

CLEAR AND CONVERSATION SPEECH PERCEPTION IN SIMULATED COCHLEAR IMPLANT AND SIMULATED ELECTRO ACOUSTIC STIMULATION

¹Ramandeep Kaur, ²Ganesh A.C, & ³Subba Rao T.A

Abstract

Perception of speech is an important goal for persons with Cochlear Implant (CI) and Electro Acoustic Stimulation (EAS). It has been argued that naturally produced clear speech is more intelligible to such persons than conversational speech. The factors responsible for clear speech advantage are increased voice intensity slower rate, frequent pauses, greater vowel duration and dynamic formant movements. The relative advantage of clear speech in CI and EAS are not yet well understood. The present study used 20 native Hindi Speaking (age range 20-30, years) participants on a listening task. The stimuli were 90 previously selected Hindi sentences processed with a noise band vocoder implemented in Matlab. The sentences were spoken by a native Hindi speaker which was recorded in both Clear and Conversational speaking styles & mixed with four talker babble at -4 dB SNR. CI processing simulated using 8-channel noise-band vocoder. The sentences were filtered to 8 bands and the envelopes were extracted from each band. The carrier noise bands were modulated by the envelopes and resynthesized to produce the processed speech. Simulation of EAS was achieved with low pass stimulus (630Hz) and upper five channels of the eight-channel vocoder. Verbatim responses of the subjects were recorded and scored for accuracy. The overall results indicated better understanding in clear speaking style than in conversational style, in simulated CI condition as well as simulated EAS condition. The data was consistent with previous studies. Clear speech had an advantage in improving speech perception whenever there were fewer cues for speech perception due to noise in CI and EAS. The details of comparison of CI and EAS conditions will be discussed.

The style in which we communicate everyday is referred to as 'conversational speech'. However, a talker may naturally adopt a distinct intelligibility enhancing style of speech production called 'clear speech' when they are aware of a speech perception difficulty that will occur when communicating in the presence of background noise, with a person with hearing impairment or in a different native language. Various perceptual studies on clear speech have shown significant improvements in intelligibility over ordinary conversational speech (Picheny, Durlach, & Braida, 1985; Gagné, Masterson, Munhall, Bilida & Querengesser, 1994; Helfer, 1998; Bradlow & Bent, 2002; Krause & Braida, 2002; Bradlow, Kraus & Hayes, 2003). Previous studies have justified clear speech advantage to be 15-20 percentage points in terms of intelligibility when compared to the conversational speaking style for various listener populations, speech materials and listening conditions (Picheny et al., 1985; Gagne et al., 1994; Helfer, 1998; Bradlow & Bent, 2002; Krause & Braida, 2002; Bradlow et al., 2003). Also Picheny, Durlach, & Braida (1986) reported vocal intensity to be 5 to 8 dB greater in clear speech along with a higher and more variable fundamental frequency relative to the conversational style.

Moreover, production of clear speech includes various articulatory/acoustic adjustments. These

include decreased speaking rate comprising of longer and frequent pauses as well as longer segments, greater sound pressure level, increased intensity of fricatives, increased intensity in the 1000 Hz to 3000 Hz range and an expanded vowel range (Picheny et al., 1986, 1989; Krause & Braida, 2004; Bradlow et al., 2003; Liu, Rio, Bradlow, & Zeng, 2004; Moon & Lindblom, 1994; Ferguson & Kewley-Port, 2002; Johnson, Flemming, & Wright, 1993). Bradlow et al. (2003) also revealed less frequent alveolar "flapping" in clear speech.

Additionally, Picheny et al. (1986), Ferguson & Kewley-Port (2002), Moon & Lindblom (1994) have showed an increase in vowel duration and enlarged acoustic vowel spaces (Bradlow et al., 2003). Also vowels have greater dynamic formant movement when spoken clearly. Ferguson et al. (2002) stated a significance of formant movement over vowel nucleus for clear speech.

Clear Speech Advantage in Cochlear implant Users

Apart from clear speech advantage in normal hearing individuals, the cochlear implants (CI), users also demonstrate this advantage. The current generation 'standard-electrode', often used, are capable of providing very high levels of speech understanding (in quiet) to many patients. However, they have tremendous

¹Lecturer, Father Muller Charitable Institution, Mangalore, ²Lecturer, Dr M.V. Shetty College of Speech and Hearing (MVSCSH), Maladi Courts, Kavoor, Mangalore-15, Email: gani0017@gmail.com, ³Prof. & Principal, MVSCSH, Mangalore-15.

difficulty in understanding speech in noise (Friesen, Shannon, Baskent & Wang, 2001; Nelson, Jin, Carney & Nelson, 2003). Normal-hearing listeners show an advantage for speech understanding in competing noise, by utilizing various auditory cues (pitch, timing and localization cues) to separate the multiple sound sources and thereby focus the target speech (Duquesnoy, 1983). CI listeners, on the other hand, are presumably unable to perceive some of these cues, making the competing talker situations particularly difficult for them to segregate the various talkers. In addition, the fluctuating and irregular nature of the competing talker background allows normal-hearing listeners to potentially listen to the target speech in the temporal and spectral 'dips' of the background. However, implant users with temporal acuity deficits reveal poor frequency resolution which worsens their ability to listen in spectral 'dips'.

Nelson et al. (2003) and Qin & Oxenham (2003) demonstrated reduced frequency resolution, in fluctuating backgrounds, to produce diminished speech recognition in normal-hearing listeners similar to cochlear implant users. Thus, the primary deficit of CI listeners in resisting the effects of competing background signals, when listening to speech, appears to be related to poor frequency resolution. Frequency resolution of normal-hearing, hearing-impaired and CI listeners was compared by Henry, Turner & Behrens (2005) in a task requiring discrimination of spectral peaks in a broadband stimulus. It was observed that Normal-hearing listeners had the best frequency resolution, which were followed by individuals with sensori-neural hearing loss, and the poorest resolution was observed in CI users.

Iverson & Bradlow (2002) reported a small clear speech advantage (5%), in quiet, for CI users and normal subjects listening to simulations. However for both the groups, a larger clear speech advantage (15%) was seen when speech was presented in Noise. This difference was seen as the clear speech, primarily, enhances the broadband amplitude modulation and not the spectral differences.

A similar study by Liu et al. (2004) compared a clear speech perceptual benefit in normal hearing adults and adults with CI. Both the groups derived a significant clear speech advantage, although listeners using CI needed somewhat better signal-to noise ratios in order to perform at the same level as normal hearing adults. Also, a high degree of individual variability was demonstrated in clear speech perception for the CI users.

Electro-Acoustic Simulations

Cochlear implantation, which was earlier recommended to only severe to profound hearing loss individuals, though recently it has become an option even for comparatively less severe (mild to moderate) hearing losses, especially with a residual hearing at the low frequencies. The implant manufacturers integrate in one device (E.g. Med-El DUET) the bimodal stimulation, which processes low frequencies of the message using an acoustic unit (classical BTE hearing aid), while high frequencies are processed by another unit which sends an electrical message (implanted electrodes).

Residual hearings, in low frequencies, provide a significant advantage for understanding speech in background noise (Turner, Gantz, Vidal, Behrens & Henry, 2004). Electro-Acoustic simulations, thus, combine the functions of CI for processing the high frequencies and hearing aid unit to amplify the low frequencies (250 to 1500 Hz). Seldran, Truy, Gallego, Berger-Vachon, Collet & Thai-Van (2008) evaluated the number of electrical channels necessary to restore the lack of speech intelligibility of a hearing impaired patient implanted with EAS. It was reported that the speech understanding was restored (>90%) with 1 channel for residual hearing till 1400Hz, 2 channel for residual hearing till 100 Hz, 3 channels for residual hearing till 700 Hz and 4 channels with residual hearing until 500 Hz. The scores of monosyllabic word recognition enhance up to 50-75% with EAS when compared to the conventional CI (Von Ilberg, Kiefer, Tillein, Pfenningdorff, Hartmann & Sturzebeche, 1999; Keifer, Pok, Adunka, Sturzebecher, Baumgartner & Schmidt et al., 2005).

The literature reveals a substantial body of research works on acoustical and perceptual characteristics of clear speech, in individuals with CI users. In Indian context, clear speech has been focused for assessing the production and perception, along with similar clear speech advantages for languages like Kannada and Indian accented English (Kumar & Kumar, 2008; Prabhu, 2009). However, insufficient data is available on perceptual characteristics of clear speech and its advantage over conversational style, in EAS condition. So the present study, thus, compares the differences in speech perception scores in clear and conversational speaking styles among Normal hearing individuals using acoustic simulations of CI and EAS. The aims of the at measuring the speech perception scores in clear and conversational speaking styles in normal hearing individuals using acoustic simulations of CI and EAS.

Method

Participants

Twenty native Hindi speakers in the age range of 20 to 30 years (mean 23 years), who had learnt this language from primary school level, participated in the study. The participants reported of no history of speech, language, or neurological problems. All the participants had their pure tone hearing thresholds ≤ 15 dB HL at octave frequencies in 250 Hz to 8000 Hz range.

Stimuli

The stimuli consisted of 90 standardized Hindi sentences, selected from the speech perception in noise test for children and adults (Kumar, 2008). The selected sentences were recorded by a 23 year old female audiologist who was a native speaker of Hindi and had an extensive experience of communicating with individuals with hearing impairment. In the first recording session, she was instructed to read the sentence list using her normal, conversational speaking style. At a second recording session on the following day, she was instructed to read the same sentences, as though talking to a hearing-impaired, pronouncing each word clearly and carefully. The talker rehearsed each style prior to recording and the experimenter monitored her productions for errors during recording. The stimuli were recorded in a sound treated room using a digital sound stereo headphone (SSD-HP-202) microphone using the Praat software (Boersma & Weenink, 2008) at a sampling rate of 44100 Hz.

Signal Processing

The recorded sentences were mixed with four-talker babble at -4 dB SNR. For the CI stimulations, the recorded speech stimuli were processed using a Noise Vocoder Signal processing was performed using MATLAB (Math works, Natick, MA). The electric stimulation was simulated with an 8-band Vocoder simulation (Shannon et al., 1995). The number of processing bands was selected based on the observation that 8-band Vocoder produces performance levels most similar to that of CI users (Friesen et al., 2001). The entire spectral range of the Vocoder processing was limited to 80 to 6000 Hz (Figure 1). The cut-off frequencies of the individual spectral bands were determined by first converting the lower and higher spectral limits in Hz into corresponding cochlear distances in mm, using the Greenwood mapping function (Greenwood, 1990), then dividing the entire cochlear length into equal distances, and finally converting the ranges in mm back to corresponding frequencies in Hz. These cutoff frequencies were determined separately for the analysis and synthesis filters. The envelopes were extracted from the output of the analysis

filters by half-wave rectification followed by a low-pass Butterworth filter (-18 dB/oct) with a cut-off frequency of 160 Hz. Filtering white noise with the synthesis filters produced the carrier noise bands. The noise carrier in each synthesis band was modulated with the envelope extracted from the corresponding analysis band. The processed speech was the sum of the modulated noise bands from all Vocoder bands.

In speech processing of Electro-Acoustic simulation strategy, the Vocoder processing and Low Pass Filter (LPF) speech was combined. In EAS map the LPF speech replaced the lower synthesis filters of the Vocoder (with the cut-offs of 630Hz). The remaining synthesis filters (5 bands) represented the limited stimulation range of electric hearing.

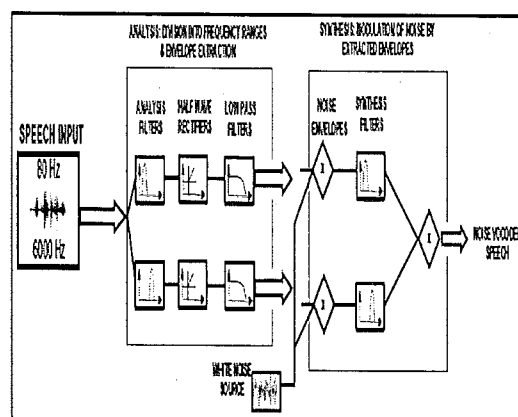


Figure 1: Representation of the signal processing of the stimuli

Procedure

A total of 80 sentences (40 in CI and 40 in EAS condition) and 10 practice trials (5 in each condition) were presented to the participants, in a sound treated room. The processed stimuli in both clear and conversational styles were mixed and presented in each simulated conditions. After presentation in each condition, the participants were released for an interval of 2 hours, in order to avoid subject bias due to fatigue. The stimulus presentation and response acquisition was controlled via 'DMDX' software, using high fidelity Tech-Com Digital Sound stereo headphones (SSD-HP 202) at the level of 50-70 dB SPL. A personal computer, HP Pavilion dv6000 was used for the same. The participants were required to repeat each word verbatim. These responses were recorded and analyzed, in each condition, for the number of correctly repeated key words. A score of '1' was given for correct repetition of each word and '0' for the incorrect response. To obtain the speech recognition scores, the total number of correctly repeated key words was summed.

Results

The comparison of speech recognition scores across clear and conversational speaking styles for the simulated CI and EAS condition has shown in figure 2 and table 1. The Clear speech advantage is clearly indicated in both the conditions relative to the conversational style. Also, the speech recognition scores suggest a better understanding in noise for simulated EAS condition compared to the simulated CI condition

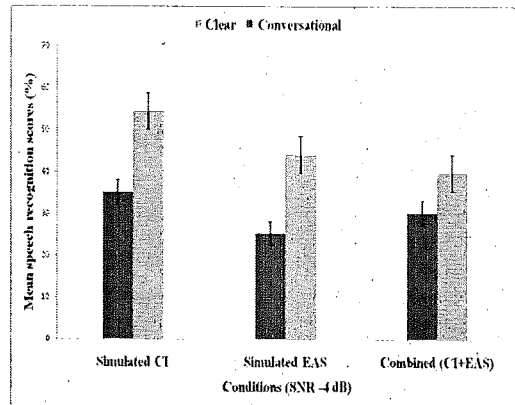


Figure 2: Speech recognition scores for Clear and Conversational speaking styles in each of the simulated conditions, namely Conventional CI and EAS.

Table 1: Speech recognition scores for Clear and Conversational speaking styles in each of the simulated conditions

Conditions	Speaking style	Mean speech recognition scores (%)	Std. Deviation
Simulated CI	Clear	35.20	4.29
	Conversational	25.39	3.78
	Total	30.29	6.37
Simulated EAS	Clear	54.46	2.41
	Conversational	44.21	2.94
	Total	39.81	11.34

Within group comparison

Simulated CI condition

The speech recognition scores for sentences presented in simulated CI condition obtained a mean percentage of 25.3% in conversational style and 35.2% in clear speaking style. Thereby, indicating better scores for clear speech relative to the conversational speech.

Simulated EAS condition

Similar to the above results, speech recognition scores for sentences presented in EAS condition attained mean percentage scores of 44.21% in conversational style and 54.46% in clear speaking style. The scores yet again indicate better speech recognition for clear speech.

Across group comparison

As evident from the figure 2, the speech recognition scores obtained in the simulated EAS condition are much higher relative to the simulated CI condition. However, both the groups showed clear speech advantage over the conversational speaking style. The statistical analyses, a two way ANOVA [$F(1, 76) = 612.86, p < .001$] showed a significant main effect on speech recognition scores across simulated CI and EAS conditions. This indicates a significantly better speech perception for sentences presented in simulated EAS condition. Moreover, similar main effect was seen for clear and conversational styles irrespective of processing conditions [$F(1, 76) = 170.05, p < 0.001$] which reveals clear speech advantage. Furthermore, Paired Sample t-test conducted between clear and conversation styles in each condition revealed significant difference between the mean values ($t = 7.65, p < 0.001$) for simulated CI and ($t = 12.03, p < 0.001$) EAS conditions.

Thereby, the overall results indicate that listeners recognized words more accurately in clear speaking style than in conversational speech, not only in simulated CI condition but also in simulated EAS condition.

Discussion

Various studies in the literature confirm the presence of clear speech advantage over conversational speaking style in normal hearing individuals (Krause & Braida, 2002; Bradlow et al., 2003; Ferguson et al., 2002). Also similar advantage in clear speech is observed in Cochlear Implant condition as well (Iverson & Bradlow, 2002; Liu et al., 2004). The present study compared speech recognition scores for stimuli presented under degraded condition across clear and conversational styles in simulated CI and simulated EAS conditions. The speech recognition scores were obtained for native speakers of the language. The present study reports clear speech advantage over the conversational style in both simulated CI condition and EAS conditions.

Moreover, the CI users show a small clear speech advantage (5%) in quiet and a greater advantage (15%) in degraded conditions (Iverson & Bradlow, 2002; Liu et al., 2004). One of the reasons for better Clear speech perception are expansion of the Vowel space which could also be associated with an intelligibility advantage (Bradlow, 2002; Ferguson & Kewley-Port, 2002; Prabhu, 2009) and slower speaking rate and larger temporal modulation (Liu et al., 2004). Following the same trend the results of the

present study suggest that naturally produced clear speech is an effective way of enhancing speech perception under adverse speaking conditions.

Sufficient amount of research work recommends EAS as an improved and more natural form of auditory rehabilitation when compared to the traditional CI, pertaining to preserve the residual hearing and further amplifying the frequencies which are at loss (Seldran et al., 2008, Gstoettner et al., 2006). Owing to the same reasons, the across group comparison of the present study confirms EAS as a better stimulation strategy compared to the traditional CI strategies, primarily in adverse listening conditions. The clear advantage in EAS mainly due to the availability of low frequency information and fine structure cues that will be available with other cues like expansion of vowel space, slow speaking rate, larger temporal modification as in CI.

Cochlear implants reduce spectral resolution while mostly leaving broadband amplitude modulation intact. Adding noise reduces amplitude modulation within the frequency bands of cochlear implant processors. The recognition of clear speech seems to be less affected by this reduction in amplitude modulation (Iverson & Bradlow, 2002). Furthermore Mathew, Kumar & Alexander (2010) indicate clear speech to facilitate a better and faster learning of foreign language words. Therefore, the clear speaking style is beneficial for training young children to learn a new language and also rehabilitation of individuals with communication disorders. The present study further extends this view towards the rehabilitation of individuals with hearing impairment which could be improved with the help of clear speech usage not only for children using CI but also with EAS users. As there are numerous differences between acoustic and electric hearing, one should be very careful about applying the results from simulations with normal-hearing listeners to actual implant users.

Conclusions

The present study aimed at comparison of speech perception scores across clear and conversational speaking situations in individuals with Normal hearing, simulated with CI and EAS. The study concludes a clear speech advantage in improving speech perception whenever there are fewer cues available for speech recognition due to noise, in both CI and EAS strategies. In addition, it plays an important role in learning of a new language, especially in pediatric Cochlear Implant group. Moreover, a critical contribution of the clear speech advantage was reported in the presence of

reduced spectral cues in noise. Recommendations for the future research include generalization of the present results into to the clinical population.

References

- Boersma, P., & Weenink, D. (2008). Praat: Doing Phonetics by computer (Version 5.0.17) [Computer Program]. Retrieved from <http://www.praat.org/>
- Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America*, 112, 272-284. Doi: 10.1121/1.1487837
- Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for learning-impaired children: Sentence perception in noise. *Journal of Speech, Language and Hearing Research*, 46, 80-97. Doi: 10.1044/1092-4388(2003/007)
- Duquesnoy, A.J., (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *Journal of the Acoustical Society of America*, 74, 739-743. Doi: 10.1121/1.389859
- Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing impaired listeners. *Journal of the Acoustical Society of America*, 112, 259-271. Doi: 10.1121/1.1482078
- Friesen, L.M., Shannon, R.V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *Journal of the Acoustical Society of America*, 110, 1150-1163. Doi: 10.1121/1.1381538
- Gagné, J.-P., Masterson, V. M., Munhall, K. G., Bilida, N., & Querengesser, C. (1994). Across talker variability in auditory, visual, and audiovisual speech intelligibility for conversational and clear speech. *Journal of the Academy of Rehabilitative Audiology*, 27, 135-158.
- Helfer, K. S. (1998). Auditory and auditory-visual recognition of clear and conversational speech by older adults. *Journal of the American Academy of Audiology*, 9, 234-242.
- Henry, B.A., Turner, C.W., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired and cochlear implant listeners. *The Journal of the Acoustical Society of America*, 118, 1111-1121. Doi: 10.1121/1.1944567
- Iverson, P., & Bradlow, A. (2002). The recognition of clear speech by adult cochlear implant users. In S. Hawkins & N. Nguyen (Eds.), *Temporal Integration in the Perception of Speech* (p. 78). Cambridge, UK: Centre for Research in the Arts, Social Sciences, and Humanities.
- Johnson, K., Flemming, E., & Wright, R. (1993). The hyperspace effect: Phonetic targets are hyper articulated. *Language*, 69, 505-528.

- Kiefer, J., Pok, M., Adunka, O., Sturzebecher, E., Baumgartner, W., Schmidt, M., et al. (2005). Combined electric and acoustic stimulation (EAS) of the auditory system: Results of a clinical study. *Audiology and Neurotology*, 10(3), 134-44. Doi: 10.1159/000084023
- Krause, J. C., & Braida, L. D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *The Journal of the Acoustical Society of America*, 112, 2165-2172. Doi: 10.1121/1.1509432.
- Kumar P. (2008). *Comparing Acoustic Parameters of Conversational and Clear Speech in Kannada*. Unpublished Master Dissertation, Mangalore University, Mangalore, India.
- Liu, S., Del Rio, E., Bradlow, A.R., & Zeng, F.G. (2004). Clear speech perception in acoustic and electric hearing. *Journal of the Acoustical Society of America*, 116, 2374-2383. DOI: 10.1121/1.1787528
- Liu, C., & Kewley-Port, D. (2004). Vowel formant discrimination for high-fidelity speech. *The Journal of the Acoustical Society of America*, 116, 1224-1233. Doi: 10.1121/1.1768958.
- Mathew, K., Alexander, S.A., & Kumar, A.U. (2010). *Foreign Word Learning: Effect of Training in Clear & Conversational Speaking Styles*. Poster presented at NYSSLHA's 50th annual convention. New York
- Moon, S. J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America*, 96, 40-55. DOI: 10.1121/1.410492
- Nelson, P., Jin, S-H., Carney, A., & Nelson, D. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, 113, 961-968. Doi: 10.1121/1.1531983
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of speech and hearing research*, 28, 96-103.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, 29, 434-446.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1989). Speaking clearly for the hard of hearing. III: An attempt to determine the contribution of speaking rate to difference in intelligibility between clear and conversational speech. *Journal of Speech and Hearing Research*, 32, 600-603.
- Prabhu, S. (2009). *Acoustic characteristics of vowels in clear & conversational speaking style in Kannada*. Unpublished Master Dissertation, Mangalore University, Mangalore, India.
- Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *Journal of the Acoustical Society of America*, 114, 446-454. Doi: 10.1121/1.1579009
- Seldran F, Truy E, Gallego S, Berger-Vachon C, Collet L, Thai-Van H (2008): Simulation of an electro-acoustic implant (EAS) with a hybrid vocoder. Retrieved from <http://www.Acoustics-o8-paris.org/>
- Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., & Henry, B. A. (2004). Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing. *Journal of the Acoustical Society of America*, 115, 1729-1735. Doi: 10.1121/1.1687425
- Von Ilberg, C., Kiefer, J., Tillein, J., Pfenningdorff, T., Hartmann, R., Sturzebecher, E., et al. (1999). Electric-acoustic stimulation of the auditory system: New technology for severe hearing loss. *ORL-Journal for Oto-Rhyno-Laryngology and its Related Specialties*, 61, 334-340.