Effect of Stimulation Rate on Speech Perception in Individuals using Nucleus Freedom Cochlear Implant

Vinaya K.C. Manchaiah[1], Julie Eyles[2] and Carl Verschuur[3]

Abstract

The aim of the present study was to investigate the effect of high electrical stimulation rates in cochlear implants (CI) on speech perception. Four participants using Nucleus Freedom CI and five adult participants with normal hearing performed a vowel consonant vowel (VCV) consonant (iCi) recognition task in quiet and a (Bamford-Kowal-Bench) BKB sentence recognition task in noise at +10 dB signal to noise ratio (SNR). The performance of the participants was measured for three stimulation rates, namely 1800 pulses per second per channel (pps/ch), 2400 pps/ch and 3500 pps/ch, with the Advanced Combination Encoder Revised i.e. ACE (RE) speech coding strategy. Participants with normal hearing were also studied to see the effect of overlap on CI acoustic modelling. The speech perception results of the participants using the Nucleus Freedom system were compared with those of participants with normal hearing on CI acoustic modelling. The VCV test scores were further analyzed for feature transmission errors in place and manner of articulation as well as in voicing features. The results showed that there was no advantage of high stimulation rates for both consonant recognition in quiet and sentence recognition in noise. The CI acoustic models were found to be useful in studying the effect of rates in those using a CI. Although CI acoustic model with overlap lead to poor performance in consonant recognition in quiet and sentence recognition in noise test, they seemed to be the better approximation of the scores among the users of CI.

Key words : *Cochlear implant, signal processing, stimulation rate, acoustic modelling, speech perception, temporal information.*

The design of the cochlear implant (CI) has developed tremendously over the years. Advancements in signal processing have improved speech processing in CIs. The current generation CIs allow manipulation of a range of processing parameters to enable optimization of the user's auditory perception (Verschuur, 2005). One such parameter which has received particular attention in recent years is electrical stimulation rate (Fu & Shannon, 2000). High stimulation rates in CIs are considered to improve temporal information, which is very useful in speech understanding. According to the basic sampling theory, the stimulation rate must be at least twice the highest frequency in the speech spectrum to convey a high fidelity signal by reducing aliasing effects (Wilson, Finley, Lawson, & Zerbi, 1997). It is clear from previous research that temporal cues play an important part in speech recognition in the absence of spectral cues, particularly in those using a CI. Studies undertaken to study the physiological effect of electrical stimulation rate suggest that "medium stimulation rates cause alteration in neural response and this is due to a complex interaction between electrical activation and neural refractory properties" (Friesen, Shannon, & Cruz, 2005, p.172). According to Wilson et al. (1997) these problems exist because low stimulation rates give rise to synchronous neural firing. The neurons get sufficient time to recover from the excitatory period and they enter the recovery cycle simultaneously. At high stimulation rates, the time between pulses is too short and none of the neurons can recover sufficiently to fire with the next pulse. However, during medium stimulation rates some neurons do recover to fire for the next pulse and some do not, which gives rise to signal aliasing. These alterations disappear only with very high stimulation rates above 2000 to 3000 pps/ch. Rubinstein, Wilson, Finely, and Abbas (1999) proposed that high stimulation rates restore the normal stochastic properties in the neural response.

Although there are a vast numbers of studies that have been conducted to examine the effects of stimulation rate on speech perception, the results are equivocal. This shows that the perceptual effects of electrical stimulation rates are not very well understood. There are also controversial results obtained from different groups of researchers (Friesen et al., 2005). One

---

1. *School of Health Science, Swansea University, United Kingdom.*
2. *South of England Cochlear Implant Centre, University of Southampton, United Kingdom.*
3. *ISVR Hearing & Balance Centre, University of Southampton, United Kingdom.*

group of researchers supports the claim that high stimulation rates improve speech perception (Au, 2003; Holden, Skinner, Holden & Demorest, 2002; Kiefer et al., 2000; Loizou, Poroy & Dorman, 2000) while other groups have shown results contradictory to this (Friesen, Shannon & Cruz, 2005; Fu & Shannon, 2000). A study by Nie, Barco and Zeng (2006) suggests that even though higher stimulation rates of up to 2000 Hz benefit speech perception, further increase in rate to 4000 Hz might lead to poor performance possibly due to channel interaction which can occur at high rates. However, in most of the studies, individual variation is seen on the rate effect. Many studies support the theory that high stimulation rates are useful for a few individuals, but not for all users of CI (Buchner, Frohne-Buchner, Battmer, & Lenarz, 2004; Psarros, Plant, Lee, Decker, Whitford, & Cowan 2002; Vandali, Whitford, Plant, & Clark, 2000; Verschuur, 2005). According to Fu and Shannon (2000), there are some individual differences seen in the effect of stimulation rate, but not all individuals are able to utilize better temporal cues resulting from higher stimulation rates. The exact reason for these individual variations and also variations in group mean scores is not very well understood. The difference in results might be due to methodological differences across studies such as variation in stimulation rates used, the speech perception tests used methods of testing, time given after mapping for acclimatization, baseline speech perception abilities of the subjects.

One method of studying the effect of rate on CI speech perception, is by using CI acoustic models. CI acoustic models are prepared by processing acoustic signals using signal processing techniques that are similar to those used in CI speech processors. These are also sometimes referred to as “vocoded” signals (Faulkner, Rosen, & Smith, 2000). In recent years researchers have used CI acoustic models to evaluate the effects of various parametric variations such as channel number, electrode insertion depth, electrical stimulation rate, etc on CI speech perception (Dorman, Loizou, & Rainey, 1997; Friesen, Shannon, Baskent, & Wang, 2001; Harczos, Katai, Klefenz, & Baljic, 2008).

The main purpose of the present study was to evaluate, if high electrical stimulation rates improved speech perception in those using the Nucleus Freedom CI. The other aim of the study was to evaluate the usefulness of CI acoustic models in studying the effect of electrical stimulation rate on speech perception.

Method

Design Overview

In order to achieve these aims, the following objectives were met. Individuals using Nucleus Freedom CI were recruited and their ability to perceive consonants in quiet and sentences in noise at three different electrical stimulation rates were measured. The same test stimuli (consonants in quiet and sentences in noise) were acoustically modelled using CI signal processing to simulate different stimulation rates. Factors such as the number of channels, type of speech processing strategy, high frequency shift due to insertion depth of electrodes, channel overlap etc. were taken into consideration.

The CI acoustic models were prepared in two ways: one with overlap (to create the effect of channel interaction) and the other without overlap, to analyze which condition best approximates the performance of individuals using a CI. The participants with normal hearing were tested at three different rate conditions simulated through the CI acoustic models.

The results of the speech perception tests on the two groups of participants (participants using CIs and participants with normal hearing on CI acoustic models) were compared.

Participants

Two groups of participants were included in this study. One group consisted of participants from the South of England Cochlear Implant Centre (SOECIC) using the Nucleus Freedom CI and the other consisted of young adults with normal hearing who were students at the University of Southampton. The following inclusion and exclusion criteria were set for recruiting subjects for the two groups.

The participants using CI were adults aged 18 years and over, with post lingual deafness, first language English, using the Nucleus Freedom CI for at least 6 months and having a score of 60 % or more on the BKB sentence test in quiet. Individuals with CI who had ossification or any other cochlear anomaly that might prevent complete insertion of the electrode array; signs of co-existing retro-cochlear or central auditory involvement; additional impairments that would prevent or restrict participation in the test; or any problem during the testing such as unwillingness or inability to comply with all the study requirements, were excluded from the study.

The participants with normal hearing were young adults between the ages of 18 and 30 years with native language English; otologically normal hearing, normal middle and external ears without any history of discharge or infection; normal hearing thresholds (pure tone average less than 20 dB HL); no family history of congenital hearing impairment; and no complaints of previous exposure to noise or oto-toxic drugs. Adults with excessive wax in the ear canals, those who believed that they were unable to familiarize themselves with the simulated speech presented even after many trials, were excluded from the study.

A sample size calculation was performed using the Sample Power software Version 1.2 which calculated that five subjects would be needed for a statistical power of 80%. The data used for this calculation were based on the findings of Vandali et al. (2000). Some demographic information was obtained from the participants using CI. Their full details can be seen in Table 1.

Stimuli

The test stimuli used were nonsense /iCi/ words spoken by a female speaker and Bamford-Kowal-Bench (BKB) sentences spoken by a male speaker. The recording of both /iCi/ words and BKB sentences are available at the website of the Institute of Hearing Research (IHR), (http://www.ihr.mrc.ac.uk/products/tests.php). The Vowel-Consonant-Vowel (VCV) test gave information about the perception of specific cues like place, manner, and voicing, without any contextual and lexical cues. Sentences were used to obtain information on auditory closure ability in the presence of background noise. The /i/ vowel environment in the /iCi/ words was chosen based on the findings of Loizou, Poroy and Dorman (2000). In their experiment they found that the effect of stimulation rate on consonant recognition was highly dependent on the vowel context. They reported a large benefit for /uCu/ or /iCi/ contexts and a small benefit in the /aCa/ context. The consonants used for the VCV test in the present study were: /b/, /ch/, /d/, /f/, /g/, /dz/, /k/, /l/, /m/, /n/, /p/, /r/, /s/, /sh/, /t/, /t/, /h/, /v/, /w/, /y/, and /z/. For the sentence test in noise, stationary

Table 1. *Demographic information of CI users*

Subject	*Age*	*Sex*	*BKB Scores in Quiet*	*Duration of Severe/ Profound Deafness*	*Duration of Implant Use*	*Speech Coding Strategy & Stimulation rate used*	*Other Info.*
MW	51 yrs	F	70 %	10 yrs	11 mths	ACE (RE) – 2400 (10)	using hearing aid in the opposite ear
JP	65 yrs	M	61 % in CUNY with Lip reading	13 yrs	6 mths	ACE (RE) – 2400 (10)	
PD	67 yrs	F	61 %	17 yrs	6 mths	ACE (RE) – 2400 (10)	
MS	65 yrs	F	90 %	22 yrs	8 mths	ACE (RE) – 2400 (10)	

speech shaped noise at a level 10 dB below the level of the speech signal was added to the BKB sentences using the Adobe Audition software. A signal-to-noise ratio (SNR) of +10 dB was chosen to eliminate the ceiling effect caused by contextual and lexical cues in sentence test and also based on the findings of a study by Friesen et al. (2001) which showed that the speech perception advantage due to higher stimulation rates is not visible in tests of speech perception in quiet.

CI acoustic modelling

For the signal processing in CI acoustic modelling, the 22 channel Advanced Combination Encoder-Revised (ACE-RE) speech coding strategy was implemented through the custom interface of a Nucleus Implant Communicator (NIC-Stream) simulation software (developed by Cochlear) and the AMO software (developed by Laneau, Moonen & Wouters, 2006) with the MATLAB software. This interface which was used to generate the CI acoustic models involved steps that mimiced the signal processing used in a CI speech processor.

To generate the model noise bands were used and high frequency pre-emphasis was added. The signal underwent a Frequency Fourier transform (FFT) and was split into filter bands yielding 128 bins. 64 real bins were re-combined into 22 channels using a channel weighting formula. A subset of filter bands was selected according to the number of desired maxima. Modulation was used to vary the noise band amplitude according to the loudness growth function (LGF). Noise bands were produced according to an overlap formula. The noise bands were then added together.

Both /iCi/ words in quiet and BKB sentences in noise were further processed in the following steps to generate the CI acoustic models for participants with normal hearing. The stimuli were first filtered using the Adobe Audition software to simulate the high frequency emphasis that is caused by the microphone characteristics of CIs. The details of the microphone characteristics of the Nucleus Freedom CI device were taken from unpublished experiments by Grasmeder and Verschuur (2003) at ISVR Hearing & Balance Centre. The filter characteristics were: +6 dB per octave up to 1800 Hz, flat from 1800 to 5000 Hz, and -24 dB per octave from 5000 to 10000 Hz and above. The stimuli were passed through a filter that was derived based on the above characteristics, before being passed through the NIC software. BKB sentences were combined with speech shaped noise at a level 10 dB below the speech signals.

The signals were generated by passing the /iCi/ words in quiet and BKB sentences in noise through the software interface, which processes the signals in a similar manner to that used in the CI speech processor. The custom interface of the software provides options to choose speech processing strategy, number of electrodes inserted, stimulation rate, high frequency shift due to incomplete insertion of the electrode array, noise band or sine wave simulation, and overlap of the signals to account for channel interaction to some extent.

All signals were generated with noise bands in the Advanced Combination Encoders (ACE) speech coding strategy with the high frequency shift, at three different stimulation rates, namely 1800 pulses per second per channel (pps/ch), 2400 pps/ch and 3500 pps/ch. Two sets of test stimuli were generated, one with overlap to account for channel interaction and the other without overlap. This was done to determine the condition which best approximates the performance of a participant with a CI. An overlap of filters or space constant of 3.3 mm was used as suggested by the findings from Laneau et al. (2006) to create the effect of masking spread. The experiments by Laneau et al. (2006) showed that the overlap of 3.3 mm (also called 'space constant') in CI acoustic modelling might be ideal to create an effect simulating the spread of masking. Incorporating this overlap in CI acoustic modelling could account for the spectral aspects of channel interaction to some extent.

A calibration noise, equivalent to the root mean square (RMS) sound pressure level (SPL) of all the stimuli, was used to ensure that the sound level was 70 dB (A) at the subject's test ear for the VCV test stimuli. This noise was produced at the Institute of Sound and Vibration Research (ISVR). A calibration noise which had been generated using the automated sentence test (AST) software, for calibration of BKB sentence list, was used to ensure that the sound level was 70 dB (A) at the subject's test ear. This level was selected in order to enable presentation of speech materials at most comfortable levels and also at levels at which the participants might obtain maximum scores. The stimuli were routed through an audiometer. Before commencement of testing, sound field calibration was undertaken at the listener's ear level using a Sound Level Meter (SLM).

Equipment

The VCV test was conducted through a PC using the Praat speech software (www.fon.hum.uva.nl/praat) and the BKB sentence test was conducted through a PC using the AST software. The stimuli were presented using a single loudspeaker and a Sound Blaster via an amplifier.

All the tests were undertaken in a sound treated room. Participants were seated in front of a touch screen on which they responded for VCV test. They were seated at 0 degree azimuth; and at 2.5mtr. from the loud speaker for the VCV test and at 1mtr. distance for the BKB sentence test. The sound level was calibrated differently for VCV and BKB tests to account for the difference in distance. The sound levels were measured using a Kamplex K4 Type 2 SLM.

Test Conditions

Participants using CI performed VCV and BKB tests under three different rate conditions, namely 1800 pps/ch with 10 maxima (20 µS/phase, 6 us IPG), 2400 pps/ch with 10 maxima (12 µS/phase, 6 us IPG) and 3500 pps/ch with 9 maxima (9.6 µS/phase, 4.8 us IPG). The ACE (RE) strategy was used in all three conditions. An experimental Freedom speech processor was used to store the MAPs for the participants using CI. MAPs were created and edited using the Nucleus Custom Sound software. Participants with normal hearing performed VCV and BKB tests under six different conditions, which involved CI acoustic models of 1800 pps/ch, 2400 pps/ch and 3500 pps/ch with overlap and CI acoustic models of 1800 pps/ch, 2400 pps/ch and 3500 pps/ch without overlap. The test conditions were presented randomly and the participants were unaware of which rate they were being tested (single blinded) on. All participants using a CI completed a questionnaire after the completion of testing. This was to obtain some information from the CI users, on preferred rate condition for speech understanding.

Procedure

The experiment was explained in detail to all the participants and informed consent was taken. All participants using a CI underwent a small tuning session before the testing. Before the participants' arrival for testing, MAPs were created for them with three different rates based on their previous MAPs. On arrival their threshold and comfort level was determined for every fourth electrode and the levels for electrodes in between these were extrapolated. They were given approximately 10 minutes to adjust to the new MAPs during which they were involved in conversation and were presented with a practice list of stimuli in order to familiarize them with the new MAPs.

Participants with normal hearing were given an opportunity to listen to a practice list of simulated speech for as long as they wanted until they felt confident with the stimuli. They took an average of 15 to 20 minutes for familiarization with the test materials. All the participants were given 10 minutes break in between testing of different rate conditions. Although they were given the option of a longer break, none of the participants required extra time.

Data Processing

Two different software programmes were used in the presentation of the stimuli, namely the Praat software for VCV test and the AST software for BKB test. Both measures gave percent correct scores for the different stimulation rate conditions. The percent correct scores were compared across the rates in both the measures. VCV test results were also analysed for perception of place, manner and voicing.

The Praat software produced tables of each subject's responses in different conditions. These were exported to Microsoft Excel and rearranged into a format that could be used by the Score Categorization Experiments (SCORE) software developed at the University College of London (UCL). SCORE converted the responses into a matrix which was analysed by the Feature

Information Xfer (FIX) program, also from UCL. The FIX software was then used to analyze the confusion matrix for place, manner and voicing transmission using the Sequential Information Analysis (SINFA) method (Wang & Bilger, 1973). This analysis is based on the feature matrices. The output from this program was the percentage transmission value for each feature which was then entered into the SPSS version 11.5 software for statistical analysis.

Results

Test results of participants using CI

Results of VCV in participants using a CI: Figure 1 displays the VCV word test scores at stimulation rates of 1800 pps/ch, 2400 pps/ch and 3500 pps/ch averaged across three participants using a CI. One out of the four users of CI performed very poorly on consonant recognition in quiet (less than 10%) and therefore the data from that participant was eliminated from the analysis. The higher simulation rates did not seem to improve consonant perception in quiet for participants using CI. The results of the analysis of variance (ANOVA) showed that the effect of stimulation rates did not reach statistical significance [F(2, 4) =0.081, p = 0.923]. Thus the consonant perception in quiet was not affected by stimulation rates.

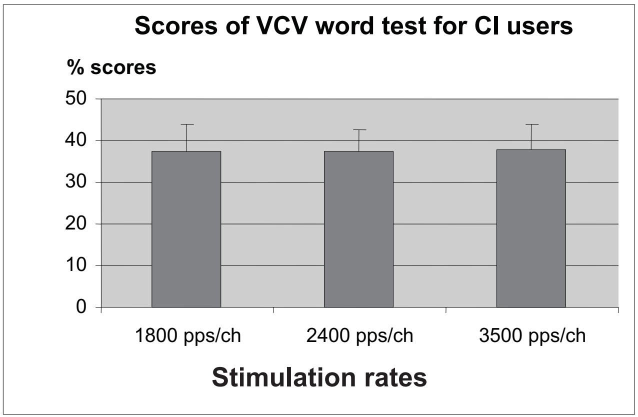

Fig.1. Total scores on the VCV word test for CI users

Figure 2 shows mean scores for feature error types for participants using CI. This shows how much information was transmitted for each of the three features of speech: place of articulation, manner of articulation and voicing. It appears that manner of articulation cues are most well preserved followed by voicing and then place of articulation cues.

The two factor repeated measures ANOVA on features and stimulation rates shows that the difference between features was significant [(F(2, 4)=8.999, p<0.05] but the effect of stimulation rates did not reach statistical significance [F(2, 4) =3.077, p=0.155]. Also, there was a significant interaction between the effect of stimulation rates and perception of different features [F(4,8)= 4.077, p<0.05]. Post-hoc t-tests were performed to investigate the effect of feature and they showed a highly significant difference among all the features. The difference between place and manner (p<0.001), between place and voicing (p<0.001) and between manner and voicing (p<0.001) were significant. Post-hoc t-test was performed among different stimulation rates within the voicing feature. The difference between scores at 1800 pps/ch and 2400 pps/ch (p=0.062), between 1800 pps/ch and 3500 pps/ch (p=0.062) and between 2400 pps/ch and 3500 pps/ch (p=0.052), showed that for the voicing feature, there was no significant difference between scores for different rates.

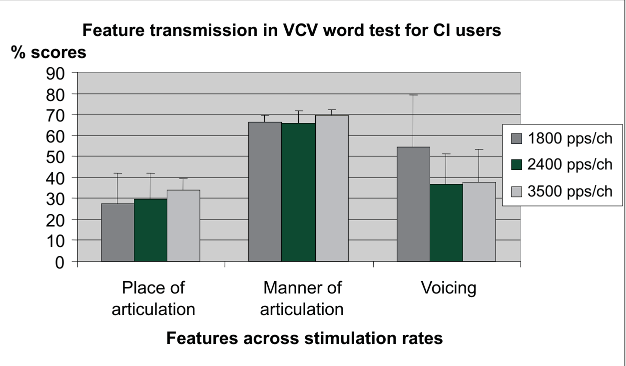

Fig. 2. Feature transmission scores on the VCV word test for CI users

BKB sentence test results of participants using CI
As shown in figure 3 the BKB sentence test scores in noise at the stimulation rates 1800 pps/ch, 2400 pps/ch and 3500 pps/ch averaged across two CI participants were 25 % (SD-7), 37 % (SD-18.3) and 29 % (SD-7) respectively. Two out of four participants using a CI performed very poorly on sentence recognition in noise (less than 10%) was and hence their data were eliminated from the analysis. The repeated measures ANOVA showed no significant difference among

the rates (F(2, 2)=1.750, p=0.364). The statistical test performed may not have been efficient enough to analyse the significance of the effect of rate as the sample was small. However, by comparing with the confidence intervals of 5 % for 50 item speech discrimination tests (Raffin & Thornton,1979) one may deduce that the 16 item speech discrimi-nation test used in this study would need a difference of more than 10 to 15 % across rates to be statistically significant. Thus, it can be assumed that there was no significant interaction between sentence recognition in noise and stimulation rates.

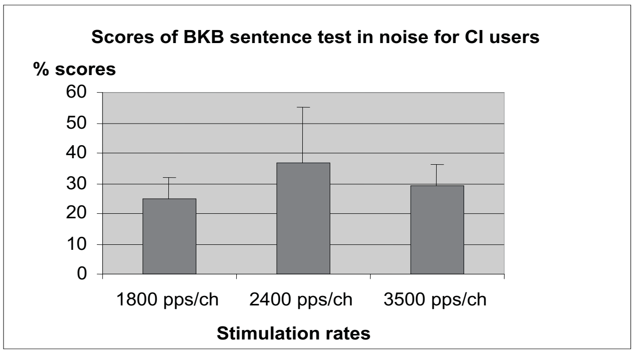

*Fig.3.*Total scores for the BKB sentence test in noise for CI users

The questionnaire used to find the preferred rate programme showed that all four CI participants preferred the MAP with the stimulation rate 2400 pps/ch. The reason behind this might be that all of them were using the MAP with 2400 pps/ch in everyday normal listening conditions. Even after the initial MAPping to adjust their loudness levels in different MAPs, most of them reported that, MAP 3 (3500 pps/ch) was very soft, MAP 1 (1800 pps/ch) was quite loud and MAP 2 was optimal.

Test results of participants with normal hearing on CI acoustic modelling

VCV word test results of participants with normal hearing : The VCV word test scores in quiet at the stimulation rates 1800 pps/ch, 2400 pps/ch and 3500 pps/ch averaged across five participants with normal hearing on CI acoustic modelling with overlap and without overlap are presented in the Figure 4. The higher simulation rates did not seem to improve consonant perception in quiet for participants with normal hearing on CI acoustic models. The ANOVA results showed that the effect of stimulation rates did not reach statistical significance [F(2, 8) =1.201 , p = 0.350]. Thus consonant perception in quiet was not affected by stimulation rates. However, the effect of overlap was highly significant [F(1, 4) =292.74, p < 0.001]. Overlap leads to poor performance on consonant recognition in quiet. The interaction effect between overlap and stimulation rates was not significant [F(2, 8) =1.497 , p = 0.280].

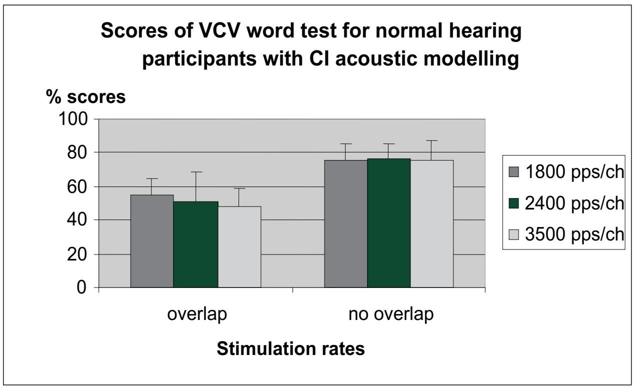

Fig. 4. Total scores for the VCV word test for normal hearing participants with CI acoustic modelling

Figure 5 shows the feature transmission scores for participants with normal hearing on CI acoustic models created with and without overlap. The feature transmission scores followed a pattern that was similar to that of participants using a CI. It can be inferred that manner of articulation cues are most well preserved followed by voicing cues and then place of articulation cues. The repeated measures ANOVA with three factors showed that the effects of overlap [F(1,4)=160.07, p<0.001] and features [F(2,8)=13.79, p<0.05] were significant. The effect of stimulation rates did not reach statistical significance [F(2,8)=0.083, p=0.921]. The interaction effect of overlap and stimulation rate [F(2,8)=0.053, p=0.948] and the interaction effect of stimulation rates and features [F(4,16)=1.611, p=0.220] were also not significant. However, the interaction effect of overlap and features [F(2,8)=5.741, p<0.05] and the combined effect of overlap, stimulation rate and features [F(4,16)=3.845, p<0.05] were significant.

Post-hoc t-tests were performed to investigate the effect of feature. The tests showed a highly significant difference between all the features. The difference between place and manner (p<0.001), between place and voicing (p<0.001) and between manner and voicing (p<0.001) were highly significant. Paired samples t-tests (Post-hoc t-tests) were performed to investigate the effect of overlap on features. The effect of overlap for features of manner (p<0.05) and voicing (p<0.05) were significant. The effect of overlap for the feature of place of articulation did not reach statistical significance (p=0.087).

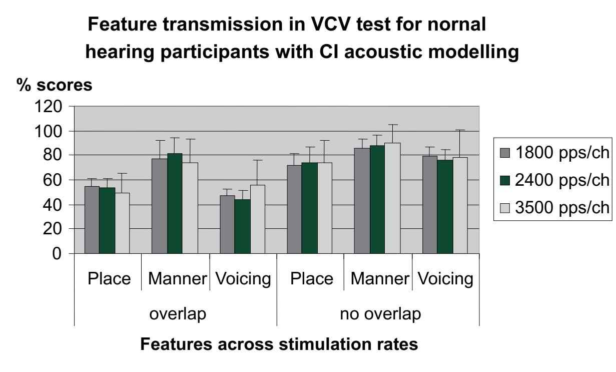

Fig. 5. Feature transmission scores in the VCV word test for normal hearing participants with CI acoustic modelling

<u>BKB sentence test results of participants with normal hearing</u> : Figure 6 shows the BKB sentence test scores in noise at different stimulation rates. These scores were averaged across five participants with normal hearing on CI acoustic modelling in two different conditions (with and without overlap). The ANOVA showed that the effect of simulation rate did not reach statistical significance [F(2,8)=0.577, p=0.583] while the effect of overlap was significant [F(1,4)=104.74, p<0.001]. The interaction effect between stimulation rate and overlap also did not reach statistical significance [F(2,8)=0.539, p = 0.603]. This shows that higher stimulation rates do not lead to better scores and overlap leads to poor sentence recognition scores in participants with normal hearing on CI acoustic modelling.

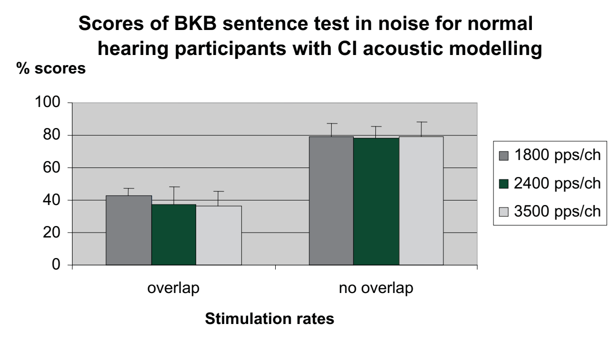

Fig. 6. Total scores for the BKB sentence test in noise for normal hearing subjects with CI acoustic modelling

Comparison of results of participants with CI and participants with normal hearing on acoustic modelling

CI acoustic models with overlap seem to result in closer approximation of the scores of users of CI, compared to models without overlap. Figure 7 shows the mean VCV scores for the users of CI and participants with normal hearing on CI acoustic modelling at different stimulation rates. The two-way repeated measures of ANOVA failed to show a significant effect of groups [F(1,2) =0.892, p=0.445] and stimulation rates [F(2,4) =2.367, p=0.210]. However, there was a significant interaction between subjects and stimulation rates [F(2, 4) =8.403, p<0.05].

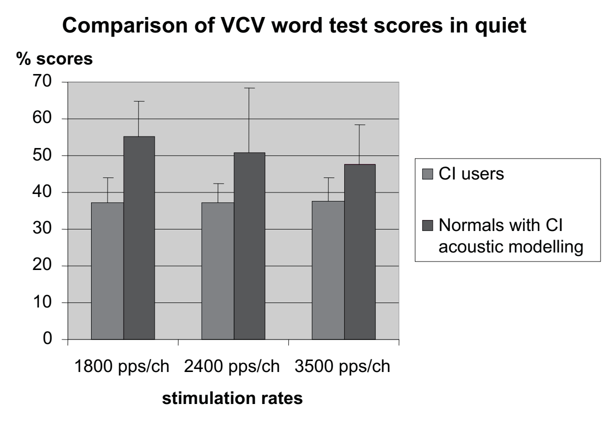

Fig.7. Comparison of scores of VCV word test

The repeated measures ANOVA with three factors showed that the effect of rate [F(2, 4) =10.83, p < 0.05] and features [F(2, 4) =18.15, p<0.05] were significant. The effect of participant groups did not reach statistical significance [F(1,2)=2.241, p=0.273]. The interaction effect between participant groups and stimulation rates [F(2, 4) =0.145, p = 0.869], participant groups and features [F(2, 4) =3.408, p = 0.137], and between stimulation rates and features [F(4, 8) =3.080, p= 0.082].The combined effect of participant group, stimulation rate and feature [F(4, 8)=2.695, p= 0.109] failed to reach the significance level.

As shown in figure 8 the feature transmission scores for participants with normal hearing on CI acoustic modelling followed pattern similar pattern that in users of CI. Manner of articulation cues were better preserved than voicing and place of articulation cues. However, the scores for participants with normal hearing on CI acoustic modelling were higher than those of participants using a CI. The scores for voicing cues at 1800 pps/ch were somewhat higher than scores at 2400 pps/ch and 3500 pps/ch for users of a CI. This might be one of the reasons for having significant effect of rates in features transmission scores. The Post-hoc t-test was performed to test for difference between voicing scores across stimulation rates in those using a CI. The results showed no significant difference among scores for voicing cues at different rates of stimulation. However, the reason for having a significant rate effect might be the very high score for voicing cues in case of one participant using a CI.

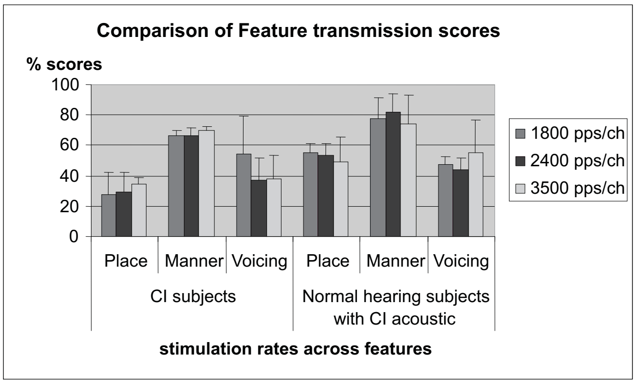

Fig. 8. Comparison of feature transmission scores

The average BKB sentence test scores in noise at different stimulation rates for those using a CI and for participants with normal hearing on CI acoustic modelling are presented in the Figure 9. The two-way repeated measures of ANOVA showed that the effect of simulation rate [F(2, 2) =0.514 , p = 0.661], the effect of participant group [F(1, 1) =0.871, p = 0.552] and also the interaction among participant group and stimulation rates [F(2, 2) =10.90, p = 0.084] failed to reach the significance level. This shows that there is no advantage of higher stimulation rates. The score of participants using a CI and participant with normal hearing on CI acoustic modelling did not vary with rates.

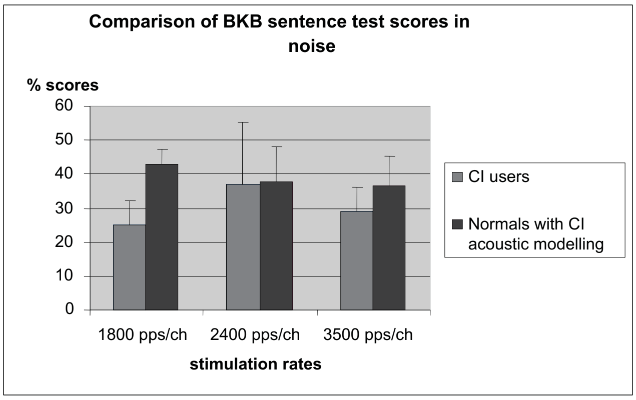

Fig. 9. Comparison of scores BKB sentence test in noise

Discussion

Speech perception in participants using a CI

There is abundant literature speech perception in users of CI; however, the relative importance of spectral and temporal cues is not very well understood. Studies which have attempted to analyze the effect of stimulation rates on speech perception have yielded equivocal results. However, the present study cannot be compared directly with previous studies due to differences in methodological issues such as different CI speech processors used, differences in participant inclusion criteria, differences in speech materials.

The average VCV test scores did not vary significantly with change in stimulation rates. Also, while the percentage feature transmission scores showed that there was no effect of rate, there was an effect of feature. The percentage feature transmission scores for manner of articulation were higher than those for voicing and place of articulation. This indicates that manner of articulation is most easily perceived and this is consistent with the findings of Friesen et al. (2001); Teoh, Neuberger, and Svirsky (2003); and Valimaa, Maata, Loponen, and Sorri, (2002). It has been reported that manner cues are mostly due to temporal information, place cues are mostly due to spectral information and voicing cues are due to both spectral and temporal information (Fu & Shannon, 1999; Xu, Thompson & Pfingst, 2005). It is believed that the temporal cues are relatively well preserved in CIs. In the present study, the average feature transmission scores at 1800 pps/ch were significantly higher than the average scores at 2400 pps/ch and 3500 pps/ch for the voicing feature. The Post-hoc tests showed no significant difference among the rates for the transmission score of voicing. This might be due to individual differences where one subject had very high feature transmission scores for voicing at 1800 pps/ch. However, the reason for this is not clear.

The average BKB sentence test scores in noise also did not vary significantly across stimulation rates. These results were quite different from the finding of Nie et al. (2006), where they found degraded performance in recognition in noise for higher stimulation rates. It is difficult to compare the results of the present study with other studies and come to a conclusion as the scores were averaged across only two users of CI. The reason for the other two participants not having good scores in this test might be because the sentence recognition test in noise is a very difficult test condition for even very good users of CI and only very few individuals are able to perform well on this test.

The effect of stimulation rates on CI speech perception : Even with little or no scientific evidence to support the high stimulation rate advantage, the CI manufacturers have advocated high stimulation rates in their current devices. It is believed that "increasing per-electrode stimulation rates from few hundred Hertz to 1000 Hz or 2000 Hz seemed to improve speech recognition under certain conditions, but increasing the rate beyond that to 4000 Hz may actually degrade CI performance" (Nie et al., 2006, p. 215). However, in the present study it appears that there is no effect of stimulation rates seen on consonant recognition in quiet or specific speech features that are more dependent on temporal information and also on sentence recognition in noise. The reason for this might be because of the way in which the Nucleus Freedom CI signal processing functions. In the Nucleus Freedom CI, to double the amount of temporal information available, the length of the FFT needs to be halved (which would halve the amount of frequency resolution). Whereas, to double the stimulation rates, the Nucleus processor simply overlaps half of the adjacent FFTs. This does add some temporal information, but it also effectively smoothes the temporal envelope. Hence, the increase in actual temporal information is theoretically much lesser than the increase in the analysis rate. With the 128-point FFT analysis with 16,000 samples per second, there are 125 FFT analyses per second. As the stimulation rate is increased, the processor increases the overlap between analyses.

From Table 1, it can be noticed that all those using a CI preferred 2400 pps/ch, which was their usual rate used in regular MAPs. This might be due to the practice effect with their regular MAP and not having sufficient time to learn to appreciate the MAPs with other rates. However, they all reported only the difference in terms of loudness (MAP with 1800 pps/ch was reported as being quite loud and the MAP with 3500 pps/ch was reported as being soft) even after the adjustment of loudness to the same level as their previous MAPs. The participants using CI reported no or little difference in terms of speech clarity across MAPs with the three different stimulation rates. The difference in loudness might be due to the different pulse width of the current in different MAPs; 1800 pps/ch with a pulse width of 20 μS/phase, 2400 pps/ch with 12 μS/phase pulses and 3500 pps/ch with 9.6 μS/phase pulses. The greater the pulse width the louder the sound and higher the stimulation rate the louder the sound.

Speech perception in participants with normal hearing on CI acoustic modelling

The average VCV test scores in quiet and BKB sentence test scores in noise did not vary significantly with change in stimulation rates for participants with normal hearing on CI acoustic modelling. Even the average feature transmission scores follow a similar trend as that of participants using a CI, where the manner of articulation was better perceived than voicing and place of articulation. The variation in average scores at different stimulation rates for both VCV and BKB tests in CI acoustic models were much less. The reason for that might be that, by using listeners with normal-hearing as participants, one rules out the individual variability such as different aetiologies of the clinical group, differences in duration of auditory deprivation leading to various amounts of degeneration of auditory functions, differences in survival of cell bodies in the spiral ganglion, differences in location of electrodes, etc which are inherent to participants with CI (Dorman, Loizou, Fitzke & Tu, 1998; Laneau et al., 2006). However, the CI acoustic models result in a more controlled situation for testing and many parameters can be changed independently.

The effect of overlap: CI acoustic models have their own set of limitations. One such limitation is that auditory simulations cannot replicate the mostly unknown dynamics of current spread in the cochlea (Dorman et al., 1998). However, the overlap is believed to create some effect of channel interaction in CI acoustic modelling to account for the current spread. In the present study the effect of overlap on consonant recognition in quiet and sentence recognition in noise were found to be significant in CI acoustic modelling. The CI acoustic models with overlap led to poorer scores which were closer to the scores of users of a CI compared to CI acoustic models without overlaps. The overlap was also significant across the features of manner and voicing but not for place of articulation. The reason for overlap not having a significant effect on place of articulation was not very clear. Overlap creates an effect of channel interaction to some extent in CI acoustic modelling, which is more realistic comparison to the performance of those using a CI.

The effect of stimulation rates on CI acoustic modelling: Different stimulation rates do not seem to have any effect on both consonant perception in quiet and sentence recognition in noise for participants with normal hearing on CI acoustic modelling. There might be many possible reasons for this. Firstly, CI acoustic models do not account for a electrical neural interface identical to that in users of a CI. Also, the stimulation rate effects are thought be only restricted to those using a CI as naturalistic stochastic responses are restricted to those using a CI (Rubinstein et al., 1999). The second possible reason might be because of the way in which the signal processing in the Nucleus Freedom CI works (discussed earlier).

Comparison of participants using a CI with participants with normal hearing on CI acoustic modelling

The speech perception scores of both participants using a CI and participants with normal hearing on CI acoustic modelling followed a similar pattern. The stimulation rates had little or no effect on consonant perception in quiet and sentence recognition in noise, in both participant groups. Even the average feature transmission scores of all the participants followed a similar trend, where the cues for manner of articulation were perceived better than voicing and place of articulation cues.

The ANOVA results on VCV scores, feature transmission scores and BKB sentence test scores showed no significant effect across participant groups. This shows that the scores of the users of a CI were not very different from scores of participants with normal hearing on CI acoustic modelling. The effect of rate was only significant in feature transmission scores and there was a significant interaction between participant group and stimulation rates. This might be attributed to individual differences and especially the voicing scores of the users of a CI at 1800 pps/ch.

By looking at these trends in score it can be assumed that CI acoustic models are useful in studying the effect of stimulation rates in those using a CI. Thus, it appears that the CI acoustic modelling is a good comparison for the performance of individuals using a CI.

The limitations of this study include the small sample size of participants using a CI which was not sufficient to reach the ideal statistical power. The time given to users of a CI for acclimatization to the new MAP was very short and the rate advantage may not have been demonstrated with such little duration of practice. Also, the time given to participants with normal hearing on CI acoustic modelling to acclimatize to the simulated signals may not be directly comparable to that of participants using a CI.

Conclusions

High electrical stimulation rates do not improve either consonant perception in quiet or sentence recognition in noise in both the groups. The feature transmission scores suggest that manner of articulation is better perceived than voicing and place of articulation in both the groups. CI acoustic models are useful in studying the rate effect on CI speech perception. CI acoustic models with overlap result in better approximation of performance of those using a CI.

Acknowledgements: The authors wish to thank John Laneau, Marc Moonen, Jan Wouters and Michael Buchler for providing the Cochlear Implant simulation software. Many thanks to the patients of the South of England Cochlear Implant Centre (SOECIC) and students of University of Southampton who gave up their precious time to participate in this study.

Address for correspondence: Mr Vinaya K. C. Manchaiah, Tutor in Audiology & Audiology Course Leader, School of Health Science Swansea University, Room 167, Glyndwr Building Swansea SA2 8PP, United Kingdom. Email: V.K.C.Manchaiah@swansea.ac.uk

References

Au, D. K. K. (2003). Effect of stimulation rate on Cantonese lexical tone perception by cochlear implant users in Hong Kong. *Clinical Otolaryngology, 28,* 533- 538.

Buchner, A., Frohne-Buchner, C., Battmer, R. D. & Lenarz, T. (2004). Two years of experience using stimulation rates 800 and 5000 pps with Clarion CII implant. *International Congress Series, 1273,* 48-51.

Dorman, M. F., Loizou, P. C., Rainey, D. (1997). Simulating the effect of cochlear-implant electrode insertion depth on speech understanding. *Journal of the Acoustical Society of America, 102* (5), 2993-2996.

Dorman, M. F., Loizou, P.C., Fitzke, J. & Tu, Z. (1998). The recognition of sentence in noise by normal-hearing listeners using simulations of cochlear-implant signal processor with 6-20 channels. *Journal of the Acoustical Society of America, 104*(6), 3583-3585.

Faulkner, A., Rosen, S., & Smith, C. (2000). Effect of the salience of pitch and periodicity information on the intelligibility of four-channel vocoded speech: implications for cochlear implants. *Journal of the Acoustical Society of America, 108,* 1877-1887.

Friesen, L. M., Shannon, R. V., Cruz, R. J. (2005). Effect of stimulation rate on speech recognition with cochlear implants. *Audiology Neurotology, 10,* 169-184.

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.

Fu, Q. J. & Shannon, R. V. (1999). Effects of electrode location and spacing on phoneme recognition with the Nucleus-22 cochlear implant. *Ear and Hearing, 20,* 321-331.

Fu, Q.J. & Shannon, R. V. (2000). Effect of stimulation rate on phoneme recognition by Nucleus-22 cochlear implant listeners. *Journal of the Acoustical Society of America,107,* 589- 597.

Grasmeder, M. L. & Verschuur, C. A. (2003). Investigation of the Front End of the Nucleus Sprint Speech Processor and Headset, Unpublished manuscript.

Harczos, T., Katai, A., Klefenz, F., Baljic, I. (2008). Effects of improper spectral delay on speech recognition with cochlear implants: a simulation study. International Conference on Biocomputation, Bioinformatics, and Biomedical Technologies, IEEE; 47-52.

Holden, L. K., Skinner, M. W., Holden, T. A. & Demorest M. E. (2002). Effect of stimulation rate with the Nucleus 24 ACE speech coding strategy. *Ear & Hearing, 23,* 463 - 476.

Kiefer, J., Ilberg, C. V., Pupprecht, V., Huber-Egener, J., Baumgartner, W., Gstoettner, W.,

Forgasi, F. & Stephan, K. (2000). Optimized speech understanding with the speech coding strategy in cochlear implants: The effect of variation in stimulus rate number of channels. Waltzman, S.B. & Cohen, N.L. (2000) (Eds.). *Cochlear Implants* (pp. 339-340). New York: Thieme Medical Publishers, Inc.

Laneau, J., Moonen, M. & Wouters, J. (2006). Factors affecting the use of noise-band vocoders as acoustic models for pitch perception in cochlear implants. *Journal of the Acoustical Society of America, 119* (1), 491- 506.

Loizou, P. C., Poroy, O. & Dorman, M. (2000). The effect of parametric variation on cochlear implant speech processor on speech understanding. *Journal of the Acoustical Society of America, 108,* 790 - 802.

Nie, K., Barco, A. & Zeng, F. (2006). Spectral and Temporal cues in cochlear implant speech processor. *Ear & Hearing, 27,* 208 -217.

Psarros, C. E., Plant, K. L., Lee, K., Decker, J. A, Whitford, L. A. & Cowan, R. S. C. (2002). Conversion from SPEAK to the ACE strategy in children using the Nucleus 24 cochlear implant system: Speech perception & speech production outcomes. *Ear & Hearing, 23,* 18S-27S.

Raffin, M. J. & Thornton, A. R. (1979). Confidence levels for differences between speech-discrimination scores: A research note. *Journal of Speech and Hearing Research, 23,* 5-18.

Rubinstein, J. T., Wilson, B. S., Finely, C. C., & Abbas, P. J. (1999). Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. *Hearing Research, 127,* 108 -118.

Teoh, S. W., Neuberger, H. S., & Svirsky, M. A. (2003). Acoustic and electrical pattern analysis of consonant perceptual cues used by cochlear implant users. *Audiology and Neuro-Otology, 8,* 269-285.

Vandali, A. E., Whitford, L. A., Plant, K. L. & Clark, G.M. (2000). Speech perception as a function of electrical stimulation rate: Using the Nucleus 24 cochlear implant system. *Ear & Hearing, 21,* 608 -624.

Valimaa, T. T., Maata, T. K., Loponen, H. J., Sorri, M. J. (2002). Phoneme recognition and confusions with multichannel cochlear implants: Vowels. *Journal of Speech Language and Hearing Research, 45,* 1039 - 1045.

Verschuur, C. A. (2005). Effect of stimulation rate on speech perception in adult users of the Med-El CIS speech processing strategy. *International Journal of Audiology, 44,* 58 - 63.

Wilson, B. S., Finley, C. C., Lawson, D. & Zerbi, M. (1997). Temporal representation with cochlear implants. *The American Journal of Otology, 18,* S30-S34.

Xu, F., Thompson, C. S. & Pfingst. B. E. (2005). Relative contributions of spectral and temporal cues for phoneme recognition, *Journal of the Acoustical Society of America, 17* (5), 3255-3267.