variable being with the percent correct. There was a significant length of hearing loss effect [$F$ (2, 10) = 5.07, $p$<.05] and noise effect [$F$ (2, 20) =4.35, $p$<.05]. There was no significant interaction between length of hearing loss and noise [$F$ (4, 20) =1.71, $p$=.19]. Paired-samples t tests were conducted to evaluate the differences among length of hearing loss and noise, using the Holm's sequential Bonferroni method to control for type I error. Performance in quiet was significantly higher than reversed 2-talker babble. A significant difference was measured between medium-HL and long-HL groups, and short-HL and long-HL groups meaning when the length of hearing loss was longer than 5 years, the performance in AV condition worsened.

### 3.3. The relationship between duration of CI use and length of hearing loss

Although each relationship between duration of CI use and speech perception in noise, and length of hearing loss and speech perception in noise were analyzed, the relationship between duration of CI use and length of hearing loss is described in this section, as shown in Figure 3.

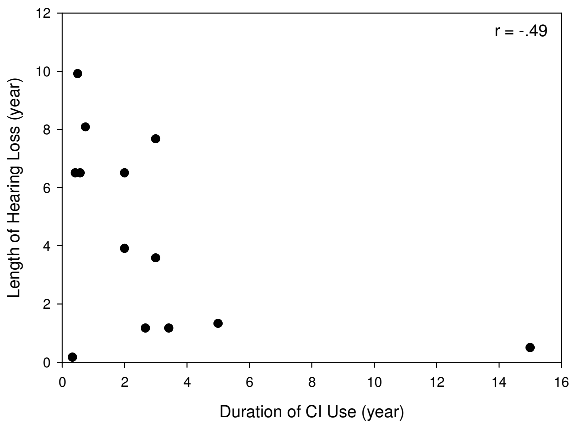


Figure 3. The duration of CI use and length of hearing loss for individual participants.

Pearson Product-Moment Correlation Coefficient was calculated. The results indicated that the correlation between duration of CI use and length of hearing loss was not significant [$r$ (1) =-.49, $p$>.05]. Therefore, we assumed that 2 factors can be analyzed independently.

## 4. Discussion

The present study was designed to investigate AV benefit in CI users for speech comprehension in noise. Statistical analysis indicated that performance of long-CI group was significantly higher than other 2 groups. In AO condition, CI users were significantly influenced by noise. In long-CI and medium-CI groups, performance was similar in all noise conditions in AV condition.

Long-term CI users performed better in quiet condition but also in reversed-2talker babble indicating long-term CI users tend to have resistance to noise. In other words, short-term CI users such as 1 to 3 years tend to be disturbed by noise in AV condition. The resistance to noise may be developed in about 4 years of CI use for AV speech perception. There was a marginal interaction between duration of CI use and noise in

AO condition [$F$ (4, 20) = 2.63, $p$ =.07]. Therefore, by increasing the number of participants, we may say that the resistance to noise is developed in about 8 years for AO speech perception (Figure 1). Moreover, statistical analysis indicated that performance of long-HL group were significantly lower than that of medium-HL group indicating when the length of hearing loss longer than 5 years, the performance in AV condition worsened.

Visual cues play an important role for their speech perception. However, surprisingly, an individual score in VO condition varied in this study (Figure 4).

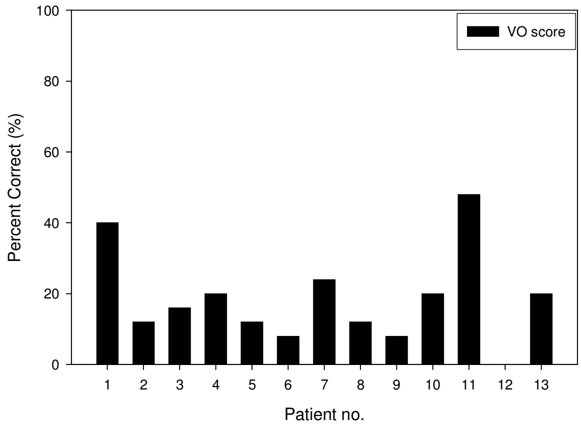


Figure 4. Distribution of correct responses in VO condition for individual patient in the order of longer duration of CI use.

CI users with higher AV score in noise were not necessary to have higher VO scores and neither were CI users with poor AO score. VO scores were not related to either length of hearing loss or duration of CI use. Visual information has a great combined effect with auditory information rather than an effect by itself, and the effect improves auditory-visual speech perception in noise compare to AO condition.

## 5. Conclusion

As from the results obtained in this study, we may conclude that the auditory speech perception in CI users is dramatically influenced by noise. When CI users do not have access to visual cues, their speech understanding ability is disturbed in noise. As the duration of CI use increases, we expect the better performance in speech perception. Specifically, the resistance to noise may be developed in 4 years of CI use for AV speech perception and perhaps in about 8 years for AO speech perception. When hearing loss is less than 5 years, we expect the better performance in speech perception thus, CI surgery may be recommended for hearing impaired listeners whose hearing loss is less than 5 years. We should encourage CI beginners to continue using CI for at least 4 years. We would like to continue this study by increasing the number of research participants and to recruit normal hearing listeners to compare the effects of noise with CI users.

## 6. Acknowledgements

This study was supported by a Grant-in-Aid for scientific Research from Japan Society for the Promotion of Science (21243049) to Dr. Kaoru Sekiyama.

## 7. References


[1] Fu, Q-J., Shannon, R. V. and Wang, X., "Effects of noise and spectral resolution on vowel and consonant recognition: acoustic and electronic hearing", Journal of the Acoustical Society of America, 104: 3586-3596, 1998.

[2] Summers, V. and Molis, M., "Speech recognition in fluctuating and continuous maskers: Effects of hearing loss and presentation level", Journal of Speech, Language, and Hearing Research, 47: 245-256, 2004.

[3] do Nascimento, L. and Bevilacqua, M., "Evaluation of speech perception in noise in cochlear implanted adults", Brazilian Journal of Otorhinolaryngology, 71; 432-438, 2005.

[4] Stickney G., Assman P., Chang J, and Zeng F-G., "Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences", Journal of the Acoustical Society of America, 122: 1069–1078, 2007.

[5] Bergeson, T., Pisoni, D., and Davis, R., "Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants", Ear and Hearing, 26: 149-164, 2005.

[6] Fetterman, B. and Domico, E., "Speech recognition in background noise of cochlear implant patients", Otolaryngology-Head and Neck Surgery, 126, 257–263, 2002.

[7] Luntz, M., Shpak, T., and Weiss, H., "Binaural-bimodal hearing: Concomitant use of a unilateral cochlear implant and a contralateral hearing aid", Acta Oto-Laryngologica, 125: 863-869, 2005.

[8] Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., and Barone, P., "Evidence that cochlear-implanted deaf patients are better multisensory integrators", PNAS, 104: 7295-7300, 2007.

[9] Desai, S., Stickney, G., Zeng, F-G., "Auditory-visual speech perception in normal-hearing and cochlear-implant listeners", Journal of the Acoustical Society of America, 123: 428-440, 2008.

[10] Schorr, E., Fox, N., van Wassenhove, V., and Knudsen, E., "Auditory-visual fusion in speech perception in with cochlear implants", PNAS, 102: 18748-18750, 2005.

[11] Doucet, M. Bergeson, F., Lassonde, M., Ferron, P., and Lepore, F., "Cross-modal reorganization and speech perception in cochlear implant users", Brain, 129: 3376-3383, 2006.

[12] Cullington, H. and Zeng, F-G., "Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simultaneous subjects", Journal of the Acoustical Society of America, 123: 450-461, 2008.

[13] Festen, J. and Plomp, R., "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing", Journal of the Acoustical Society of America, 88: 1725-1736, 1990.