Novel speech processing strategies in cochlear implants: real improvements competing with learning effects

Jong, M.A.M. de

Author: Jong, M.A.M. de
Title: Novel speech processing strategies in cochlear implants: real improvements competing with learning effects
Issue Date: 2019-06-13

Novel Speech Processing Strategies in Cochlear Implants

Real improvements competing with learning effects

Monique Anna Maria de Jong

ISBN: 978-94-6380-334-2

Printed by: ProefschriftMaken || www.proefschriftmaken.nl

The publication of this thesis was financially supported by: Advanced Bionics, ALK, Beter Horen, Chipsoft, Dos Medical B.V., EmiD audiologische apparatuur, Entercare B.V., Meditop Medical Products B.V., Mylan Healthcare B.V., Oticon Medical

Novel Speech Processing Strategies in Cochlear Implants

Real improvements competing with learning effects

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker

volgens besluit van het College voor Promoties

te verdedigen op donderdag 13 juni 2019

klokke 15.00 uur

door

MONIQUE ANNA MARIA DE JONG

geboren te Gouda in 1988

Promotor:	Prof. dr. ir. J.H.M. Frijns
Copromotor:	Dr. ir. J.J. Briaire
Leden promotiecommissie:	Prof. dr. A. van Wieringen (KU Leuven) Prof. dr. D. Başkent (UMC Groningen) Prof. dr. P.P.G. Van Benthem

CONTENTS

1.	Introduction	7
2.	Learning Effects in Psychophysical Tests of Spectral and Temporal Resolution EAR AND HEARING, 2018	19
3	Take-Home Trial Comparing Fast Fourier Transformation-Based and Filter Bank-Based Cochlear Implant Speech Coding Strategies BIOMEDICAL RESEARCH INTERNATIONAL, 2017	35
4	Dynamic Current Focusing: A Novel Approach to Loudness Coding in Cochlear Implants EAR AND HEARING, 2019	51
5	Dynamic Current Focusing for Loudness Encoding in Cochlear Implants: A Take-Home Trial INTERNATIONAL JOURNAL OF AUDIOLOGY, 2019	77
6	Effectiveness of Phantom Stimulation in Shifting the Pitch Percept in Cochlear Implant Users Under review	103
7	General Discussion	127
8	Nederlandse samenvatting	147
A	Appendix	153
	Abbreviations	155
	Contributing Authors	157
	List of Publications	159
	About the Author	161
	Dankwoord	163

CHAPTER 1

General Introduction

THE (PATHO)PHYSIOLOGY OF HEARING

Hearing, or auditory perception, is the transformation of sound vibrations into nerve impulses that are conveyed to the brain, where they are interpreted as sounds. The auditory system that facilitates this process is anatomically divided in three parts: (1) the outer ear, (2) the middle ear, and (3) the inner ear (Figure 1). The outer ear consists of the auricle and the external ear canal, which ends blindly at the tympanic membrane. The tympanic membrane separates the outer ear from the small air-filled cavity of the middle ear, where the three auditory ossicles (the malleus, incus and stapes) are located. The stapes is placed on the oval window, which separates the middle from the inner ear. The two functional units of the inner ear are the vestibular system with its semicircular canals and the cochlea, which contains the sensory organ of hearing. The process of hearing starts when sound waves enter the outer ear and cause the tympanic membrane to vibrate. Because the malleus is connected to the tympanic

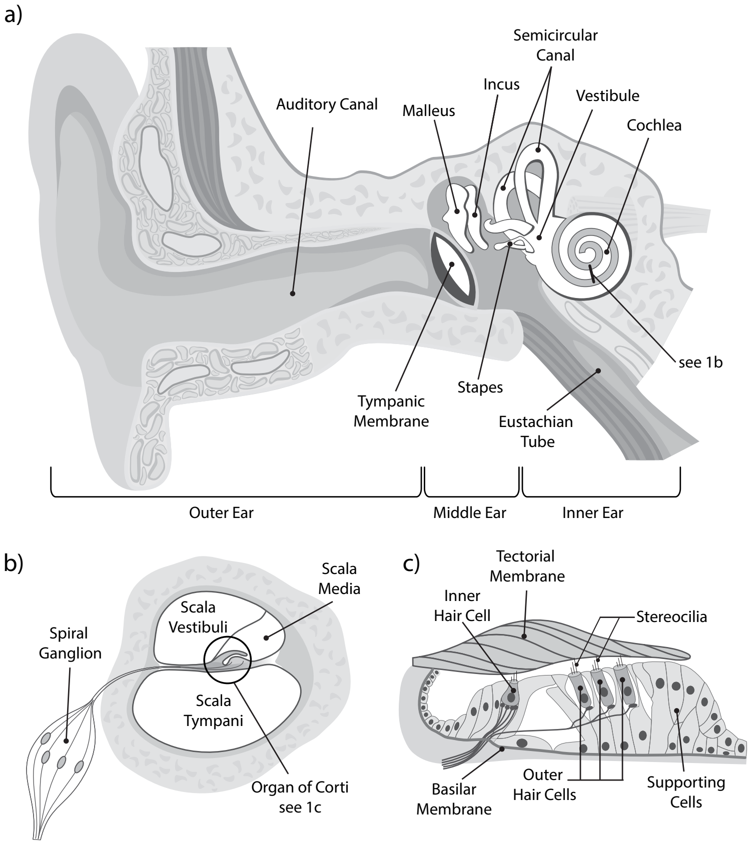


FIGURE 1. SCHEMATIC ILLUSTRATION OF THE ANATOMY OF THE HUMAN EAR.

membrane, the ossicle chain starts to vibrate also, amplifying the vibration pressure roughly 20 times. The stapes transmits sound waves to the fluid filled inner ear through the oval window, which is a flexible membrane between the middle and the inner ear. These fluid vibrations set up traveling waves along the basilar membrane that stimulate the hair cells of the organ of Corti, which is located in the cochlea (Figure 1c). These hearing cells convert the sound vibrations into action potentials in the fibers of the cochlear nerve. There are four types of hearing loss: conductive, sensorineural, retrocochlear, and mixed hearing loss. In conductive hearing loss the outer and/or middle ear is the organic substrate of the impairment. In sensorineural hearing loss the inner ear is damaged and in retrocochlear hearing loss the auditory nerve or the central auditory system is affected. Diseases that underlie conductive hearing loss can often be treated either medically or with surgery. As medical procedures often do not fully reverse the conducive hearing loss, hearing is regularly amplified with hearing aids. When the hearing loss is of sensorineural or retrocochlear nature, and it is severe, hearing aids are often not sufficient. In the case of sensorineural hearing loss, Cochlear Implants (CIs) can offer a solution.

COCHLEAR IMPLANTS

A CI is a surgically implanted electronic device that is developed to restore hearing of severely hearing impaired and deaf individuals. It is the first example of a neural prosthesis that can actually substitute a sensory organ and has become the standard care for the rehabilitation of severely hearing impaired adults and children. The CI bypasses the malfunctioning auditory periphery and cochlea to directly stimu-

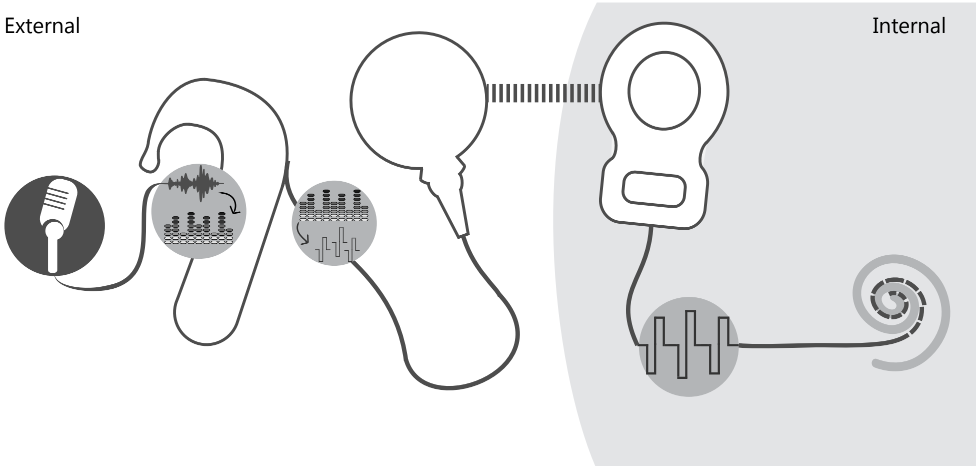


Figure 2. Schematic representation of the basic components of a Cochlear Implant system.

late the auditory nerve. Due to the remarkable success in providing speech perception to this hearing impaired population, CI users can even participate in a hearing society. A CI consists of an external and a surgically implanted internal part (Figure 2). The external part of the device consists of a microphone, a speech processor, and a transmitter coil to send the auditory signal to the internal part of the implant. The microphone captures the auditory signal and the processor divides the signal into separate frequency bands, one for each active channel of the CI. The envelope of the signal of each frequency band is extracted and used to determine the amplitudes. The stimulation pattern, which is also dependent on the type of speech coding strategy, is then sent to the internal part of the CI via a radio-frequency link. The internal part of this link decodes the signal into an electrical current that is send to the implanted electrode array. This is an array of 12 to 22 electrode contacts, depending on the type of device, which is inserted approximately 1.5 turns into the scala tympani of the cochlea via a round window insertion or a cochleostomy. CIs make use of the frequency-to-place representation of the cochlea called 'tonotopy'. Tonotopy is the spatial arrangement of where sounds of different frequencies are processed in the cochlea, where the higher frequencies are coded in the base of the cochlear spiral, and the lower frequencies in the apex[1]. In a CI each electrode contact is located near auditory nerve fibers coding for different frequencies, mimicking this tonotopic arrangement. Each activated electrode contact ideally depolarizes a separate population of nerve fibers and thus cause a distinct pitch percept. The administered pulses are always charge-balanced and usually return via a ground electrode contact that is placed under the temporal muscle (monopolar stimulation).

THE DEVELOPMENT OF COCHLEAR IMPLANTS

Considerable technological and scientific advances have been made since the first attempts to restore hearing through electrical stimulation in 1957 by Djourno and Eyries[2]. In the following decades, thinner electrode arrays with multiple electrode contacts were developed, improving performance significantly (Figure 3). The Nucleus multichannel CI was the first that resulted in an 'open set' speech understanding with electrical stimulation alone[3]. Also, multiple processing strategies were investigated in that time, of which the so called 'Continuous Interleaved Sampling' (CIS) resulted in a breakthrough in this field. While earlier algorithms presented the stimuli simultaneously to all electrode contacts, with CIS pulses are presented in a non-overlapping sequence, preventing electrical interaction between the channels and, therefore, resulting in major improvements in speech understanding[4]. Since then, a lot of effort has been put into improving the unraveling of sound and translating it into efficient neural stimulation patterns. Nevertheless, the majority of today's speech-processing methods are still based on CIS. Figure 3 depicts average sentence recognition scores in quiet with the different devices and speech coding strategies that have been implemented during

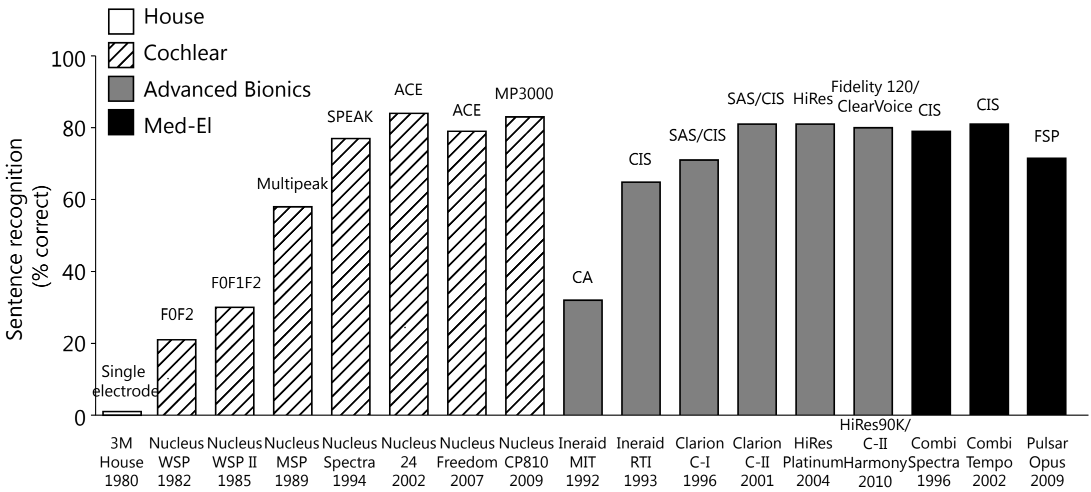


FIGURE 3. SENTENCE RECOGNITION SCORES IN QUIET AS A FUNCTION OF TIME FOR THE DIFFERENT COCHLEAR IMPLANT SYSTEMS AND PROCESSING STRATEGIES. The x-axis labels show the type of device, the processor model, and the year the study was published. The labels above the scores show the used speech coding strategies. Note that the used tests can differ between studies, but the outcome measures (sentence recognition in quiet) was the same across studies. The scores until 2004 are copied from Zeng (2004). The following results were extracted from the following papers: Nucleus Freedom38, Nucleus CP810[39], HiRes system[40], HiRes/CII+Harmony[41], Med-El Opus Device[42]. Abbreviations; ACE: advanced combinational encoder, CA: compressed analog, CIS: continuous interleaved sampling, FSP: Fine structure processing, SAS: simultaneous analog stimulation, SPEAK: spectral peak speech coding.

the previous decades. The figure reveals that the improvement has reached a plateau after the introduction of CIS. The reason why no improvement is observed is twofold: 1) despite the recent technological developments we are still incapable to transmit more detailed information, and 2) speech tests have their restrictions in that a long adaptation time to new speech coding strategies is required to measure the final hearing outcome. Moreover, comparing speech tests across research centers is problematic due to language differences and different test setups.

THE EVALUATION OF TECHNOLOGICAL DEVELOPMENTS

As we are dealing with this ceiling effect in terms of speech understanding in quiet, the focus of research has shifted towards the improvement of speech understanding in noise and music perception. Yet, improvements in this area are expected to come in smaller steps than the previous developments. While traditional speech perception outcomes are most relevant to the CI users themselves, examination of the fundamental principles of hearing (e.g. spectral and temporal resolution) enables detecting smaller changes in performance. Moreover, it provides researchers with more detailed information about the effect of new speech coding strategies. Spectral resolution is the ability to detect the multiple frequency components of a complex sound. Temporal resolution is the ability to identify variations in intensity of sound in time. Both spectral and temporal resolution play a key role in speech understanding, because speech contains numerous frequency and timing cues. These abilities can be tested with several psychophysical tests like spectral ripple[5,6] and temporal modulation detection[7,8] tests. These tests can be used instantaneously (without adaptation to new speech coding strategies) and are language independent, facilitating comparison of outcomes across countries. The psychophysical measures are extensively studied and improved over the years. For example, Henry & Turner (2003) developed a spectral ripple tests in which subjects had to discriminate between two rippled noise stimuli in which the frequency positions of the peaks and valleys are reversed[9]. Azadpour & McKay (2012) concluded that the discrimination of these stimuli can be influenced by other factors than spectral resolution, such as differences in loudness, spectral centroid, and changes to the spectral edges[10]. For that reason, the spectral-temporally modulated ripple test (SMRT) was developed to avoid that CI listeners could make use of other cues than the intended spectral ripple cues[11]. By eliminating all other cues from the stimuli, the interpretation of the outcomes is much easier. This is the test that is regularly used in the Leiden University Medical Center (LUMC), although still ambivalence exists about which measure of spectral resolution (e.g. spectral ripple density, spectral modulation depth) should be used. Nevertheless, these psychophysical tests are widely used in multiple CI research centers around the world, and those CI users that take part in CI research programs execute these tests several times. It is well known from other fields that psychophysical measures are prone to learning effects[12], while this has not been well documented about the tests used in CI research. Although some studies investigated the effect of performing the tasks multiple times on one day, repeated testing over several days or weeks has not yet been studied. As the expected effect sizes in CI research are only small, also small learning effects could drastically change the interpretation of results and are, therefore, important to take into account. Chapter 2 describes the effect of repeated testing with such psychophysical tests when a relatively short period between tests days was chosen, as often is the case in CI research.

NEW SPEECH CODING STRATEGIES

It was assumed that especially the poor frequency information that is provided with a CI (only 12-22 frequency channels) causes a relatively poor quality of sound. To be able to increase the spectral resolution, first more detailed frequency information should be obtained from the audio signal. As increasing the number of filter banks requires too much computing power, a finite impulse response (FIR) filter bank based on fast Fourier transformation (FFT) processing was introduced[13]. This FFT filter bank has a 14.7 ms sliding window and recalculates the FFT every 1.1 ms[14]. This method results in a more accurate spectral analysis, and it is also a more energy efficient manner of analyzing the incoming sound. Secondly, additional frequency channels are required to transmit this information to the auditory nerve. One approach to increase the number of frequency channels is to use so-called 'current steering'[14–16]. Current steering facilitates stimulation of auditory nerve regions that are located in between physical electrode contacts by simultaneous stimulation of two adjacent contacts, thereby creating virtual electrode contacts (Figure 4). In theory, more frequency information could be taken up by CI users when current steering is applied. In practice, however, the effect is somewhat disappointing as some studies report beneficial effects of current steering[17–19], but an equal number of studies could not demonstrate these effects[16,20–22]. It was hypothesized that the introduction of FFT-based filters had a negative effect on speech perception[21]. The time windows of FFT filters are relatively broad, possibly deteriorating the transmission of time domain information (temporal resolution). As temporal resolution is as important for understanding speech as spectral resolution, it could be that a beneficial effect on spectral resolution was counteracted by the decline in temporal resolution. The effect of FFT-based filter banks on CI performance is studied in **Chapter 3.**

Nevertheless, even if FFT processing would have a disadvantageous effect, it could not be the only explanation of the minimal gain in performance resulting from the addition of (virtual) frequency channels. The fact that Friesen et al. (2001) showed that speech understanding in quiet does not improve when the active number of spectral channels is increased above eight suggests that the electrode-neuron interface is a limiting factor[23]. Also, Biesheuvel et al. (2018) showed in an extensive pitch discrimination experiment that subjects with an Advanced Bionics (Sylmar, California) CI can only discriminate between 11 of the 16 implanted electrode contacts on average[24]. While deficits on the neuronal side are to be expected in CI users[25], which could explain the relatively poor pitch discrimination, also the stimulation side comes with shortages. As previously mentioned, in most stimulation strategies monopolar stimulation is used. In monopolar mode, the extra cochlear electrode contact, which is placed under the temporal muscle, is used as a return pathway. As a consequence the spread of electrical stimulation is broad, especially at high stimulation levels[26]. The fact that the spatial selectivity is poor causes individual channels to interact, thereby decreasing the actual number of independent frequency channels. Therefore, several methods to improve spatial selectivity with CIs have been developed. In tripolar stimulation the two adjacent electrode contacts serve as the return pathway instead of the extra-cochlear ground electrode (Figure 4). This so-called 'current focusing' technique creates a narrower intra-cochlear electrical field and hence increases the spatial selectivity[27,28]. This increase in spatial selectivity has indeed shown to decrease channel interaction in CI users[29]. One disad-

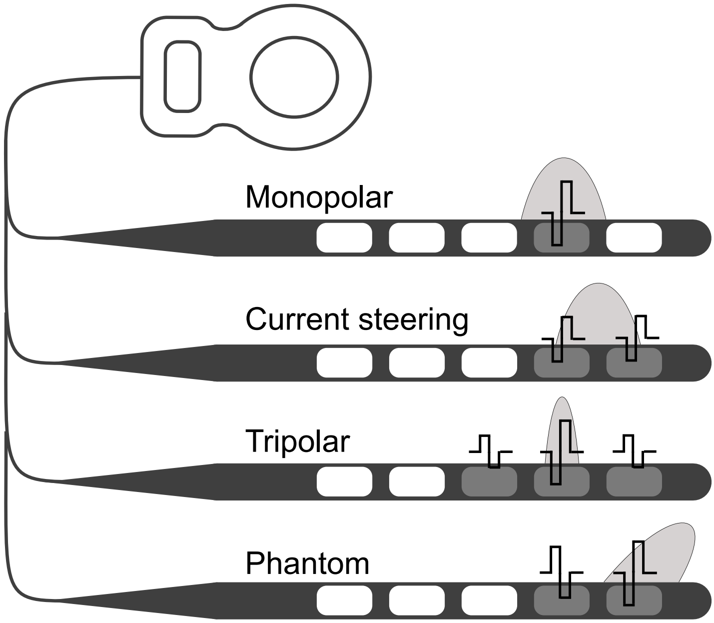


Figure 4. Illustration of the different stimulation strategies. The grey electrode contacts depict the active ones and the grey shades illustrate the expected electrical field following the stimulation strategy.

vantage is the excessively increased power consumption, to such an extent that maximum comfortable levels cannot always be reached with the tripolar configuration[29–31]. Moreover, at the highest loudness levels, the effect of current focusing is relatively small because loudness level partly determines the amount of current spread[32]. Therefore, a speech coding strategy was developed in the LUMC that uses high degrees of current focusing at the lower loudness levels, and lower degrees of current focusing at higher loudness levels. In this way, the beneficial effects of tripolar stimulation are well utilized, while the disadvantages are kept to a minimum. The strategy is called 'dynamic current focusing', and its qualities and possibilities will be further studied in Chapters 4 and 5.

Another way to enhance the number of discriminable pitches is with phantom stimulation, which was first introduced by Wilson et al. (1993)[33]. In phantom stimulation two adjacent electrode contacts are simultaneously stimulated with opposite polarity (Figure 4). Depending on the settings the electrical field is shifted either towards the apex or to the base. This stimulation mode makes it possible to stimulate auditory nerve fiber regions that are located outside the usual stimulation area of the implanted electrode contacts. These nerve regions normally cannot be addressed because of the limited insertion depths of the electrode arrays. Inserting an electrode array too deep into the cochlea can lead to trauma to the spiral ganglion cells and is, therefore, undesirable[34–36]. With the use of phantom stimulation, the CI can make use of a greater range of auditory nerve fibers without causing damage. Because the frequency channels then virtually lie further away from each other, less channel interaction would occur, leading to an improved channel discrimination and better transmission of frequency information. Multiple phantom configurations are proposed, but it is unclear which results in the greatest pitch shift. Chapter 6 describes the different phantom configurations and explores which of these would fit best in a speech coding strategy.

An overall discussion of the major results and conclusions of all chapters is presented in Chapter 7, followed by some clinical implications and future perspectives. A summary of this thesis in Dutch can be found in Chapter 8.

REFERENCES

1. Von Békésy, G. (1960) Experiments in Hearing. ***New York McGraw-Hill***
2. Djourno, A. & Eyries, C. (1957) Auditory prosthesis by means of a distant electrical stimulation of the sensory nerve with the use of an indwelt coiling. ***Presse Med. 65,*** 1417.
3. Dowell, Mecklenburg & Clark. (1986) Speech recognition for 40 patients receiving multichannel cochlear implants. ***Arch. Otolaryngol. Head Neck Sugergy 112,*** 1054–1059.
4. Wilson, B.S., Finley, C.C., Lawson, D.T. ***et al.*** (1991) Better speech recognition with cochlear implants. ***Nature 352,*** 236–238.
5. Henry, B., Turner, C. & Behrens, A. (2005) Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. ***J. Acoustal Soc. Am. 118,*** 1111–1121.
6. Won, J.H., Drennan, W.R. & Rubinstein, J.T. (2007) Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. ***J. Assoc. Res. Otolaryngol. 8,*** 384–392.
7. Bacon, S.P. & Viemeister, N.F. (1985) Temporal Modulation Transfer Functions in Normal-Hearing and Hearing-Impaired Listeners. ***Audiology 24,*** 117–134.
8. Won, J.H., Drennan, W.R., Nie, K., Jameyson, E.M. & Rubinstein, J.T. (2011) Acoustic temporal modulation detection and speech perception in cochlear implant listeners. ***J. Acoust. Soc. Am. 130,*** 376–388.
9. Henry, B.B. a. & Turner, C.W.C. (2003) The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. ***J. Acoustal Soc. Am. 113,*** 2861–2873.
10. Azadpour, M. & McKay, C.M. (2012) A psychophysical method for measuring spatial resolution in cochlear implants. ***J. Assoc. Res. Otolaryngol. 13,*** 145–57.
11. Aronoff, J.M. & Landsberger, D.M. (2013) The development of a modified spectral ripple test. ***J. Acoust. Soc. Am. 134,*** 217–222.
12. Fine, I. & Jacobs, R.A. (2002) Comparing perceptual learning tasks: a review. ***J. Vis. 2,*** 190–203.
13. Cochran, W.T., Cooley, J.W., Favin, D.L., ***et al.*** (1967) What is the Fast Fourier Transform? ***Trans. Audio Electroacoust. 15,*** 45–55.
14. Nogueira, W., Litvak, L.M., Edler, B., Ostermann, J. & Büchner, A. (2009) Signal Processing Strategies for Cochlear Implants Using Current Steering. ***EURASIP J. Adv. Signal Process. 2009,*** 1–21.
15. Litvak, L.M., Krubsack, D.A. & Overstreet, E.H. (2003) Method and system to convey the within-channel fine structure with a cochlear implant. ***Adv. Bionics Corp. United States Patent, p.17***
16. Snel-bongers, J., Briaire, J.J., Vanpoucke, F.J. & Frijns, J.H.M. (2011) Spread of Excitation and Channel Interaction in Single- and Dual-Electrode Cochlear Implant Stimulation. ***Ear Hear. 32,*** 1–10.
17. Koch, D.B., Downing, M., Osberger, M.J. & Litvak, L. (2007) Using Current Steering to Increase Spectral Resolution in CII and HiRes 90K Users. ***Ear Hear. 28,*** 38–41.
18. Firszt, J., Holden, L., Reeder, R.M. & Skinner, M.W. (2009) Speech recognition in cochlear implant recipients: comparison of standard HiRes and HiRes 120 sound processing. ***Otol. Neurotol. 30,*** 146–152.
19. Park, H.J., Lee, S.C., Chun, Y.M. & Lee, J.Y. (2009) HiRes with Fidelity 120 benefit in native speakers of Korean. ***Cochlear Implants Int. 10,*** 85–88.
20. Berenstein, C.K., Mens, L.H.M., Mulder, J.J.S. & Vanpoucke, F.J. (2008) Current Steering and Current Focusing in Cochlear Implants: Comparison of Monopolar, Tripolar, and Virtual Channel Electrode Configurations. ***Ear Hear. 29,*** 250–260.
21. Drennan, W., Won, J., Nie, K., Jameyson, E. & Rubinstein, J.T. (2010) Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. ***Hear. Res. 262,*** 1–8.

22. Melo, T.M. De, Bevilacqua, M.C., Costa, O.A. & Moret, A.L.M. (2013) Influence of signal processing strategy in auditory abilities. ***Braz. J. Otorhinolaryngol. 79,*** 629–35.
23. 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. ***J. Acoust. Soc. Am. 110,*** 1150.
24. Biesheuvel, J., Briaire, J.J., de Jong, M.A.M. & Frijns, J.H.M. (2018) Clinical assessment of channel discernibility along the cochlear implant electrode array and its relation to speech perception. ***Submitted***
25. Nadol, J.B. (1997) Patterns of neural degeneration in the human cochlea and auditory nerve: Implications for cochlear implantation. ***Otolaryngol. - Head Neck Surg. 117,*** 220–228.
26. Bierer, J.A. & Middlebrooks, J.C. (2002) Auditory Cortical Images of Cochlear-Implant Stimuli: Dependence on Electrode Configuration. ***J. Neurophysiol. 87,*** 478–492.
27. Bierer, J.A. & Middlebrooks, J.C. (2004) Cortical responses to cochlear implant stimulation: channel interactions. ***J. Assoc. Res. Otolaryngol. 5,*** 32–48.
28. Vellinga, D., Briaire, J.J., van Meenen, D.M.. & Frijns, J.H.M. (2017) Comparison of multipole stimulus configurations with respect to loudness and spread of excitation. ***Ear Hear. 38,*** 487–496.
29. Bierer, J.A. (2007) Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration. ***J. Acoust. Soc. Am. 121,*** 1642–1653.
30. Kral, A., Hartmann, R., Mortazavi, D. & Klinke, R. (1998) Spatial resolution of cochlear implants: The electrical field and excitation of auditory afferents. ***Hear. Res. 121,*** 11–28.
31. Mens, L.H.M. & Berenstein, C.K. (2005) Speech Perception with Mono- and Quadrupolar Electrode Configurations: A Crossover Study. ***Otol. Neurotol. 26,*** 957–964.
32. Khalili Moghaddam, G., Lovell, N.H., Wilke, R.G.H., Suaning, G.J. & Dokos, S. (2014) Performance optimization of current focusing and virtual electrode strategies in retinal implants. ***Comput. Methods Programs Biomed. 117,*** 334–342.
33. Wilson, B.S., Lawson, D.T. & Zerbi, M. (1993) Third Quarterly Report: Speech processors for auditory prostheses. In. Research triangle park: Research triangle inst.
34. Wardrop, P., Whinney, D., Rebscher, S.J., Luxford, W. & Leake, P. (2005) A temporal bone study of insertion trauma and intracochlear position of cochlear implant electrodes. II: Comparison of Spiral Clarion and HiFocus II electrodes. ***Hear. Res. 203,*** 68–79.
35. Finley, C.C. & Skinner, M.W. (2008) Role of electrode placement as a contributor to variability in cochlear implant outcomes. ***Otol. Neurotol. 29,*** 920–928.
36. Kalkman, R.K., Briaire, J.J. & Frijns, J.H.M. (2014) Current focussing in cochlear implants: An analysis of neural recruitment in a computational model. ***Hear. Res. 322,*** 89–98.
37. Zeng, F.-G. (2004) Trends in cochlear implants. ***Trends Amplif. 8,*** 1–34.
38. Balkany, T., Hodges, A., Menapace, C., ***et al.*** (2007) Nucleus Freedom North American clinical trial. ***Otolaryngol. - Head Neck Surg. 136,*** 757–762.
39. Muller-Deile, J., Kortmann, T., Hoppe, U., Hessel, H. & Morsnowski, A. (2009) Improving speech comprehension using a new cochlear implant speech processor. ***HNO 57,*** 567–574.
40. Koch, D.B., Osberger, M.J., Segel, P. & Kessler, D. (2004) HiResolution™ and conventional sound processing in the HiResolution™ bionic ear: Using appropriate outcome measures to assess speech recognition ability. ***Audiol. Neuro-Otology 9,*** 214–223.
41. Frohne-Büchner, C., Brendel, M., Saalfeld, H., Büchner, A. & Lenarz, T. (2010) Evaluation of the Built-In TCoil of the Behind-the-Ear Sound Processor Harmony. ***Cochlear Implants Int. 11,*** 412–415.
42. Riss, D., Arnoldner, C., Reiß, S., Baumgartner, W.D. & Hamzavi, J.S. (2009) 1-year results using the Opus speech processor with the fine structure speech coding strategy. ***Acta Otolaryngol. 129,*** 988–991.

CHAPTER 2

Learning Effects in Psychophysical Tests of Spectral and Temporal Resolution

Monique A.M. de Jong, Jeroen J. Briaire, and Johan H.M. Frijns

Ear and Hearing, 2018

ABSTRACT

Objectives

Psychophysical tests of spectral and temporal resolution, such as the spectral-ripple discrimination task and the temporal modulation detection test, are valuable tools for evaluation of cochlear implant performance. Both tests correlate with speech intelligibility and are reported to show no instantaneous learning effect. However, some of our previous trials have suggested there is a learning effect over time. The aim of this study was to investigate the test-retest reliability of the 2 tests when measured over time.

Design

Ten adult cochlear implant recipients, experienced with the HiResolution speech coding strategy, participated in this study. Spectral ripple discrimination and temporal modulation detection ability with the HiResolution strategy were assessed both before and after participation in a previous trial that evaluated 2 research speech coding strategies after 2 weeks of home-usage. Each test was repeated six times on each test day.

Results

No improvement was observed for same-day testing. However, comparison of the mean spectral ripple discrimination scores before and after participation in the take-home trial showed improvement from 3.4 to 4.8 ripples per octave ($p<0.001$). The mean temporal modulation detection thresholds improved from -15.2 dB to -17.4 dB ($p=0.035$).

Conclusions

There was a clear learning effect over time in the spectral and temporal resolution tasks, but not during same-day testing. Learning effects may stem from perceptual learning, task learning or a combination of those two factors. These results highlight the importance of a proper research design for evaluation of novel speech coding strategies, where the baseline measurement is repeated at the end of the trial to avoid false positive results as a consequence of learning effects.

INTRODUCTION

Psychophysical tests of spectral and temporal resolution, such as the spectral-ripple test[1–6] and the temporal modulation detection test[7–9], are valuable tools for the evaluation of cochlear implant (CI) hearing during clinical trials. The extent to which the implementation of novel technologies affects the performance of CI recipients is often too mild to detect with traditional speech or music outcome measures. Psychophysical measures are more sensitive to processor changes as they allow for the evaluation of basic abilities, such as spectral and temporal resolution[2,10,11], which are fundamental aspects of how well people hear. Both tests have been shown to correlate independently with vowel, consonant, and speech recognition in CI recipients[1,12–16]. It is generally assumed that the evaluation of basic psychophysical capabilities yields a measure of hearing that does not change over time[1,2]. Previous studies have investigated potential 'task learning effects' of spectral and temporal measurements, that is, improvement in performance caused by practice with the task rather than actual improvement in spectral and/or temporal resolution. No task learning effect was found in an acute setting, when tasks were repeated up to nine times[1,2,17]. To the best of our knowledge, only 1 study examined the test-retest reliability of the spectral-ripple threshold measurement over a longer period of time in experienced CI users 1. No task learning effect was found when repeating the task on separate test days, although, no time interval between the measurements was reported. Drennan *et al*. (2015) studied learning in both spectral and temporal modulation tests in newly implanted CI users and, on average, did not find a significant improvement over the first 12 months after activation[18]. However, 20% of the individuals significantly improved on both tasks and another 20% significantly deteriorated.

The previously mentioned studies suggest that spectral and temporal testing serve as useful, and most probably also reliable diagnostic tools for assessment of CI outcome in a research setting. However, we have observed somewhat different outcomes in our research center. The modified spectral ripple test (SMRT), developed by Aronoff & Landsberger (2013)[5], and the modulation detection threshold (MDT) test, adapted from Bacon & Viemeister (1985)[19], are frequently used in the evaluation of novel processing strategies in the Leiden University Medical Center (LUMC), the Netherlands. As a limited number of CI users are available for research purposes, many of our subjects have participated in multiple studies over the last few years. As a result, these subjects have had substantial practice on the SMRT and MDT test with several different speech coding strategies. We noticed higher SMRT and MDT scores in these more practiced CI users and therefore hypothesize that the psychophysical performance among these CI recipients improved because of this practice. It is well known that CI recipients improve performance in the first few months after

TABLE 1. SUBJECT CHARACTERISTICS

Subject	Gender	Age (y)	Aetiology	Deafness (y)	CI side	CI experience (mo)	CVC (Ph%)	Implant
S1	Male	60	Familial	46	Left	160	93	Clarion CII
S2	Male	51	Progressive	25	Right	124	76	HiRes 90K
S3	Female	74	Familial	22	Left	114	93	HiRes 90K
S4	Male	58	Unknown	11	Right	31	93	HiRes 90K
S5	Male	67	Familial	67	Right	96	86	HiRes 90K
S6	Male	57	Familial	45	Right	174	96	Clarion CII
S7	Female	67	Meningitis	6	Left + Right	73	88	HiRes 90K
S8	Female	43	Meningitis	5	Left + Right	59	95	HiRes 90K
S9	Female	66	Unknown	35	Right	105	84	HiRes 90K
S10	Female	59	Type-II Usher	4	Right	46	89	HiRes 90K

CVC, Dutch phonetically balanced monosyllabic Consonant-Vowel-Consonant words; Ph%, percentage phonemes correct implantation[20–22]. It is plausible that this improvement is caused by 'perceptual learning', which is a process by which the ability of the auditory system to process stimuli is improved through experience. Also Moberly ***et al***. (2015) suggested that CI users could learn from new speech cues, which might be present in novel speech coding strategies[23]. Repeated testing with the SMRT and MDT test in a research setting with multiple novel speech coding strategies could, therefore, lead to both task and perceptual learning and consequently to improved SMRT and MDT performance. The present study assessed performance on the SMRT and MDT test before and after participation in a previous take-home trial, in which 2 experimental speech coding strategies were evaluated.

MATERIALS AND METHODS

Subjects

A group of 10 adult cochlear implant (CI) recipients who had been implanted with a HiRes90K device with HiFocus 1J or a CII device with HiFocus with a positioner electrode array (Advanced Bionics, Sylmar, CA) at the LUMC were recruited for this study. All had used the Harmony processor programmed with the HiResolution (HiRes) speech coding strategy for multiple years. Subject demographics are shown in Table 1. Ages ranged from 43 to 74 years with a mean of 60.2 years. The average duration of deafness was 26.6 years (range 4-67 years) and average implant experience was 98 months (range 31-174 months). Mean phoneme scores on open set Dutch monosyllabic consonant-vowel-consonant (CVC) words in quiet conditions at 65 dB were 89.3% (range 76-96%).

Protocol and speech coding strategies

Spectral ripple discrimination and temporal modulation transfer functions were assessed at baseline (week 0) and 2, 4, and 6 weeks after the baseline measures. Participants were tested twice (at t=0 weeks and t=6 weeks) with their standard clinical speech coding strategy, HiRes. This is a bandpass filter based strategy that uses a traditional processing approach in which channel-specific temporal envelopes are extracted and delivered with interleaved, high-rate pulse trains. More detailed information about this speech coding strategy is provided by Firszt (2003)[24]. The examinations at week 2 and 4 were part of a separate take-home trial, in which 2 variations of the HiRes speech coding strategy, which applied different filtering techniques, were evaluated. The 2 experimental strategies, HiRes FFT and HiRes Optima[25], utilize a finite impulse response filter in conjunction with Fast Fourier Transformation (FFT) processing. HiRes Optima also uses current steering to create up to 135 virtual spectral channels. In fact, it is a more energy-efficient variation of HiRes Fidelity 120[26]. As the number of excitable channels is increased with HiRes Optima, an improved performance on the SMRT is expected as compared to both HiRes and HiRes FFT, as was demonstrated for HiRes Fidelity 120 by Drennan ***et al***. (2010). However, these authors also argued that FFT processing potentially decreases temporal resolution, due to spectral smearing. Therefore, both HiRes FFT and HiRes Optima might decrease performance on the MDT test. To eliminate order and practice effects, the participants received the 2 experimental take-home strategies in randomized order and had the chance to adapt to the strategies during the 2 weeks prior to the testing. In this paper, the randomization allows for the evaluation of test date effects (between week 2 and 4) while minimizing the effects of processing strategy. In other words, in the paired comparison between performance at week 2 and 4, half of the participants was using HiRes FFT and half was using HiRes Optima at each test session. Because the order of strategy was randomized, the effect of strategy was minimal. The current study was approved by the Medical Ethical Committee of the LUMC (ref. P02.106.Y).

Psychophysical testing

During all psychophysical tasks, the listeners were seated in a double-walled sound-attenuating booth. Sounds were presented via a single loudspeaker, placed 1 meter from the listener at 0 degrees and level with the listener's head. Subjects received instructions for the psychophysical tests and then practiced the tasks six times or more if necessary, to avoid learning in the actual test setting. Listeners responded using a mouse with a custom computer interface, or they responded verbally when they were unable to use the mouse (for example subject 10 was visually impaired). All stimuli were presented at 65 dB (SPL).

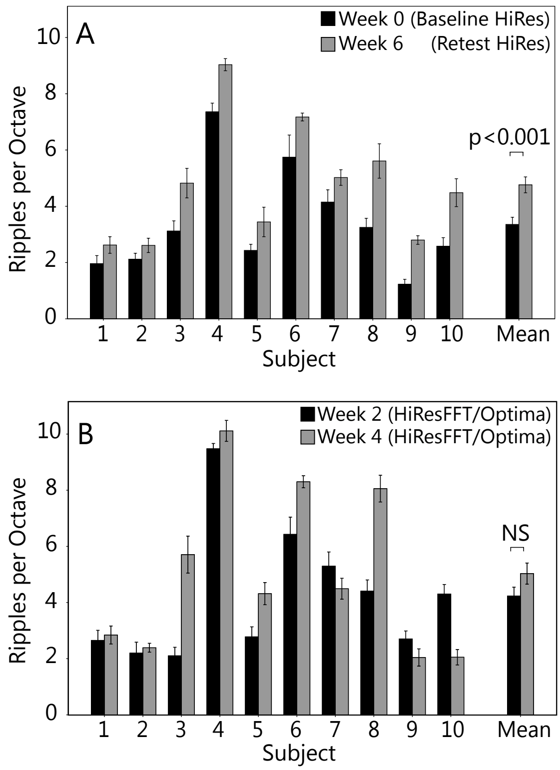


Figure 1. A, Individual and mean spectral ripple thresholds for 10 subjects (HiRes) B, The same as A, now for the HiResFFT and Optima strategies. Error bars represent one standard error of the mean.

Spectral resolution was examined with the spectral-temporally modulated ripple test (SMRT) as developed by Aronoff & Landsberger (2013)[5]. In this adaptive 3-alternative forced choice task, listeners were asked to discriminate a spectrally rippled stimulus, that is, a stimulus that is amplitude modulated in the frequency domain, from a reference stimulus. The reference stimuli had fixed ripple densities of 20 ripples per octave (RPO), whereas the ripple density of the target stimulus was modified until the listener was unable to discriminate between the stimuli. The SMRT differs from previous spectral ripple tests (e.g. Henry & Turner 2003) in that the ripple stimuli are modified[27]. The SMRT uses a spectral ripple with a modulation phase that drifts with time (See fig. 1a in Aronoff & Landsberger 2013)[5], thereby avoiding loudness cues and edge effects. No feedback about the correct answer was given. The procedure was repeated six times per testing run, and the estimated thresholds were averaged as the final SMRT score.

The temporal modulation transfer function (TMTF) test, a 2-alternative forced choice measure of temporal resolution, was used to determine the modulation depth detection threshold (MDT)[8]. Two 1-second intervals consisting of wide band noise were presented to the listener. While the reference stimulus was unmodulated, the target stimulus was amplitude modulated in the time domain with a frequency of 100 Hz and a starting modulation depth of 100%, because these conditions, when combined with spectral ripple thresholds, accounted for the highest amount of speech variance in previous studies[8]. Subjects were instructed to choose the interval that contained the modulated noise after which feedback of the correct answer was provided. A 2-down, 1-up adaptive procedure was used to obtain MDTs in dB relative to 100% modulation [$20\log_{10}$(modulation depth)]. The average of six tracking histories provided the final MDT score.

Statistical analysis

A 2-way repeated measures analysis of variance (ANOVA) using within-subject factors of 'visit' (Week numbers) and 'repetition number' (Repetition number 1-6) were used to determine if there was a main effect of visit, repetition number, and an interaction between those two factors. Because two different strategies were examined in ran-

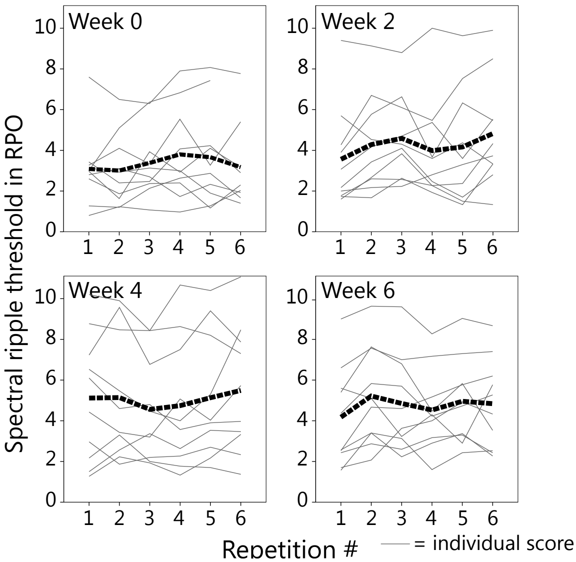


Figure 2. Effects of instantaneous learning for the spectral ripple task. The figure shows individual and mean spectral ripple thresholds as a function of trial number based on data from 10 subjects at 4 test intervals.

domized order in week 2 and 4, those weeks could not be compared to week 0 or 6. Therefore, only week 0 and 6 were compared to each other and week 2 was compared to week 4. SPSS Statistics Version 20 was used for calculations.

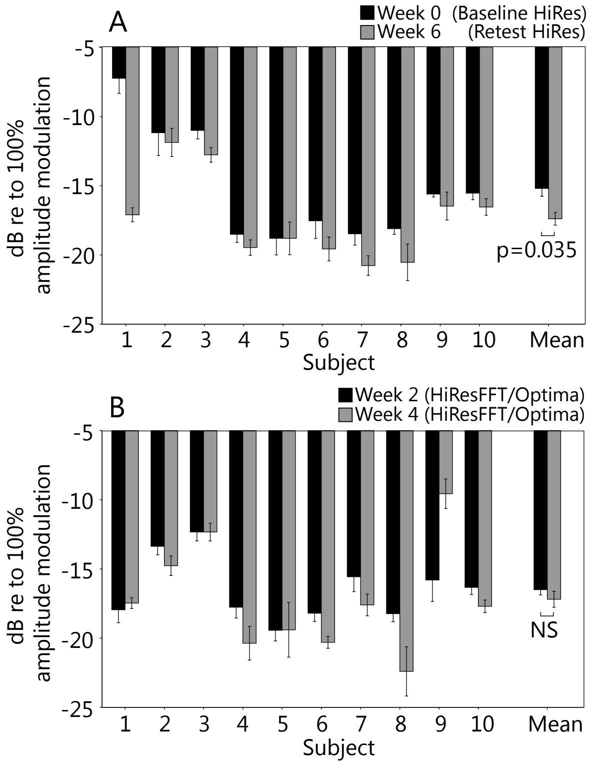


Figure 3. A, Individual and mean modulation detection thresholds for 10 subjects (HiRes) B, The same as A, now for the HiResFFT and Optima strategies. Error bars represent one standard error of the mean.

RESULTS

Individual and mean SMRT scores per test day are demonstrated in Figures 1A and 1B. The average scores of the six repetitions was 3.4 RPO at baseline and 4.2 RPO, 5.0 RPO, and 4.8 RPO at weeks 2, 4, and 6, respectively. The results from a 2-way repeated-measures ANOVA indicated a highly significant improvement between baseline and 6-week SMRT thresholds ($F_{1,9}=52.2$, $p<0.001$), which was present for all ten subjects (Fig.1). No significant effect of repetition number ($F_{5,45}=1.5$, $p=0.195$) or interaction between visit and repetition number ($F_{5,45}=1.398$, $p=0.243$) was observed. There was no significant difference between SMRT scores at week 2 and 4 ($F_{1,9}=1.755$, $p=0.218$) (Fig. 1B). Figure 2 shows the individual and mean SMRT thresholds as a function of trial number at instantaneous testing, i.e. repeating the task on the same test day. A 2-way repeated-measures ANOVA using the Greenhouse-Geisser correction revealed no learning over the course of the six repeated runs on a given test day when all 4 test days were included ($F_{2.4,21.5}=2.347$, $p=0.112$). When comparing the first with the last measurements in the sequence of six, a borderline significant improvement of 0.7 RPO was found ($F_{1,9}=5.012$, $p=0.052$).

Individual and mean MDT scores are shown in Figures 3A and 3B. The mean MDT scores at weeks 0, 2, 4, and 6 were -15.2 dB, -16.5 dB, -17.2 dB, and -17.4 dB, respectively, relative to 100% amplitude modulation. A 2-way repeated-measures ANOVA showed that there was a significant improvement between the first and second six repetitions with the HiRes speech coding strategy, i.e. week 0 versus week 6 ($F_{1,9}$=6.108, p=0.035) (Fig. 3A). No effect of repetition number ($F_{5,45}$=0.965, p=0.449) or interaction between visit and repetition number ($F_{5,45}$=0.483, p=0.787) was observed. As 1 outlier was observed in this analysis (subject 1), the repeated-measures ANOVA was repeated while excluding the outlier, resulting in mean scores of -16.1 dB at baseline and -17.4 dB at 6 weeks. The improvement appeared to still be highly statistically significant ($F_{1,8}$=23.7, p=0.001), and still revealed no effect of repetition number ($F_{5,40}$=1.018, p=0.420) or interaction between visit and repetition number ($F_{5,40}$=0.709, p=0.620). The MDT scores at weeks 2 and 4 were not significantly different from each other ($F_{1,9}$=0.608, p=0.456) (Fig. 3B) and a 2-way repeated-measures ANOVA using the Greenhouse-Geisser correction to adjust for non-sphericity revealed no effect of repetition number when all 4 test days were included ($F_{2.2, 19.9}$=0.967, p=0.405). Moreover, no improvement was observed between the first and last of the six repetitions ($F_{1,9}$=2.289, p=0.165). Altogether, these findings indicated no instantaneous learning effect (Fig. 4). A Pearson product-moment correlation coefficient was computed to assess the relationship between the SMRT and MDT scores, and revealed a significant correlation between the two measures, R^2=0.298, p<0.001.

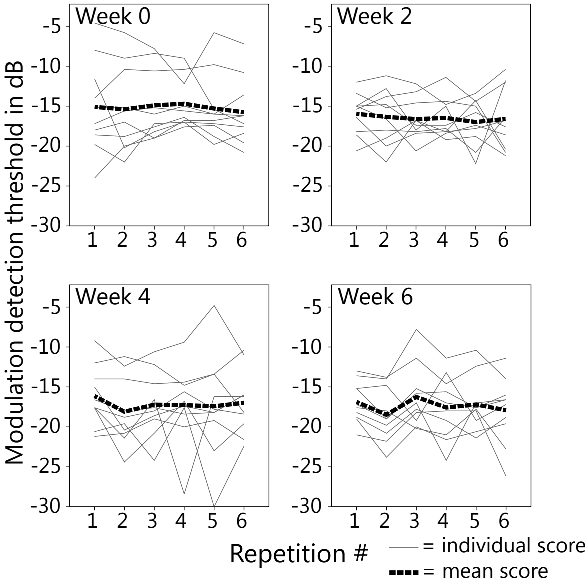


FIGURE 4. EFFECTS OF INSTANTANEOUS LEARNING FOR THE TEMPORAL MODULATION TRANSFER FUNCTION TEST. The figure shows individual and mean Modulation detection thresholds as a function of trial number based on data from 10 subjects at 4 test intervals.

DISCUSSION

The present study demonstrated a clear significant learning effect over time for both the SMRT and the MDT test after repeated examination with the use of different speech coding strategies. Group spectral-ripple discrimination ability improved from 3.4 RPO at baseline to 4.8 RPO at the retest measurement after participation in a clinical trial. The difference was significant on the individual level in five out of ten subjects. The MDT results improved from -15.2 dB to -17.4 dB, a difference of -2.2 dB group-wide in the same time interval. Two out of ten individual participants improved significantly. None of the listeners deteriorated in performance and, in line with previous literature[1,2,17], no learning was observed for either task in an acute setting. Although no instantaneous learning effect was detected in this or previous studies, there are also studies that conclude that there is no learning over time. However, this previous research on learning effects mainly focused on acute settings, and if long-term learning was assessed, either the duration was poorly reported or learning effects after longer time intervals (at least 2 months) were investigated[1,18,28]. During clinical take-home trials, when the presence of a potential learning effect is essential for the interpretation of results, participants are typically exposed to multiple speech coding strategies and execute the psychophysical tasks relatively frequently, e.g., every 2-4 weeks. This makes it essential to identify learning effects in these time frames.

Multiple practice sessions with rather short time intervals introduce two potential risks; perceptual and task learning. Exposure to new speech-coding strategies, and therefore novel speech cues, leads to perceptual learning. Because the population in this study participated in a clinical trial between baseline and retest measurements, they did get a chance to adapt to different speech cues and learn new auditory percepts. This perceptual learning could have potentially been used in the spectral and/or temporal discrimination tasks[23]. On the other hand, speech scores were also assessed during this trial, for which no improvement was observed ($F_{1,9}$=0.826, p=0.387). The lack of a correlation between improvement of the speech scores and the SMRT or MDT scores (R=0.039 and R=0.073 respectively), suggests that the potential effect of perceptual learning is limited. It is reasonable to assume that repeated psychophysical testing in a short period of time could cause task learning. Moreover, it is well-known that perceptual learning amplifies this task learning because of a so-called "carryover effect" [29–31]. A carryover effect is an effect, or ability, that carries over from one experimental condition to another. When time intervals between test sessions are sufficient, like in the study of Drennan ***et al***. (2015)[18], a carryover effect can be considered as (at least partially) extinguished. In other words, a so-called "wash-out period" of sufficient duration compensates for the carryover effect. Moreover, as the purpose of Drennan ***et al***. (2015) was to examine whether basic spectral and temporal discrimination abilities would change over the first year of implant use, they did not vary speech coding strategies. Hence, no perceptual learning induced by the use of novel speech coding strategies could occur. Given the frequency of test intervals in the current study, which was comparable to many other take-home studies[32–34], it is possible that the duration between visits was shorter than the wash-out period and therefore a carryover effect cannot be ruled out. Although it is clear that a learning effect was present for both tests, the current study cannot identify the exact mechanism for this effect. It could be due to task learning, perceptual learning by participating in a clinical trial, or a combination of both factors.

Our results could also partially be explained by the upward trend in motivation of participants, and the placebo effect of any new speech coding strategy. Moreover, the contrasting results found in this study compared with previous work, could, although unlikely, be attributed to the use of different versions of the psychoacoustic measures. For example, Drennan ***et al***. (2014) used a non-adaptive clinical version of the spectral ripple test, which differed considerably from the spectral ripple task that was used in the current study[28]. For example, the current spectral ripple task implemented a temporal effect to avoid potential loudness cues. This resulted in a significant, though fairly low, correlation between SMRT and MDT scores, implying that the SMRT does not purely measure spectral resolution, but is also influenced by temporal effects. This emphasizes the need for an improved measure of spectral resolution, that is less influenced by both loudness and temporal cues.

Because the order in which the two experimental strategies were examined in this study was randomized, an extra analysis between the second and third test day, irrespective of the speech coding strategy, could be performed. No significant difference was found between the test days for either task, implying that no learning, or too little effect size to reach sufficient power, is present after two blocks of testing on separate days. Unfortunately, the number of practice sessions that are necessary for the learning effect to be completely extinguished is unclear and information about the effect size of learning in the MDT test and SMRT is not provided. In that light, it would have been helpful if basic HiRes scores were evaluated at each session, regardless of what condition the subjects were sent home with, although the fatigue that comes with multiple test sessions on one day introduces another bias. A placebo controlled trial, in which participants perform the psychophysical tasks multiple times, on separate test days, with the same speech coding strategy (with a "fake" remapping in which the subject may think that the strategy is different, but in fact is not), would provide us with more specific information. Nevertheless, this study provides us with sufficient evidence that a learning effect is present in the two tasks.

These results do not diminish the value of psychophysical testing for the evaluation of newly developed speech coding strategies; rather, they emphasize the importance of a proper, randomized research design. As a carryover effect could be the cause of learning, it should be dealt with by allowing sufficient time between test dates to "wash-out" the effect of the previous test. Moreover, it is important to incorporate a second testing phase with the baseline speech coding strategy at the end of the trial, in addition to the initial baseline measurement, if one wants to conclude that one of the coding strategies under test is really improving speech perception. In line with this, Donaldson ***et al***. (2011) found a significant improvement in vowel recognition the second time the baseline strategy was evaluated and used these results for comparison with the research strategy[31]. Another example is the study of Vermeire *et al*. (2010)[35], where an experimental strategy was examined acutely and after one, three, six and twelve months of usage. They found a significant improvement in speech intelligibility in noise with the experimental strategy over time. However, switching back to the baseline strategy resulted in a similar improvement (see fig.1. of Vermeire *et al*. (2010)), underlining the importance of comparing speech perception results with a second baseline measurement. This helps to avoid misinterpretation of improvements due to learning effects as true differences between strategies.

CONCLUSIONS

The SMRT and MDT tasks show a clear learning effect over time when examined relatively frequently in a clinical trial. Although an unmistakable explanation has not been shown, these results emphasize the vigilance with which these psychophysical test should be used in clinical trials, for the explicit reason that they are assumed to not change over time. Moreover, great caution with respect to (specifically long-term) learning effects is advised for the development of new psychophysical measures.

ACKNOWLEDGEMENTS

We thank all our participants for their time and effort during this trial. The work was financially supported by Advanced Bionics.

REFERENCES

1. Won, J.H., Drennan, W.R. & Rubinstein, J.T. (2007) Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. ***J. Assoc. Res. Otolaryngol. 8***, 384–392.
2. Drennan, W., Won, J., Nie, K., Jameyson, E. & Rubinstein, J.T. (2010) Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. ***Hear. Res. 262***, 1–8.
3. Won, J.H., Drennan, W.R., Kang, R.S. & Rubinstein, J.T. (2010) Psychoacoustic abilities associated with music perception in cochlear implant users. ***Ear Hear. 31***, 796–805.
4. Anderson, E.S.E., Oxenham, A.A.J., Nelson, P.B. & Nelson, D.A. (2012) Assessing the role of spectral and intensity cues in spectral ripple detection and discrimination in cochlear-implant users. ***J. Acoust. Soc. Am. 132,*** 3925–3934.
5. Aronoff, J.M. & Landsberger, D.M. (2013) The development of a modified spectral ripple test. ***J. Acoust. Soc. Am. 134***, 217–222.
6. Jones, G.L., Drennan, W.R. & Rubinstein, J.T. (2013) Relationship between channel interaction and spectral-ripple discrimination in cochlear implant users. ***J. Acoust. Soc. Am. 133***, 425–433.
7. Shannon, R. V. (1992) Temporal modulation transfer functions in patients with cochlear implants. ***J. Acoust. Soc. Am. 91***, 2156–2164.
8. Won, J.H., Drennan, W.R., Nie, K., Jameyson, E.M. & Rubinstein, J.T. (2011) Acoustic temporal modulation detection and speech perception in cochlear implant listeners. ***J. Acoust. Soc. Am. 130***, 376–388.
9. Fraser, M. & McKay, C.M. (2012) Temporal modulation transfer functions in cochlear implantees using a method that limits overall loudness cues. ***Hear. Res. 283***, 59–69.
10. Büchner, A., Brendel, M. Krueger, ***et al***. (2008) Current steering and results from novel speech coding strategies. ***Otol. Neurotol. 29***, 203–207.
11. Brendel, M., Büchner, A., Krüger, B., Frohne-Büchner, C. & Lenarz, T. (2008) Evaluation of the Harmony soundprocessor in combination with the speech coding strategy HiRes 120. ***Otol. Neurotol. 29***, 199–202.
12. Fu, Q.-J. (2002) Temporal processing and speech recognition in cochlear implant users. ***Neuroreport 13***, 1635–1639.
13. Henry, B., Turner, C. & Behrens, A. (2005) Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. ***J. Acoustal Soc. Am. 118***, 1111–1121.
14. Won, J.H., Jones, G.L., Drennan, W.R., Jameyson, E.M. & Rubinstein, J.T. (2013) Evidence of across-channel processing for spectral-ripple discrimination in cochlear implant listeners. ***J. Acoust. Soc. Am. 130***, 2088–2097.
15. Holden, L.K., Firszt, J.B., Reeder, R.M., ***et al***. (2016) Factors Affecting Outcomes in Cochlear Implant Recipients Implanted With a Perimodiolar Electrode Array Located in Scala Tympani. ***Otol. Neurotol. 37***, 1662–1668.
16. Zhou, N. (2017) Deactivating stimulation sites based on low-rate thresholds improves spectral ripple and speech reception thresholds in cochlear implant users. ***J. Acoust. Soc. Am. 141***, 243–248.
17. Drennan, W.R., Longnion, J.K., Ruffin, C. & Rubinstein, J.T. (2008) Discrimination of Schroeder-phase harmonic complexes by normal-hearing and cochlear-implant listeners. ***J. Assoc. Res. Otolaryngol. 9***, 138–149.
18. Drennan, W.R., Won, J.H., Timme, A.O. & Rubinstein, J.T. (2015) Nonlinguistic Outcome Measures in Adult Cochlear Implant Users Over the First Year of Implantation. ***Ear Hear. 37***, 1–11.
19. Bacon, S.P. & Viemeister, N.F. (1985) Temporal Modulation Transfer Functions in Nor-

2

mal-Hearing and Hearing-Impaired Listeners. ***Audiology 24***, 117–134.
20. Staller, S., Menapace, C., Domico, E., ***et al***. (1997) Speech perception abilities of adult and pediatric Nucleus implant recipients using the Spectral Peak (SPEAK) coding strategy. ***Otolaryngol. - Head Neck Surg. 117***, 236–242.
21. Rouger, J., Lagleyre, S., Fraysse, B., ***et al***. (2007) Evidence that cochlear-implanted deaf patients are better multisensory integrators. ***Proc. Natl. Acad. Sci. 104***, 7295–7300.
22. Ruffin, C., Tyler, S., Witt, S., ***et al***. (2007) Long-Term Performance of Clarion 1.0 Cochlear Implant Users. ***Laryngoscope 117***, 1183–1190.
23. Moberly, A.C., Bhat, J. & Shahin, A.J. (2015) Acoustic Cue Weighting by Adults with Cochlear Implants : A Mismatch Negativity Study. ***Ear Hear. 37***, 465–472.
24. Firszt, J.B. (2003) HiResolution ™ Sound Processing. ***Adv. Bionics White Pap. Sylmar, Calif.*** 1–4.
25. Advanced Bionics. (2012) HiRes™ Optima Clinical Results. ***Advanced Bionics White Paper. Sylmar, California*** 1–2.
26. Firszt, J., Holden, L., Reeder, R.M. & Skinner, M.W. (2009) Speech recognition in cochlear implant recipients: comparison of standard HiRes and HiRes 120 sound processing. ***Otol. Neurotol. 30***, 146–152.
27. Henry, B. & Turner, C. (2003) The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. ***J. Acoustal Soc. Am. 113***, 2861–2873.
28. Drennan, W.R., Anderson, E.S., Won, J.H. & Rubinstein, J.T. (2014) Validation of a clinical assessment of spectral ripple resolution for cochlear-implant users. ***Ear Hear. 35***, 92–98.
29. Liu, Z. (1999) Perceptual learning in motion discrimination that generalizes across motion directions. ***Proc. Natl. Acad. Sci. U.S.A. 96***, 14085–14087.
30. Liu, Z. & Weinshall, D. (2000) Mechanisms of generalization in perceptual learning. ***Vision Res. 40***, 97–109.
31. Donaldson, G., Dawson, P. & Borden, L. (2011) Within-subjects comparison of the HiRes and Fidelity120 speech processing strategies: speech perception and its relation to place-pitch sensitivity. ***Ear Hear. 32***, 238–250.
32. Holden, L., Brenner, C., Reeder, R.M. & Firszt, J.B. (2013) Postlingual adult performance in noise with HiRes 120 and ClearVoice Low, Medium, and High. ***Cochlear Implants Int. 14***, 1–22.
33. Neben, N., Lenarz, T., Schuessler, M., Harpel, T. & Büchner, A. (2013) New cochlear implant research coding strategy based on the MP3(000™) strategy to reintroduce the virtual channel effect. ***Acta Otolaryngol. 133***, 481–90.
34. Frijns, J.H.M., Snel-Bongers, J., Vellinga, D., ***et al***. (2013) Restoring speech perception with cochlear implants by spanning defective electrode contacts. ***Acta Otolaryngol. 133***, 394–9.
35. Vermeire, K., Punte, A.K. & Van De Heyning, P. (2010) Better speech recognition in noise with the fine structure processing coding strategy. ***J. Oto-Rhino-Laryngology, Head Neck Surg. 72***, 305–311.

CHAPTER 3

Take-Home Trial Comparing Fast Fourier Transformation-Based and Filter Bank-Based Cochlear Implant Speech Coding Strategies

Monique A.M. de Jong, Jeroen J. Briaire, and Johan H.M. Frijns

Biomed Research International, 2017

ABSTRACT

Previous studies have demonstrated no improved nor deteriorated speech intelligibility with the HiResolution Fidelity 120 speech coding strategy (HiResF120) over the original HiRes strategy. Improved spectral and deteriorated temporal sensitivity has been shown, making it plausible that the beneficial effect in the spectral domain was offset by the worsened temporal sensitivity. We hypothesize that the implementation of Fast Fourier Transformation (FFT) processing, instead of the traditionally used bandpass filters, explains the reduction of temporal sensitivity. In this study, spectral ripple discrimination, temporal modulation detection and speech intelligibility in noise was assessed in a two-week take-home trial with 3 speech coding strategies; one with conventional bandpass filters (HiRes), one with FFT-based filters (HiRes FFT) and one with FFT-based filters and current steering (HiRes Optima). One participant dropped out due to discomfort with both research programs. The 10 remaining participants performed equally well on all tasks with all three speech coding strategies, implying that FFT processing does not change the ability of CI recipients to discriminate spectral or temporal information, nor speech understanding.

The study was approved by the Medical Ethical Committee of the LUMC (ref. P02.106.Y).

INTRODUCTION

In an attempt to boost cochlear implant (CI) performance, the cochlear implant sound coding strategy "HiResolution Fidelity 120™" (Advanced Bionics, Valencia, CA) (HiResF120) was developed[1]. This strategy implemented "current steering", which facilitates stimulation of auditory nerve regions that are located in between physical electrode contacts. By simultaneously stimulating 2 adjacent electrode contacts with different weights, the peak of excitation shifts between the 2 contacts, creating an intermediate pitch percept[2–4]. Theoretically, this strategy generates up to 120 tonotopic positions, although psychophysical data reveal that most CI users are unable to discriminate such small differences in place pitch[4–7]. Although some studies reported improved speech understanding with HiResF120[8–10], most were not able to demonstrate this[11–16]. Drennan *et al.* compared HiResF120 with the traditional HiRes processing strategy and observed an improved spectral and a decreased temporal resolution, but no benefit for speech intelligibility in noise for HiResF120 users[17]. Also other studies reported an improved spectral resolution with the HiResF120 strategy[5,6,18], which could be attributed to the higher tonotopic precision of stimulation. As temporal cues are important for speech intelligibility in noisy environments[19–22], we hypothesize that the unchanged speech intelligibility in noise is because the beneficial effect in the spectral domain is offset by the reduced temporal sensitivity.

3

The cause of the detrimental effect on temporal discrimination ability with HiResF120 is not known, but the way the frequency analysis is performed to enable current steering, may be involved[17]. In the standard HiRes processing strategy, filter banks are implemented as 6th-order Butterworth band-pass filters in which spectral updating occurs at the pulse rate. To facilitate current steering, a filter bank based on fast Fourier transformation (FFT) is used in HiResF120. These FFT filters provide a detailed spectral profile and are computationally efficient[23], making them of great interest in the implementation of speech-processing designs. However, the 14.7 ms sliding window of these filters (256pts Hamming Window) might cause temporal smearing, resulting in a decrease in temporal resolution. The present study examined the effect of FFT-based filter banks on temporal resolution, spectral resolution and speech perception in noise.

MATERIALS AND METHODS

Subjects

Eleven adults with post-lingual deafness who had been implanted with a HiRes90K device with HiFocus1J or a CII HiFocus with positioner electrode array (Advanced Bionics, Valencia, CA) at the Leiden University Medical Centre (LUMC) participated in this study. All participants clinically used a Harmony processor with the HiRes speech coding strat-

TABLE 1. Subject characteristics

Subject	Gender	Age (y)	Etiology	Deafness (y)	CI side	CI experience (mo)	CVC (Ph%)	Implant
S1	Male	60	Familial	46	Left	160	93	Clarion CII
S2	Male	51	Progressive	25	Right	124	76	HiRes 90K
S3	Female	74	Familial	22	Left	114	93	HiRes 90K
S4	Male	58	Unknown	11	Right	31	93	HiRes 90K
S5	Male	67	familial	67	Right	96	86	HiRes 90K
S6	Male	57	Familial	45	Right	174	96	Clarion CII
S7	Female	67	Meningitis	6	Left + Right	73	88	HiRes 90K
S8	Female	43	Meningitis	5	Left + Right	59	95	HiRes 90K
S9	Female	66	Unknown	35	Right	105	84	HiRes 90K
S10	Female	59	Type-II Usher	4	Right	46	89	HiRes 90K
S11	Female	65	Meniere	25	Right	149	92	Clarion CII

CI, Cochlear Implant; CVC, Dutch phonetically balanced monosyllabic Consonant-Vowel-Consonant words; Ph%, percentage phonemes correct egy. The mean age was 60.6 (range 43 to 74) years, the average duration of deafness was 26.4 (range 4 to 67) years, and the average implant experience was 103 (range 31 to 174) months. The mean phoneme score for open set Dutch monosyllabic (CVC) words during quiet conditions at 65 dB SPL (sound pressure level) was 89.6% (range 76 to 96%) (Table 1). Subject 11 dropped out because of difficulty with the acceptance of the research speech processing strategies and due to a poor attention span.

Speech coding strategies and programming

Participants were tested with 3 different speech-coding strategies, all programmed on a Harmony processor. Strategy 1 (reference) was their standard clinical program, HiRes (Advanced Bionics, Valencia, CA), which is a band pass filter-based strategy. More detailed information about this speech coding strategy is provided by Firszt[24]. The research strategies were HiRes FFT (strategy 2) and HiRes Optima (strategy 3). HiRes Optima is the current clinical standard strategy for Advanced Bionics implants, which is an energy efficient version of HiResF120. It saves energy by limiting current steering to only half of the area between 2 physical electrode contacts[25]. The distribution of current is expressed in alpha (α), where all current is delivered to the most apical electrode at α=0, and to the basal electrode contact at α=1. At α=0.5 current is equally distributed. HiResF120 applies current steering between α=0 and α=1, while HiRes Optima steers between α=0.25 and α=0.75. HiRes FFT (strategy 2) was identical to HiRes Optima, without the implementation of current steering and it uses 16 instead of 15 channels for the FFT (see Table 2 for strategy characteristics).

TABLE 2. Speech coding strategy characteristics

Strategy	Filter Bank	Envelope Extraction	Stimulation mode	Range of Alpha	Spacing of filters
HiResolution	Butterworth	HWR + LPF	Monopolar	-	Logarithmic
HiResolution FFT	FFT-based Filters	Hilbert Envelope	Monopolar	0	Logarithmic
HiResolution Optima	FFT-based Filters	Hilbert Envelope	Dual-electrode	0.25-0.75	Logarithmic

FFT, Fast Fourier Transform; HWR, Half wave rectifier; LPF, low-pass filter

The HiRes MAPs (MAP refers to programmed settings including T- and M-levels, stimulation rate, as well as other parameters) were transferred and adapted from the clinical software Soundwave to the research tool BEPS+ (Bionic Ear Program System+, Advanced Bionics, Valencia, CA), with which the 2 research strategies were programmed. Both strategies 2 and 3 were optimized by applying a pre-set gain profile, in which the signal is progressively attenuated with increasing electrode contact numbers (i.e., more basal electrode contacts). This gain profile results in a less sharp overall sound, thereby increasing the perceptual similarity with the clinical strategy. If the participant reported poor sound quality individual MAPs were adjusted minimally, as is done in clinical practice. In Table 3 the fitting parameters for all subjects are shown. Three subjects (S3, S7 and S9) had up to four electrode contacts switched off in their HiRes MAP due to the clinical practice in our center at the time of hook up. If this was the case, this pattern was copied to the HiRes FFT program. As it is impossible to copy this pattern to the HiRes Optima strategy, and impedances on those electrodes were within normal ranges, the full electrode array was used for HiRes Optima fitting. Subject 1 had clinically switched off electrodes 3 and 4 because of relative high impedances and electrode 6 and 9 according to clinical practice. Only the high impedance electrodes were switched off for the research strategies. Subject S8 (bilaterally implanted) had clinically switched off electrode contacts 14-16 on the right side and 1-3 on the left side to compensate for interaural frequency mismatch caused by different intra-cochlear positions of the 2 electrode arrays. The same electrodes were used for the HiRes FFT and Optima strategies. Bilateral users (S7 and S8) were tested bilaterally.

Protocol

The participants were randomly assigned into 2 groups, which participated in the study in a different order to avoid potential influence from auditory experience with the CI. To avoid outcomes due to learning effects rather than differences in strategy, the psychophysical test protocol was first completed with the HiRes strategy. These results were discarded here, but used in a companion paper on learning effects. Subsequently, 2 weeks of at-home adjustment time was offered with one of the research strategies. When the subject returned, the test-battery was repeated and the other research strategy was fitted on the processor. After another 2 weeks of practice at home, the second research strategy was evaluated. Final measurements with strategy 1 (the HiRes strategy) were obtained 2 weeks after finishing the trial.

Psychophysical testing

All tests were conducted in a double-walled sound-attenuating booth. The sounds were presented at 65dB SPL via a single loudspeaker, placed approximately 1 m from the listener at a straight angle that was in level with the listener's head. A Flemish sentence test (LIST) was used to measure speech reception thresholds (SRT) in speech shaped noise[26]. The standard LIST protocol was followed, but the level of the speech was held constant at 65 dB SPL to avoid loudness effects on speech discrimination. The noise level was adapted via a one-down, one-up procedure with step sizes of 2 dB, starting at 69 dB SPL. Five runs were obtained to determine the average SRT in dB signal to noise ratio (SNR).

TABLE 3. Fitting parameters

	HiRes			HiRes FFT			HiRes Optima		
Subject	# Channels	Pulse width	Pulse Rate (PPS)	# Channels	Pulse width	Pulse Rate (PPS)	# Channels	Pulse Width	Pulse rate (PPS)
S1	12	21.6	1547	12	10.8	1547	12	25.1	1657
S2	16	21.6	1450	16	21.6	1450	15	18.0	1856
S3	12	21.6	1933	12	21.6	1933	15	26.0	1280
S4	16	21.6	994	16	31.4	994	15	43.1	773
S5	16	21.6	1450	16	21.6	1450	15	19.8	1687
S6	16	21.6	1450	16	21.6	1450	15	23.3	1428
S7	12/12	43.1/21.6	967/1933	12/12	43.1/21.6	967/1933	15/15	49.4/21.6	675/1856
S8	13/13	21.6/21.6	1785/1785	13/13	21.6/21.6	1785/1785	12/12	23.3/26.0	1785/1477
S9	12	10.8	1547	12	10.8	1547	15	29.6	750
S10	16	21.6	1450	16	21.6	1450	15	23.3	1428
S11	8	21.6	1450	-	-	-	-	-	-

PPS, pulses per second

To test spectral resolution, the Spectral-temporally Modulated Ripple Test (SMRT) as developed by Aronoff and Landsberger[27] was used. This 1-up, 1-down adaptive, 3 alternative, forced choice task, determines the maximum number of ripples per octave (RPO), e.g. the ripple density, that the listener can distinguish from 20 RPO. In the present study the test was repeated 6 times to determine the average ripple density threshold.

Information about temporal sensitivity was obtained with a 2-down, 1-up adaptive forced choice task as adapted from Won ***et al.***[28]. The modulation frequency of the amplitude-modulated wide band noise was 100 Hz, as this modulation frequency, when combined with ripple thresholds, accounts for the highest amount of speech variance[28]. Six tracking histories were conducted to determine the average modulation detection thresholds (MDTs) in dB relative to 100% modulation.

Subjective assessment

To evaluate the subjective rating of speech coding strategies, the Speech, Spatial and Qualities of Hearing Scale (SSQ) was used[29]. The SSQ questionnaire is a measure for evaluating various aspects of hearing disability, of which the domains 'quality of hearing' and 'speech understanding' were assessed.

Statistical analysis

A 2-way repeated measures ANOVA with within factors 'strategy' (HiRes, HiRes FFT and HiRes Optima) and 'repetition number' (Repetition number 1-5 or 1-6) was used to determine if there was a main effect of strategy, repetition number, and an interaction between those 2 factors. SPSS Statistics Version 20 was used for calculations. A post hoc power analysis was conducted using the software package, G*Power[30]. The alpha level used for this analysis was $p < 0.05$ and the observed correlation among repeated measures were 0.8, 0.5 and 0.75 for the SMRT, MDT task and LIST, respectively. Effect sizes f for the SMRT and MDT task were 0.28 and 0.58, based on data from the study of Drennan ***et al***. (2010)[18]. For the speech in noise task no effect was found by Drennan ***et al***. (2010). Therefore, an effect size of 0.25, which is considered a moderate/clinically relevant effect, was chosen. The analysis revealed that the statistical power to detect the expected effect for the SMRT, MDT and LIST results were 0.89, 0.96 and 0.80, respectively. From these results, we concluded that the statistical power with 10 subjects was sufficient.

RESULTS

The results of the speech in noise test are shown in Figure 1A. Mean SRTs were 1.3 dB SNR for HiRes, 0.96 dB SNR for HiRes FFT, and 1.4 dB SNR for HiRes Optima. The 2-way

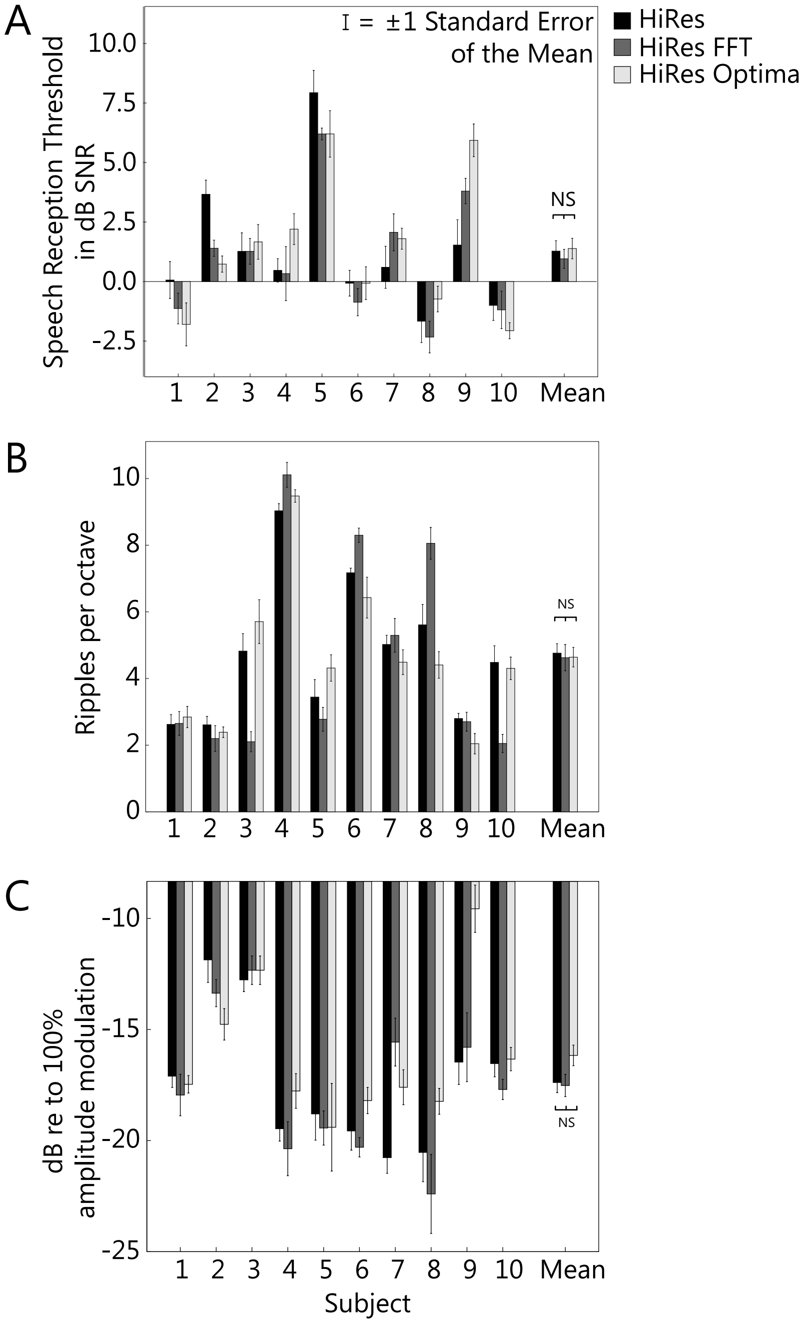


FIGURE 1. INDIVIDUAL AND MEAN PSYCHOPHYSICAL RESULTS. Error bars indicate 1 SD.
A. Speech in Noise intelligibility (LIST)
B. Spectral-ripple discrimination thresholds (SMRT).
C. Amplitude modulation thresholds.

repeated measures ANOVA failed to detect a statistically significant difference between speech coding strategies [Greenhouse-Geisser corrected F(1.23,11.1)=0.396, p=0.585]. Also no significant effect of repetition number [F(4,36)=2.2, p=0.09] or interaction between strategy and repetition number [F(8,72)=0.819, p=0.589] was observed.

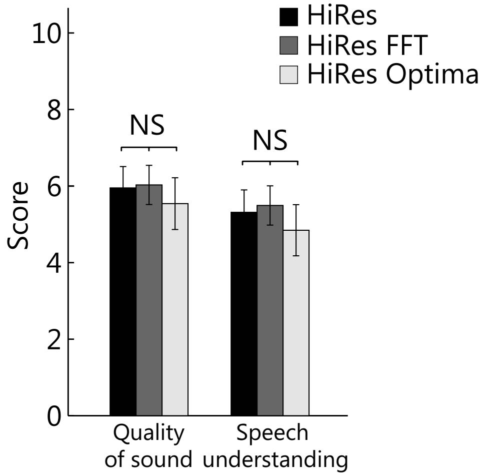


Figure 2: Subjective rating of processing strategies (SSQ) concerning quality of sound (left panel) and speech understanding in different listening conditions (right panel).

The individual and mean results of the spectral ripple test are shown in Figure 1B. Mean SMRT scores were 4.76, 4.63 and 4.64 RPO for HiRes, HiRes FFT and HiRes Optima, respectively. SMRT scores were not statistically significant different across speech coding strategies, [Greenhouse-Geisser corrected F(1.1, 9.9)=0.046, p=0.86]. A significant effect of repetition number [F(45,5)=2.862, p=0.025] was found, whereas no interaction between strategy and repetition number [F(10,90)=0.910, p=0.527] was observed.

Individual and mean results of the MDT test are shown in Figure 1C. The MDTs in dB relative to 100% modulation were -17.38, -17.52, and -16.17 dB for HiRes, HiRes FFT and HiRes Optima, respectively. Although the results were numerically higher (worse performance) with HiRes Optima, there was no statistically significant effect of speech coding strategy, F(2,18)=1.93, p=0.175. No effect of repetition number [F(5,45)=0.973, p=0.445] or interaction between strategy and repetition number [F(10,90)=1,519, p=0.145] was found. An additional paired t-test, comparing the average MDTs of HiRes FFT and HiRes Optima to final HiRes scores was performed, but also this direct comparison between FFT and bandpass filter based strategies could not demonstrate a significant effect (p=0.403). Similarly, no significant effect of current steering on SMRT scores was found when comparing the average SMRT scores for HiRes and HiRes FFT to the HiRes Optima scores with a paired t-test (p=0.882).

The means of the subjective ratings based on a 10-point scale are shown in Figure 2, separated in the quality of sound and speech understanding in different listening situations. On average, subjective quality of sound was rated 5.95, 6.03 and 5.54 with HiRes, HiRes FFT and HiRes Optima [F(2,16)=1,295, p=0.3]. Speech understanding was rated as 5.32, 5.49 and 4.85 respectively [F(2,16)=1.43, p=0.268].

DISCUSSION

This study evaluated 3 sound processing strategies, which used bandpass filters (HiRes), FFT filters (HiRes FFT) or FFT filters and current steering (HiRes Optima), to examine if there is an effect of FFT processing. Speech intelligibility in noise was not statistically significantly different for the 3 speech coding strategies, implying there was minimal influence from the combined changes to the type of filter bank, envelope extraction technique, or use of current steering. Considering the notion that prolonged experience with new strategies increases performance[15], one might argue that the optimal effect was not reached after 2 weeks of exposure to the strategies. Although no benefit has been seen with HiRes FFT and HiRes Optima, it is good to notice that also no acute detriment was observed when switching to these speech coding strategies. Moreover, many other research groups found no or only minor improvements on clinical abilities with HiResF120 as compared to HiRes[8,17], which is in line with our results.

To study the sound processing strategies in more detail, more specific tests were needed. The SMRT and MDT test are tests for spectral and temporal resolution, respectively. Both can be used in an acute setting and are correlated with speech recognition scores over time [28,31–33]. No statistically significant benefit over standard HiRes was observed for spectral ripple discrimination with the HiRes Optima or HiRes FFT strategies, even while more electrode contacts were switched on with the HiRes Optima strategy in some subjects. This is in contrast with previous research, where improved spectral ripple discrimination was observed with HiResF120[17,18]. Also, Firszt *et al*. [2007] reported a decrease in just noticeable difference in pitch[5]. An explanation for our contradictory results might be that we used HiRes Optima, a more energy efficient version of HiResF120. Where HiResF120 applies current steering to the full area between 2 pairs of physical electrode contacts (between α=0, and α=1) HiRes Optima only steers current along part of this area (between α=0.25 and α=0.75). This might explain the decrease in benefit in the spectral domain with HiRes Optima as compared to HiResF120, although no differences in speech understanding between these two strategies was found in a clinical study[25]. This could be explained by the fact that speech in noise tests are not sensitive enough to detect small differences between strategies and fine spectral detail may not be needed to achieve those levels of performance. To confirm the latter explanation, these 2 sound processing strategies (HiResF120 and HiRes Optima) should be investigated more extensively by comparing spectral ripple thresholds.

Although it seemed plausible that temporal smearing, caused by the wider time window of FFT processing, would lead to more difficulties in the temporal domain[17], our results do not confirm this hypothesis. Temporal modulation detection is not statistically significant different between the speech coding strategies tested, although performance was numerically worse with HiRes Optima relative to HiRes FFT (p=0.175). Interestingly, this study showed a significant effect of repetition number within each SMRT test session, contrary to the companion study on learning effects. There, only a borderline significant effect (p=0.052) was observed when comparing the first and last measurements in a sequence of six. However, in that paper, comparison of baseline and 6-week SMRT and TMTF scores revealed a clear learning effect over time. Therefore, baseline HiRes scores were discarded in the present study, and only final HiRes scores were used as a reference for HiRes FFT and HiRes Optima. Nevertheless, it turned out that even if baseline HiRes scores would have been used, no significant effect of speech coding strategy on both SMRT (p=0.071) and MDT (p=0.126) scores could be demonstrated.

CONCLUSION

The present study compared the influence on several aspects of CI performance of FFT-based filter banks and the traditional bandpass filters as used in the HiRes speech processing strategy. Neither detrimental nor beneficial effects were found in spectral and temporal resolution, or speech intelligibility in noise. The known benefits of FFT filters, e.g., their computational efficiency, encourage their implementation in future speech coding strategies.

ACKNOWLEDGMENTS

This research was supported by non-restrictive research funding from Advanced Bionics.

REFERENCES

1. Litvak, L.M., Krubsack, D.A. & Overstreet, E.H. (2003) Method and system to convey the within-channel fine structure with a cochlear implant. ***Adv. Bionics Corp. United States Patent, p.17***
2. Townshend, N., Van Compernolle, D. & White, R.L. (1987) Pitch perception by cochlear implant subjects. ***J. Acoust. Soc. Am. 82,*** 106–115.
3. Frijns, J.H.M., Kalkman, R.K., Vanpoucke, F.J., Bongers, J.S. & Briaire, J.J. (2009) Simultaneous and non-simultaneous dual electrode stimulation in cochlear implants: evidence for two neural response modalities. ***Acta Otolaryngol. 129,*** 433–9.
4. Snel-bongers, J., Briaire, J.J., Vanpoucke, F.J. & Frijns, J.H.M. (2011) Spread of Excitation and Channel Interaction in Single- and Dual-Electrode Cochlear Implant Stimulation. ***Ear Hear. 32,*** 1–10.
5. Firszt, J.B., Koch, D.B., Downing, M. & Litvak, L. (2007) Current Steering Creates Additional Pitch Percepts in Adult Cochlear Implant Recipients. ***Otol. Neurotol. 28,*** 629–636.
6. Koch, D.B., Downing, M., Osberger, M.J. & Litvak, L. (2007) Using Current Steering to Increase Spectral Resolution in CII and HiRes 90K Users. ***Ear Hear. 28,*** 38–41.
7. Donaldson, G.S., Kreft, H. a & Litvak, L. (2005) Place-pitch discrimination of single- versus dual-electrode stimuli by cochlear implant users. ***J. Acoust. Soc. Am. 118,*** 623–626.
8. Firszt, J., Holden, L., Reeder, R.M. & Skinner, M.W. (2009) Speech recognition in cochlear implant recipients: comparison of standard HiRes and HiRes 120 sound processing. ***Otol. Neurotol. 30,*** 146–152.
9. Park, H.J., Lee, S.C., Chun, Y.M. & Lee, J.Y. (2009) HiRes with Fidelity 120 benefit in native speakers of Korean. ***Cochlear Implants Int. 10,*** 85–88.
10. Melo, T., Bevilacqua, M. & Costa, O. (2012) Speech perception in cochlear implant users with the HiRes 120 strategy: a systematic review. ***Braz. J. Otorhinolaryngol. 78,*** 129–133.
11. Brendel, M., Büchner, A., Krueger, B., Frohne-Büchner, C. & Lenarz, T. (2008) Evaluation of the Harmony soundprocessor in combination with the speech coding strategy HiRes 120. ***Otol. Neurotol. 29,*** 199–202.
12. Büchner, A., Brendel, M., Krueger, B., ***et al.*** (2008) Current steering and results from novel speech coding strategies. ***Otol. Neurotol. 29,*** 203–207.
13. Berenstein, C.K., Mens, L.H.M., Mulder, J.J.S. & Vanpoucke, F.J. (2008) Current Steering and Current Focusing in Cochlear Implants: Comparison of Monopolar, Tripolar, and Virtual Channel Electrode Configurations. ***Ear Hear. 29,*** 250–260.
14. Donaldson, G., Dawson, P. & Borden, L. (2011) Within-subjects comparison of the HiRes and Fidelity120 speech processing strategies: speech perception and its relation to place-pitch sensitivity. ***Ear Hear. 32,*** 238–250.
15. Büchner, A., Lenarz, T., Boermans, P.P., ***et al.*** (2012) Benefits of the HiRes 120 coding strategy combined with the Harmony processor in an adult European multicentre study. ***Acta Otolaryngol. 132,*** 179–187.
16. Melo, T.M. De, Bevilacqua, M.C., Costa, O.A. & Moret, A.L.M. (2013) Influence of signal processing strategy in auditory abilities. ***Braz. J. Otorhinolaryngol. 79,*** 629–35.
17. Drennan, W., Won, J., Nie, K., Jameyson, E. & Rubinstein, J.T. (2010) Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. ***Hear. Res. 262,*** 1–8.
18. Won, J.H., Nie, K., Drennan, W.R. & Rubinstein, J.T. (2012) Maximizing the spectral and temporal benefits of two clinically used sound processing strategies for cochlear implants. ***Trends Amplif. 16,*** 201–10.
19. Lorenzi, C., Gilbert, G., Carn, H., Garnier, S. & Moore, B.C.J. (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. ***Proc. Natl. Acad. Sci. U. S. A. 103,*** 18866–18869.

3

20. Hopkins, K., Moore, B.C.J. & Stone, M. a. (2008) Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. ***J. Acoust. Soc. Am. 123,*** 1140–1153.
21. Hopkins, K. & Moore, B.C.J. (2009) The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. ***J. Acoust. Soc. Am. 125,*** 442.
22. Moon, I.J. & Hong, S.H. (2014) What is temporal fine structure and why is it important? ***Korean Journal of Audiology 18,*** 1–7.
23. Cochran, W.T., Cooley, J.W., Favin, D.L., ***et al.*** (1967) What is the Fast Fourier Transform? ***Trans. Audio Electroacoust. 15,*** 45–55.
24. Firszt, J.B. (2003) HiResolution ™ Sound Processing. ***Adv. Bionics White Pap. Sylmar, Calif.*** 1–4.
25. Advanced Bionics. (2012) HiRes™ Optima Clinical Results. ***Advanced Bionics White Paper. Sylmar, California*** 1–2.
26. Van Wieringen, A. & Wouters, J. (2008) LIST and LINT: sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands. ***Int. J. Audiol. 47,*** 348–355.
27. Aronoff, J.M. & Landsberger, D.M. (2013) The development of a modified spectral ripple test. ***J. Acoust. Soc. Am. 134,*** 217–222.
28. Won, J.H., Drennan, W.R., Nie, K., Jameyson, E.M. & Rubinstein, J.T. (2011) Acoustic temporal modulation detection and speech perception in cochlear implant listeners. ***J. Acoust. Soc. Am. 130,*** 376–388.
29. Gatehouse, S. & Noble, W. (2004) The speech, spatial and qualities of hearing scale (SSQ). ***Int. J. Audiol. 43,*** 85–99.
30. Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. ***Behav. Res. Methods 39,*** 175–191.
31. Won, J.H., Drennan, W.R. & Rubinstein, J.T. (2007) Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. ***J. Assoc. Res. Otolaryngol. 8,*** 384–392.
32. Henry, B. & Turner, C. (2003) The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. ***J. Acoustal Soc. Am. 113,*** 2861–2873.
33. Saoji, A. a, Litvak, L.M. & Hughes, M.L. (2009) Excitation patterns of simultaneous and sequential dual-electrode stimulation in cochlear implant recipients. ***Ear Hear. 30,*** 559–567.

CHAPTER 4

Dynamic Current Focusing:

A Novel Approach to Loudness Coding

in Cochlear Implants

Monique A.M. de Jong, Jeroen J. Briaire, and Johan H.M. Frijns

Ear and Hearing, 2019

ABSTRACT

Objectives

In an attempt to improve spectral resolution and speech intelligibility, several current focusing methods have been proposed to increase spatial selectivity by decreasing intra-cochlear current spread. For example, tripolar (TP) stimulation administers current to a central electrode and uses the two flanking electrodes as the return pathway, creating a narrower intra-cochlear electrical field and hence increases spectral resolution as compared to monopolar (MP) stimulation. However, more current is required and in some patients, specifically the ones with high electrode impedances, full loudness growth cannot be supported because of compliance limits. The present study describes and analyses a new loudness encoding approach, which uses TP stimulation near threshold and gradually broadens the excitation (by decreasing compensation coefficient σ) to increase loudness without the need to increase overall current. It is hypothesized that this dynamic current focusing (DCF) strategy increases spatial selectivity, especially at lower loudness levels, while maintaining maximum selectivity at higher loudness levels, without reaching compliance limits.

Design

Eleven postlingually deafened adult CI recipients, with at least 9 months of experience with their HiRes90K implant, were selected to participate in this study. Baseline performance regarding speech intelligibility in noise (Dutch matrix sentence test), spectral ripple discrimination at 45 and 65 dB and temporal modulation detection thresholds were assessed using their own clinical program, fitted on a Harmony processor. Subsequently, the DCF strategy was fitted on a research Harmony processor. Threshold levels were determined with σ=0.8, which means 80% of current is returned to the flanking electrodes and the remaining 20% to the extra-cochlear ground electrode. Instead of increasing overall pulse magnitude, σ was decreased to determine most comfortable loudness. After 2-3 hours of adaptation to the research strategy, the same psychophysical measures were taken.

Results

At 45 dB, average spectral ripple scores improved significantly from 2.4 ripples per octave (RPO) with their clinical program to 3.74 RPO with the DCF strategy (p=0.016). Eight out of eleven participants had an improved spectral resolution at 65 dB. Nevertheless, no significant difference between DCF and MP was observed at higher presentation levels. Both speech in noise and temporal modulation detection thresholds were equal for MP and DCF strategies. Subjectively, two participants preferred the DCF strategy over their own clinical program, two preferred their own strategy, while the majority of the participants had no preference. Battery life was decreased and ranged

from 1.5 - 4 hours.

Conclusions

The DCF strategy gives better spectral resolution, at lower loudness levels, but equal performance on speech tests. These outcomes warrant for a longer adaptation period to study long term outcomes and evaluate if the outcomes in the ripple tests transfer to the speech scores. Further research, e.g., with respect to fitting rules and reduction of power consumption, is necessary to make the DCF strategy suitable for routine clinical application.

The study was approved by the medical ethical committee of the LUMC (REF. P02.106. AA)

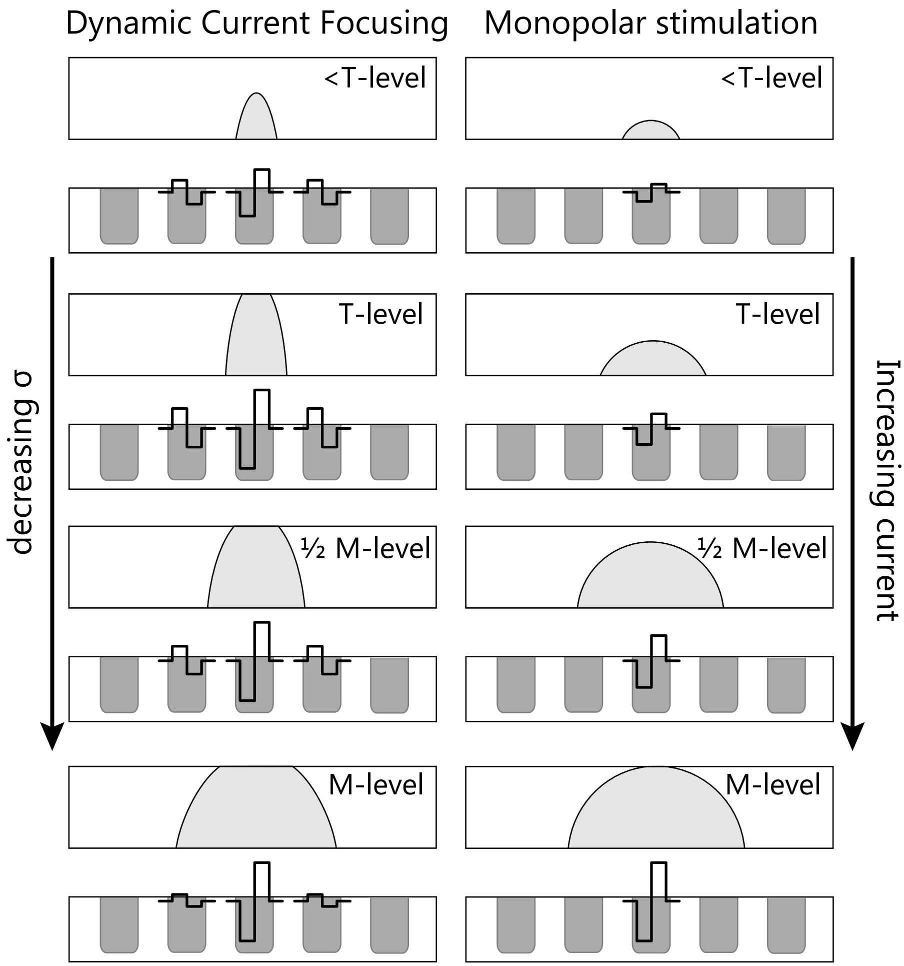


FIGURE 1. THE CONCEPT OF LOUDNESS CODING IN THE DYNAMIC CURRENT FOCUSING (DCF) STRATEGY AND IN MONOPOLAR (MP) STIMULATION MODE. The upper bar for each loudness step shows the auditory nerve with the excitation shown pattern in grey. The lower bars show the implanted electrode array with the electrode contacts in grey. In the DCF strategy, the amplitudes of the main and neighbouring electrode contacts are increased equally up to the threshold level (T-level). To increase the loudness from the T-level, σ is decreased as a function of the stimulus level, resulting in a broader excitation pattern and, accordingly, in a higher loudness level. In MP mode, the amplitude of the main electrode contact is increased as a function of the stimulus level, resulting in broad current spreads at all loudness levels. M-level, most comfortable level.

INTRODUCTION

Although average speech understanding has improved in cochlear implant (CI) users in recent decades due to improved CI technology, patients who are implanted with the same device show large variability in speech understanding[1]. This is, at least in part, due to differences in the abilities of the CI users to resolve spectral contrast, also termed spectral resolution. Spectral resolution can be measured with spectral ripple tests, like the recently developed spectral-temporally modulated ripple test (SMRT)[2]. Several studies revealed that performance on the SMRT is correlated with speech understanding, specifically in difficult listening conditions[3–5]. Moreover, this relation between spectral resolution and speech perception in noise seems to hold across different spectral ripple tests[6–8].

Spectral resolution can be limited by poor spatial selectivity, the degree of spread of neural activity across cochlear place[9]. Spatial selectivity is influenced by the electrode-neuron interface i.e. by how individual electrode contacts interact with the auditory nerve[10]. Notably, there is large variability in measures of spatial selectivity, both between and within subjects[11]. Two components of the electrode-neuron interface are thought to underlie these inter-subject differences: (1) the electrode-to-neuron distance and (2) spiral ganglion survival[12,13]. The distance between electrode contacts and their corresponding neurons is influenced by the electrode design, the surgical placement of the implant[14,15], the insertion depth[16], and bone and tissue growth within the cochlea. Spiral ganglion cell count is determined mainly by the underlying cause and duration of the hearing loss[17].

CIs have multiple electrode contacts along the scala tympani, and each electrode is potentially capable of electrically stimulating a different sub-population of the surviving auditory neurons in the cochlea. These contacts are usually stimulated in so-called monopolar (MP) mode in which the current is returned to a far-field electrode contact. As a result, the electrical potential field patterns are broad. This causes neighbouring electrodes to activate overlapping populations of neurons, especially if the electrode-to-neuron distance is substantial, since this decreases spatial selectivity and reduces the number of spectral channels that can be distinguished[12]. Poor spiral ganglion survival can be addressed in part by increasing the current amplitude, although this in turn increases the current spread and can therefore exacerbate the issue described above.

In an attempt to improve spatial selectivity, and therefore spectral resolution and speech intelligibility, several current focusing methods have been proposed that increase spatial selectivity by reducing intra-cochlear current spread. Computer modelling data[18], as well as animal[19] and human[12,20–25] data, show that current focusing creates a narrower intra-cochlear electrical field and hence increases spectral resolution compared to MP stimulation. For example, tripolar (TP) stimulation administers current to a central electrode and uses two flanking electrodes as the return pathways. This current focusing strategy improves speech understanding, with some researchers suggesting that especially poor performers would benefit from TP stimulation[26,27]. However, this gain comes at the expense of an increased amount of current required to achieve a given loudness. In some patients, specifically in patients with high electrode impedances, only part of the dynamic range can be covered within the compliance limit of the implant[19,20,28]. Moreover, because the stimulus level partly determines the current spread, the benefit in the spectral domain may be compromised at higher loudness levels[29].

The problem of limited loudness growth was addressed by the introduction of a partial tripolar (pTP) strategy in which only a fraction σ (called the compensation coefficient) of the current is returned to the flanking electrodes[21,28]. In fact, this strategy uses a superposition of the MP and TP stimulation strategies. Relative to the TP strategy, the pTP strategy results in greater loudness at the expense of less selective stimulation. Despite the reduced level of current focusing, pTP stimulation improves spectral ripple discrimination[21,24], while speech perception showed to be improved in some[30,31], but not in all studies[21,27]. Nogueira ***et al***. (2017) developed a stimulation mode called the Dynamically Compensated Virtual Channel (DC-VC) in which four adjacent electrodes are stimulated simultaneously to decrease power consumption[32]. Although this quadrupolar strategy saves power, it also generates broader electrical fields, specifically at higher loudness levels. To compensate for this, current focusing is applied by sending current of opposite polarity to the two outer electrode contacts. Loudness balancing experiments with different degrees of current focusing revealed that higher degrees of current focusing result in significantly higher current levels that are required to maintain equal loudness.

The present study describes and analyses a new approach to loudness encoding that is called “dynamic current focusing” (DCF). Previous research showed that loudness is increased at fixed current levels by lowering the degree of current focusing[33]. The DCF uses pTP stimulation near the threshold, and it gradually broadens the excitation by decreasing the compensation coefficient σ in order to increase loudness without the need to increase the overall current (see Figure 1 for a schematic overview). In practice, this means that more current is consumed at low than at high loudness levels, because the current levels for the flanking electrodes are lowered with increasing loudness, while the level on the center contact remains the same. It is hypothesized that the DCF strategy increases spatial selectivity, especially at lower loudness levels, while maintaining the most optimal selectivity possible at higher loudness levels and staying within device compliance limits. This optimal spatial selectivity across the dynamic range

TABLE 1. Characteristics of the study subjects

Subject	Gender	Age (y)	CI side	Duration of implant use (y)	CVC (Ph%)	Electrode array	Clinical processor	Clinical strategy	Clinical settings		DCF settings			
									Rate	Pulse width	Rate	Pulse width	T-level σ	M-level σ [range]
S01	Male	60	Left	9	90	HiFocus 1J	Harmony	Optima-S	2475	26.9	479	74.5	1.0	0.68 [0.6-0.8]
S02	Female	65	Right	9	85	HiFocus 1J	Naida CI Q70	Optima-S	3093	21.6	612	58.4	0.9	0.74 [0.65-0.68]
S03	Male	79	Left	5	86	HiFocus 1J	Harmony	HiRes-S-F120	3093	21.6	510	70.0	0.8	0.34 [0.24-0.44]
S04	Male	70	Left	5	95	HiFocus 1J	Neptune	HiRes-S-F120	3093	21.6	594	60.2	0.8	0.40 [0.31-0.54]
S05	Male	60	Right	7	81	HiFocus 1J	Naida CI Q70	HiRes-S-F120	1485	44.9	710	50.3	0.9	0.46 [0.39-0.49]
S06	Male	48	Left	3	99	HiFocus MS	Neptune	HiRes-S-F120	3093	21.6	1657	21.6	1.0	0.21 [0-0.57]
S07	Female	75	Right	7	88	HiFocus 1J	Naida CI Q70	Optima-S	2007	33.2	594	60.2	0.9	0.53 [0.47-0.64]
S08	Female	76	Left	9	78	HiFocus 1J	Naida CI Q70	Optima-S	3093	26.0	710	50.3	0.9	0.73 [0.69-76]
S09	Male	65	Right	2	90	HiFocus MS	Naida CI Q70	Optima-S	2750	24.2	510	70.0	0.9	0.54 [0-0.80]
S10	Female	71	Right	1	95	HiFocus MS	Naida CI Q70	Optima-S	3093	21.6	1985	18.0	0.8	0.29 [0.07-0.42]
S11	Male	75	Right	15	96	HiFocus 1J	Neptune	HiRes-S	773	43.1	621	57.5	0.8	0.46 [0.36-0.54]

CVC, Dutch phonetically balanced monosyllabic Consonant-Vowel-Consonant words; DCF, dynamic current focusing; Ph%, percentage phonemes correct on a standard monosyllabic (CVC) word test at 65 dB; T-level σ, σ level at threshold level; M-level σ, average σ and range at most comfortable level over all 14 electrode contacts.

4

would then lead to an improved spectral resolution, primarily at lower loudness levels. Here we evaluated the novel DCF loudness coding strategy in 11 subjects in an acute setting in which all tasks were performed with both their conventional clinical strategy and the DCF strategy on a single day. Spectral resolution was assessed with the SMRT, which measures the spectral ripple density threshold. In addition, temporal modulation detection thresholds, speech intelligibility in noise and loudness growth functions were assessed.

MATERIALS AND METHODS

Listeners

Eleven postlingually deafened adults who were unilaterally implanted with an Advanced Bionics HiRes90K implant for at least 9 months were selected to participate in this study. None of the subjects had functional hearing in the contralateral ear. Because the DCF strategy uses multipolar stimulation, only CI users with all 16 electrode contacts working were included in the study. The study included 4 women and 7 men aged 48 to 79 years. Of these, 1 subject used HiResolution (HiRes)[34] in the clinical setting, 4 subjects used HiRes Fidelity 120 (HiResF120)[35] and 6 subjects used HiRes Optima[36]. Table 1 shows the patients' clinical characteristics.

After assessment of baseline performance with their clinical strategy fitted on a dedicated Harmony sound processor in the laboratory, the DCF strategy was fitted on this research processor. The subjects had 2 to 3 hours to adjust to the experimental strategy by taking a break in a busy restaurant, during which they were actively communicating with the researcher. After this, their performance with DCF was evaluated on the same day. The only exception was S01, who was unable to perform all psychophysical tasks in a single day due to fatigue. For this subject, the DCF strategy was evaluated after 10 extra minutes of adaptation on a separate test day. The subjects were not blinded to the tested speech coding strategy as the subjects could easily detect their normal strategy, and the order of the tested strategy was not randomized.

Fitting procedures

Clinical strategy

The data from each subject's last clinical visit was copied from the SoundWave™ program (Advanced Bionics, Valencia, CA, USA) and fitted on a Harmony research processor. If the noise cancelation features were active in their every day program, they were turned off for the testings.

Research strategy

The concept of the DCF strategy is schematically displayed and compared to MP stimulation in Figure 1. The DCF program was created for each subject using BEPS+ software (Advanced Bionics, Valencia, CA, USA). Threshold levels (T-levels) were determined for each electrode contact by gently increasing the total amount of current at σ=0.8, meaning that 80% of the current is returned equally to the two flanking electrodes and 20% is returned to the extra-cochlear electrode. To ensure that sufficient loudness is achieved if σ is reduced to zero (i.e., MP stimulation on the central electrode contact), current on the central electrode contact had to be at least 300 clinical units (CU). CU represent constant charge, which means that automatic adjustments in pulse width also result in automatic adjustments in pulse amplitude in order to maintain constant charge: (amplitude (in μA) · pulse width (in μs) · k (scaling constant)). A level of 300 CU was chosen because in the clinical fittings (in MP stimulation mode) all subjects' most comfortable level (M-level) values were below 300 CU. It was therefore expected that the full dynamic range could be covered by varying the compensation coefficient σ, and not increasing the current level on the center electrode contact above 300 CU. If the total current necessary for the T-level was below 300 CU at σ=0.8, the T-level was determined again, using a compensation coefficient σ=0.9. If total current was still below 300 CU, T-levels were determined using σ=1.0. The used σ values at T-level are depicted in Table 1.

4

Next, the M-levels for the DCF strategy were determined by gradually decreasing σ in steps of 0.01, while the current levels on the central electrode contacts were kept constant, thereby broadening the excitation pattern. As a result, the dynamic range is defined by variations in σ and the subjective loudness at T- and M-level is perceived similar to that with their clinical strategy. To verify the latter, subjects were asked if the loudness level with the speech program turned on was similar to that with their regular strategy. If this was not the case, the loudness level was adjusted accordingly. No loudness balancing per electrode contact was applied. Low power modes (e.g. automated power management, and reduced maximum power mode) were turned off to avoid potential difficulties with power management, as well as noise reduction algorithms. Because 3 physical electrode contacts are required to create 1 current focusing channel, the 2 outer electrodes could not be used. Therefore, the DCF strategy had only 14 effective channels while the clinical strategy had 16 (HiRes) or 120 theoretical (current steering) spectral channels.

Psychophysical tasks

Loudness growth functions

For both stimulation strategies, a loudness scaling experiment was performed at three different locations along the electrode array (electrodes 3, 9 and 14). Because the two stimulation strategies use different mechanisms to achieve M-level, it is impossible to compare them in the same quantity when the electrode contacts are directly stimulated.

In MP stimulation the driving factor to achieve M-level is current, while there is not one individual current or electrode contact to which you can link loudness growth in DCF stimulation. Therefore, the "acoustic" stimuli, that were generated using a custom MATLAB program, were presented via a direct input system into the Harmony speech processor. In this way, the same stimuli (sine waves with frequencies corresponding to the center frequencies of electrode 3, 9 or 14, and with amplitude V) could be generated for the two stimulation strategies, after which they were processed with the use of either the MP or DCF speech coding strategy. The stimulus levels were calculated as follows:

$$\text{Loudness Level (in dB)} - 20 \cdot \log_{10}\left(\frac{V}{V_{ref}}\right) \qquad \text{(eq. 1)}$$

with V_{ref} = 10 µV. The level was slowly increased in step sizes of 5 dB, starting from 0 dB and never exceeding 100 dB to avoid overstimulation. Loudness was subjectively rated on an 8-point loudness scale as used in our previous current focusing experiments [33]. This loudness scale ranged from the threshold level (1), to the most comfortable loudness (5) to the upper limit of comfortable loudness (8); after this, the experiment was terminated [37]. The experiment was repeated three times. To quantify the slope of the loudness scaling curves, the areas under the curves (AUCs) were calculated as follows:

$$\text{AUC} = \sum_{i=2}^{4}\left(\Delta LL_i \frac{SL_i + SL_{i+1} - SL_2}{2}\right) \qquad \text{(eq. 2)}$$

With i = subjective loudness level, which was ranging from 2 to 5 (from 'very soft sound' to 'most comfortable loudness') as these levels were considered to be the most important for understanding speech. SL = Subjective loudness, ΔLL = difference in loudness level (in dB) between SL_i and SL_{i+1}. Differences in AUCs (ΔAUC) between the two strategies were expressed as a percentage relative to the AUC for the clinical strategy ($AUC_{clinical}$):

$$\Delta\text{AUC}[\%] = \frac{AUC_{DCF} - AUC_{clinical}}{AUC_{clinical}} \cdot 100\% \qquad \text{(eq. 3)}$$

with AUC_{DCF} = the AUC for the DCF strategy. The offset of the loudness growth functions was measured at the level (in dB) at subjective loudness level 2 for each electrode contact and stimulation strategy.

$$\Delta\text{offset[dB]} \sum_{e \in S} (\text{offset}_{DCF\,e} - \text{offset}_{clinical\,e})/3 \qquad \text{(eq. 4)}$$

With S = 3, 9 or 14, referring to electrode contacts (e), and where Δoffset was calculated per stimulation strategy as the average difference in offset of the three electrode contacts.

Because of the pilot-like nature of this study, the research protocol was fine-tuned during the trial. Accordingly, the first three subjects (S01, S02 and S03) had a slightly different research set-up than subsequent subjects. For these subjects, loudness curves were obtained as a function of the current in μA for the clinical strategy and as a function of σ in the DCF strategy, using a direct connection to the implant. Because this made it impossible to compare the loudness growth functions between the strategies, we switched to the direct connection to the speech processer described above, and discarded the loudness scaling data for the first three subjects.

Psychophysical tasks - sound booth testing

The tasks described below were performed in the free field, with subjects seated 1 meter away from the front of a single loudspeaker in a double-walled sound-attenuating booth. To minimize the impact of learning effects on the results of the psychophysical tasks, the subjects went through a dry run before each actual test run.

Spectral ripple test

The Spectral-temporally Modulated Ripple Test (SMRT)[2] was used to determine spectral ripple thresholds at 45 dB and 65 dB. It is a three-alternative forced choice (3AFC) task that determines the maximum ripples per octave (RPO) that a listener can differentiate from 2 reference stimuli that have ripple densities of 20 RPO. Study subjects were asked to indicate the deviant stimulus from the 3 sounds without receiving feedback about the correct answer. This spectral ripple density test was chosen because, as opposite to the previously existing spectral ripple test[6,8], the SMRT was designed to avoid a number of potential confounders, like cues that are related to local loudness or the spectral center of gravity[2]. Moreover, the SMRT has been shown to correlate with speech recognition in noise in a variety of CI users[4]. The procedure was repeated 6 times per condition and the average thresholds were calculated.

Temporal modulation detection test

Temporal resolution was assessed with the temporal modulation detection test, adapted from Won ***et al***. (2011)[38]. The two-alternative adaptive measure has two wideband noise stimuli, one without amplitude modulation and one with a modulation frequency of 100 Hz that was adaptive in modulation depth. A modulation frequency of 100 Hz was chosen because this task, along with spectral ripple thresholds, accounts for the highest amount of variance in CNC word scores[38]. The subjects were asked to identify the interval that contained the modulated noise and were then given feedback about whether this was the correct answer. The task was performed at 65 dB and repeated 6 times per condition, then the average modulation detection thresholds (MDTs) were calculated in dB relative to 100% modulation ($20 \cdot \log_{10} \cdot$ modulation depth).

TABLE 2. Differences in the area under the loudness curves (ΔAUC) between the strategies for each subject at the indicated electrodes. The (absolute) average AUC of the three electrode contacts are also shown and are expressed as the percentages of the AUC for the MP strategy ($\Delta AUC[\%] = AUC_{DCF} – AUC_{clinical} / AUC_{clinical} \cdot 100\%$)

Subject	Electrode 3	Electrode 9	Electrode 14	Absolute average
S04	-20%	+17%	+13%	17%
S05	-58%	-26%	-28%	37%
S06	+23%	+20%	-8%	17%
S07	-14%	-10%	+45%	23%
S08	-23%	-19%	-1%	14%
S09	-82%	-60%	+161%	101%
S10	+72%	-26%	+117%	72%
S11	+4%	+7%	+15%	9%

Speech-in-Noise test

The Dutch matrix sentence test is an adaptive speech-in-noise test that uses 50 unique words that are combined into 200 grammatically equivalent sentences, which are grouped into 10 balanced lists[39]. At each test round, the sentences are randomly selected from the subset. The task was carried out using the APEX 3 program (Leuven, Belgium)[40], installed on a personal computer. After the presentation of a sentence, the subjects are asked to repeat the 5 words and to guess if they are not sure. Testing was done at a fixed speech level of 65 dB and with the adaptive speech-shaped noise starting from -4 dB signal-to-noise ratio (SNR). The outcome measure of the matrix test is the speech reception threshold (SRT), which is defined as the SNR at which 50% of the words are repeated correctly. An average SNR was calculated over 3 repetitions to determine the final SRT score. During the matrix task practice session, subjects were exposed to all possible words.

Subjective rating

The subjective quality of the incoming sound in terms of overall loudness, loudness growth, sound clarity, speech understanding, etc. was discussed with the subjects. In addition, all subjects were asked whether they would be able to function normally with this new program in their home situation and if their overall rating of the DCF strategy was better, equal to or worse than their clinical program.

Statistical analysis

Repeated measurements were obtained in all experiments; therefore, two-way repeated measures analysis of variance (ANOVA) was used for the statistical analysis. The factors, 'speech coding strategy' (clinical and DCF) and 'repetition number' (1—3 for

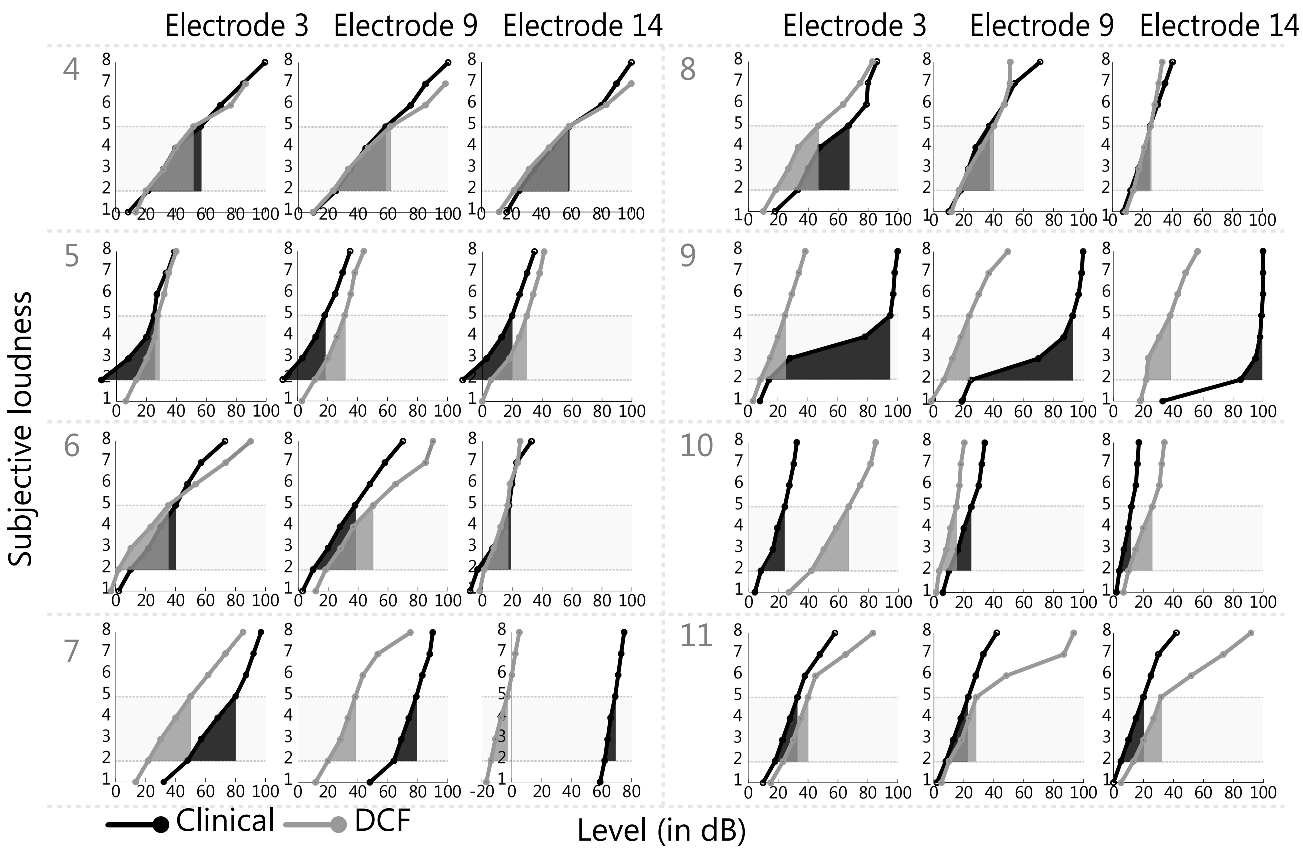


FIGURE 2. THE INDIVIDUAL LOUDNESS GROWTH FUNCTIONS OF S04 TO S11 ON AN 8-POINT LOUDNESS SCALE (Y-AXES). The stimulus levels were calculated as follows: Loudness Level (in dB) - $20 \cdot \log_{10}\left(\frac{V}{V_{ref}}\right)$, with $V_{ref} = 10\ \mu V$. For understanding speech, the most important loudness levels are considered to range is from "2" (very soft sound) to "5" (the most comfortable loudness). This range is shaded grey. The areas under the curves (AUCs) are patterned for the two strategies. DCF, dynamic current focusing.

the speech in noise test and 1-6 for the spectral ripple and the temporal modulation detection test) were used to determine if there was a main effect of the speech coding strategy, repetition number and interaction between the two factors. IBM® SPSS Statistics for Windows, Version 23.0 was used for calculations.

RESULTS

Loudness growth functions

The individual loudness growth functions of S04 to S11 were obtained for the clinical and the DCF strategies. The loudness scores are plotted as a function of the presented stimulus level in dB (Figure 2). The AUCs are depicted by the filled areas under the curves. Both the slope (expressed in AUC) and the offset of the loudness growth curves with the DCF strategy showed considerable deviations from the corresponding values with the clinical strategy in some subjects. The individual differences in AUCs (ΔAUC) per electrode as calculated from the data are presented in Table 2. S05, S09 and S10 showed the largest ΔAUCs, while the loudness growth functions of S04, S06, S07, S08 and S11 were quite similar in terms of the AUCs. Table 3 shows the differences in offsets (Δoffset) of the loudness growth curves per electrode contact, where the offset at loudness level 2 with the clinical program was subtracted from the one with DCF. It turned out that S05, S07, S09 and S10 showed large discrepancies between the offsets with the two strategies, when compared to the other subjects. For S05 and S10 Δoffset was negative, meaning that for DCF a higher input was required to reach the offset.

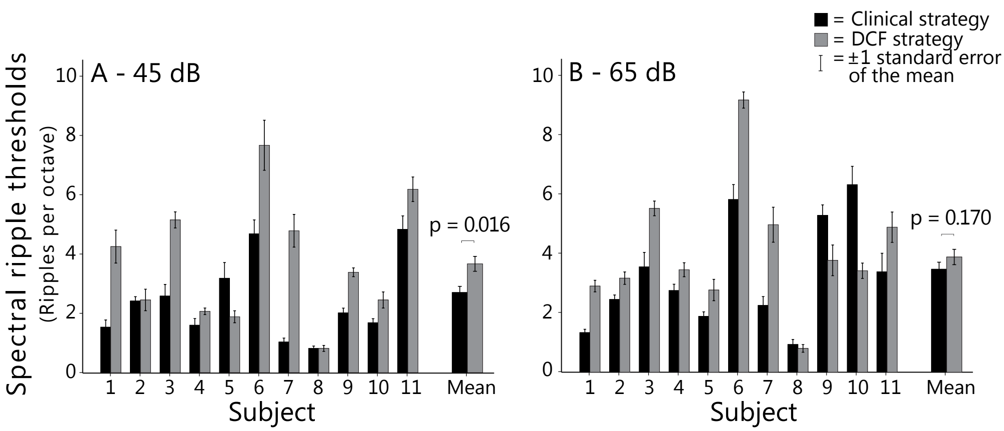


Figure 3. Individual and mean spectral ripple thresholds at 45 dB SPL (A) and at 65 dB SPL (B) for the 11 study subjects using their clinical strategy and the dynamic current focusing (DCF) strategy. The error bars represent ± 1 standard error of the mean.

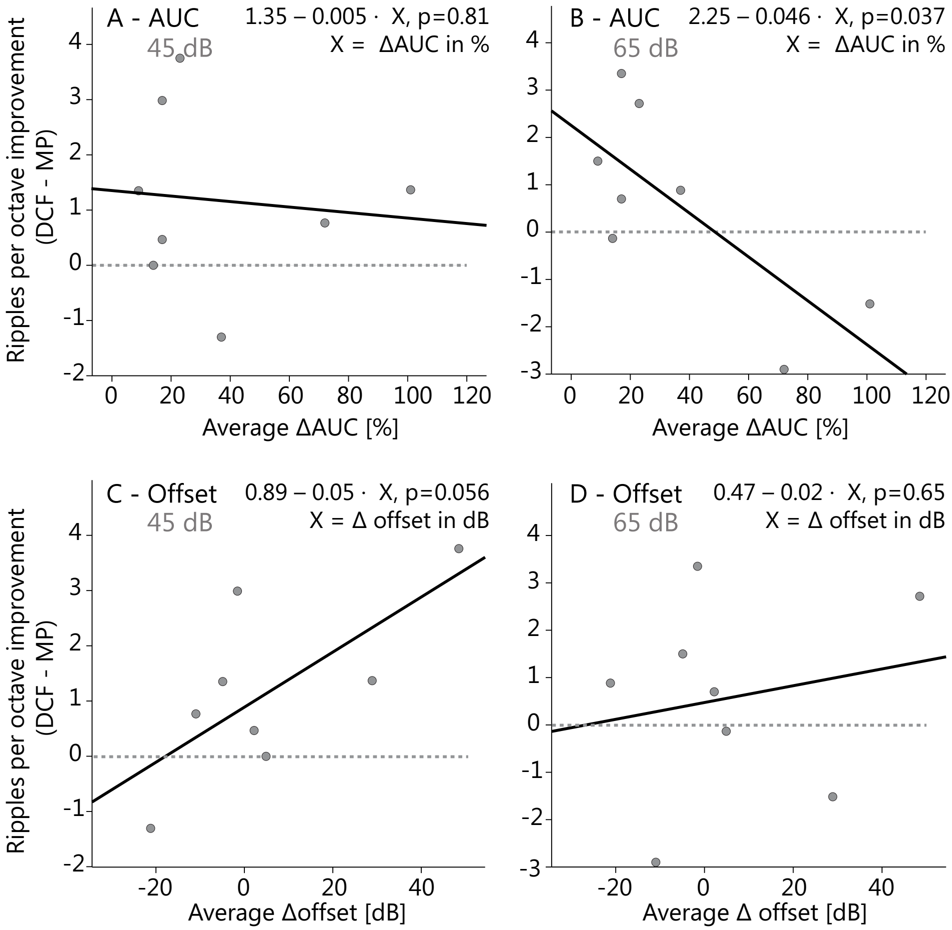


Figure 4. Correlation between deviating loudness growth and improvement on the spectral ripple task at 45 dB SPL and 65 dB SPL. Deviating loudness growth was expressed as the absolute average delta area under the cure of the three electrode contacts ($\Delta AUC[\%] = AUC_{DCF} - AUC_{clinical}/AUC_{clinical} \cdot 100\%$) (A, B) and the average difference in offset of the loudness growth curves ($\Delta offset[dB] = \sum_{e \in S}(offset_{DCF\,e} - offset_{clinical\,e})/3$) with dynamic current focusing (DCF) and the clinical program (C, D).

Spectral ripple test

The individual and mean spectral ripple discrimination thresholds are shown in Figure 3. At 45 dB, the mean thresholds for the clinical and the DCF strategies were 2.40 RPO and 3.74 RPO, respectively; at 65 dB, the values were 3.27 RPO and 4.07 RPO, respectively. As shown in Figure 3A, 8 of 11 subjects showed an improved spectral resolution with the DCF strategy relative to their clinical strategy at 45 dB. A two-way repeated measures ANOVA showed that the average difference was statistically significant at 45 dB for 1.34 RPO: $F(1,10)=8.369$, $p=0.016$. No effect was observed for the repetition number, $F(5,50)=0.080$, $p=0.995$, or for the interaction between strategy and repetition number, $F(5,50)=1.069$, $p=0.389$). At 65 dB, also 8 of 11 subjects showed improved spectral ripple thresholds (Figure 3B). No significant difference was observed between DCF and MP at 65 dB ($F(1,10)=2.186$, $p=0.170$). This result may be explained by the less focused stimulation provided by DCF at these presentation levels. Further, there was no effect for the repetition number ($F(5,50)=0.227$, $p=0.949$) or for the interaction between strategy and repetition number (Greenhouse-Geisser corrected $F2.8,27.7=2.581$, $p=0.078$). All subjects improved on spectral ripple discrimination at either 45 dB or at 65 dB or both. In a post-hoc analysis, where the subjects with the largest ΔAUC (S05, S09 and S10) were excluded, there was a highly significant improvement from 2.8 RPO to 4.4 RPO at 65 dB: $F(1,7)=14.862$, $p=0.006$. Although there was no significant correlation between the ΔAUC and spectral ripple performance at 45 dB (Figure 4A), exclusion of the same three AUC outliers resulted in a highly significant improvement from 2.4 RPO to 4.2 RPO: $F(1,7)=11.264$, $p=0.012$. Linear regression revealed that (when including S05, S09 and S10 in the analysis) the ΔAUC could significantly predict spectral ripple performance with the DCF strategy at 65 dB: $F(1,6) = 7.079$, $p=0.037$. The average ΔAUC accounted for 54.1% of the variation in spectral ripple scores with an adjusted $R^2=46.5\%$. The regression equation was as follows:

$$P_{PRO\_65dB} = 2.25 - 0.046 \cdot \Delta AUC \qquad \text{(eq. 5)}$$

where P_{PRO_65dB} is the predicted improvement in RPO with the DCF strategy as compared to the clinical strategy, at 65 dB (Figure 4B).

When the 2 outliers that drove the direction of the correlation (S09 and S10) were excluded from the linear regression, however, the correlation completely disappeared ($F(1,4)=0.001$, $p=0.979$). These subjects drive the correlation because they were poor performers with the DCF strategy, and had large differences in their ΔAUC's. By excluding these subjects from this analysis only the better performing subjects were included. The differences in the offset of the loudness curves were also analysed, but the correlation between Δoffset and spectral ripple performance at 45 dB was not statistically significant ($F(1,6) = 5.6$, $p=0.056$) (Figure 4C) and also no correlation at 65 dB ($p=0.65$) (Figure 4D) was found.

Temporal modulation detection test

The individual and mean MDT's in dB relative to 100% amplitude modulation, that were measured at comfortable level, are shown in Figure 5. The mean MDT's were -9.35 dB for the clinical speech coding strategies and -8.73 dB for the DCF strategy. A two-way repeated measures ANOVA revealed no significant differences in performance between the two speech coding strategies: $F(1,10) = 12.611$, $p=0.497$, in terms of amplitude modulation detection. There were no correlations between the ΔAUC or Δoffset and MDT scores, and the exclusion of the AUC or Δoffset outliers did not change the MDT results.

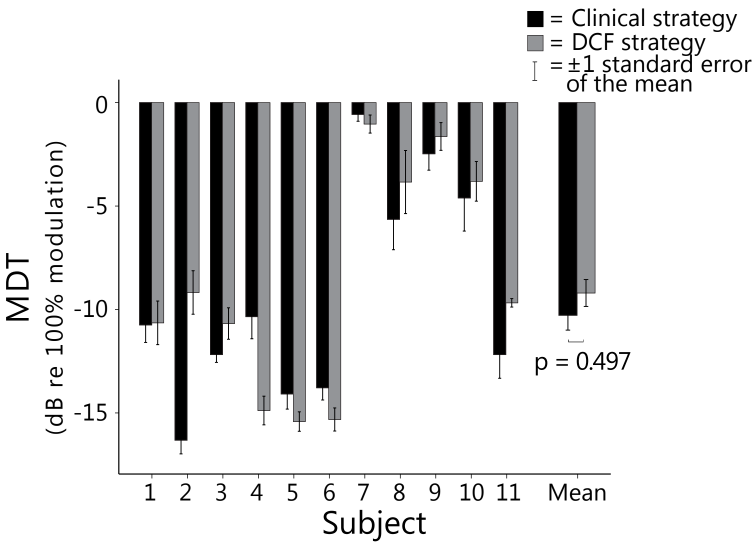


FIGURE 5. INDIVIDUAL AND MEAN MODULATION DETECTION THRESHOLDS (MDTS) FOR THE 11 STUDY SUBJECTS AT 65 dB SPL. The error bars represent ± 1 standard error of the mean.

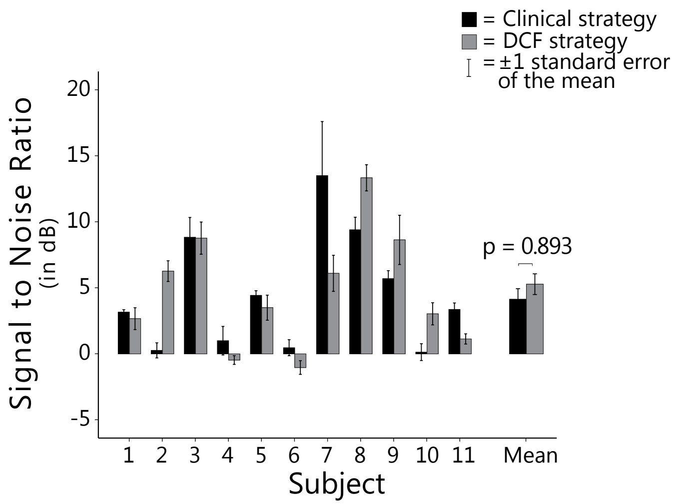


FIGURE 6. SPEECH INTELLIGIBILITY IN NOISE (MATRIX) WITH FIXED SPEECH AT 65 dB. SNR, the speech-to-noise ratio, for which 50% of the words was repeated correctly. The error bars represent ± 1 standard error of the mean.

4

Speech-in-Noise test

The individual and mean SRTs, as measured at 65 dB, are shown in Figure 6. The mean SRTs were 4.57 dB SNR and 4.72 dB SNR for the clinical and DCF strategies, respectively. The standard errors were quite large for some subjects, and a two-way repeated measures ANOVA revealed that the difference in SRTs was not statistically significant: $F(1,10)=0.019$, $p=0.893$. No correlations were found between the ΔAUC or Δoffset and SRT values, and exclusion of the AUC or Δoffset outliers did not change the speech-in-noise results.

Subjective rating

Subjectively, 2 subjects preferred the DCF strategy over their own clinical strategy and 7 had no preference. The 2 remaining subjects preferred their own strategy over the acutely-tested DCF research strategy. Some subjects noted that the DCF strategy resulted in a 'richer' sound with greater pitch perception and that the sound was clear. Others felt that the DCF strategy resulted in some background noise. The battery life of the PowerCel Slim battery decreased from an average of 9 hours with the clinical strategies to 1.5 to 4 hours with the DCF strategy.

DISCUSSION

The DCF strategy showed promising results in this initial study, despite the large disparity in experience with the DCF strategy versus the individuals' clinical strategies. Use of the DCF strategy improved spectral ripple discrimination by 1.34 RPO at lower loudness levels, which is a large improvement compared to reports in the literature using the same test. For example, Zhou (2017) found a 1.05 RPO improvement with an experimental strategy in which five high-threshold stimulation sites were deactivated[5], and Aronoff *et al*. (2016) showed that interleaved processors improve SMRT scores by 1.0 RPO[41]. Moreover, a study where indiscriminable electrode contacts were deactivated in n-of-m strategies even found deteriorated SMRT scores[42]. The comparison of the DCF strategy results with other current focusing strategies is complex because often different spectral ripple measures are used. For example, Smith *et al*. (2013) found a 5.7 dB improvement in a spectral ripple phase discrimination experiment at 2.0 cycles/octave when weighted TP stimulation was compared to MP stimulation[24]. In our study, the most striking improvement was at 45 dB, which was in accordance with our hypothesis, since the DCF strategy is set up such that higher levels of current focusing are achieved at lower loudness levels. Multiple studies have shown that the narrowing effect on current spread is negligible at $\sigma \leq 0.5$[22,43,44], while the mean degree of current focusing in this study was 0.88 [0.8–1.0] at the T-level and 0.49 [0.21–0.74] at the M-level. Thus, one could expect a greater benefit at lower loudness levels. Although most of the subjects benefited from current focusing across their entire dynamic range, some required fairly low σ values to reach the M-level. The DCF strategy therefore did not enhance spatial selectivity at M-level, which explains why no significant difference was observed between DCF and MP stimulation in spectral resolution.

One potential disadvantage of current focusing techniques is that they require wider pulse widths, and thus lower stimulation rates, to reach sufficient loudness[44]. This reduces temporal resolution and therefore speech perception. Although the stimulation rates decreased from 2550 pulses per second (pps) on average with the clinical strategy to 817 pps with the DCF strategy, subjects performed equally well on the temporal modulation detection task using the two strategies. The DCF strategy was fitted without difficulties, and the subjects gave predominantly positive feedback on the sound quality. This is in accordance with the literature, as previous studies also report positive results concerning the quality of sound with current focusing techniques[22,25]. Another disadvantage for the DCF strategy is the decrease in spectral channels (from 16 with HiRes or 120 theoretical channels with current steered strategies, to only 14 with DCF) that can be used, because 3 physical electrode contacts are required to create 1 current focusing channel. Therefore, a smaller portion of the auditory nerve can be used for stimulation, possibly leading to a decrease in spectral resolution. Moreover, most subjects were clinically fitted with a speech coding strategy that uses current steering, which creates additional (virtual) spectral channels[45], while current steering was not implemented in the DCF strategy. Although many studies found a beneficial effect of the implementation of current steering[46,47], others were unable to find this improved performance[21,48,49]. Nevertheless, the subjects in the current study may have benefited from current steering, that was only implemented in their clinical speech coding strategy.

Notably, acutely measured perception of speech in noise was as good with the DCF strategy as with the clinical strategy, even though the subjects had at least 9 months of experience with their clinical strategy but just a few hours of experience with the DCF strategy. Spectral ripple tests were added, as they are reported to be ideally suited for acute testing and correlate with long-term speech perception[4], while speech tests need adaptation time. Thus, the significantly improved spectral ripple thresholds strengthen our conviction that the DCF strategy shows promise for improving the perception of speech in noise long-term. While previous research did reveal significant correlations between SMRT scores and speech understanding[4], we were not able to demonstrate this when the clinical 65 dB measures were used ($F(1,9)=1.8$, $p=0.211$. $R=0.41$). This is probably due to the relatively small research group. Srinivasan ***et al***. (2013) demonstrated improved speech understanding in an acute setting with partial TP stimulation in 6 CI listeners after only 20 minutes of adaptation time[30]. They found that SRTs were improved by 3 dB compared to an experimental MP strategy. However, they compared 2 strategies that were new to the subjects, whereas we compared the novel DCF strategy,

4

to each subject's established clinical strategy.

Although the overall results with the DCF strategy were encouraging, not all subjects benefited from the novel loudness encoding strategy. Three subjects scored better on the spectral ripple test with the DCF strategy only at 45 dB or 65 dB but not at both levels. The results of the present study suggest that a change in loudness growth, from the clinical to the DCF strategy, could cause this lack of improved performance. The three subjects (S05, S09, S10) who showed large discrepancies between the slopes of the loudness curves for the 2 strategies performed worse on the psychophysical tasks. Because a significant negative correlation was found between ΔAUC and performance on the spectral ripple task (at 65 dB), a post-hoc analysis was performed leaving these three subjects (with the largest ΔAUCs) out, resulting in a greater improvement with the DCF strategy and a higher statistical significance ($p=0.006$). While one would expect similar offsets of the loudness growth curves for the two speech coding strategies, remarkable deviations were found in some subjects, even across electrode contacts. The correlation between this Δoffset and spectral ripple scores at lower loudness levels was positive, although this was not statistically significant. It makes sense that subjects with negative Δoffsets performed worse at 45 dB, as it could be that the sounds were inaudible, or at least very soft, at this loudness level. Altogether, this suggests that the way loudness growth is achieved could be of importance for CI performance. Nevertheless, previous research found only minor effects of loudness growth on speech performance[50]. This is consistent with our data, as we only found a detrimental effect on spectral ripple performance and not on speech perception. Moreover, it could be that longer adaptation to different loudness growth cancels out a detrimental effect. Interestingly, S09 did not show unnatural loudness growth with the DCF strategy, but with the clinical Optima strategy. This subject had probably adapted to the aberrant loudness growth with the used clinical program and experienced difficulty adjusting to the (more regular) loudness growth of the DCF strategy. This observation highlights the beneficial effects of having a longer period of time to adapt to novel speech coding strategies, which might have resolved this issue. No clinical reasons were observed for the unexpected loudness growth with the clinical strategy for this subject (S09), such as an aberrant return pathway due to, for example, otosclerosis.

Bierer and Litvak (2016) suggested that especially poor performers benefit from strategies that reduce channel interactions, presumably because they suffer from more channel interaction in the first place[27]. If there is a relatively poor electrode-neuron interface due to a large electrode-to-neuron distance[12,13], the DCF strategy would theoretically greatly impact the overall performance. More laterally positioned electrodes benefit more from current focusing, as the efficacy of multipolar stimulation depends on interactions in the far field[18,51]. This study mostly included subjects implanted with a HiFocus 1J electrode array, which is designed to be in an outer wall position, which in turn favours electrical field interaction. On the other hand, the study population comprised relatively good performers (CVC phoneme scores of 78% or more at 65 dB), leaving relatively little room for improvement. The beneficial effects of the DCF strategy in a larger population of CI users with higher variation in performance might even be greater than in the current study.

If degeneration of the spiral ganglion cells underlies poor CI results, the beneficial effects of using a DCF strategy are likely to be less prominent. Several studies reported channel-to-channel threshold variability across the electrode array to be highly correlated with poorer performance on speech tests[20,52–54]. High thresholds are believed to be caused, at least in part, by degeneration of the spiral ganglia or so-called dead regions, which presumably results in ineffective channels. As focused configuration leads to stimulation of a more localized region of the spiral ganglion, one could expect even more variability in focused stimulation mode[12], which could offset the beneficial effects of current focusing. Among the current study population, some subjects showed great variability in T-levels (most prominently S02 and S09). There are two possible solutions to this problem: (1) turn these electrode contacts off to improve speech intelligibility, as done by Bierer and Litvak (2016)[27] or (2) switch the contact from TP to MP so that no auditory information is sent to ineffective channels. It was not possible to study possibly positive adaptations to the fitting, like switching off contacts with great variability, within the context of the present study, because this would have introduced cofactors influencing the comparison of the two strategies.

We recently found that learning effects might interfere with the results of psychophysical measures used in this study (especially the SMRT and the MDT test)[55]. However, this has only been shown in long-term studies, not in acute settings as in the present study[56]. In addition, the practice tests that were provided before the actual testing are likely to cancel out any minor acute learning effect[4]. Moreover, due to logistics, the DCF strategy was always tested last, at the end of a long day of testing. It seems likely that the decline in performance due to fatigue probably overcompensated for any potential learning effects. So we speculate that the gain in spectral resolution is underestimated in our study. Only S10, who was evaluated with the DCF strategy on a separate test day, may have benefitted from learning during the psychophysical tasks.

In conclusion, the main finding in this study was that the DCF strategy for loudness encoding significantly improved spectral resolution in an acute setting compared to the current clinically used stimulation strategies, in soft but not at higher presentation levels. As the battery life was considerably reduced, changes to the power scheme should be made to make the strategy suitable for at-home usage. A decrease in the T-level σ to

0.8 and deactivation of ineffective electrode contacts could potentially increase battery life and therefore the clinical suitability of this novel loudness encoding strategy. Another potential solution to the high energy consumption of the DCF strategy might be to add parallel channels, as was done in the research of Langner ***et al***. (2017)[57]. Although previous research in our clinic[58] showed that paired pulsatile stimulation might have a detrimental effect in MP mode, it has potential in current focused stimulation[59]. Also the use of n-of-m strategies might be a valid option to decrease battery consumption and is therefore of interest for future research. As benefits, particularly for speech intelligibility, of new speech coding strategies are generally greater after longer adaptation periods, the next step is to find out whether a greater improvement occurs over time in a take-home trial.

CONCLUSION

The present study showed that the DCF strategy gives better spectral resolution at lower loudness levels after only a few hours of adaptation to the strategy. Subjects had months of experience with the comparative MP speech coding strategies. Equal performance on speech and temporal modulation tests was found. Future research will reveal whether long term usage of the DCF strategy also gives improved speech scores.

REFERENCES

1. Lazard, D.S., Vincent, C., Venail, F., ***et al.*** (2012) Pre-, Per- and Postoperative Factors Affecting Performance of Postlinguistically Deaf Adults Using Cochlear Implants: A New Conceptual Model over Time. ***PLoS One 7,*** 1–11.
2. Aronoff, J.M. & Landsberger, D.M. (2013) The development of a modified spectral ripple test. ***J. Acoust. Soc. Am. 134,*** 217–222.
3. Holden, L.K., Firszt, J.B., Reeder, R.M., ***et al.*** (2016) Factors Affecting Outcomes in Cochlear Implant Recipients Implanted With a Perimodiolar Electrode Array Located in Scala Tympani. ***Otol. Neurotol. 37,*** 1662–1668.
4. Lawler, M., Yu, J. & Aronoff, J.M. (2017) Comparison of the Spectral-Temporally Modulated Ripple Test With the Arizona Biomedical Institute Sentence Test in Cochlear Implant Users. ***Ear Hear. 38,*** 760–766.
5. Zhou, N. (2017) Deactivating stimulation sites based on low-rate thresholds improves spectral ripple and speech reception thresholds in cochlear implant users. ***J. Acoust. Soc. Am. 141,*** 243–248.
6. Henry, B.B. a. & Turner, C.W.C. (2003) The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. ***J. Acoustal Soc. Am. 113,*** 2861–2873.
7. Litvak, L.M., Spahr, A.J., Saoji, A. a & Fridman, G.Y. (2007) Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. ***J. Acoust. Soc. Am. 122,*** 982–991.
8. Won, J.H., Drennan, W.R. & Rubinstein, J.T. (2007) Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. ***J. Assoc. Res. Otolaryngol. 8,*** 384–392.

9. Azadpour, M. & McKay, C.M. (2012) A psychophysical method for measuring spatial resolution in cochlear implants. ***J. Assoc. Res. Otolaryngol. 13,*** 145–57.
10. Jones, G.L., Drennan, W.R. & Rubinstein, J.T. (2013) Relationship between channel interaction and spectral-ripple discrimination in cochlear implant users. ***J. Acoust. Soc. Am. 133,*** 425–433.
11. Bierer, J.A. & Faulkner, K.F. (2010) Identifying cochlear implant channels with poor electrode-neuron interface: partial tripolar, single-channel thresholds and psychophysical tuning curves. ***Ear Hear. 31,*** 247–258.
12. Bierer, J.A. (2010) Probing the electrode-neuron interface with focused cochlear implant stimulation. ***Trends Amplif. 14,*** 84–95.
13. Bierer, J.A., Deeks, J.M., Billig, A.J. & Carlyon, R.P. (2015) Comparison of Signal and Gap-Detection Thresholds for Focused and Broad Cochlear Implant Electrode Configurations. ***J. Assoc. Res. Otolaryngol. 284,*** 273–284.
14. Wardrop, P., Whinney, D., Rebscher, S.J., ***et al.*** (2005) A temporal bone study of insertion trauma and intracochlear position of cochlear implant electrodes. I: Comparison of Nucleus banded and Nucleus Contour™ electrodes. ***Hear. Res. 203,*** 54–67.
15. Wardrop, P., Whinney, D., Rebscher, S.J., Luxford, W. & Leake, P. (2005) A temporal bone study of insertion trauma and intracochlear position of cochlear implant electrodes. II: Comparison of Spiral Clarion and HiFocus II electrodes. ***Hear. Res. 203,*** 68–79.
16. Finley, C.C. & Skinner, M.W. (2008) Role of electrode placement as a contributor to variability in cochlear implant outcomes. ***Otol. Neurotol. 29,*** 920–928.
17. Nadol, J. jr, Young, Y. & RJ, G. (1989) Survival of spiral ganglion cells in profound sensorineural hearing loss: implications for cochlear implantation. ***Ann Otol Rhinol Laryngol 98,*** 411–416.
18. Kalkman, R.K., Briaire, J.J. & Frijns, J.H.M. (2014) Current focussing in cochlear implants: An analysis of neural recruitment in a computational model. ***Hear. Res. 322,*** 89–98.
19. Kral, A., Hartmann, R., Mortazavi, D. & Klinke, R. (1998) Spatial resolution of cochlear implants: The electrical field and excitation of auditory afferents. ***Hear. Res. 121,*** 11–28.
20. Bierer, J.A. (2007) Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration. ***J. Acoust. Soc. Am. 121,*** 1642–1653.
21. Berenstein, C.K., Mens, L.H.M., Mulder, J.J.S. & Vanpoucke, F.J. (2008) Current Steering and Current Focusing in Cochlear Implants: Comparison of Monopolar, Tripolar, and Virtual Channel Electrode Configurations. ***Ear Hear. 29,*** 250–260.
22. Landsberger, D.M., Padilla, M. & Srinivasan, A.G. (2012) Reducing current spread using current focusing in cochlear implant users. ***Hear. Res. 284,*** 16–24.
23. Zhu, Z., Tang, Q., Zeng, F.-G., Guan, T. & Ye, D. (2012) Cochlear-implant spatial selectivity with monopolar, bipolar and tripolar stimulation. ***Hear. Res. 283,*** 45–58.
24. Smith, Z.M., Parkinson, W.S. & Long, C.J. (2013) Multipolar current focusing increases spectral resolution in cochlear implants. ***Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*** 2796–2799. doi:10.1109/EMBC.2013.6610121
25. Padilla, M. & Landsberger, D.M. (2016) Reduction in spread of excitation from current focusing at multiple cochlear locations in cochlear implant users. ***Hear. Res. 333,*** 98–107.
26. Srinivasan, A.G., Shannon, R. V. & Landsberger, D.M. (2012) Improving virtual channel discrimination in a multi-channel context. ***Hear. Res. 286,*** 19–29.
27. Bierer, J. a. & Litvak, L. (2016) Reducing Channel Interaction Through Cochlear Implant Programming May Improve Speech Perception: Current Focusing and Channel Deactivation.

4

Trends Hear. ***20,*** 1–12.

28. Mens, L.H.M. & Berenstein, C.K. (2005) Speech Perception with Mono- and Quadrupolar Electrode Configurations: A Crossover Study. ***Otol. Neurotol. 26,*** 957–964.
29. Khalili Moghaddam, G., Lovell, N.H., Wilke, R.G.H., Suaning, G.J. & Dokos, S. (2014) Performance optimization of current focusing and virtual electrode strategies in retinal implants. ***Comput. Methods Programs Biomed. 117,*** 334–342.
30. Srinivasan, A.G., Padilla, M., Shannon, R. V & Landsberger, D.M. (2013) Improving speech perception in noise with current focusing in cochlear implant users. ***Hear. Res. 299,*** 29–36.
31. Wu, C.-C. & Luo, X. (2013) Current steering with partial tripolar stimulation mode in cochlear implants. ***J. Assoc. Res. Otolaryngol. JARO 14,*** 213–231.
32. Nogueira, W., Litvak, L., Landsberger, D.M. & Büchner, A. (2017) Loudness and Pitch Perception using Dynamically Compensated Virtual Channels. ***Hear. Res. 344,*** 223–234.
33. Vellinga, D., Briaire, J.J., van Meenen, D.M.. & Frijns, J.H.M. (2017) Comparison of multipole stimulus configurations with respect to loudness and spread of excitation. ***Ear Hear. 38,*** 487–496.
34. Firszt, J.B. (2003) HiResolution ™ Sound Processing. ***Adv. Bionics White Pap. Sylmar, Calif.*** 1–4.
35. Büchner, A., Lenarz, T., Boermans, P.P., ***et al.*** (2012) Benefits of the HiRes 120 coding strategy combined with the Harmony processor in an adult European multicentre study. ***Acta Otolaryngol. 132,*** 179–187.
36. Advanced Bionics. (2012) HiRes™ Optima Clinical Results. ***Advanced Bionics White Paper. Sylmar, California*** 1–2.
37. Potts, L.G., Skinner, M.W., Gotter, B.D., Strube, M.J. & Brenner, C.A. (2007) Relation Between Neural Response Telemetry Thresholds, T- and C-Levels, and Loudness Judgments in 12 Adult Nucleus 24 Cochlear Implant Recipients. ***Ear Hear. 28,*** 495–511.
38. Won, J.H., Drennan, W.R., Nie, K., Jameyson, E.M. & Rubinstein, J.T. (2011) Acoustic temporal modulation detection and speech perception in cochlear implant listeners. ***J. Acoust. Soc. Am. 130,*** 376–388.
39. Houben, R., Koopman, J., Luts., H., ***et al.*** (2014) Development of a Dutch matrix sentence test to assess speech intelligibility in noise. ***Int. J. Audiol. 53,*** 760–763.
40. Francart, T., van Wieringen, A. & Wouters, J. (2008) APEX 3: a multi-purpose test platform for auditory psychophysical experiments. ***J. Neurosci. Methods 172,*** 283–293.
41. Aronoff, J.M., Stelmach, J., Padilla, M. & Landsberger, D.M. (2016) Interleaved processors improve cochlear implant patients' spectral resolution. ***Ear Hear. 37,*** e85–e90.
42. Vickers, D., Degun, A., Canas, A., Stainsby, T. & Vanpoucke, F.J. (2016) Deactivating Cochlear Implant Electrodes Based on Pitch Information for Users of the ACE Strategy. ***Adv Exp Med Biol. 894,*** 115–123.
43. Middlebrooks, J. & Bierer, J. (2002) Auditory cortical images of cochlear-implant stimuli: coding of stimulus channel and current level. ***J. Neurophysiol. 87,*** 493–507.
44. Bonham, B.H. & Litvak, L.M. (2008) Current focusing and steering: modeling, physiology, and psychophysics. ***Hear. Res. 242,*** 141–153.
45. Koch, D.B., Downing, M., Osberger, M.J. & Litvak, L. (2007) Using Current Steering to Increase Spectral Resolution in CII and HiRes 90K Users. ***Ear Hear. 28,*** 38–41.
46. Firszt, J., Holden, L., Reeder, R.M. & Skinner, M.W. (2009) Speech recognition in cochlear implant recipients: comparison of standard HiRes and HiRes 120 sound processing. ***Otol. Neurotol. 30,*** 146–152.

47. Park, H.J., Lee, S.C., Chun, Y.M. & Lee, J.Y. (2009) HiRes with Fidelity 120 benefit in native speakers of Korean. ***Cochlear Implants Int. 10,*** 85–88.
48. Donaldson, G., Dawson, P. & Borden, L. (2011) Within-subjects comparison of the HiRes and Fidelity120 speech processing strategies: speech perception and its relation to place-pitch sensitivity. ***Ear Hear. 32,*** 238–250.
49. Snel-bongers, J., Briaire, J.J., Vanpoucke, F.J. & Frijns, J.H.M. (2011) Spread of Excitation and Channel Interaction in Single- and Dual-Electrode Cochlear Implant Stimulation. ***Ear Hear. 32,*** 1–10.
50. Fu, Q.J. & Shannon, R. V. (1998) Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners. ***J. Acoust. Soc. Am. 104,*** 2570–2577.
51. Litvak, L.M., Spahr, A.J. & Emadi, G. (2007) Loudness growth observed under partially tripolar stimulation: model and data from cochlear implant listeners. ***J. Acoust. Soc. Am. 122,*** 967–981.
52. Pfingst, B.E., Xu, L. & Thompson, C.S. (2004) Across-Site Threshold Variation in Cochlear Implants: Relation to Speech Recognition. ***Audiol. Neurotol. 9,*** 341–352.
53. Pfingst, B.E. & Xu, L. (2004) Across-Site Variation in Detection Thresholds and Maximum Comfortable Loudness Levels for Cochlear Implants. ***JARO - J. Assoc. Res. Otolaryngol. 5,*** 11–24.
54. Long, C.J., Holden, T.A., McClelland, G.H., ***et al.*** (2014) Examining the electro-neural interface of cochlear implant users using psychophysics, CT scans, and speech understanding. ***JARO - J. Assoc. Res. Otolaryngol. 15,*** 293–304.
55. de Jong, M.A.M., Briaire, J.J. & Frijns, J.H.M. (2017) Learning Effects in Psychophysical Tests of Spectral and Temporal Resolution. ***Ear Hear. 39,*** 475–481.
56. Drennan, W., Won, J., Nie, K., Jameyson, E. & Rubinstein, J.T. (2010) Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. ***Hear. Res. 262,*** 1–8.
57. Langner, F., Saoji, A. a., Büchner, A. & Nogueira, W. (2017) Adding simultaneous stimulating channels to reduce power consumption in cochlear implants. ***Hear. Res. 345,*** 96–107.
58. Bonnet, R.M., Boermans, P.-P.B., Avenarius, O.F., Briaire, J.J. & Frijns, J.H.M. (2012) Effects of pulse width, pulse rate and paired electrode stimulation on psychophysical measures of dynamic range and speech recognition in cochlear implants. ***Ear Hear.*** 489–496.
59. Vellinga, D., Bruijn, S., Briaire, J.J., Kalkman, R.K. & Frijns, J.H.M. (2017) Reducing interaction In simultaneous paired stimulation with CI. ***PLoS One 12,*** 1–11.

CHAPTER 5

Dynamic Current Focusing for Loudness Encoding in Cochlear Implants: A Take-Home Trial

Monique A.M. de Jong, Jeroen J. Briaire, Séline van der Woude, and Johan H.M. Frijns

International Journal of Audiology, 2019

ABSTRACT

Objectives

This study aimed to evaluate a more energy-efficient dynamic current focusing (DCF) speech-processing strategy after long-term listening experience. In DCF, tripolar stimulation is used near the threshold and loudness is controlled by the compensation coefficient σ. A recent acute pilot study showed improved spectral-temporally modulated ripple test (SMRT) scores at low loudness levels, but battery life was reduced to 1.5-4 hours.

Design

Within-subject comparisons were made for the clinical vs. DCF strategy after 5 weeks of at-home usage. Speech intelligibility in noise, spectral ripple discrimination, temporal modulation detection, loudness growth, and subjective ratings were assessed. Study sample: Twenty HiRes90K (Advanced Bionics, Valencia, USA) cochlear implant (CI) users.

Results

Average battery life was 9 hours with the newly implemented DCF compared to 13.4 hours with the clinical strategy. Compared with measurements made at the beginning of the study, SMRT-scores and speech intelligibility in noise were significantly improved with DCF. However, both measures suffered from unexpected learning effects over time. The improvement disappeared and speech intelligibility in noise declined significantly relative to the final control measurement with the clinical strategy.

Conclusions

Most CI users can adapt to the DCF strategy in a take-home setting. Although DCF has the potential to improve performance on the SMRT test, learning effects complicate the interpretation of the current results.

INTRODUCTION

Although speech perception is relatively good with current cochlear implants (CIs), especially in a quiet environment, there can be great variability in performance between subjects using the same device[1]. Specifically, performance declines when listening conditions become more difficult, such as in noisy environments. In addition to patient-specific factors, this deterioration is likely caused by a large current spread throughout the cochlea[2]. This broad current spread, particularly with monopolar (MP) stimulation, leads to channel interactions, which decrease the number of functional spectral channels and spectral resolution[3]. These aspects are essential for sound perception in difficult listening conditions[4]. As a result, current focusing stimulation modes have been developed to avert this broad current spread by shaping the electric current field[5–7]. In focused stimulation schemes, the active electrode delivers the intended current waveform and flanking electrode contacts, which carry the opposite polarity of current to close the circuit loop, serve as the return electrodes. One example is partial tripolar (TP) stimulation, which decreases the current spread, improves spatial selectivity, and improves spectral resolution and speech performance in many CI patients[8–11]. A major challenge in the clinical applicability of these stimulation modes is power consumption. Since the opposite polarities of the center and adjacent electrodes act to cancel each other, large currents are required to reach threshold and comfortable listening levels[8,12,13]. Full loudness growth is not always accomplished within the compliance limits of the implant, especially in patients with high electrode impedances.

To resolve this issue, a novel loudness encoding strategy called dynamic current focusing (DCF) was developed at Leiden University Medical Center (LUMC) in the Netherlands[14]. The DCF strategy was designed to retain the advantageous effects of current focusing with partial TP stimulation near threshold, while remaining energy efficient and optimally focused at the most comfortable level (M-level) (Figure 1). The amplitudes of the active and flanking electrode contacts of the partial tripole are increased equally up to the threshold level (T-level) with a constant level of current focusing, expressed as current compensation coefficient σ (equal to 0.8, 0.9 or 1.0), which denotes the fraction of the return current going through the flanking electrodes. To further increase the loudness from the T-level, σ is gradually decreased, resulting in a broader excitation pattern and, consequently, a higher perceived loudness level. In a pilot study of 11 CI users, spectral-temporally modulated ripple test (SMRT) scores at lower presentation levels were significantly improved with the DCF strategy compared with the scores using their regular clinical speech coding strategy[14]. In each subject, the SMRT scores improved at one or both of the presentation levels. Speech intelligibility in noise remained the same even though the subjects had at least 9 months of experience with their clinical strategy and just a few hours of experience with the DCF strategy. In addition, the clinical and DCF strategies were rated equally when the subjects were asked if they preferred the DCF strategy, their own clinical program, or had no preference. This is promising, as CI users often prefer their known strategy over a novel (for them) speech coding strategy in an acute setting. Moreover, in a recently published study with another dynamic TP strategy, in which T-levels were fitted with a σ of 0.8 and M-levels with a σ of 0.5, a beneficial effect was found on vowel identification in noise[15]. Considering the data from our pilot study and the available literature, we hypothesized that a beneficial effect on speech perception and SMRT scores would be achieved if CI users employed the DCF strategy long term.

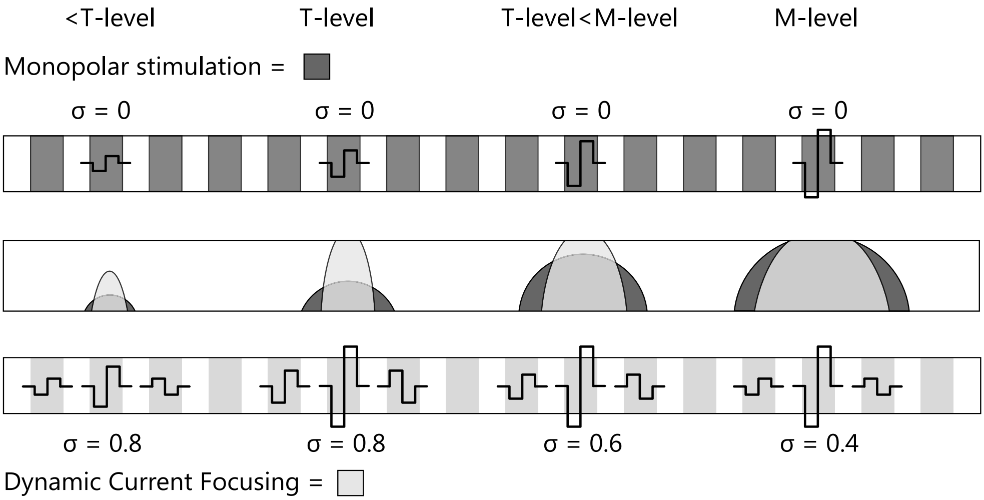


Figure 1. The concept of loudness coding with dynamic current focusing (DCF) and monopolar (MP) stimulation. The upper bar for each loudness step shows the auditory nerve with the excitation pattern in grey. The lower bars show the implanted electrode array with the electrode contacts in grey. In the DCF strategy, the amplitudes of the main and neighboring electrode contacts are increased equally up to the threshold level (T-level). To increase the loudness from the T-level, σ is decreased as a function of the stimulus level, resulting in a broader excitation pattern and higher loudness level. In MP mode, the amplitude of the main electrode contact is increased as a function of the stimulus level, resulting in broad current spreads at all loudness levels. M-level, most comfortable level.

To allow subjects to acclimatize to the DCF strategy in a take-home study, further development was needed. Although the power consumption of the DCF strategy was decreased relative to full TP stimulation and DCF stimulation always remained within the compliance limits of the device, the battery usage was still significantly increased compared with MP stimulation in the pilot study. To increase the clinical applicability of this strategy, power saving adjustments were made as described in the Materials and Methods. One relevant change was that the degree of current focusing at the T-level was decreased to a sigma of 0.8 for all electrode contacts in all participants. This decreased the battery consumption, though it also potentially weakened the beneficial effects of the novel speech coding strategy.

The current study was a take-home trial in which participants had 5 weeks to adjust to the optimized version of the DCF strategy. Performances regarding speech intelligibility in noise, spectral ripple discrimination, and temporal modulation detection were examined at different presentation levels and compared with the participants' clinical speech coding strategies. Although previous research has shown that learning effects in psychophysical measures are extinguished after 5-week test intervals[16], a measurement was made with the clinical speech coding strategy after completing the trial to determine if any learning effect was present.

MATERIALS AND METHODS

Subjects

Twenty postlingually deaf adults (12 women and 8 men, aged 34 to 70 years) participated in this study. All were unilaterally implanted with an Advanced Bionics CI (Sylmar, CA) at least 9 months before inclusion. The speech coding strategies used clinically were the HiResolution (HiRes; n=1)[17], HiRes Fidelity 120 (HiResF120; n=2)[18], and HiRes Optima (n=17)[19]. The patients' clinical characteristics are presented in Table 1.

The DCF strategy uses multipolar stimulation; therefore, only CI users with 16 active electrode contacts in their normal program were included in the study. Subjects 7 and 10 voluntarily dropped out due to difficulties adjusting to the novel strategy and underestimating the burden of participation in a clinical trial. Subject 14 completed the trial but was excluded from the analysis because of technical issues with the CI that interfered with the results, but were unrelated to the trial. Due to time constraints and fatigue, the baseline spectral ripple measurement at 45 dB was missing for Subject 11. The modulation detection threshold (MDT) measurements were also missing for Subjects 11 and 16. Subject 1 did not fill in the quality section of the SSQ and Subject 20 did not fill in the entire SSQ questionnaire. No other data were missing.

The study protocol was approved by the Committee for Medical Ethics of the Leiden University Medical Center (P02.106).

Study design

The current trial consisted of three sessions with an interval of 5 weeks between test days. In session 1, the subject's clinical strategy and the DCF strategy were fitted on a Harmony sound processor dedicated to research purposes. The baseline performance was then assessed with their usual clinical strategy. In the first 5 weeks, the subjects became familiar with the DCF strategy. If the subjects had complaints, an extra fitting was offered in the first week. During this extra fitting, the programmed settings or MAPs (including T- and M-levels, stimulation rate, and other parameters) were adjusted. Eight of the 17 subjects requested this extra fitting, after which most of them considered the subjective quality of sound to be improved. During the trial, subjects were allowed to use their clinical speech processor but were instructed to use the DCF strategy as much as possible. Thirteen of the 17 subjects reported using the DCF strategy in the same way as their clinical strategy, two other subjects started using the DCF strategy full-time after the extra fitting in the first week (Subjects 2 and 6), Subject 16 used the strategy every other day because of dissatisfaction, and Subject 1 barely used the DCF strategy during the 5-week accommodation period because of dissatisfaction. In session 2, subjects performed the psychophysical tasks with the DCF strategy (after the 5-week accommodation period). Then, their clinical strategy was fitted on a Harmony research processor for the next 5 weeks. In session 3 (after five weeks of home usage), the psychophysical measures were performed again with the clinical strategy to check for learning effects. The subjects were not blinded to the tested speech coding strategy, as they could easily detect their normal strategy.

Fitting procedures

Clinical Strategy. The MAPs from each subject's last clinical visit were copied from the SoundWave™ program (Advanced Bionics, Valencia, CA, USA) and fitted on a Harmony research processor. If the noise cancelation features were active, they were turned off throughout the study.

Research Strategy

The concept of the DCF strategy is clarified in Figure 1. The DCF program was created for each subject using BEPS+ software (Advanced Bionics, Valencia, CA, USA). As the pilot study revealed a decrease in the DCF battery life to 1.5-4 hours[14], some alterations to the fitting parameters were necessary to make the strategy suitable for the current take-home trial. Instead of aiming for a high σ at T-level, T-levels were determined for each electrode contact by slowly increasing the total amount of current with a σ of 0.8, which means that 80% of the current was returned equally to the two flanking electrodes and 20% was returned to the extra-cochlear ground electrode. In the current study this was true for all subjects. In the pilot study, however, σ at T-level was increased to 0.9 or 1.0 when the T-level was below 300 clinical units to ensure full loudness growth before σ was reduced to zero. As with the pilot DCF version, the M-levels were determined by gradually decreasing σ in steps of 0.01. This means that the current level on each main contact was kept constant while the current levels on the flanking electrode contacts were decreased. In this way, the excitation pattern was broadened and the loudness level increased. As a result, the dynamic range of the DCF strategy was defined by variations in σ. Another relevant difference to the pilot study was that a negative σ value was allowed to ensure full loudness growth (i.e., flanking electrodes could have the same polarity as the center electrode contact). Subjects 1, 4, and 20 needed negative σ values at multiple electrode contacts to reach sufficient loudness. The mean σ at M-level for all subjects and electrode contacts was 0.17 [-0.22 – 0.39] (Table 1).

TABLE 1. Characteristics of the study subjects

Subject	Gender	Age (yrs)	CI side	Duration of implant use (yrs)	CVC (Ph%)	Electrode array	Clinical strategy	Clinical settings		DCF settings		
								Rate (pps)	Pulse width (µs)	Rate (pps)	Pulse width (µs)	Mean M-level σ
S01	Male	64	R	3	89	HiRes90K HiFocus MS	Optima	1217	54.8	599	38.6	-0.22
S02	Male	68	L	12	88	HiRes90K HiFocus 1J	Optima	2560	26.0	1020	35.0	0.32
S03	Male	61	R	2	89	HiRes90K HiFocus MS	Optima	3093	21.6	1020	35.0	0.10
S04	Female	64	R	1	85	HiRes90K HiFocus MS	Optima	2007	33.2	1020	35.0	-0.08
S05	Female	59	R	5	87	HiRes90K HiFocus 1J	HiResF120	3093	21.6	1020	35.0	0.28
S06	Female	70	L	8	90	HiRes90K Hifocus 1J	Optima	2395	27.8	765	44.9	0.27
S07	Female	58	L	1	80	HiRes90K HiFocus MS	Optima	2970	22.4	829	43.1	0.43
S08	Male	70	L	7	90	HiRes90K HiFocus MS	Optima	3039	21.6	1020	35.0	0.36
S09	Female	47	R	4	84	HiRes90K Hifocus 1J	HiResF120	3039	21.6	1020	35.0	0.32
S10	Female	42	R	3	80	HiRes90K HiFocus MS	Optima	2750	24.2	765	44.9	
S11	Female	66	L	2	85	HiRes90K HiFocus MS	Optima	2970	22.4	1020	35.0	0.15
S12	Female	61	R	3	90	HiRes90K HiFocus MS	Optima	3039	21.6	1020	35.0	0.32
S13	Male	68	L	1	92	HiRes90K HiFocus MS	Optima	3039	21.6	1020	35.0	0.07
S14	Male	67	L	1	74	HiRes90K HiFocus MS	Optima	2750	24.2	865	38.6	0.39
S15	Male	64	R	1	88	HiRes90K HiFocus MS	Optima	3039	21.6	1020	35.0	0.01
S16	Female	73	L	2	88	HiRes90K HiFocus MS	Optima	3039	21.6	1020	35.0	0.22
S17	Male	34	L	3	87	HiRes90K HiFocus MS	Optima	1768	37.7	925	38.6	0.16
S18	Female	54	R	10	78	HiRes90K Hifocus 1J	Optima	1865	35.9	1020	35.0	-0.01
S19	Female	50	R	14	86	Clarion CII HiFocus 1J	Optima	1428	46.7	750	47.6	0.35
S20	Female	68	R	10	88	HiRes90K Hifocus 1J	HiRes	725	43.1	846	42.2	-0.03

L, left ear; R, right ear; CVC, consonant-vowel-consonant; Ph%, percentage phonemes correct on a standard monosyllabic (CVC) word test at 65 dB; DCF, dynamic current focusing; M-level, most comfortable level; HiFocus MS, HiFocus Mid-Scala; Mean M-level σ, mean σ at most comfortable level over all 14 electrode contacts. Excluded subjects are shaded in grey.

5

To optimize the subjective quality of the sound, M-levels were manually adjusted per electrode contact after turning on the speech mode and, therefore, no loudness balancing per electrode contact was performed. The overall loudness with the DCF strategy was adjusted until the subjects indicated that it had the same loudness as their clinical speech coding strategy. In contrast to the pilot study, low power modes were turned on (the target tank value was automated and the maximum power that can be supplied to the device was lowered) and the target voltage was set to 6.5 V instead of 7.3 V to avoid system voltage fluctuations. Noise reduction algorithms were switched off. Because three physical electrode contacts are required to create one current focusing channel, the two outer electrodes of the array could not be used. Therefore, the DCF strategy had only 14 effective spectral channels, whereas the clinical strategy had 16 for the HiRes program and 120 virtual spectral channels for the current steered speech coding strategies. The stimulation rates were significantly lower with the DCF strategy (929 pulses per second (pps) on average) than the clinical ones (2491 pps on average) (see Table 1).

Psychophysical Tasks

Loudness Growth Functions. As in the pilot study, a loudness scaling experiment was performed in this take-home trial at three different locations along the electrode array (electrodes 3, 9, and 14) for both stimulation strategies. The stimuli for the loudness growth experiment were acoustically administered via a direct connection to the speech processor and, therefore, processed with either the DCF or clinical strategy. The stimuli were sine waves with frequencies corresponding to the center frequencies of electrodes 3, 9, and 14, which were generated by a custom MATLAB program. The stimulus levels were calculated as follows:

$$\text{Loudness Level (in dB) - } 20 \cdot \log_{10}\left(\frac{V}{V_{ref}}\right) \qquad \text{(eq. 1)}$$

where V was the administered voltage and V_{ref} was calibrated for each electrode using the peak calculation of BEPS+ software so the digital stimuli approximately matched the microphone output with the same sound pressure level. To avoid overstimulation at the initiation of the measurement, only ascending scaling was performed in steps of 2 dB, starting from 0 dB to a maximum of 100 dB. As in previous studies[7,20,21], loudness was subjectively rated on an 8-point loudness scale that ranged from the T-level (1) to the most comfortable loudness (5) and the upper limit of comfortable loudness (8). When the upper limit was reached, the experiment was terminated[22]. The procedure was repeated three times per electrode contact, and the average voltage per loudness level was calculated.

As in the pilot study, the areas under the curves (AUCs) from loudness points 2 to 5 (from 'very soft sound' to 'most comfortable loudness') were calculated to quantify the slope of the loudness growth curves. Loudness points 2 to 5 were chosen because they correspond to regular speech levels and were considered the most important for understanding speech.

$$AUC = \sum_{i=2}^{4}\left(\Delta LL_i \frac{SL_i + SL_{i+1} - SL_2}{2}\right) \qquad \text{(eq. 2)}$$

In Equation 2, i is the subjective level from Equation 1, SL is the subjective loudness level, and ΔLL_i is the difference in loudness level between SL_i and SL_{i+1}(in dB). Differences in the AUC (ΔAUC) between the two strategies were calculated as follows:

$$\Delta AUC[\%] = \frac{AUC_{DCF} - AUC_{clinical}}{AUC_{clinical}} \cdot 100\% \qquad \text{(eq. 3)}$$

where $AUC_{clinical}$ is the AUC for the clinical strategy and AUC_{DCF} is the AUC for the DCF strategy.

Psychophysical Tasks - Sound Booth Testing

The following tasks were performed in the free field in a double-walled sound-attenuating booth. Subjects were facing a single loudspeaker at a distance of 1 m. To avoid learning effects of the psychophysical tasks, the subjects went through at least one dry run before the actual test runs were performed.

Spectral Ripple Test

The Spectral-temporally Modulated Ripple Test (SMRT)[23] was used to determine spectral ripple density thresholds at 45 dB and 65 dB. This spectral ripple test was chosen because it deals with potential confounders, such as loudness cues. Although it is unclear whether the SMRT scores represent frequency resolution or other perceptual abilities[24], the scores correlate with speech understanding under multiple listening conditions[25–27]. The SMRT is a three-alternative forced choice (3AFC) task that determines the maximum ripples per octave (RPO) that a listener can distinguish from two reference stimuli with unresolvable high ripple densities of 20 RPO. The listeners were asked to discriminate the adaptive spectrally rippled stimulus from the two reference stimuli, and no feedback about the correct answer was provided. The task was repeated six times per test condition and the average SMRT score was computed.

Temporal Modulation Detection Test

Since stimulation rates decrease with the DCF strategy (Table 1), the effect on temporal resolution was measured. The temporal modulation detection test adapted from Won *et al*. (2011) was used to examine the MDTs with each speech coding strategy[28]. In this two-alternative, forced choice, adaptive measure, two 1-second-wide band noise stimuli were presented; one was amplitude modulated with a frequency of 100 Hz. Subjects were instructed to choose the amplitude-modulated stimulus, the modulation depth of which was adapted using a 2-down 1-up procedure. To help the subjects remember which of the two intervals was the target stimulus, visual feedback about the correct answer was provided. A modulation frequency of 100 Hz was chosen because this task, along with the spectral ripple thresholds, accounts for the highest amount of variance in consonant nucleus consonant word scores[28]. The task was performed at 65 dB and repeated six times, and average MDTs were calculated in decibels relative to 100% modulation ($20 \cdot \log_{10} \cdot$ modulation depth).

Speech-in-Noise Test

A Dutch version of the matrix sentence test was used to measure the speech reception thresholds (SRTs) for each test condition. This speech-in-noise test uses 50 unique words combined into 200 grammatically equivalent sentences[29]. Ten balanced lists with 20 randomly selected sentences are available for testing. The task was carried out using the APEX 3 program (Leuven, Belgium)[30] installed on a personal computer. After the presentation of each sentence, the subjects were instructed to repeat the five words and to guess if they were unsure. Testing was done at a fixed speech level of 65 dB or 45 dB and with an adaptive speech-shaped noise starting from a -4 dB signal-to-noise ratio (SNR). Two training runs of 20 sentences each were performed prior to each test condition, as recommended by Kollmeier *et al*. (2015)[31]. An average SNR was calculated over three repetitions to determine the final SRT score for each loudness level and speech coding strategy.

Subjective Ratings

To evaluate the subjective ratings of the clinical and DCF speech coding strategies, the Speech, Spatial, and Qualities of Hearing Scale (SSQ) was used[32]. This is a measure for evaluating various aspects of hearing disabilities, and the domains ‘quality of hearing’ and ‘speech understanding’ were assessed. In addition, all subjects were asked whether their overall rating of the DCF strategy was better, equal to, or worse than their clinical program. Subjects also kept a daily log about the time (in hours) they had turned on the DCF strategy each day.

Statistical Analysis

Repeated measurements within subjects were obtained in all experiments. Therefore, two-way repeated measures analysis of variance (ANOVA) and linear mixed effects models were run using IBM® SPSS Statistics for Windows, Version 23.0. Both models control for the within-subject nature of the tasks by including random effects for

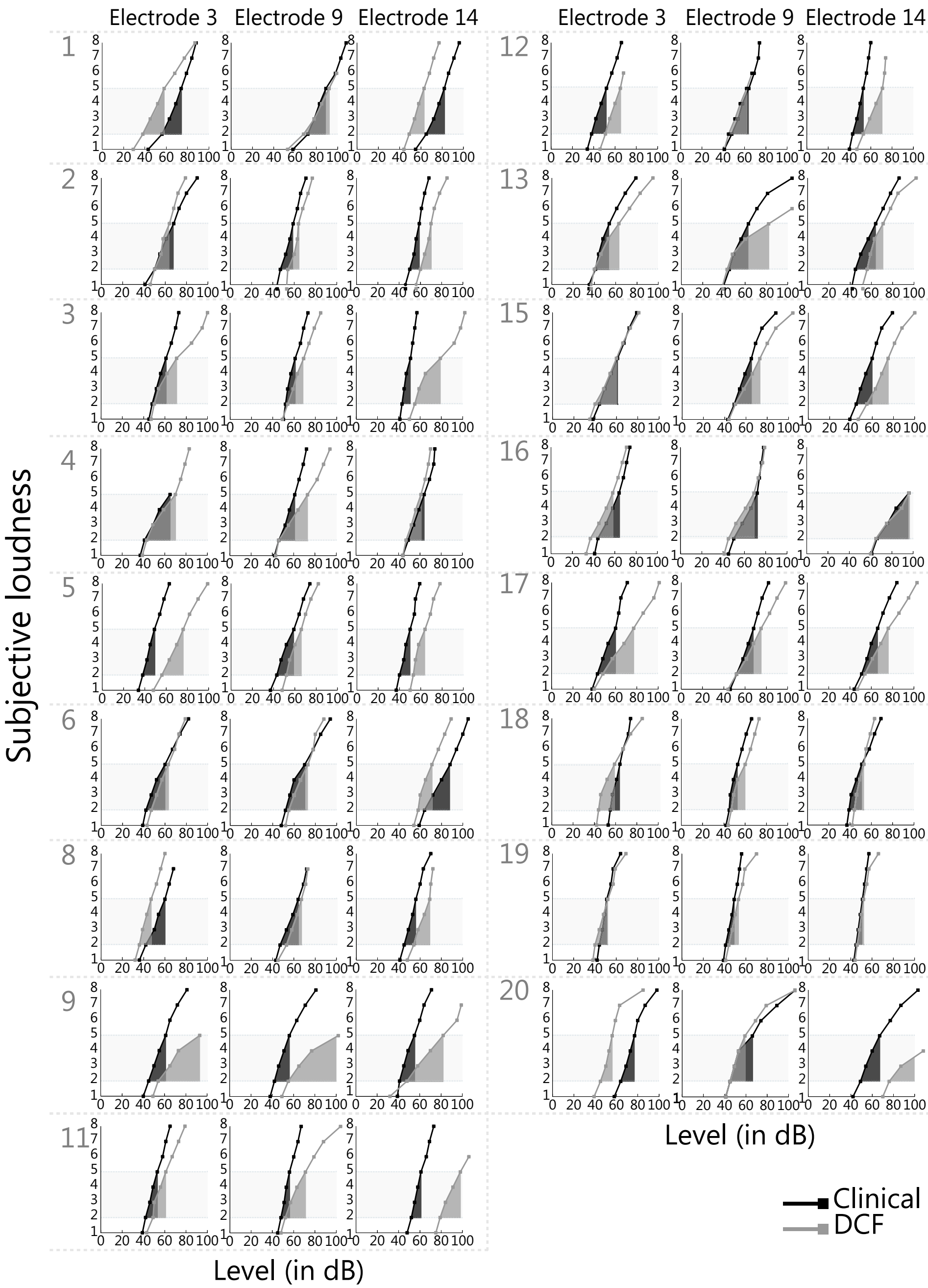


FIGURE 2. INDIVIDUAL LOUDNESS GROWTH FUNCTIONS ON AN 8-POINT LOUDNESS SCALE (Y-AXES). The stimulus levels were calculated as follows: loudness level (in dB) = 20 · $\log_{10}(V/V_{ref})$, where V_{ref} was calibrated for each electrode contact using BEPS+ software. For understanding speech, the loudness levels '2' (very soft sound) to '5' (the most comfortable loudness) were considered the most relevant. This loudness range is shaded in grey. The areas under the curves (AUCs) are indicated for the two strategies. DCF, dynamic current focusing.

subject and subject-task interactions. Moreover, the mixed-model design accounts for missing data. The fixed factors 'test session' and 'presentation level' were included in the linear mixed model, and 'test session' and 'repetition number in each test session' (1-3 for the Dutch Matrix test and 1-6 for the SMRT and the MDT) were used in the two-way repeated measures ANOVA to determine whether repetition number and the interaction between test session and repetition number had main effects.

RESULTS

Loudness Growth Functions

The individual loudness growth functions with the clinical and DCF strategies are shown in Figure 2. The loudness scores are plotted as a function of the presentation level in decibels, and the AUCs are represented by the filled areas. The average AUC was 22.1 for the clinical strategies and 31.0 (i.e., more shallow curves) for the DCF strategy. A 2-way repeated measures ANOVA with 'strategy' and 'electrode number' as factors revealed that this difference was significant ($F(1,16)=9.645$, $p=0.007$). Individual ΔAUCs per electrode contact are shown in Table 2. In all but one subject and one electrode contact (Subject 20, electrode 14), sufficient loudness was achieved to reach the most comfortable level. In 8 of the 17 subjects (Subjects 4, 8, 9, 11, 12, 13, 16, and 20), we were unable to reach the highest acceptable loudness with the DCF strategy, within the limits described in the Material and Methods.

TABLE 2. Differences in the areas under the loudness curves (ΔAUCs) between the two strategies for each subject at the indicated electrodes. The (absolute) average AUC of the three electrode contacts is also shown and expressed as the percentage of the AUC for the MP strategy: $\Delta AUC[\%] = AUC_{DCF} - AUC_{clinical} / AUC_{clinical} \cdot 100\%$

Subject	Electrode 3	Electrode 9	Electrode 14	Absolute average
S01	+29%	+28%	-5%	21%
S02	-26%	-50%	+12%	30%
S03	+78%	+59%	+306%	147%
S04	+20%	+89%	-13%	41%
S05	+68%	-13%	+39%	40%
S06	-18%	-11%	-33%	21%
S08	-33%	-20%	+61%	38%
S09	+170%	+282%	+123%	192%
S11	+39%	+150%	+126%	105%
S12	+6%	+7%	+95%	36%
S13	+68%	+161%	-14%	81%
S15	+11%	+40%	+15%	22%
S16	+1%	+24%	+208%	78%
S17	+63%	+45%	+46%	51%
S18	+97%	+80%	-33%	70%
S19	+79%	+8%	+11%	33%
S20	-14%	-45%	-28%	29%

Spectral Ripple Test

The individual and mean SMRT results at 45 and 65 dB are depicted in Figure 3. A linear mixed model with the fixed factors 'test session', 'presentation level', and

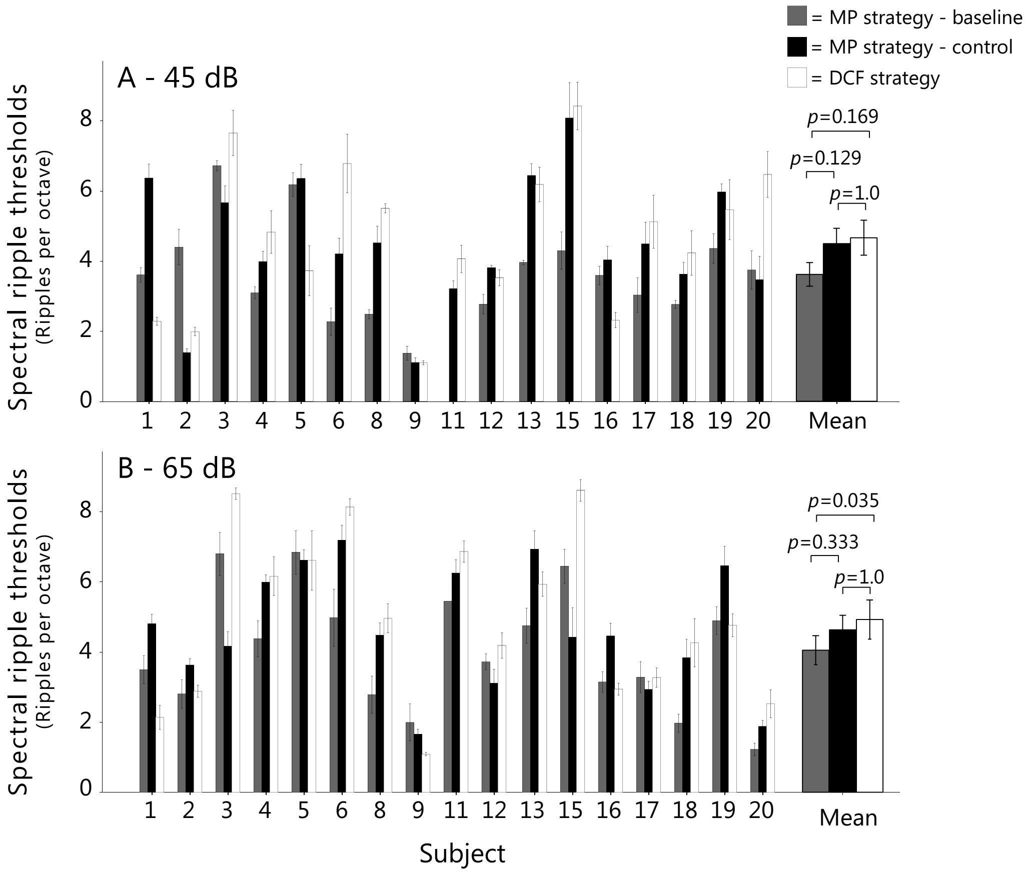


FIGURE 3. INDIVIDUAL AND MEAN SPECTRAL RIPPLE THRESHOLDS AT 45 dB (A) AND 65 dB (B) for the 17 study subjects using their monopolar (MP) clinical strategy (baseline and control measurements) and dynamic current focusing (DCF). Error bars represent ± 1 standard error of the mean.

'test session · presentation level' showed that the overall performance on the SMRT (45 and 65 dB combined) was significantly influenced by the test session ($F(2,16) = 4.260$, p=0.033). A pairwise comparison specified that the performance was significantly improved from 3.86 RPO at baseline to 4.80 RPO with the DCF strategy (Bonferroni-corrected p=0.035). However, a marginal improvement also occurred from baseline to the control clinical measurement (4.56 RPO, with a nominal p-value of 0.042). Because DCF stimulation was expected to improve spectral resolution at lower loudness levels, the two scores at the two loudness levels were also evaluated separately.

At 45 dB, the mean SMRT scores were 3.62, 4.49, and 4.66 RPO for the baseline, control clinical, and DCF strategies, respectively. A linear mixed model showed that test session had no effect on the SMRT scores at this loudness level ($F(2,16=2.873$, p=0.086). At 65 dB, the mean SMRT scores were 4.05, 4.63, and 4.92 RPO for the baseline, control clinical, and DCF strategies, respectively. The linear mixed model showed a significant effect of test session ($F(2,16)=5.062$, $p=0.02$), and the pairwise comparison showed that the mean improvement with the DCF strategy compared with the clinical baseline measurement was significant (Bonferroni-corrected $p=0.035$). However, when the DCF SMRT scores were compared with the clinical control measurement, the difference was not significant (Bonferroni corrected $p=1.0$). No significant improvement over time (baseline versus control clinical measurement) could be demonstrated (Bonferroni-corrected $p=0.333$). The repetition number had a significant effect within test sessions ($F(5,80)=3.58$, $p=0.006$), but there was no interaction between the repetition number and test session by two-way repeated measures ANOVA ($F(10,160)=0.88$, $p=0.552$).

In the pilot study, a small ΔAUC (i.e., a small difference between AUCs with the clinical and DCF strategies) predicted a better performance on the SMRT at 65 dB. In the current study, however, linear regression revealed that the ΔAUC did not significantly predict improvement in SMRT scores with the DCF strategy relative to both the baseline and control clinical measurements at 45 dB ($p=0.58$ and $p=0.48$, respectively) or 65 dB ($p=0.61$ and $p=0.42$, respectively).

Temporal Modulation Detection Test

Figure 4 presents the individual and mean MDTs in decibels relative to 100% amplitude modulation. The mean MDTs for the clinical speech coding strategies were -12.30 dB at baseline and -13.45 dB at the control measurement. The mean MDT with the DCF strategy was -11.88 dB. The linear mixed model with the fixed factor 'test session' showed

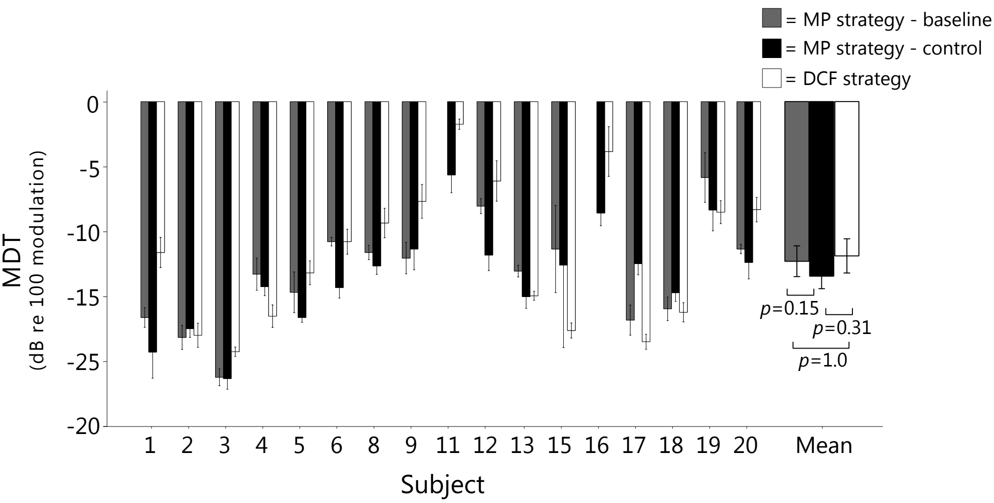


Figure 4. Individual and mean modulation detection thresholds (MDTs) for the 17 study subjects at 65 dB SPL, using their monopolar (MP) clinical strategy (baseline and control measurements) and dynamic current focusing (DCF). Error bars represent ± 1 standard error of the mean.

that the test session had no significant effect ($F(2,14)=2.442$, $p=0.122$). Linear regression revealed that the ΔAUC could not significantly predict improvement in MDT with the DCF strategy relative to both the baseline ($p=0.2$) and control clinical measurements ($p=0.75$).

Speech-in-Noise Test

The individual and mean SRTs are shown per presentation level in Figure 5. The mean SRT (45 and 65 dB combined) was 2.42 dB SNR for baseline, -0.21 dB SNR for the control, and 1.36 dB SNR for the DCF strategy. A linear mixed model with the fixed factors 'test session', 'presentation level', and 'test session*presentation level' revealed test session had a significant effect ($F(2,16)=17.983$, $p=0.001$). A pairwise comparison showed that only the difference between the baseline and control clinical measurement was significant (Bonferroni-corrected $p<0.001$), as the DCF SRT did not significantly differ from the baseline (Bonferroni-corrected $p=0.234$) and control measurements (Bonferroni-corrected $p=1.0$). As performance was expected to improve at lower loudness levels, the two measured loudness levels were also analyzed separately.

At 45 dB, the linear mixed model with the fixed factors 'test session', 'presentation level', and 'test session*presentation level' showed that the mean score was significantly affected by test session ($F(2,16)=15.256$, $p<0.001$). The SNR was improved with the DCF strategy (+0.92 dB SNR) when compared with the baseline measurement (+2.38 dB SNR; Bonferroni-corrected $p=0.052$). When DCF was compared with the control clinical measurement (-0.32 dB SNR), the numerical decline in performance was not significant (Bonferroni-corrected $p=0.227$). Despite the training sessions, the performance with the clinical strategy showed highly significant improvement from baseline to control testing (Bonferroni-corrected $p<0.001$). A two-way repeated measures ANOVA showed no effect for repetition number ($F(2,32)=2.03$, $p=0.148$) or the interaction between test session and repetition number ($F(4,64)=0.59$, $p=0.674$).

At 65 dB, the speech-in-noise results were also significantly influenced by test session ($F2,16)=14.656$, $p<0.001$). The DCF strategy SNR (+1.80 dB) was not significantly different from the baseline clinical measurement (+2.45 dB SNR; Bonferroni-corrected $p=1.0$). However, the control clinical measurement (-0.10 dB SNR) was better than the DCF results (Bonferroni-corrected $p=0.036$). In addition, at 65 dB there was a highly significant improvement from baseline to control (Bonferroni-corrected $p=0.001$). A two-way repeated measures ANOVA revealed there was no effect from repetition number per test session ($F(2,32)=2.07$, $p=0.115$) or the interaction between test session and repetition number ($F(4,64)=0.26$, $p=0.905$). The improvement in speech understanding in noise at 45 dB with the DCF strategy compared with baseline ($p=0.35$) and control ($p=0.20$) was not predicted by the ΔAUC. However, linear regression analysis showed that ΔAUC

5

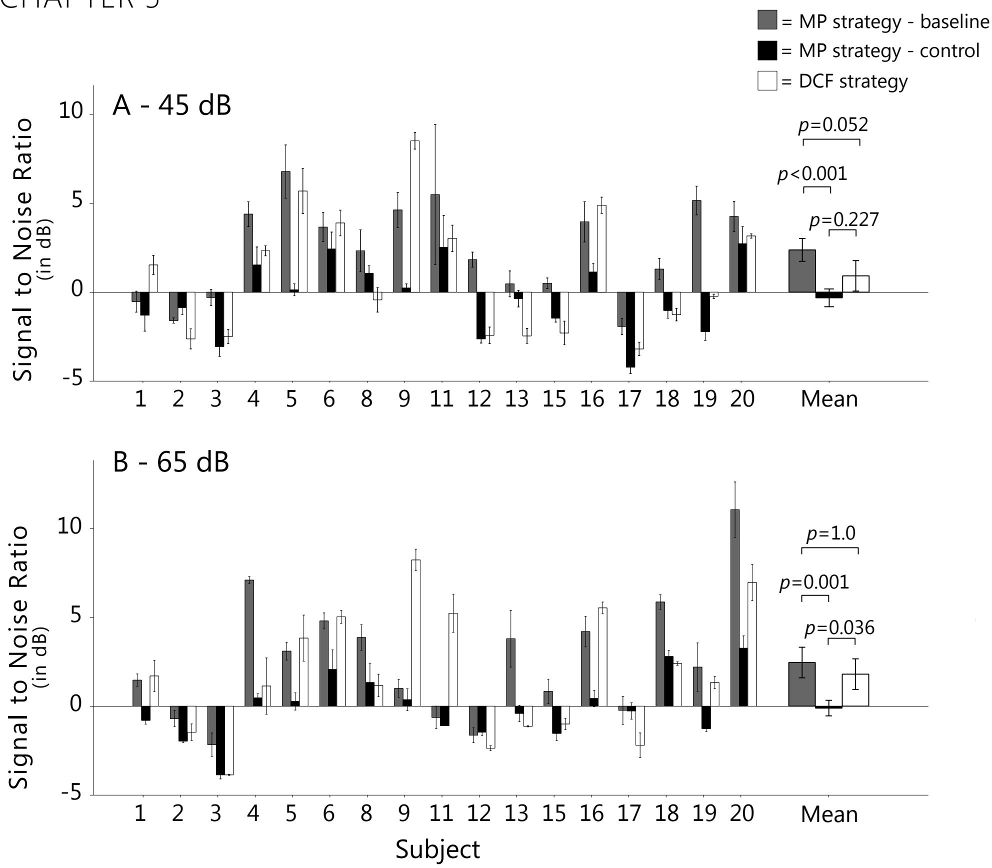


Figure 5. Individual and mean speech intelligibility in noise (Dutch Matrix test) with fixed speech at 45 dB SPL (a) and 65 dB SPL (b) for the 17 study subjects using their monopolar (MP) clinical strategy (baseline and control measurement) and dynamic current focusing (DCF). The y-axis is the speech-to-noise ratio for which 50% of the words were repeated correctly. Error bars represent ± 1 standard error of the mean.

could (borderline) significantly predict improvement in speech understanding in noise at 65 dB, with a weak correlation (R^2= 0.24, p=0.045).

This study intended to examine the effect of using the DCF strategy chronically; therefore, an extra analysis was performed where the subjects who did not use the strategy for at least 4 weeks were excluded (Subjects 1 and 16). In this post-hoc analysis of the SMRT data (45 dB and 65 dB combined), the improvement with DCF relative to the baseline measurement lost significance (from p=0.035 to p=0.254, both Bonferroni corrected). For the speech-in-noise data, the improvement with the DCF strategy compared with the baseline measurement at 45 dB gained significance (from p=0.052 to p=0.013), and the deterioration when compared with the control measurement (at 65 dB) lost significance (from p=0.036 to p=0.144). All other comparisons remained the same when Subject 1 and 16 were excluded from the analysis.

Subjective Ratings

Subject 1 did not complete the quality section of the SSQ and Subject 20 did not complete the entire questionnaire (Figure 6). A paired t-test revealed that, on average, the subjective rating of speech did not significantly differ between the clinical (5.2) and DCF (4.7) strategies (t(15) = 1.522, p=0.149). However, the subjective quality significantly deteriorated from 6.2 to 5.5 (t(14) = 2.279, p=0.039). Five of the 17 subjects preferred the DCF strategy over the clinical strategy, four had no preference, and eight preferred their regular clinical speech coding strategy. Seven subjects reported a continuous background noise at activation of the DCF strategy. In two of these subjects, the noise disappeared after a few minutes of listening (due to adaptation); whereas, in four subjects, the problem was solved by lowering the current at the T-level. A consequence of lowering the T-level current was that the loudness percept at the M-level also decreased. Therefore, the σ at M-level had to be decreased to maintain equal loudness. Subject 1 kept complaining about the continuous background noise, despite MAP adjustments, and was unable to use the DCF strategy on a daily basis. The eight subjects that asked

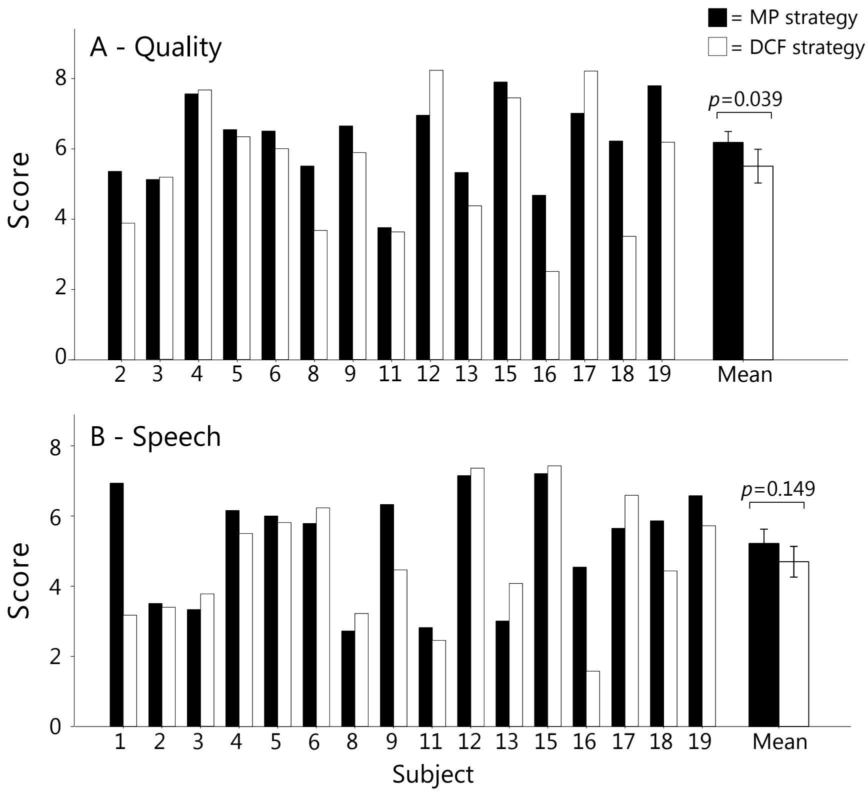


Figure 6. Subjective rating of processing strategies (SSQ) concerning quality of sound (a) and speech understanding under multiple listening conditions (b). Error bars represent ± 1 standard error of the mean.

for an extra fitting session reported that the incoming sound with the DCF strategy was shallow and/or echoing. This relative decrease in sound quality developed over the first few days after the fitting procedure and was easily solved by adjusting the MAPs by increasing the levels on the lower frequency electrodes and decreasing those at high frequency electrodes. Average battery life decreased from 13.4 hours with the clinical strategy to 9.0 hours with the DCF strategy, which is clinically acceptable.

DISCUSSION

The current study evaluated a novel loudness encoding strategy in a 5-week take-home trial with 20 CI users. A recently published pilot study of the DCF strategy[14] showed there was better performance on the SMRT with DCF, especially at lower loudness levels. Because the SMRT correlates with speech understanding[25–27], speech perception in noise was also expected to improve after some adaptation time. As the pilot study revealed a decrease in DCF battery life to 1.5-4 hours, some alterations to the fitting parameters were required to enhance clinical suitability. These energy-saving adjustments helped increase battery life to 9.0 hours on average. When the DCF scores were compared with measurements made at the beginning of the study, a marginally significant improvement was found in both speech intelligibility in noise at lower loudness levels (-1.5 dB SNR, Bonferroni-corrected p=0.052) and SMRT scores at louder levels (+0.9 RPO, Bonferroni-corrected p=0.035). However, when comparing the DCF results to the clinical control measurements, a significant deterioration was observed in speech intelligibility in noise at higher loudness levels (+1.9 dB SNR, Bonferroni-corrected p=0.036), suggesting there was a learning effect. In line with our expectations, no significant differences were found for the MDT results.

Despite the 5-week interval between test sessions and multiple practice sessions on each test day, a learning effect over time was observed for the SMRT (at 45 dB, marginally significant with a nominal p-value of 0.042) and the speech in noise task (at 45 and 65 dB) while using the clinical speech coding strategy. A recent study in our clinic revealed that a learning effect was present for the SMRT when it was repeated every 2 weeks[33]. However, previous research did not find this effect after longer time intervals between tests[16] and we assumed this learning effect fades out after longer time intervals. Yet, it still existed after a 5-week test interval, which complicates the interpretation of our results and signifies the importance of including a control measurement. Notably, repetition number within each test day had a significant effect on the SMRT, suggesting an acute learning effect. This finding is surprising, as multiple practice sessions were performed and such effects were not described previously[16,34,35]. There was no interaction between strategies and repetitions on each test day; therefore, performances can be directly compared on different test days. The learning effect could be caused by the transfer of speech cues over the course of the trial, as described previously[33]. This means that subjects adapted to new speech cues while using the experimental strategy and learned new auditory percepts. This so-called 'perceptual learning' may have played a major role during the extra control measurements. Perceptual learning has been extensively studied in the field of vision[36]. In the field of CIs, Irvine (2018) argued that the improvements in speech perception of CI users over the months and years following implantation are a form of perceptual learning[37]. If perceptual learning played a role in this study, the study subjects learned new percepts that they can also use in their everyday life. Thus, participating in clinical trials might be beneficial for CI users, even when the involved experimental strategies have no beneficial effect.

It may be that an actual difference in performance was measured and not a learning effect for subjects who performed better or worse with the DCF strategy than both the baseline and control clinical measurements. Although the current study design does not support statistical analyzes on the individual level, approximately half the subjects improved and one fourth performed worse on the SMRT. In the pilot study, the mean improvement was more striking (+1.34 RPO at 45 dB and +0.8 RPO at 65 dB), but more importantly, only 9% and 18% of the subjects deteriorated at 45 dB and 65 dB, respectively[14]. On the speech-in-noise test at 45 dB, five of the 17 subjects performed better, eight performed the same, and four performed worse with the DCF when compared with both clinical strategy measurements (the measurements made at the beginning and end of the trial); while at 65 dB, five subjects performed better, five the same, and seven worse. Although these results are below expectations, the current study showed that CI listeners can adapt to the present DCF strategy and that it potentially leads to a beneficial effect on SMRT scores, though smaller than expected based on the pilot study. However, four of the 20 included subjects were unable to adapt to this new speech coding strategy, and this was clear at the fitting session. This indicates it may be easy to detect patients who might not benefit from the DCF strategy.

Unfortunately, the possible SMRT advantage did not translate into improved speech perception in noise. This could be because the temporal resolution, measured with the MDT test, deteriorated in half the subjects. These relatively poor MDT results are comparable to those in the pilot study and might be a consequence of using lower pulse rates (58% lower with the DCF strategy on average), which are necessary to achieve sufficient loudness with the DCF strategy. Büchner ***et al***. (2012) showed that Advanced Bionics CI users perform better with higher pulse rates[18]. Among the nine subjects with deteriorated MDT scores in this study, five also had deteriorated speech intelligibility at the same loudness level and only two improved, supporting this hypothesis. On the other hand, some subjects in the current study (e.g., Subject 4) had improved SMRT scores and MDTs, but did not improve or deteriorate in speech understanding at 45 and 65 dB. This was unexpected, as both psychophysical measures are known to correlate independently with speech understanding[16,25,27,38,39]. The adaptation period of 5 weeks may not have been long enough to completely adjust to the novel strategy, so the final performance was still not reached[26]. This study intended to examine the effect of using the DCF strategy chronically; thus, an extra analysis was performed excluding subjects who did not use the strategy for at least 4 weeks. The exclusion of these subjects minimally improved the speech-in-noise results with the DCF strategy at both loudness levels, but did not drastically change the interpretation of our results.

The comparison of the results from this study with other current focusing strategies is complex because different spectral ripple measures are often used. For example, Smith ***et al***. (2013) found a 5.7 dB improvement in spectral ripple phase discrimination experiments at 2.0 cycles per octave when weighted TP stimulation was compared with MP stimulation[40]. The effects of using the DCF strategy (measured in RPOs) and the weighted TP strategy are therefore hard to compare.

During the current trial, two subjects complained of continuous background noise, which faded after the strategy was turned on for a few minutes. Interestingly, other subjects stated that the overall loudness level decreased after a few days of using the DCF, resulting in decreased sound quality in certain cases. These subjects had an extra fitting, in which the M-levels were increased, solving the problem. These phenomena imply that neural adaptation occurred that could be a consequence of the below-threshold stimulation inherent to the DCF strategy. Previous studies have shown that continuous high rate stimulation of the auditory nerve leads to different refractory states of the individual nerve fibers[41]. Though the electrical stimulation of deafened auditory nerves produces highly synchronized responses[42], this continuous electrical noise can cause desynchronization of the responses of the different nerve fibers[41,43]. This mimics the spontaneous activity of a healthy auditory nerve and may increase the dynamic range[44]. Although the loudness level of the conditioning noise signals used in previous research was above threshold, the high current levels administered right below T-level (with high levels of current focusing) could have resulted in a similar effect. This would explain the shallower loudness growth curve, i.e., the increase in dynamic range, found with the DCF strategy in the current study.

Loudness growth is also influenced by the 'interaction component K', which is a DCF parameter that determines the rate of change for σ depending on the input level of the signal. Both Litvak ***et al***. (2007) and Arenberg ***et al***. (2018) thoroughly describe this K value[15,45]. K is based on the degree of interaction between the three involved electrode contacts in DCF stimulation; K = 0 means there is only little interaction and K = 1 indicates the maximum amount of interaction. For the DCF strategy, this means that at high K values loudness growth is achieved by small decreases in σ, and at low K values larger decreases in σ are required to achieve full loudness growth. In the current study, K was assumed to be 1.0 for each electrode contact and subject; although, this does not necessarily correspond to the actual interaction since, for example, K is influenced by electrode-to-tissue distance[45]. It is hypothesized that loudness growth can be disturbed if the assumed K deviates from the actual K. To check this in the current study, K was calculated in six subjects using the following formula derived from Litvak ***et al***. (2007):

$$K = \frac{1 - TL_{MP} - TL_{0.9}}{0.9}$$

where TL_{MP} is the number of clinical units (CUs) required to reach T-level in MP mode, and $TL_{0.9}$ in TP mode when σ=0.9. The average K was 0.83 (SD=0.16), which is similar to the results of Arenberg ***et al***. (2018)[15]. This discrepancy between assumed K and actual K could have led to different loudness growth curves. It is hypothesized that loudness growth is improved when the assumed and actual K values are equal; thus, K values should be predicted for each individual electrode contact before fitting the DCF strategy.

The current implementation of the DCF strategy, although more energy efficient, has some characteristics that potentially minimized its beneficial effects. The decrease in σ at T-level caused a significant decline in σ at M-level from 0.49 on average in the pilot study to 0.17 in the current trial. A previous study demonstrated that current focusing coefficients below 0.5 resemble MP stimulation[46], implying that there would be no beneficial effect from DCF around the low coefficients at M-level in the current study. In the pilot study, however, the entire dynamic range was covered with effective current focusing, probably resulting in greater beneficial effects. In line with this theory, the beneficial effects found in this study were mostly present in the tasks at lower loudness levels (45 dB) and, thus, at higher levels of current focusing. The future focus of this strategy will be on the quality of the sound instead of energy efficiency. The DCF MAPs of both the pilot and take-home trial subjects revealed that, if current levels at T-level 400 CU or higher, none of the electrode contacts had a negative σ at M-level, and the average σ at M-level was 0.54 [0.12-0.86]. Therefore, future research should aim for current levels ≥ 400 CU at T-level for the DCF strategy. This will be achieved by increasing the σ at T-level if current levels are too low.

Moreover, the DCF strategy may only be suitable for a specific group of CI users, particularly subjects who are implanted with a lateral wall electrode array. In this study, most subjects were implanted with a HiFocus Mid-Scala (MS) electrode array, whereas in the pilot study most subjects were implanted with a HiFocus 1J (lateral wall) electrode array. The HiFocus MS electrode array is designed to enable closer placement

5

to the modiolus compared with the HiFocus 1J electrode array, requiring less current while achieving more specific stimulation of the auditory nerve[47]. For current focusing techniques, however, a certain distance between the electrodes and auditory neurons is necessary to enable electrical field shaping. In a computational model of the human cochlea, Kalkman ***et al***. (2014) predicted that current focusing does not achieve increased spatial selectivity to the same degree for perimodiolar electrodes as it does for lateral wall electrodes[20]. In addition, the close proximity of the electrodes to the auditory neurons can cause neuronal excitation by the non-center contacts in multipolar stimulation, resulting in so-called side lobes, which can negatively affect performance. This mechanism is expected in the DCF strategy tested here because of the allowance for negative σ values at the M-level, i.e., positively stimulating the non-center contacts. Nevertheless, no correlation was found between the type of electrode array and performance for any of the tasks (R^2 ranged from 0.003 to 0.11). Sufficient data were not collected in the current study to confirm this theory; therefore, the hypothesis should be studied in more detail in a larger cohort.

CONCLUSION

Most CI users can adapt to the DCF strategy and use it on a daily basis. The strategy has the potential to improve performance on the SMRT test, though the present implementation resulted in insufficient levels of current focusing across the dynamic range. In future DCF research, the aim will be to cover the entire dynamic range with effective levels of current focusing for each electrode contact to increase the advantageous effects of DCF stimulation. Because learning effects in psychophysical measurements are prominent, even after 5-week intervals, a randomized study design is essential for future research utilizing these measurements.

ACKNOWLEDGEMENTS

We would like to thank all the CI patients who participated in these experiments. This research was supported by non-restrictive research funding from Advanced Bionics.

REFERENCES

1. Firszt, J.B. *et al.* (2004) Recognition of Speech Presented at Soft to Loud Levels by Adult Cochlear Implant Recipients of Three Cochlear Implant Systems. ***Ear Hear.*** *25*, 375–387.
2. Bierer, J.A. (2010) Probing the electrode-neuron interface with focused cochlear implant stimulation. ***Trends Amplif.*** *14*, 84–95.
3. Jones, G.L., Drennan, W.R. & Rubinstein, J.T. (2013) Relationship between channel interaction and spectral-ripple discrimination in cochlear implant users. ***J. Acoust. Soc. Am***. *133*, 425–433.

4. Shannon, R. V, Fu, Q.J. & Galvin, J. (2004) The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. ***Acta Oto-Laryngologica*** *124*, 50–54.
5. Bonham, B.H. & Litvak, L.M. (2008) Current focusing and steering: modeling, physiology, and psychophysics. ***Hear. Res.*** *242*, 141–153.
6. Srinivasan, A.G., Landsberger, D.M. & Shannon, R. V. (2010) Current focusing sharpens local peaks of excitation in cochlear implant stimulation. ***Hear. Res.*** *270*, 89–100.
7. Vellinga, D., Bruijn, S., Briaire, J.J., Kalkman, R.K. & Frijns, J.H.M. (2017) Reducing interaction In simultaneous paired stimulation with CI. PLoS One *12*, 1–11.
8. Bierer, J.A. (2007) Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration. ***J. Acoust. Soc. Am***. *121*, 1642–1653.
9. Zhu, Z., Tang, Q., Zeng, F.-G., Guan, T. & Ye, D. (2012) Cochlear-implant spatial selectivity with monopolar, bipolar and tripolar stimulation. ***Hear. Res***. *283*, 45–58.
10. Srinivasan, A.G., Padilla, M., Shannon, R. V & Landsberger, D.M. (2013) Improving speech perception in noise with current focusing in cochlear implant users. ***Hear. Res.*** *299*, 29–36.
11. Padilla, M. & Landsberger, D.M. (2016) Reduction in spread of excitation from current focusing at multiple cochlear locations in cochlear implant users. ***Hear. Res***. *333*, 98–107.
12. Frijns, J.H.M., Dekker, D.M.T. & Briaire, J.J. (2011) Neural excitation patterns induced by phased-array stimulation in the implanted human cochlea. ***Acta Otolaryngol.*** *131*, 362–70.
13. van den Honert, C. & Kelsall, D.C. (2007) Focused intracochlear electric stimulation with phased array channels. ***J. Acoust. Soc. Am***. *121*, 3703–16.
14. de Jong, M.A.M., Briaire, J.J. & Frijns, J.H.M. (2019) Dynamic Current Focusing: A Novel Approach to Loudness Coding in Cochlear Implants. ***Ear Hear.*** 40, 34-44.
15. Arenberg, J.G. *et al.* (2018) A Dynamically Focusing Cochlear Implant Strategy Can Improve Vowel Identification in Noise. ***Ear Hear***. *39*, *1136-1145*
16. Won, J.H., Drennan, W.R. & Rubinstein, J.T. (2007) Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. ***J. Assoc. Res. Otolaryngol***. *8*, 384–392.
17. Firszt, J.B. (2003) HiResolution ™ Sound Processing. ***Adv. Bionics White Pap***. *Sylmar, Calif.* 1–4.
18. Büchner, A. *et al.* (2012) Benefits of the HiRes 120 coding strategy combined with the Harmony processor in an adult European multicentre study. ***Acta Otolaryngol.*** *132*, 179–187.
19. Advanced Bionics. (2012) HiRes™ Optima Clinical Results. ***Advanced Bionics White Paper***. *Sylmar, California* 1–2.
20. Kalkman, R.K., Briaire, J.J. & Frijns, J.H.M. (2014) Current focussing in cochlear implants: An analysis of neural recruitment in a computational model. ***Hear. Res.*** *322*, 89–98.
21. Vellinga, D., Briaire, J.J., van Meenen, D.M.. & Frijns, J.H.M. (2017) Comparison of multipole stimulus configurations with respect to loudness and spread of excitation. ***Ear Hear.*** *38*, 487–496.
22. Potts, L.G., Skinner, M.W., Gotter, B.D., Strube, M.J. & Brenner, C.A. (2007) Relation Between Neural Response Telemetry Thresholds, T- and C-Levels, and Loudness Judgments in 12 Adult Nucleus 24 Cochlear Implant Recipients. ***Ear Hear.*** *28*, 495–511.
23. Aronoff, J.M. & Landsberger, D.M. (2013) The development of a modified spectral ripple test. ***J. Acoust. Soc. Am.*** *134*, 217–222.
24. DiNino, M. & Arenberg, J.G. (2018) Age-Related Performance on Vowel Identification and the Spectral-temporally Modulated Ripple Test in Children With Normal Hearing and With Cochlear Implants. ***Trends Hear***. *22*, 1–20.
25. Holden, L.K. *et al.* (2016) Factors Affecting Outcomes in Cochlear Implant Recipients Implanted With a Perimodiolar Electrode Array Located in Scala Tympani. ***Otol. Neurotol.*** *37*, 1662–1668.
26. Lawler, M., Yu, J. & Aronoff, J.M. (2017) Comparison of the Spectral-Temporally Modulated Ripple Test With the Arizona Biomedical Institute Sentence Test in Cochlear Implant Users.

5

Ear Hear. *38*, 760–766.
27. Zhou, N. (2017) Deactivating stimulation sites based on low-rate thresholds improves spectral ripple and speech reception thresholds in cochlear implant users. ***J. Acoust. Soc. Am.*** *141*, 243–248.
28. Won, J.H., Drennan, W.R., Nie, K., Jameyson, E.M. & Rubinstein, J.T. (2011) Acoustic temporal modulation detection and speech perception in cochlear implant listeners. ***J. Acoust. Soc. Am.*** *130*, 376–388.
29. Houben, R. *et al.* (2014) Development of a Dutch matrix sentence test to assess speech intelligibility in noise. ***Int. J. Audiol.*** *53*, 760–763.
30. Francart, T., van Wieringen, A. & Wouters, J. (2008) APEX 3: a multi-purpose test platform for auditory psychophysical experiments. ***J. Neurosci. Methods*** *172*, 283–293.
31. Kollmeier, B. *et al.* (2015) The multilingual matrix test: Principles, applications, and comparison across languages: A review. ***Int. J. Audiol.*** *54*, 3–16.
32. Gatehouse, S. & Noble, W. (2004) The speech, spatial and qualities of hearing scale (SSQ). ***Int. J. Audiol.*** *43*, 85–99.
33. de Jong, M.A.M., Briaire, J.J. & Frijns, J.H.M. (2017) Learning Effects in Psychophysical Tests of Spectral and Temporal Resolution. ***Ear Hear.*** *39*, 475–481.
34. Drennan, W.R., Longnion, J.K., Ruffin, C. & Rubinstein, J.T. (2008) Discrimination of Schroeder-phase harmonic complexes by normal-hearing and cochlear-implant listeners. ***J. Assoc. Res. Otolaryngol.*** *9*, 138–149.
35. Drennan, W., Won, J., Nie, K., Jameyson, E. & Rubinstein, J.T. (2010) Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. ***Hear. Res.*** *262*, 1–8.
36. Liu, Z. & Weinshall, D. (2000) Mechanisms of generalization in perceptual learning. ***Vision Res.*** *40*, 97–109.
37. Irvine, D.R.F. (2018) Auditory perceptual learning and changes in the conceptualization of auditory cortex. ***Hear. Res.*** 1–14.
38. Fu, Q.-J. (2002) Temporal processing and speech recognition in cochlear implant users. ***Neuroreport*** *13*, 1635–1639.
39. Won, J.H., Jones, G.L., Drennan, W.R., Jameyson, E.M. & Rubinstein, J.T. (2013) Evidence of across-channel processing for spectral-ripple discrimination in cochlear implant listeners. ***J. Acoust. Soc. Am.*** *130*, 2088–2097.
40. Smith, Z.M., Parkinson, W.S. & Long, C.J. (2013) Multipolar current focusing increases spectral resolution in cochlear implants. ***Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*** 2796–2799.
41. Rubinstein, J.T. & Hong, R. (2003) Signal coding in cochlear implants: exploiting stochastic effects of electrical stimulation. ***Ann. Otol. Rhinol. Laryngol. Suppl.*** *191*, 14–19.
42. Liberman, M.C. (1978) Auditory-Nerve Response From Cats Raised in a Low-Noise Chamber. ***J. Acoust. Soc. Am.*** *63*, 442–455.
43. Matsuoka, A.J., Abbas, P.J., Rubinstein, J.T. & Miller, C.A. (2000) The neuronal response to electrical constant-amplitude pulse train stimulation: Additive Gaussian noise. ***Hear. Res.*** *149*, 129–137.
44. Hong, R.S. & Rubinstein, J.T. (2003) High-rate conditioning pulse trains in cochlear implants: dynamic range measures with sinusoidal stimuli. ***J. Acoust. Soc. Am.*** *114*, 3327–3342.
45. Litvak, L.M., Spahr, A.J. & Emadi, G. (2007) Loudness growth observed under partially tripolar stimulation: model and data from cochlear implant listeners. ***J. Acoust. Soc. Am.*** *122*, 967–981.
46. Landsberger, D.M., Padilla, M. & Srinivasan, A.G. (2012) Reducing current spread using current focusing in cochlear implant users. ***Hear. Res.*** *284*, 16–24.
47. Van Der Jagt, M.A., Briaire, J.J., Verbist, B.M. & Frijns, J.H.M. (2016) Comparison of the HiFocus Mid-Scala and HiFocus 1J electrode array: Angular insertion depths and speech per-

ception outcomes. ***Audiol. Neurotol. 21***, 316–325.

CHAPTER 6

Effectiveness of Phantom Stimulation to Shift the Pitch Percept in Cochlear Implant Users

Monique A.M. de Jong, Jeroen J. Briaire, Jan Dirk Biesheuvel, Jorien Snel-Bongers, Stefan Böhringer, Guy R. F. M. Timp, and Johan H.M. Frijns

Under review

ABSTRACT

Objectives

Phantom electrodes were developed in an attempt to transmit more low-frequency information through cochlear implant (CI) systems, without inserting the electrode array deeper into the cochlea. Phantom stimulation involves simultaneously stimulating a primary and a compensating electrode with opposite polarity, thereby shifting the electrical field towards the apex and eliciting a lower pitch percept. The current study is the first to compare the effect sizes (in pitch shifts) of multiple phantom configurations by matching the perceived pitch with phantom stimulation to that perceived with monopolar stimulation. Additionally, the effects of electrode location, type of electrode array, and loudness on the perceived pitch were investigated.

Design

Fifteen adult Advanced Bionics CI users participated in this study, which included four experiments to eventually measure the pitch shifts with five different phantom configurations. The proportions of current delivered to the compensating electrode, expressed as σ, were 0.5, 0.6, 0.7, and 0.8 for the symmetrical biphasic pulses ($SBC_{0.5}$, $SBC_{0.6}$, $SBC_{0.7}$, and $SBC_{0.8}$) and 0.75 for the pseudo-monophasic pulse shape ($PSA_{0.75}$). A pitch discrimination experiment was first completed to determine which basal and apical electrode contacts should be used for the subsequent experiments. An extensive loudness balancing experiment followed where both the threshold level (T-level) and most comfortable level (M-level) were determined to enable testing at multiple levels of the dynamic range. A pitch matching experiment was then performed to roughly estimate the pitch shift at the chosen electrode contacts. These rough pitch shifts were then used in the subsequent experiment, where the pitch shifts were determined more accurately.

Results

Reliable data were obtained from 20 electrode contacts. The average pitch shifts were 0.39, 0.53, 0.64, 0.76, and 0.53 electrode contacts towards the apex for $SBC_{0.5}$, $SBC_{0.6}$, $SBC_{0.7}$, $SBC_{0.8}$, and $PSA_{0.75}$, respectively. When only the best configurations per electrode contact were included, the average pitch shift was 0.92 electrode contacts (range: 0.25-2.0). While $PSA_{0.75}$ lead to equal results as the SBC configurations in the apex, it did not result in a significant pitch shift at the base. The pitch shift was significantly larger at the apex and with lateral wall electrode contacts. Loudness did not affect the pitch shift.

Conclusions

Phantom stimulation results in significant pitch shifts, especially at the apical part of the electrode array. The phantom configuration that leads to the largest pitch shift differs between subjects. Therefore, the settings of the phantom electrode should be individualized so that the phantom stimulation is optimized for each CI user. The real added value to the sound quality needs to be established in a take-home trial.

INTRODUCTION

Cochlear implants (CIs) are electronic devices that partially restore hearing in severely hearing-impaired and deaf individuals. Although CI users could score up to 100% correct on speech tests in quiet[1], their understanding of speech in noisy environments declines drastically compared with normal hearing subjects. In addition, CI users report limited perceived sound quality and music appreciation[2]. This is not surprising, as a lot of the frequency information from acoustic sound is lost during CI processing. The sound is filtered between approximately 200 and 7,500 Hz; then, the envelope of the signal is extracted and delivered to the auditory nerve via 12-22 frequency bands, depending on the type of device. The CI users often describe the perceived sound as very high-pitched and sharp, implying that the lower frequencies are underrepresented. It is well known that the transmission of low-frequency information, either electrically or acoustically, is important for speech perception and music appreciation[3–6]. The current study investigated whether the transmission of low-frequency information can be improved using phantom stimulation.

In natural hearing, low-frequency sounds are coded in both place (place pitch) and time (rate pitch) in the apical region of the cochlea. While it is well known that place pitch is coded across the complete cochlea, there is only little evidence that rate pitch is coded at the base[7,8]. CIs make use of the tonotopic organization of the cochlea (i.e., place pitch) by delivering low-frequency signals via the apically located electrode contacts. To achieve proper transmission of low-frequency signals, CI electrode arrays are ideally inserted all the way up to the apex in the cochlea. This deep insertion, however, is limited because of the anatomy of the human cochlea, which becomes narrower towards the apex. In addition, other variables such as the characteristics of the electrode array itself (e.g., stiffness, length, shape, thickness), the experience of the surgeon, and anatomic abnormalities contribute to the relatively shallow insertion depth of the currently available electrode arrays[9]. For example, the HiFocus1J electrode array (Advanced Bionics, Valencia, USA) has a mean angular insertion depth of only 405-480 degrees[10-12]. On the other hand, there is some evidence that when deep insertion is achieved, such as with the MED-EL CI System standard electrode array (Innsbruck, Austria)[10], CI users hear better when the most apical electrodes are deactivated[11]. This implies that the beneficial effect of the deeper inserted electrode array is limited. This could be a consequence of trauma to the spiral ganglion cells due to the deeper insertion[12–14].

As deep insertion of electrode arrays is complex and some studies suggest that it might have a negative effect on understanding speech, alternative methods to deliver low-frequency information to the auditory nerve have been developed. Recent studies have shown that phantom stimulation can produce pitch percepts lower than that of the most apical electrode contact, without needing to further insert the electrode array[15–17]. In phantom stimulation, two electrode contacts, one primary and one compensating, are simultaneously stimulated with opposite polarity. The current directed to the compensation electrode contact is a fraction (denoted as the current compensation coefficient σ) of that administered to the primary electrode contact (Fig. 1b); for example, $\sigma = 1$ means that the amplitude at the compensating contact is equal to the amplitude at the primary contact (i.e., bipolar stimulation), $\sigma = 0.5$ means that the amplitude is 50% of that at the primary contact, and $\sigma = 0$ equals monopolar (MP) stimulation. It is thought that the center of the electrical field, and therefore the perceived pitch, is steered towards the apex because of electrical field shaping with phantom stimulation. Multiple studies demonstrated significant pitch shifts towards the apex with phantom stimulation that used symmetric biphasic (SB) pulse shapes. For example, Saoji and Litvak (2010) found pitch shifts of 0.5-2.0 electrode contacts towards the apex when using the σ values that led to the greatest pitch shift for each subject[18]. These best σ values varied greatly from 0.38 to 0.88, implying that, in some subjects, the pitch shift is smaller at higher σ values. Moreover, one of the subjects heard two distinct pitches at $\sigma = 1$. It was hypothesized that, at these higher σ values, the compensating electrode contact also generates a secondary peak that causes excitation of fibers near the compensating electrode[15,17]. The extra peak, or side lobe, can counteract the effect of lowering the pitch, and could also explain the reported dual pitched tone. This hypothesis was confirmed in our 3D computer model of the human cochlea[18].

To avoid this side lobe phenomenon, pseudo-monophasic (PS) pulses were introduced[16]. These pulses consist of a short- and high-amplitude phase, followed by a long- and low-amplitude phase (Fig. 1c). When the pulse of the primary electrode contact starts with an anodic phase, it is expected that this contact excites the auditory nerve, while the compensating contact does not. This hypothesis is based upon the finding that anodic current and high amplitudes are more effective in exciting the human auditory nerve than cathodic pulses[19]. In line with this hypothesis, Macherey et al. (2011) showed that anodic-first pulses elicit a lower place-pitch than cathodic first pulses when the PS pulse shape was used[16].

Although previous studies showed that phantom stimulation can result in a pitch shift towards the apex, the size of the effect is unknown. Saoji and Litvak (2010) compared multiple SB stimulation modes in pitch-ranking experiments to find the configuration that led to the largest pitch shift for each individual subject[15]. The pitch shift resulting from this best configuration was quantified with a 2-interval, forced-choice procedure. However, because only the best configurations were examined, their results do not include the average pitch shift per phantom configuration. Macherey et al. (2011 and 2012) compared both SB and PS configurations at multiple pulse rates in pitch-ranking experiments[16,17]. Such experiments provide information about the direction of the pitch shift relative to the other configurations and which of the tested configurations lead to the largest pitch shift, but no information about the size of the pitch shift. This and the low number of subjects limited the interpretation of the results. Therefore, the current study quantified the pitch shift after phantom stimulation with multiple configurations by pitch-matching the phantom stimuli to MP (or current steered) stimuli in 2-alternative forced choice tasks in 15 subjects. To improve the reliability of the pitch matching experiments, only the electrode contacts for which the CI users had a relatively high pitch discrimination were tested.

When incorporating phantom stimulation in a speech coding strategy, pitch shifts due to variations in intensity could interfere with the perceptual outcome; thus, the effect of loudness must be studied. While previous studies about phantom stimulation focused on the effect of different configurations, to the best of our knowledge, no data are available about the effect of loudness on the perceived pitch. Previous studies report contradictory results about the effect of the stimulus level on pitch perception for MP stimuli. Arnoldner et al. (2006)[20] and Carlyon et al. (2010)[21] demonstrated an increased pitch with increasing stimulus level, while others found a significant decrease in pitch when the MP stimulus level increased[7,22,23]. The pitch shift following phantom stimulation was modeled as a function of stimulus intensity in our computational model of the human cochlea[18]. The model predicted that the pitch percept was dependent on the stimulation level indeed, with higher stimulation levels leading to a larger pitch shift towards the apex. To study this, the pitch-matching experiments were repeated at multiple loudness levels. The current study is the first to compare the effect size (in pitch shift) of multiple phantom configurations by matching the perceived pitch to that from MP stimulation. Additionally, the effects of the electrode location, the type of electrode array, and the stimulus level on the perceived pitch shift were investigated.

MATERIALS AND METHODS

Subjects

All 15 subjects were unilaterally implanted with an Advanced Bionics HiRes90K or Clarion CII device, with either the HiFocus 1J, HiFocus Mid-scala, or HiFocus 1J with positioner electrode array (Sylmar, CA). Only post-lingually deaf CI recipients were included in this study. The mean phoneme score on open set Dutch monosyllabic (CVC) word tests taken in quiet conditions at 65 dB was 88.7% (range: 78-98%). All subjects had all 16 active electrode contacts in their clinical strategy. Table 1 displays the characteristics of each subject. The study protocol was approved by the committee for Medical Ethics of the Leiden University Medical Center (P02.106 AC).

Phantom stimulation

The concept of both MP and phantom stimulation is schematically illustrated in Figure 1. In the current study, the apical electrode contact was designated as the primary electrode and the basal electrode as the compensating one. The SB pulse shape had a cathodic phase first on the primary electrode (therefore denoted as SBC, Fig. 1b), and the primary and compensating electrode contacts were adjacent to each other. The PS pulse shape was asymmetrical, so that the amplitude of the first phase (which was anodic, therefore denoted as PSA) was 4 times as high as the second phase, and the duration of the second phase was 4 times longer than the first (Fig. 1c). The amplitude of the second PSA phase was reduced 4-fold to maintain charge balancing. The

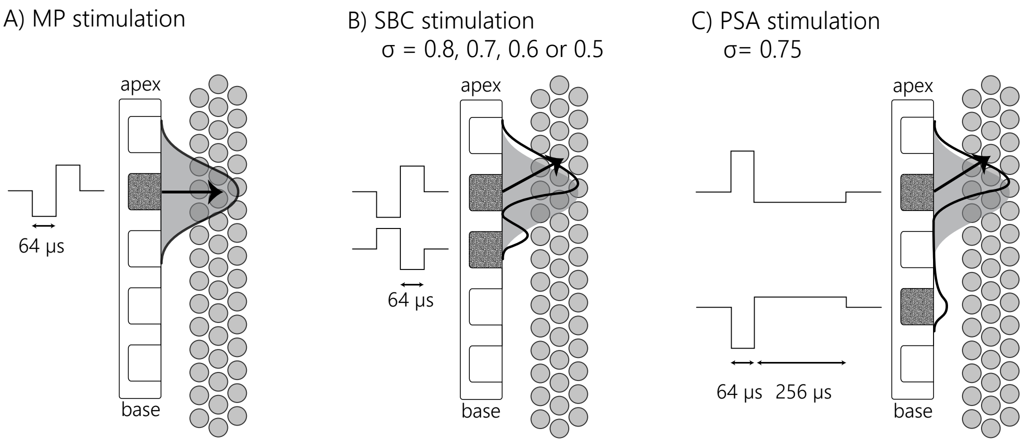


Figure 1. Schematic illustration of different stimulation techniques. Spiral ganglion cells are illustrated as circles. The activated electrode contacts are filled, and the electrical field from monopolar (MP, A) stimulation is shaded grey. The shape of the expected electrical field of the phantom (PE) symmetric biphasic cathodic first pulses (SBC) (B) and pseudo-monophasic anodic first pulse (PSA) (C) are displayed as black lines. The direction of the expected pitch with the different stimulation modes is indicated by the arrow.

TABLE 1. Characteristics of the study subjects

Subject	Gender	Age (y)	CVC (Ph%)	Aetiology	Duration of deafness (y)	CI experience (y)	Implant	CI side	Tested electrode contacts	Pulse width (µs)
S3	Male	59	78	Unknown	4	2	HiRes 90K HiFocus MS	Right	4, 12	64.6
S6	Male	59	93	Hereditary	30	16	Clarion CII HiFocus with positioner	Right	4, 12	64.6
S7	Female	66	87	Unknown	48	10	HiRes 90K HiFocus 1J	Right	12	64.6
S8	Male	72	91	Otosclerosis	23	9	HiRes 90K HiFocus 1J	Right	5, 13	64.6
S9	Female	68	94	Unknown	24	16	Clarion CII HiFocus I with positioner	Left	11	64.6
S10	Male	70	80	Meniere	21	4	HiRes 90K HiFocus 1J	Right	12	75,4
S11	Female	71	83	Hereditary	18	3	HiRes 90K HiFocus 1J	Right	6	97.0
S12	Female	74	88	Unknown	13	8	HiRes 90K HiFocus 1J	Right	9	64.6
S13	Female	49	83	Meningitis	1	14	Clarion CII HiFocus I with positioner	Right	7	64.6
S14	Female	64	91	Hereditary	8	7	HiRes 90K HiFocus 1J	Left	5, 13	86.2
S15	Female	55	98	Unknown	19	16	Clarion CII HiFocus I with positioner	Left	7, 10	64.6
S16	Female	64	96	Unknown	15	16	Clarion CII HiFocus I with positioner	Left	5, 12	86.2
S17	Male	62	92	Noise-Induced	4	13	HiRes 90K HiFocus 1J	Right	5	75.4
S19	Male	66	90	Hereditary	2	3	HiRes 90K HiFocus MS	Right	7,14	64.6
S20	Female	66	86	Hereditary	14	2	HiRes 90K HiFocus MS	Left	8	64.6

CVC, consonant-vowel-consonant; Ph%, percentage phonemes correct on a standard monosyllabic (CVC) word test at 65 dB

6

two stimulated electrodes were spaced one electrode from one another, identical to the configuration used in Macherey and Carlyon (2012)[17]. The compensation coefficient (σ) values used in this study are depicted in Table 2 and were chosen because these configurations were assumed to result in the largest pitch shift without causing pitch reversal[15,16]. The default pulse widths were 64 μs for the first phase (in both SBC and PSA modes), 64 μs for the second phase in SBC, and 256 μs in PSA mode. Stimulus level was measured in clinical units (CUs), according to the formula CU = pulse width (μs) · amplitude (μA) / 78.7. If the stimulus level tended to exceed the compliance level, the phase duration was increased by 10.8 μs per phase to increase the charge. Stimuli were 300 ms long, with pulse rates of 1400 pulses/s, and the time between stimuli was 500 ms.

TABLE 2. Configurations used in this study

Pulse shape	Compensation coefficient σ	denotation
Symmetrically biphasic	0.5, 0.6, 0.7, 0.8	$SBC_{0.5}$, $SBC_{0.6}$, $SBC_{0.7}$, and $SBC_{0.8}$
Pseudo-monophasic	0.75	$PSA_{0.75}$

SBC, Symmetrically biphasic cathodic first pulse shape; PSA, pseudo-monophasic anodic first pulse shape

Experiments

Four experiments were conducted using a custom-made MATLAB (Mathworks, Inc., Natick, MA) interface and the Advanced Bionics' research tool BEDCS (Bionic Ear Data Collection System, Advanced Bionics, Sylmar, CA). (1) A pitch discrimination experiment was first completed to determine which basal and apical electrode contacts should be used for the subsequent experiments. (2) An extensive loudness balancing experiment followed to determine the threshold level (T-level) and most comfortable level (M-level) and enable testing at multiple well-defined levels of the dynamic range. (3) A pitch matching experiment was then performed to roughly estimate the pitch shift at the chosen electrode contacts. (4) These rough pitch shifts were then used in the final experiment to more accurately determine the pitch shifts.

Pitch discrimination experiment

In this 3-alternative forced-choice task, adapted from Biesheuvel et al. (2018)[24], subjects were asked to differentiate a target stimulus from two identical reference stimuli. Both the target and reference stimuli consisted of 300 ms pulse trains with biphasic pulses. Pulse widths were 32 μs per phase and pulse rates were 1400 pulses/s. All stimuli were presented at the M-level, which was determined using the 8-point loudness scale described by Potts et al. (2007)[25]. If the M-level (in CU) exceeded the saturation current, the pulse width was increased in increments of 10.78 µs. The M-level on each electrode contact was loudness balanced with the apically adjacent electrode contact. The target stimulus was based on current steering, which involves the simultaneous stimulation of two adjacent electrode contacts, thereby creating an intermediate pitch percept[26]. Snel-Bongers et al. showed that current steered and MP are equivalent with regard to spread of excitation, channel interaction and threshold levels[27,28]. The proportion of the total current directed to the basal contact was denoted as α, and the proportion to the apical contact as 1 – α. The target stimulus had α values ranging from 0.25 to 1, whereas the α for the reference stimuli was 0 (apical electrode only). Initially, the experiment was repeated 5 times for each electrode pair, with the target stimulus having an α = 1 (basal electrode only), i.e., the spatial difference between the target and reference stimuli was one electrode contact. If the percentage correct exceeded 66% for a certain electrode pair, the test was repeated at a more difficult ratio: with the distance between the target and reference stimuli halved (α = 0.5). If the score at this ratio was still > 66%, α = 0.25 was tested. The final pitch discrimination score for each electrode pair was calculated as follows:

$$\text{Pitch discrimination score} = K - (\%\ \text{correct} \cdot L) \qquad \text{(eq. 1)}$$

with K = the lowest α at which the score was ≥ 66.6% and L = the lowest measured α. The percentage correct refers to the score with the lowest measured α. The pitch shift caused by phantom stimulation was expected to be approximately one electrode contact apical to the main electrode contact[15,29]. Thus, the electrode contacts basally from the best apical and best basal electrode contacts, i.e. those with the lowest α (with a maximum of α = 1), were chosen for further testing. Only electrode pairs with an α ≤ 1.0 were used for further testing to ensure that the subjects were capable of undergoing subsequent testing.

Loudness balancing experiment

To estimate the dynamic range, both threshold levels (T-levels) and M-levels were determined for all electrode contacts, in all the configurations depicted in Table 2 and in MP mode. First, the impedances of all 16 electrode contacts were measured to determine the voltage compliance limit. At higher σ values, the pulse widths may need to be increased instead of the current level to reach equal loudness within the compliance limits of the device [15]. To keep pulse widths equal across all configurations, the highest σ level ($SBC_{\sigma = 0.8}$) pulse width was set as the standard pulse width for all experiments for each subject (see Table 1). The T- and M-levels were determined using the same 8-point loudness scale used in the pitch discrimination experiment[24,25]. Two ascending and 2 descending trials per electrode contact were performed and the average M-levels were calculated. The SBC and PSA M-levels were balanced with the MP M-levels by sequentially presenting the two stimuli, and adjusting the MP current level until equal loudness was achieved.

Pitch matching experiment

The pitch matching experiment, which was a 2-up-2-down procedure, was conducted to roughly estimate the pitch shift for all the SBC and PSA configurations at 2 locations along the electrode array, 1 basal and 1 apical electrode contact determined in the pitch discrimination experiment. Phantom and MP stimuli were administered alternately. At the initiation of the experiment, the MP stimulus was delivered 3 electrode contacts apical or 1 electrode contact basal to the main electrode contact, in random order, while the location of the phantom stimulus remained constant. The MP stimulus gradually changed pitch towards the main electrode contact (in step sizes of 0.05 α) using current steering, until the 2 stimuli were perceived as equal in pitch. Then, the pitch of the MP signal was shifted beyond the main contact until the pitch was distinctive again. Next, the MP stimulus was shifted back to where the MP and phantom signal were equal in pitch again, after which the experiment was terminated. Ten percent loudness roving was applied to prevent the loudness from influencing the results. The experiment was repeated 4 times per configuration, and the average was used as the reference electrode contact in the final pitch shift experiment.

Pitch shift experiment

In this 2-alternative forced-choice task, based on experiment 4 in the paper by Saoji and Litvak (2010)[15], the pitch percept of a phantom stimulus was compared with that of current steered MP stimuli. The MP stimuli were presented at the (virtual) electrode contacts that were 0.0, 0.25, 0.5, 0.75, 1.0, 2.0, and 4.0 contacts from the reference electrode contact, determined in the pitch matching experiment. The phantom and MP stimuli were sequentially presented in a random order, after which the subjects were asked to indicate which of the two stimuli was higher pitched, with no feedback about the right answer provided. Each phantom configuration was compared with the 13 MP stimuli and repeated 15 times, resulting in blocks of 195 trials. The stimuli were presented at the M-level, determined in the loudness balancing experiment, with loudness roving of 10%. To test the effect of loudness on pitch shift, the experiment was repeated at 75% and 50% of the M-level. Due to time constraints, this effect of loudness was only evaluated with the PSA configuration and the SBC configuration that resulted in the largest pitch shift for that specific subject.

Data analysis

To quantify the degree of pitch shift, each subject's data were fit with a cumulative Gaussian psychometric function using the "***psignifit***" algorithm[30,31]. To assure only true pitch shifts were measured, the reliability of the psychometric functions was assessed. It was assumed that subjects are able to discriminate stimuli that are spatially separated two electrode contacts from each other, and this was confirmed in the pitch discrimination experiment. The MP stimuli at the upper and lower limit of the psychometric function were separated 2 and 4 electrode contacts from the electrode location at which the phantom stimulus was expected to be perceived (the reference electrode contact). Therefore, subjects should be able to correctly indicate if the MP stimulus was higher pitched than the phantom stimulus, or not. If the subject was unable to reach a correct score of at least 66.6% in these comparisons, the measurement was considered unreliable and was discarded from the analysis. This was the case for three phantom configurations for (subject number-electrode number) S07-E12, S11-E06, and S12-E09, two configurations for S19-E7, and one configuration for S03-E04, S06-E04, S06-E12, and S16-E12. All data for S10-E12 and S19-E14 were excluded because 6/9 and 5/9 measurements, respectively, did not meet the reliability rules. Ten out of the 15 discarded measurements were obtained at lower loudness levels (50% or 75% of the M-level). This was in line with our expectations, as the task difficulty increases at lower loudness levels. In total, data were obtained for 14 subjects, 6 of whom were measured at two locations along the electrode array. This means that data were obtained for 20 electrode contacts. There were 9 measurements per electrode contact, resulting in a total of 180 measurements, 15 of which were discarded because of the reliability rules described above.

Statistical analysis

The *psignifit* software package for MATLAB provides a pitch shift value with a confidence interval per tested setting. To account for the uncertainty in the measurement process and the reliability of each measurement, multiple imputations were made. Measurements were imputed independently from the normal distribution corresponding to the confidence interval. All further analyses were performed on 10 imputed data sets and final results were based on pooling, using Rubin's rule implemented in SPSS[33]. For the pitch matching experiment, a linear mixed model analysis was used. A linear mixed model takes into account that measurements taken on an individual are more similar than measurements taken on different individuals. Furthermore, it corrects for missing data. Pitch shift values were checked for normality using histograms and did not show deviations. All data were analyzed with SPSS 23 (Statistical Package for the Social Sciences, SPSS Inc., Chicago, IL) software.

RESULTS

Pitch discrimination

The results of the pitch discrimination experiment are shown in Table 3. The grey shaded electrode contacts were those that were located basally from the best functioning electrode contacts and were chosen for subsequent testing. Only electrode contacts with a score of 1.0 or less, or that had at least 1 adjacent electrode contact with a maximum α=1.0 were selected for subsequent testing. The far right column in Table 3 lists the selected electrodes. Due to time constraints, E5-E9 and E13-E16 were not tested for subjects 3 and 6.

Reliability of the experiments

The individual pitch matching results are combined with the individual results of the pitch shift experiment in Figure 2. The pitch matching reference point usually fell within the error bars of the pitch shift experiment when there was SBC stimulation, showing that the pitch matching experiment had additional value for the reliability of the pitch shift experiment. However, with PSA stimulation, the results of the two methods of measuring the pitch shift differed considerably. The reference pitch was lower than the results from the pitch shift experiment during 6 of the 11 apical measurements.

Effect of phantom configuration on the mean pitch shift

Figure 3 depicts the mean pitch shift per phantom configuration and electrode location at the M-level, calculated with a linear mixed model analysis. Subjects confirmed that all of the presented stimuli were perceived as one clear single tone. The mean pitch shifts were 0.39 (SE=0.14), 0.53 (SE=0.14), 0.64 (SE=0.15), and 0.76 (SE=0.14) electrode contacts towards the apex for $SBC_{0.5}$, $SBC_{0.6}$, $SBC_{0.7}$, and $SBC_{0.8}$, respectively; i.e., the pitch shift increased with increasing σ value for the SBC configurations. For $PSA_{0.75}$, the mean pitch shift was 0.53 (SE=0.14). A linear mixed model with electrode location and phantom configuration as factors showed that only the pitch shifts of $SBC_{0.5}$ and $SBC_{0.8}$ were statistically significantly different from one another (p=0.005). All configurations were statistically significantly different from 0, where phantom stimulation was not used (p<0.01 for all configurations).

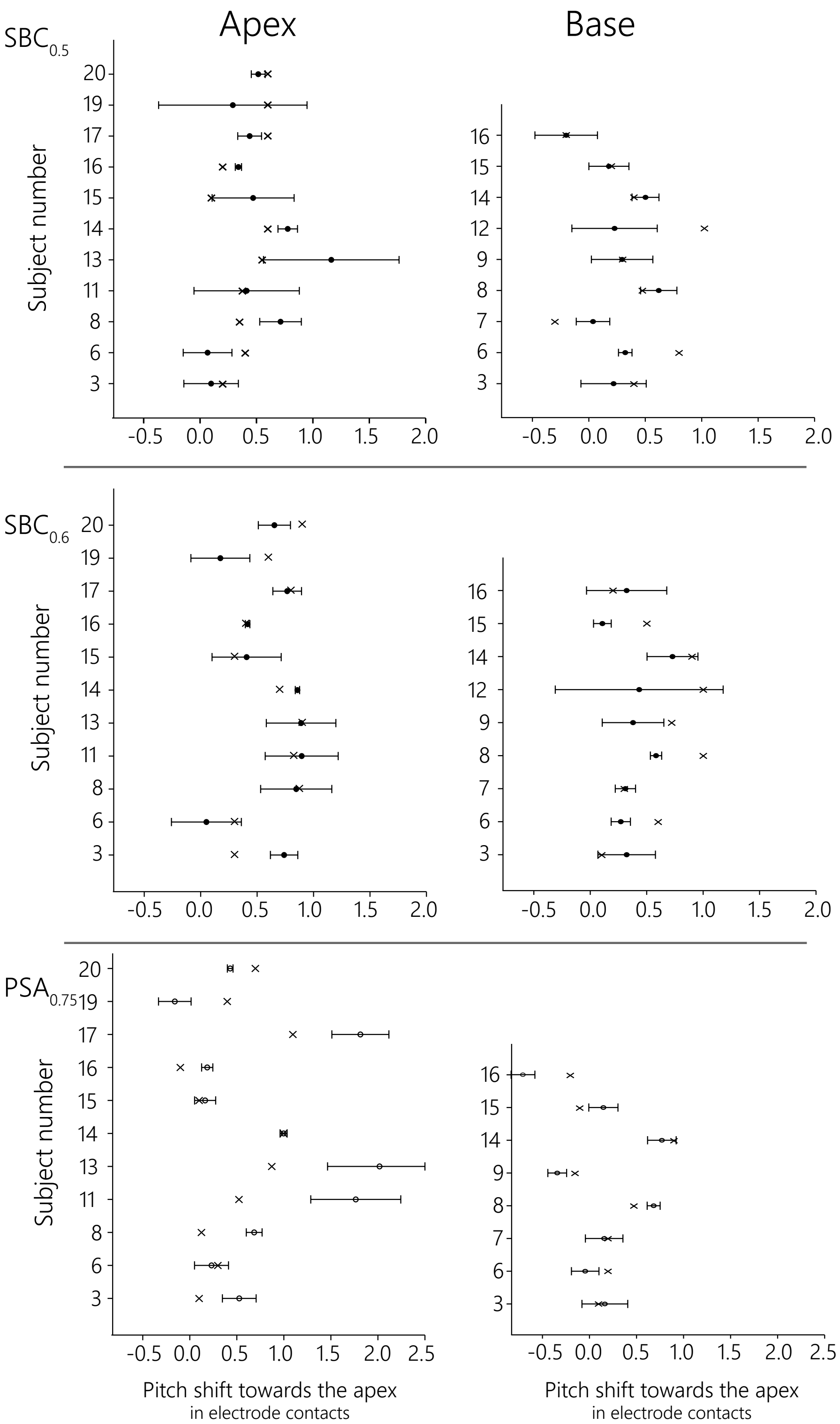
The largest pitch shifts per electrode contact were most often obtained with $SBC_{0.7}$ and $SBC_{0.8}$ (both 6 times), followed by $SBC_{0.6}$ and $PSA_{0.75}$ (both 3 times), and least often with $SBC_{0.5}$ (2 times). When only these results were included in the analysis, as in Saoji and Litvak (2010), the mean pitch shift was 0.92 electrode contacts towards the apex, with a minimum of 0.25 and a maximum of 2.0 electrode contacts. The mean pitch shift at the apex was significantly larger (0.66, SE=0.12) than at the basal part of the electrode array (0.47, SE=0.12) (p=0.04) (Figure 3b). Therefore, the effect of configuration was

TABLE 3. Pitch discrimination scores per electrode pair

	Electrode pair															Selected primary electrodes
Subject	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11	11-12	12-13	13-14	14-15	15-16	
S3	1.4	0.4	0.8	1.4					1.4	1.4	0.7	2				4, 12
S6	0.8	0.35	0.8	0.35					0.35	0.4	0.3	0.7				4, 12
S7	1.6	1.8	1.6	1.8	1.4	1.6	1.6	0.9	1.4	0.8	0.8	1.4	0.7	1.4	0.45	12
S8	1.6	0.8	0.8	0.45	0.7	0.8	0.4	0.7	0.7	1.4	0.4	0.45	0.35	0.7	1.4	5, 13
S9	1.4	2	0.8	1.4	1.6	0.8	1.6	1	0.8	0.45	1.4	1.4	0.7	1.8	0.8	11
S10	1.6	1.4	1.4	1.8	1.4	0.7	1.4	0.8	1.4	1	0.4	0.9	1.4	0.7	0.4	12
S11	0.8	1.6	1.4	1.4	0.8	0.8	1.6	1.4	0.8	1.4	1.6	1.8	1.4	1.4	0.8	6
S12	1.6	1.4	0.8	0.8	1.4	1.4	0.9	0.4	1.4	0.9	0.9	1.4	1.4	0.9	1.4	9
S13	0.8	1.6	1.4	1.6	1.8	0.4	1.6	1.8	1.6	0.8	1.8	1.6	1.6	1.6	1.6	7
S14	0.35	0.35	0.45	0.35	0.4	0.4	0.45	0.35	0.7	0.8	0.4	0.3	0.35	0.7	0.8	5, 13
S15	0.8	1.8	1.4	1.4	0.7	0.4	1.4	0.9	0.4	1.6	0.7	1.4	1.8	1.8	0.7	7, 10
S16	1.6	0.8	0.4	0.3	0.35	0.4	0.7	0.7	0.3	0.4	0.4	0.3	0.45	0.7	0.8	5, 12
S17	1.6	1.4	1.4	0.8	1.4	1.4	1.6	2	1.4	1.4	1.6	1.6	1.8	1.6	0.9	5
S19	1.6	0.7	1.4	1.6	0.8	0.7	0.9	1.4	0.9	1.6	1.4	0.8	0.8	1.4	1.8	7,14
S20	0.8	0.45	1.4	1.4	1.4	0.7	0.5	0.9	0.9	2	1.4	1.8	1.4	0.7	2	8

The electrode pairs around the electrode contacts that were selected for further assessment are shaded grey. Missing data is shaded light grey.

6


Apex
Base
SBC0.5
SBC0.6
PSA0.75
Subject number
20
19
17
16
15
14
13
12
11
9
8
7
6
3
-0.5 0.0 0.5 1.0 1.5 2.0
-0.5 0.0 0.5 1.0 1.5 2.0 2.5
Pitch shift towards the apex
in electrode contacts

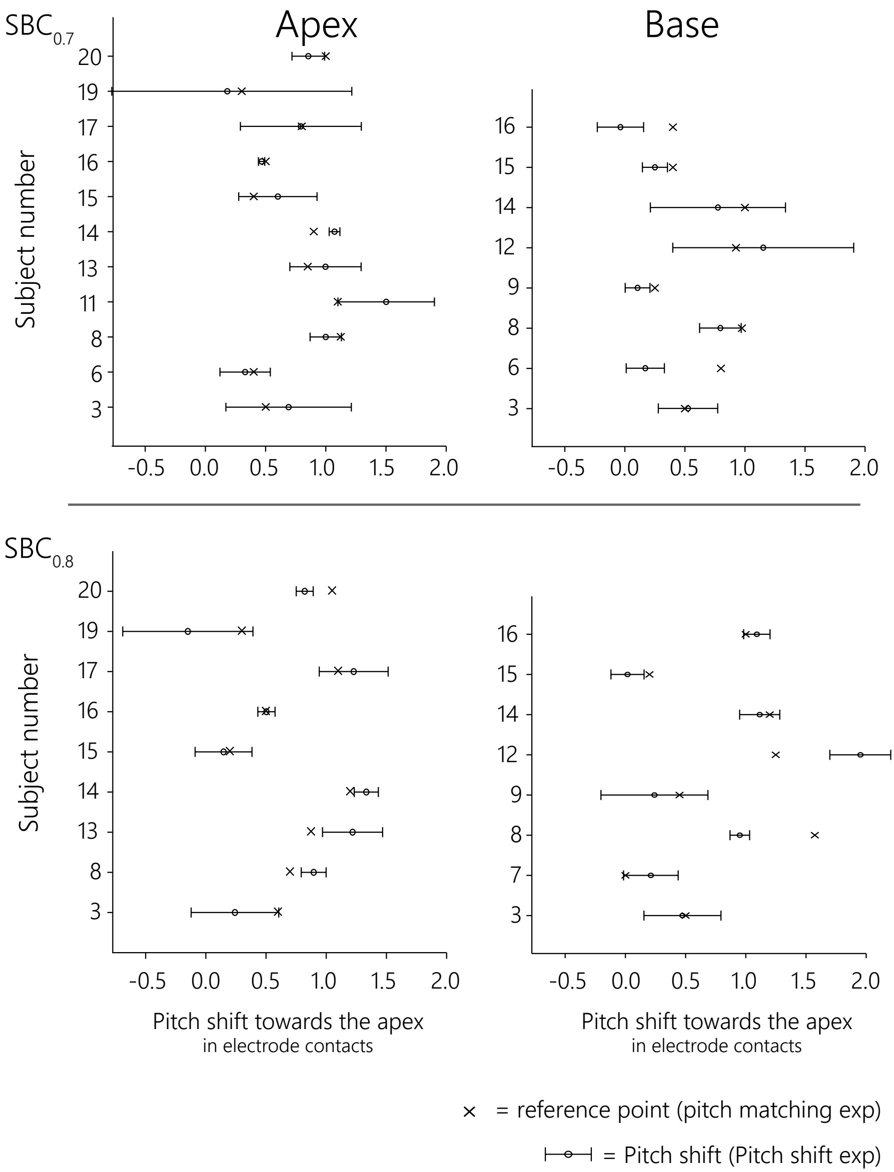


FIGURE 2. INDIVIDUAL PITCH MATCHING AND PITCH SHIFT EXPERIMENT RESULTS FOR THE SBC AND THE PSA CONFIGURATIONS AT THE MOST COMFORTABLE LEVEL. The X is the reference point measured in the pitch-matching experiment. The circles with error bars show the mean and standard errors of the pitch shift measured in the pitch shift experiment. Abbreviations: SBC, symmetric biphasic pulse shape, cathodic first; PSA, pseudo-monophasic pulse shape, anodic first.

6

analyzed separately for the apical and basal electrode contacts (Figure 3c and 3d). A linear mixed model revealed that the differences in pitch shifts between the phantom configurations are especially prominent when phantom stimulation is applied to the more basally located electrode contacts. Both $SBC_{0.5}$ and $SBC_{0.8}$ (p=0.005) and $SBC_{0.8}$ and $PSA_{0.75}$ (p=0.001) differed significantly at the base, while no significant differences were found at the apex. Interestingly, $PSA_{0.75}$ stimulations did not result in a significant pitch shift at the base (p=0.47) but led to a substantial pitch shift (p<0.001) at the apical electrode contacts. Also, no significant pitch shift was achieved at the base (p=0.13) with $SBC_{0.5}$. All other configurations were statistically significantly different from MP stimulation at both the apical and basal electrode contacts.

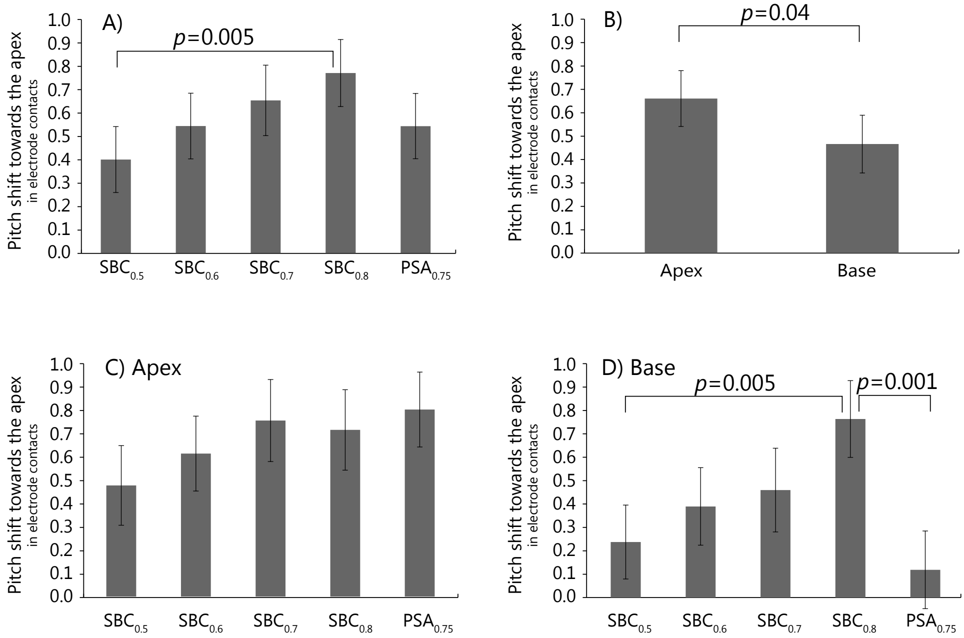


FIGURE 3. MEAN PITCH SHIFT, IN ELECTRODE CONTACTS, FROM THE MAIN ELECTRODE CONTACT TOWARDS THE APEX AT THE M-LEVEL. The effects of different phantom configurations (A), electrode location (B), and different phantom configurations per electrode location (C, D) on the mean pitch shift. All mean and standard errors were calculated using a linear mixed model analysis with either "configuration", "electrode location", or "configuration and electrode location" as factors. The error bars represent the standard error. The displayed means were compared in a more extensive linear mixed model, as described in the Results section, and only statistically significant p-values (<0.05) are displayed. Abbreviations: SBC, symmetric biphasic pulse shape, cathodic first; PSA, pseudo-monophasic pulse shape, anodic first.

As the electrode location influences the pitch shift and the electrode type is known to influence the location of the electrode contacts, an additional analysis of the effect of electrode type on pitch shift was performed. Electrode types were divided into (semi-) medial wall (HiFocus 1 with positioner or HiFocus MS, N=8) and lateral wall

(N=7) electrodes. Because the HiFocus 1 with positioner electrode array is positioned closer to the modiolus than without positioner[34], subjects implanted with this electrode array were assigned to the (semi-) medial wall group. The results are displayed in Figure 4.

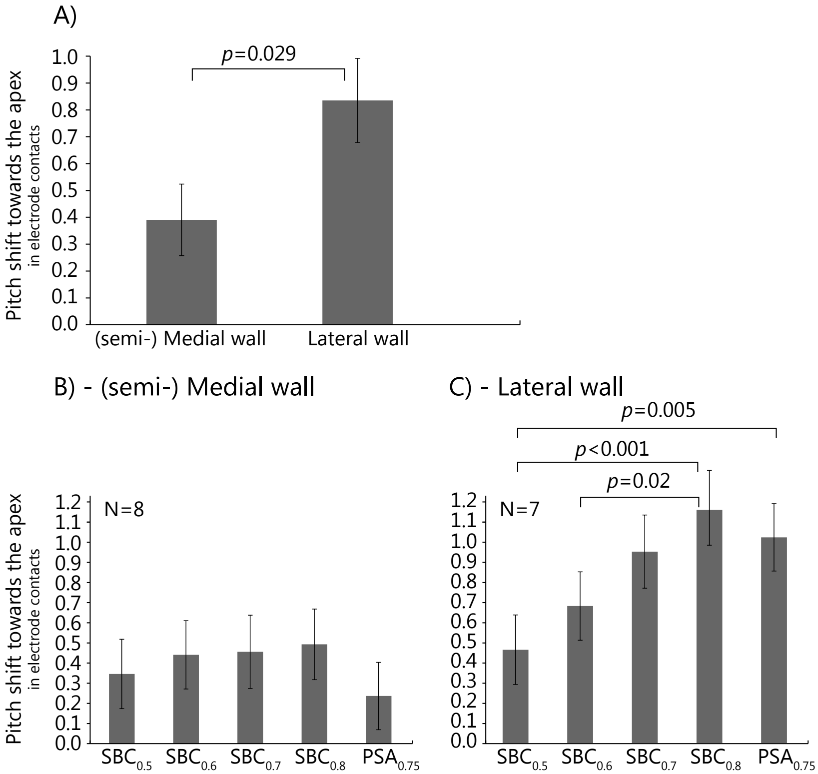


Figure 4. Effect of electrode type on pitch shift. Mean pitch shift from the main electrode contact for the (semi-)medial wall (i.e. mid-scalar or peri-modiolar) and lateral wall electrodes at the M-level (A). The effects of the different phantom configurations are displayed for the (semi-)medial wall electrodes (B) and lateral wall electrodes (C) separately. All displayed mean and standard errors were calculated using a linear mixed model analysis with either "electrode type" or "configuration" as factors. The error bars represent the standard error. The displayed means were compared in a more extensive linear mixed model, as described in the Results section, and only the statistically significant p-values (<0.05) are displayed. Abbreviations: SBC, biphasic cathodic first pulse shape; PSA, pseudo-monophasic pulse shape, anodic first.

A linear mixed model with the factors "electrode type" and "phantom configuration" showed that the mean pitch shift with lateral wall electrode arrays (0.83, SE=0.16) was significantly larger (p=0.029) than that with (semi-) medial wall electrode arrays (0.39, SE=0.13). While the lateral wall electrode arrays significantly differed from MP stimulation (p<0.001), the (semi-) medial wall arrays did not (p=0.12). When comparing the two electrode types separately, the configuration type did not influence the pitch shift in (semi-) medial wall electrode arrays, while higher degrees of phantom stimulation (higher σ values) did lead to larger pitch shifts in lateral wall electrodes. Specifically, $SBC_{0.8}$ (1.16, SE=0.19) led to a significantly larger pitch shift than both $SBC_{0.5}$ (0.46, SE=0.19, $p<0.001$) and $SBC_{0.6}$ (0.68, SE=0.19, $p=0.02$). Also $PSA_{0.75}$ (1.02, SE=0.2) showed a significantly larger pitch shift ($p=0.005$) than $SBC_{0.5}$.

The mean pitch shifts at three stimulus levels are displayed in Figure 5 for both the SBC and PSA configurations. For SBC stimulation only, the configuration that led to the largest pitch shift was tested at loudness levels other than the M-level. A linear mixed model with the factors “loudness” and “configuration” revealed no significant differences between the pitch shifts at the M-level versus 75% and 50% of the M-level.

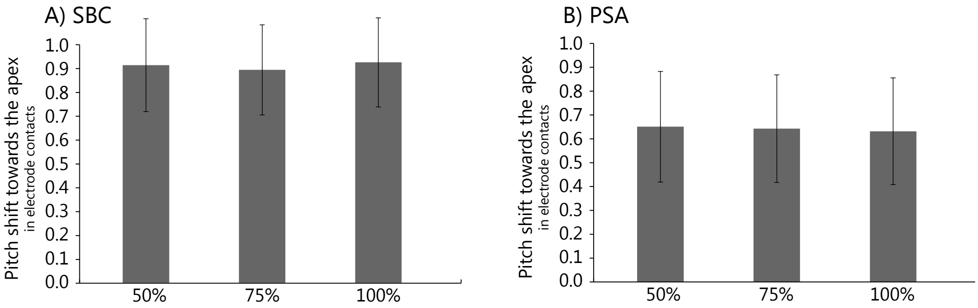


Figure 5. The effect of loudness on pitch shift. Mean pitch shift from the main electrode contact towards the apex at different loudness levels for the biphasic phantom configuration (A) and the pseudo-monophasic configuration (B), both calculated using a linear mixed model with “loudness level” as the main effect. The loudness levels are calculated as percentages of the amplitudes at the most comfortable loudness. For each subject, only the biphasic configurations that had the largest pitch shift at 100% of the most comfortable level was measured at all loudness levels. There was no significant difference in pitch shift between loudness levels. Error bars represent the standard error. Abbreviations: SBC, symmetric biphasic pulse shape, cathodic first; PSA, pseudo-monophasic pulse shape, anodic first.

DISCUSSION

This study showed that both the SBC and PSA phantom stimulation modes cause a statistically significant shift of the pitch percept towards the apex. Of the tested phantom configurations, $SBC_{0.8}$ caused the largest average pitch shift of 0.76 (SE=0.14) electrode contacts. However, there was great variability within and between subjects. A higher degree of phantom stimulation (higher σ value) does not always cause a bigger pitch shift, which is in line with previous studies that describe pitch reversals following phantom stimulation[15,17]. When only the configurations that resulted in the largest pitch shift for each specific electrode contact were considered, the mean pitch shift was 0.92 (range: 0.25-2.0) electrode contacts, which is comparable to previous research[15,16]. The pitch shift was significantly larger at the apex of the electrode array, and also for lateral wall electrode arrays versus (semi-) medial wall electrode arrays.

The current study is an addition to the existing literature about phantom stimulation because of its relatively high number of study subjects (15) and more accurate testing of the perceived pitch shift following phantom stimulation. The direct comparison with the pitch perceived with MP stimulation is advantageous because it enables more accurate quantification of the pitch shift. Nevertheless, it is extremely difficult for CI users to distinguish pitches and indicate what the higher-pitched sound is, and this could lead to uncertain results on psychophysical tests concerning pitch, especially at lower loudness levels. To compensate for this in the current study, only electrode contacts with high pitch discrimination scores were selected for further testing, and an initial pitch matching test was performed to increase the reliability of the pitch shift test in high performing subjects. A significant correlation was found between pitch discrimination scores and variations in psychometric functions (R^2=0.157, n=165, p<0.001). This implies that the selection of high performing electrodes leads to less variation in the pitch shift test, signifying the importance of the pitch discrimination experiment. The pitch matching results were similar to the measured pitch shifts for all SBC configurations at all electrode contacts and for the $PSA_{0.75}$ results taken at the more basally located electrode contacts. Interestingly, the pitch matching and pitch shift measurements did not match for the $PSA_{0.75}$ configuration measured at the more apically located electrode contacts. This could be a consequence of cross-turn stimulation in the apex, as described by Finley and Skinner (2008)[13] and Frijns et al. (2001)[35]; although, we cannot explain why this is specifically the case for $PSA_{0.75}$ stimulation and not for the other tested stimulation modes. The configurations were tested in a random order, to exclude the role of fatigue or learning.

When comparing the pitch shifts with the different phantom configurations, we specifically looked at the apical measurements because a phantom electrode contact would be implemented at the most apical location of the electrode array. The $SBC_{0.8}$, $SBC_{0.7}$, and $PSA_{0.75}$ configurations showed the largest pitch shifts, with approximately the same average pitch shift of 0.75 electrode contacts towards the apex. Nevertheless, the variation between subjects differed for the three best performing configurations with the $SBC_{0.8}$ configuration resulting in the most variation between subjects, followed by $PSA_{0.75}$, while $SBC_{0.7}$ was the most constant across subjects (Figure 2). For that reason, one could conclude that $SBC_{0.7}$ would be the most convenient configuration to implement in a speech coding strategy, in a one-fits-all construction. On the other hand, the greatest pitch shifts were achieved with $PSA_{0.75}$ (up to 2 electrode contacts), implying that a larger gain could be achieved for some subjects. Therefore, we recommend an individual fitting, in which a pitch ranking experiment is performed with all three configurations to determine which should be implemented in the final speech coding strategy. The reason for the greater variation between subjects with $SBC_{0.8}$ and $PSA_{0.75}$ probably has to do with individual differences in electrode location. The $SBC_{0.8}$ setting has a relatively high amplitude on the compensating electrode contact. If this contact lies relatively close to the spiral ganglion cells, it could stimulate the auditory nerve on its own, counteracting the electrical field shaping with phantom stimulation in some subjects. The $PSA_{0.75}$ setting has a larger distance between the two involved electrode contacts, which could also influence the effectiveness of electrical field shaping.

Subjects with a lateral wall electrode achieved higher pitch shifts than those with a (semi-) medial wall electrode array, which was also predicted from our computational model of the human cochlea[18]. In (semi-) medial wall electrode arrays, the electrode contacts are placed relatively close to the auditory nerve[36], while a certain distance between the electrodes and the spiral ganglion cells is necessary to effectuate electrical field shaping. This has also been shown in other strategies that make use of electrical field shaping, for example, current focusing that shapes the electrical field to increase spatial selectivity. In a computational model of the human cochlea, Kalkman et al. (2014b) demonstrated that closer proximity to the spiral ganglion cells results in less effective current focusing[14]. Moreover, the close proximity can cause neuronal excitation by the compensating electrode already at low degrees of phantom stimulation[37], canceling out the pitch shift towards the apex. This also explains why an increase in the degree of phantom stimulation does not increase the pitch shift for the (semi-) medial wall electrodes, but that a clear trend is visible for the lateral-wall electrodes (Figure 4b and 4c).

Interestingly, the effect of electrode location was also significant, as a larger pitch shift was observed at the apex than at the base. While this is beneficial for the implementation of phantom stimulation in a speech coding strategy, as one would use phantom stimulation specifically on the most apical electrode contact, the rationale behind it is

uncertain. It could be that the electrode contacts at the apex have a larger distance to the auditory nerve than the basal electrodes, although this was not seen in the current dataset. Another hypothesis that could not be confirmed in the current study was that the neural survival at the apex is better than at the base of the cochlea[38], resulting in lower thresholds that could decrease the chance for side lobe activation[14]. Nevertheless, it is likely that neural survival plays an important role in the effect of phantom stimulation. For example, if the compensating electrode is in a dead region[39], it may not excite nearby auditory nerve fibers; thus, potential side lobes will not cause neural excitation and will have no detrimental effect, even for high σ values. Another potential reason for this benefit at the apex is that the spread of excitation (and therefore channel interactions) might be greater at the apex[23], enhancing electrical field shaping.
In contrast to previous studies, loudness did not affect the perceived pitch in the current study. The computational model of the human cochlea predicted a decrease in pitch shift at higher loudness levels. The hypothesis was that the steering of the center of excitation is caused by a suppression of the excitation on the basal side (compensating electrode), while the fibers at the main contact are still excited. At low levels, the number of fibers that can be suppressed on one side is smaller; thus, the phantom effect is diminished. Although this is a plausible hypothesis, the current study cannot confirm this, and the effect of loudness was not clear from previous reports. If loudness has a limited effect on pitch, indeed, this leads to easier implementation of phantom stimulation in speech coding strategies.

Future perspectives

The phantom electrode technique works. Previous studies reported positive effects on speech perception when incorporating it in a speech coding strategy[6,41]. However, to minimize the risk of unwanted excitation near the compensating electrode contact that may cause a pitch reversal or dual-tone, only phantom-based speech coding strategies that use relatively low σ values were studied. The results of the current study imply that the beneficial effect can be even greater when higher σ values are used. Because the best configuration is different for each individual subject and electrode contact, it might be helpful to perform a pitch ranking or pitch discrimination test before fitting subjects with a phantom strategy. This individualization of speech coding strategies might be advantageous not only for phantom stimulation but also for other speech coding strategies. Moreover, the perceived pitch shift following phantom stimulation is greatest in CI users with a lateral wall electrode array, and this should be considered during clinical implementation.

REFERENCES

1. Rak, K. *et al.* (2017) Stable Longitudinal Performance of Adult Cochlear Implant Users for More Than 10 Years. *Otol. Neurotol. 38,* e315–e319.
2. Kong, Y.Y., Cruz, R., Jones, J.A. & Zeng, F.G. (2004) Music Perception with Temporal Cues in Acoustic and Electric Hearing. *Ear Hear. 25,* 173–185.
3. Hochmair, I. *et al.* (2003) Deep electrode insertion in cochlear implants: Apical morphology, electrodes and speech perception results. *Acta Otolaryngol. 123,* 612–617.
4. Qi, B. *et al.* (2010) The contribution of apical stimulation to Mandarin speech perception in users of the MED-EL COMBI 40 cochlear implant. *Acta Otolaryngol. 131,* 52–58.
5. Von Ilberg, C.A., Baumann, U., Kiefer, J., Tillein, J. & Adunka, O.F. (2011) Electric-acoustic stimulation of the auditory system: A review of the first decade. *Audiol. Neurotol. 16,* 1–30.
6. Munjal, T., Roy, A.T., Carver, C., Jiradejvong, P. & Limb, C.J. (2015) Use of the Phantom Electrode strategy to improve bass frequency perception for music listening in cochlear implant users. *Cochlear Implants Int. 16,* S121–S128.
7. Townshend, N., Van Compernolle, D. & White, R.L. (1987) Pitch perception by cochlear implant subjects. *J. Acoust. Soc. Am. 82,* 106–115.
8. Carlyon, R.P., Lynch, C. & Deeks, J.M. (2010) Effect of stimulus level and place of stimulation on temporal pitch perception by cochlear implant users. *J. Acoust. Soc. Am. 127,* 2997–3008.
9. Cosetti, M.K. & Waltzman, S.B. (2012) Outcomes in cochlear implantation: Variables affecting performance in adults and children. *Otolaryngol. Clin. North Am. 45,* 155–171.
10. Buchman, C., Dillon, M., King, E., Adunka, M. & Pillsbury, H. (2014) Influence of cochlear implant insertion on performance: a prospective randomized. *Otol neurotol 35,* 1773–9.
11. Gani, M., Valentini, G., Sigrist, A., Kós, M.I. & Boëx, C. (2007) Implications of deep electrode insertion on cochlear implant fitting. *JARO - J. Assoc. Res. Otolaryngol. 8,* 69–83.
12. Wardrop, P., Whinney, D., Rebscher, S.J., Luxford, W. & Leake, P. (2005) A temporal bone study of insertion trauma and intracochlear position of cochlear implant electrodes. II: Comparison of Spiral Clarion and HiFocus II electrodes. *Hear. Res. 203,* 68–79.
13. Finley, C.C. & Skinner, M.W. (2008) Role of electrode placement as a contributor to variability in cochlear implant outcomes. *Otol. Neurotol. 29,* 920–928.
14. Kalkman, R.K., Briaire, J.J. & Frijns, J.H.M. (2014) Current focussing in cochlear implants: An analysis of neural recruitment in a computational model. *Hear. Res. 322,* 89–98.
15. Saoji, A. a & Litvak, L.M. (2010) Use of 'phantom electrode' technique to extend the range of pitches available through a cochlear implant. *Ear Hear. 31,* 693–701.
16. Macherey, O., Deeks, J.M. & Carlyon, R.P. (2011) Extending the limits of place and temporal pitch perception in cochlear implant users. *J. Assoc. Res. Otolaryngol. 12,* 233–51.
17. Macherey, O. & Carlyon, R.P. (2012) Place-pitch manipulations with cochlear implants. *J Acoust Soc Am 131,* 2225–2236.
18. Snel-Bongers, J. (Leiden University, The Netherlands, 2013) Dual electrode stimulation in cochlear implants: From concept to clinical application (Doctoral thesis). Available at: https://openaccess.leidenuniv.nl/handle/1887/22291.
19. Macherey, O., Carlyon, R.P., Van Wieringen, A., Deeks, J.M. & Wouters, J. (2008) Higher sensitivity of human auditory nerve fibers to positive electrical currents. *JARO - J. Assoc. Res. Otolaryngol. 9,* 241–251.
20. Arnoldner, C., Kaider, A. & Hamzavi, J. (2006) The role of intensity upon pitch perception in cochlear implant recipients. *Laryngoscope 116,* 1760–1765.
21. Carlyon, R.P., Lynch, C. & Deeks, J.M. (2010) Effect of stimulus level and place of stimulation on temporal pitch perception by cochlear implant users. *J. Acoust. Soc. Am. 127,*

2997–3008.
22. Pijl, S. (1997) Pulse Rate Matching by Cochlear Implant Patients Effects of Loudness Randomization and Electrode Position. *Ear Hear. 18,* 316–325.
23. Kalkman, R.K., Briaire, J.J., Dekker, D.M.T. & Frijns, J.H.M. (2014) Place pitch versus electrode location in a realistic computational model of the implanted human cochlea. *Hear. Res. 315,* 10–24.
24. Biesheuvel, J., Briaire, J.J., de Jong, M.A.M. & Frijns, J.H.M. (2018) Clinical assessment of channel discernibility along the cochlear implant electrode array and its relation to speech perception. *Submitted*
25. Potts, L.G., Skinner, M.W., Gotter, B.D., Strube, M.J. & Brenner, C.A. (2007) Relation Between Neural Response Telemetry Thresholds, T- and C-Levels, and Loudness Judgments in 12 Adult Nucleus 24 Cochlear Implant Recipients. *Ear Hear. 28,* 495–511.
26. Firszt, J.B., Koch, D.B., Downing, M. & Litvak, L. (2007) Current Steering Creates Additional Pitch Percepts in Adult Cochlear Implant Recipients. *Otol. Neurotol. 28,* 629–636.
27. Snel-Bongers, J., Briaire, J.J., Vanpoucke, F.J. & Frijns, J.H.M. (2012) Spread of Excitation and Channel Interaction in Single- and Dual-Electrode Cochlear Implant Stimulation. *Ear Hear. 33,* 367–376.
28. Snel-Bongers, J., Briaire, J.J., Van Der Veen, E.H., Kalkman, R.K. & Frijns, J.H.M. (2013) Threshold levels of dual electrode stimulation in cochlear implants. *JARO - J. Assoc. Res. Otolaryngol. 14,* 781–790.
29. Carlyon, R.P., Monstrey, J., Deeks, J.M. & Macherey, O. (2014) Evaluation of a cochlear-implant processing strategy incorporating phantom stimulation and asymmetric pulses. *Int. J. Audiol. 53,* 871–879.
30. Wichmann, F. a. & Hill, N.J. (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys. 63,* 1293–1313.
31. Wichmann, F.A. & Hill, N.J. (2001) The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Percept. Psychophys. 63,* 1314–1329.
32. Harel, O. & Zhou, X.-H. (2007) Multiple imputation: review of theory, implementation and software. *Stat. Med. 26,* 3057–3077.
33. Van Buuren, S. (Chapman and Hall/CRC, New York., 2018) *Flexible Imputation of Missing Data, Second Edition*.
34. Van Der Beek, F.B., Boermans, P.P.B.M., Verbist, B.M., Briaire, J.J. & Frijns, J.H.M. (2005) Clinical evaluation of the Clarion CII HiFocus 1 with and without positioner. *Ear Hear. 26,* 577–592.
35. Frijns, J.H., Briaire, J.J. & Grote, J.J. (2001) The importance of human cochlear anatomy for the results of modiolus-hugging multichannel cochlear implants. *Otol. Neurotol. 22,* 340–349.
36. van der Jagt, M.A., Briaire, J.J., Verbist, B.M. & Frijns, J.H.M. (2016) Comparison of the HiFocus Mid-Scala and HiFocus 1J electrode array: Angular insertion depths and speech perception outcomes. *Audiol. Neurotol. 21,* 316–325.
37. Litvak, L.M., Spahr, A.J. & Emadi, G. (2007) Loudness growth observed under partially tripolar stimulation: model and data from cochlear implant listeners. *J. Acoust. Soc. Am. 122,* 967–981.
38. Bierer, J.A. (2007) Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration. *J. Acoust. Soc. Am. 121,* 1642–1653.
39. Moore, B.C.J. (2004) Dead Regions in the Cochlea: Conceptual Foundations, Diagnosis, and Clinical Applications. *Ear Hear. 25,* 98–116.
40. Nogueira, W., Litvak, L.M., Saoji, A. a. & Büchner, A. (2015) Design and evaluation of a cochlear implant strategy based on a 'Phantom' channel. *PLoS One 10,* 1–25.

CHAPTER 7

General Discussion

Cochlear Implant (CI) users generally achieve high performance on speech tests in quiet[1,2]. Yet, real world speech is rarely presented without background sounds, and understanding speech in a noisy environment is much more challenging for CI users. Moreover, the auditory world consists of more complex sounds, like music, of which the perception through conventional CI technology is rarely attained at a high level[3]. The purpose of this thesis was to explore the possibilities of new sound coding strategies that potentially make listening in noisy environments and listening to more complex stimuli easier for CI users. This discussion will elaborate on two speech coding strategies that use electrical field shaping to improve the transmission of spectral detail, while the current methods to evaluate these new strategies will also be critically reviewed. Besides the actual performance of a CI user, who is listening using a certain speech coding strategy, many factors also influence performance on psychophysical tests. Examples of these include fatigue, motivation, hearing performance itself, and experience with the tests taken. The following section will focus on the effect of experience with the tests, specifically looking at learning effects both acutely and over time. The potential mechanism behind these learning effects will be discussed and the importance of handling them correctly will be highlighted.

LEARNING EFFECTS

Learning effects over time

While it has been well documented in other fields that repeated psychophysical testing can cause learning[4,5], learning over time has not been properly studied in CI research. The current thesis provides us with new and pertinent data on this matter. Table 1 gives an overview of the learning effects observed in this thesis. In Chapters 2 and 5 psychophysical tasks were repeatedly performed while using the same speech coding strategy, with relatively long time intervals between test sessions. This enabled evaluation

TABLE 1. Learning effects across the different chapter of this thesis

Study	SMRT		MDT		Speech in Noise	
	Acute	Long-term	Acute	Long-term	Acute	Long-term
Chapter 2/3	NS	p<0.001	NS	p=0.035	NS	NS
Chapter 4	NS	X	p=0.05	X	p=0.017	X
Chapter 5	p=0.006	45dB: p=0.042	NS	NS	NS	45dB: p<0.001
		65dB: NS				65dB: p=0.012

SMRT, Spectrally Temporally Modulated Ripple Test, MDT, Temporal Modulation Detection Test. Note that the speech in noise tests were the Flemish LIST in Chapters 2 and 3, and the Dutch Matrix test in Chapters 4 and 5.

of learning effects over time for multiple psychophysical measures. In both chapters a learning effect over time was found with the spectrally temporally modulated ripple test (SMRT). For the temporal modulation detection test (MDT) this could only be demonstrated in Chapter 2, but not in Chapter 5. Two different speech in noise tests were used in this thesis: a Flemish sentence test (LIST)[6] in Chapter 2 and 3, and the Dutch Matrix test[7] in Chapters 4 and 5. Of those speech in noise tests, only the Dutch Matrix test showed a significant learning effect over time in Chapter 5.

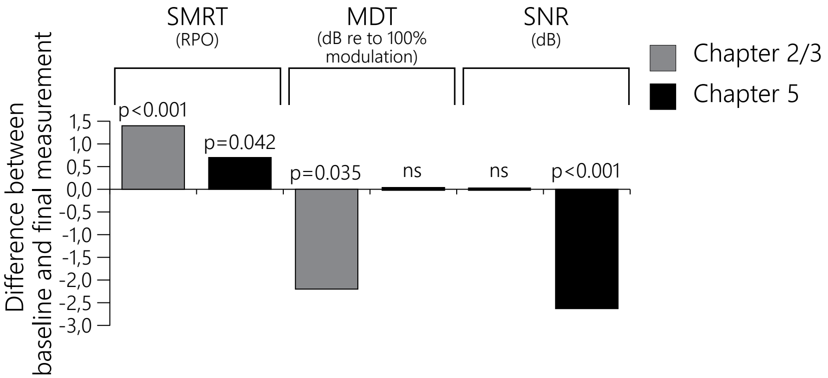


FIGURE 1. LEARNING EFFECTS OVER TIME. NOTE THAT THE TEST-RETEST INTERVAL WAS TWO WEEKS IN CHAPTER 2 AND FIVE WEEKS IN CHAPTER 5. SMRT: Spectrally Temporally Modulated Ripple Test, MDT: Amplitude Modulation Detection Threshold, SNR: Signal-to-Noise ratio. In Chapter 2 and 3 the Flemish LIST speech-in-Noise test was used, and in Chapter 5 the Dutch Matrix test. The units for the SMRT are ripples per octave, for the MDT dB relative to 100% modulation, and for the SNR speech-to-noise ratio in dB.

Figure 1 depicts the effect sizes of the learning effects found over time. Because two different types of speech in noise tests were used across this thesis, these results are incomparable. However, the same SMRT and MDT tests were used throughout the thesis, which allows for comparing the results across studies. Interestingly, the observed learning effects of these two measures are approximately twice as large in Chapter 2 as in Chapter 5. In both studies subjects performed a baseline measurement with their regular clinical speech coding strategy, followed by measurements with one or two experimental strategies (HiRes FFT, HiRes Optima, or DCF), and lastly a re-measurement with their clinical strategy. Next to the experimental speech coding strategies that were used between test-retest sessions, the most relevant difference between these studies was the period between test sessions: only two weeks in chapter 2 and five weeks in chapter 5. A time-interval of five weeks was chosen to avoid false positive results as a consequence of learning, as explained below. In chapter 2 it was argued that a so-called "wash-out period" of sufficient duration between test sessions compensates for learning effects. This was partly based on the study of Drennan *et al*. (2015), who found no learning effects when psychophysical measures were repeated with test-retest intervals of two months. To increase the feasibility of the take-home trial, a shorter test- retest interval was chosen than in the study of Drennan and colleagues, although it was considerably longer than the time-interval in Chapter 2. This resulted in smaller learning effects, which unfortunately still confounded the results of the study significantly. However, the effect size of the learning effects decreased with longer test-intervals, thereby supporting the hypothesis about wash-out periods.

Acute learning effects

While multiple studies show that there is no effect of repeated testing on the same test day, i.e., these are no acute learning effects, this thesis shows opposing results. For example, in the take-home trial of Chapter 5 an acute learning effect was found in the spectrally temporally modulated ripple test (SMRT), while previous studies[9–11], including the other studies in this thesis (Chapters 2, 3, and 4), did not demonstrate this. In Chapter 4, where two speech coding strategies were extensively evaluated on one test day, an acute learning effect was observed in the amplitude modulation detection (MDT) test ($F5,10=2.4$, $p=0.05$). Nevertheless, this effect was absent in the other studies of this thesis in which the same measure was used (Chapter 2, 3, and 5). The Dutch Matrix speech in noise test showed a significant effect of repetition number in Chapter 4 ($F2,20=5.05$, $p=0.017$), but not in Chapter 5. This inconsistent presence of acute learning effects emphasizes the caution with which the psychophysical tests should be used. One method to minimize the risk of acute learning is to introduce a preliminary task-practice phase, where the subjects are trained using the task. In this task-practice phase sufficient practice rounds should be offered. For example, subjects were offered only one practice session for the Matrix tests in Chapter 4, but two in Chapter 5. This resulted in a highly significant acute learning effect in Chapter 4, which was not present in chapter 5. From the above observations it is clear that learning effects are present, but the question remains what the subjects have actually learned.

The mechanisms of learning effects

Two types of learning are considered: perceptual and task learning. Perceptual learning is the improvement in perceptual performance as a consequence of sensory interaction with the environment as well as through practice in performing specific sensory tasks[12]. In other words, it is improvement in the sensory performance itself. In task learning, however, subjects improve performance on the task after repeated administrations, in the absence of an underlying general improvement of the listening ability.

Task learning

Task learning is divided into content and procedural learning[13]. Content learning means that the subject becomes familiarized with the test material and therefore, increases his or her ability to gain information. Procedural learning concerns improved handling of the task, e.g. by holding your breath during testing so that the sound of your own breathing does not interfere with the test. Yund & Woods (2010) revealed that content learning in particular occurs in the Hearing in Noise Test (HINT), where the same sentences are repeated over multiple test sessions[13]. They showed that the procedural learning reached its plateau already after the first half of the first test session. The fact that content learning occurs when the same words/sentences are often repeated explains the discrepancy between learning on the LIST and the Dutch Matrix tests. The LIST uses a dataset of sentences with different grammatical structures, while the Matrix test uses sentences of identical grammatical structure and selects words from a closed set of alternatives[6,7]. In the LIST test the subjects can benefit from procedural learning, but because the content is different every test session, the task has to be repeated frequently to be able to learn from it. However, in the Matrix material, words and grammar are repeated continuously, enabling content learning. Kollmeier ***et al.***, (2015) stated that 2 practice sessions are required to achieve such a plateau in learning[14].

Perceptual learning

While task learning is an important issue and should be kept to a minimum by developing smarter psychophysical tasks, perceptual learning is a more interesting phenomenon from a clinical perspective. There is a lot of evidence that CI users can improve their performance without changing the electrical input. For example, auditory perception clearly improves following initial activation of the CI and also the speech perception of experienced CI users improves with new speech coding strategies during approximately 3 to 6 months after the transition[15–18]. Apparently, considerable auditory plasticity exists in CI users, and they need time to learn how to effectively use the patterns of activation produced by electrical stimulation. Although passively using the CI can improve performance up to a certain level, auditory training can accelerate perceptual learning and increase the final performance of a CI user. For example, training with speech-based stimuli have shown to result in 10-15% improvements, even in experienced users[19]. Examples of these improvements also come from the field of visual perceptual learning. For example, Liu & Weinshall (2000) showed that perceptual skills that were learned in one visual task can transfer to other visual tasks as both an acceleration of learning the task and an immediate improvement in performance. Similarly, Moberly ***et al.*** (2015) found that CI users can learn new speech cues, and can use this ability in other auditory psychophysical tasks[21]. Because the gains in performance found in this thesis are relatively large (e.g. 2.63 dB signal-to-noise ratio (SNR) improvement in Chapter 5), it is unlikely that the learning effects are based on task learning alone. A study with bimodal CI users, that was conducted in our clinic, also repeatedly tested performance on the Dutch Matrix test. The occurrence of perceptual learning was less likely in that study because the CI users were very experienced with their hearing situation and no extra auditory training was offered. Nevertheless, a significant (probably task) learning effect of -2.0 dB SNR was found between the 1st and 3th test session (Figure 2, unpublished data). This is somewhat smaller than the learning effect found in chapter 5. Moreover, the average test-retest intervals were only 2-3 weeks, which is insufficient for a wash-out period. In other words, if the same test-retest interval had been used as was done in Chapter 5, the learning effect would probably have been smaller. This would make the difference in effect size between the bimodal study (probably mostly task learning) and chapter 5 (both task and perceptual learning) larger. This suggests that the CI users in chapter 5 perceptually learned from listening with a, for them, new speech coding strategy for a few weeks. Although no active training programs were offered, CI users were instructed to subjectively examine their hearing in multiple listening situations. It could be that the new speech cues from the experimental speech coding strategies and the extra focus on their hearing ability led to the perceptual learning. If this is the case, it might be beneficial for CI users to take part in clinical trials, or change their everyday speech coding strategy once in a while, to gain performance. The study designs in the current thesis, however, are unable to confirm this hypothesis. Therefore, a placebo controlled study of perceptual learning during clinical trials should be conducted. Here one group actually uses a new speech coding strategy, while in another arm the experimental strategy is actually is exactly the same as their regular clinical strategy.

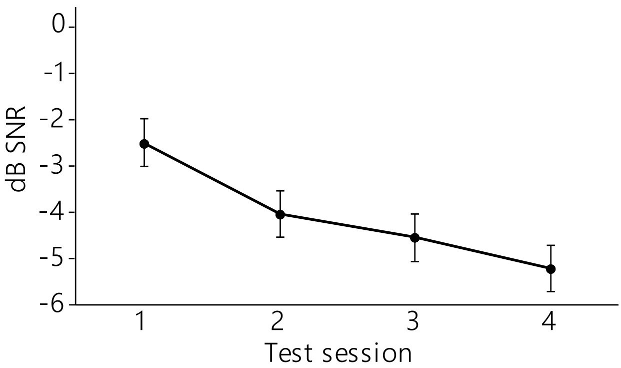


Figure 2. Learning effects in the dutch matrix test during a study with bimodal CI users (unpublished data). The average speech reception thresholds of fifteen bimodal CI users are displayed per test session. Time intervals between test sessions were 2-3 weeks on average, but differed between and within subjects with outliers to 3 months. Error bars represent 1 standard error of the mean.

How to handle learning effects

As in all research, the studies discussed in this thesis have their limitations. One of these concerns is the handling of learning effects which were present in the psychophysical measures used in the different studies. What we've learned is that sufficient practice sessions, for example at least two lists in the Dutch Matrix test[14], are essential to avoid acute learning effects, and that longer time intervals between test sessions, called wash-out periods, help to decrease learning over time. Moreover, speech in noise tests should use a wide range of speech materials to avoid content learning. Nevertheless, it is impossible to completely eradicate both task and perceptual learning effects, and therefore, the study design itself should compensate for this by randomizing the order of the tested speech coding strategies under test. Only then, one can be sure to measure a real increase in performance instead of a gain as a consequence of learning.

Performance on psychophysical tasks

Listening itself is a demanding task for CI users, as they have to do this actively. Alhanbali *et al*. (2017) showed that CI users report an increased listening effort and fatigue compared to normal hearing individuals[22]. It is well known that this extra effort to achieve perceptual success comes at the cost of processing resources that might otherwise be available for additional cognitive processes[23,24]. To the best of our knowledge, the effect of fatigue on performance on psychophysical measures has not yet been investigated. Nevertheless, based on the evidence described above, one can imagine that fatigue has a negative influence. This is something we experienced during the trials conducted for this thesis. Therefore, it is necessary to take enough breaks and/or to divide the study protocol over multiple test days, especially because research protocols often encompass long test-days. Another factor that obviously influences performance on psychophysical testing is the hearing performance itself, irrespective of which speech coding strategy is used. This is problematic because we are dealing with both ceiling and floor effects. A test that is too difficult for a certain CI user would result in excessively poor performance, and may suffer from floor effects. Conversely, a test that is too easy may suffer from ceiling effects and not be able to adequately monitor performance improvements[25]. In the studies of this thesis, inclusion criteria were based on monosyllabic consonant-vowel-consonant (CVC) word phoneme scores. However, it is unknown which level of phoneme score in quiet is required to be able to meaningfully undertake, for example, speech in noise tests. Perhaps it would be more valid to use sentence in quiet scores (with the same speech material as the speech in noise test that would be used) as an inclusion criterion. Future research should show whether there is a correlation between CVC scores and reliability on other psychophysical measures.

THE EFFECT OF CHANGING THE FILTERING OF SOUND

As discussed in the introduction of this thesis, there are many steps that must be completed before an actual stimulation pattern is presented to the auditory nerve of CI users. All these processing steps are important for the final outcome. Nevertheless, research on new speech coding strategies is often focused on the last step: which stimulation patterns are the most useful for understanding speech in all kinds of environments? The comparison of strategies that differ in multiple dimensions is ubiquitous in CI research. For example, the fine structure processing (FSP) coding strategy was developed for the Med-El CI system (Innsbruck, Austria)[26]. The goal of this strategy is to represent fine structure in the low-frequency channels, thereby improving speech understanding. This new strategy has been compared to CIS strategies in multiple studies, with mixed results[27–30]. A weakness of these studies was that the FSP strategy did not only differ in the presentation of fine structure, but also had different frequency filter settings than the CIS strategy. Therefore, the reported differences in speech perception between the FSP and the CIS coding strategies may have resulted from changes in the frequency-to-filter assignment instead of the representation of fine structure information[31]. Thus, the research questions of these studies have not been properly answered. This also happened in the research into current steering[11,32]. The purpose of this new speech coding strategy was to provide more frequency information by stimulating auditory nerve regions that are located in between two adjacent physical electrode contacts. To achieve this, more frequency information should be obtained from the incoming signal, for which a more sophisticated manner of filtering the incoming sound was required. Fast Fourier transformation (FFT) based filters provide a detailed spectral profile, are computationally efficient[33], and thus are implemented in the current steering speech coding strategy. It could be that the conversion from traditional band-pass filters to FFT-based filter banks caused a change in performance on its own. Drennan *et al*. (2010) speculated that the implementation of FFT-based filter banks led to a decreased temporal resolution, which counteracted the potential improvement due to current steering. Chapter 3 of this thesis investigated the effect of FFT-based filter banks on temporal resolution, spectral resolution, and speech perception in noise. The study showed that the implementation of FFT-based filters does not change the performance on any of the tests used, so the known benefits of FFT filters encourage their implementation in future speech coding strategies. Although the current study revealed that the FFT-based filters themselves had no effect on performance, researchers should keep in mind that often multiple factors are changing in speech coding strategies and that studying each individual modification is essential for the correct interpretation of the results obtained.

ELECTRICAL FIELD SHAPING AND ITS EFFECT

In the following paragraphs two experimental speech coding strategies will be reviewed. The strategies were intended to increase the spatial selectivity or to extend the frequency range for CI users, through the use of electrical field shaping. As discussed in the introduction, there are multiple methods available to shape the electrical field. First, the effects of Dynamic Current Focusing (DCF) and phantom stimulation techniques will be elaborated. Subsequently, the overarching and contradictory characteristics of these strategies will be discussed. Lastly, some future perspectives will be provided.

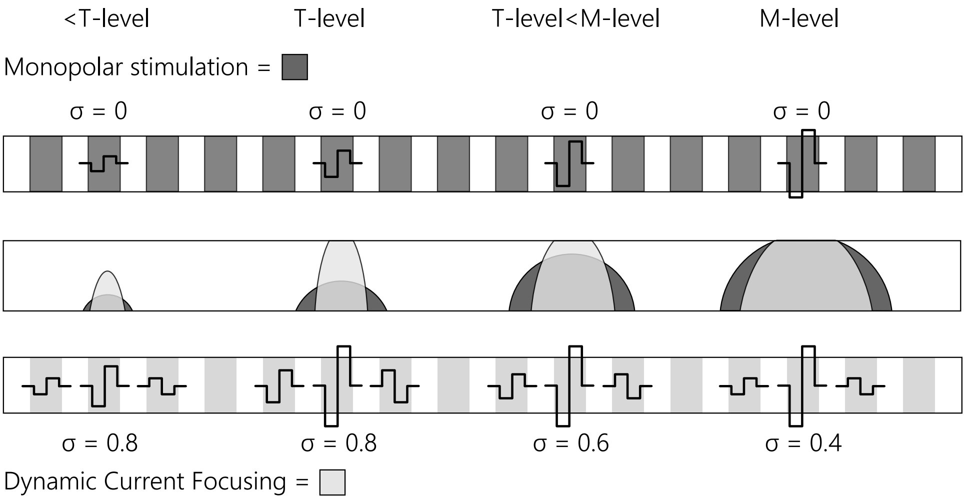


FIGURE 3. THE PRINCIPLE OF DYNAMIC CURRENT FOCUSING. In the middle of the figure the auditory nerve is schematically illustrated with the expected electrical fields following from either monopolar stimulation (top panel, dark grey) or stimulation with the dynamic current focusing strategy (bottom panel, light grey) at multiple subjective loudness levels. T-level: threshold level, M-level: most comfortable level.

DYNAMIC CURRENT FOCUSING

The DCF strategy is a loudness encoding strategy that uses high degrees of current focusing at the lower loudness levels, and lower degrees of current focusing at higher loudness levels. The degree of current focusing is expressed in a current compensation coefficient, σ, which describes the percentage of the main electrode contact's current that is returned via the two compensating electrode contacts. A sigma of 1 means that all of the administered current is returned via the two compensating contacts, σ = 0.5 means that 50% is returned via these contacts and σ = 0 means that none of the current is returned via the compensating electrode but rather it is all returned via the extra-cochlear ground electrode (i.e., monopolar stimulation, Figure 3). The effect and development of the DCF strategy were extensively studied and described in Chapters 4 and 5, and an overview of the results is given in Table 2. In the acute study (Chapter 4), where relatively high degrees of current focusing were used (the average σ was 0.88 at T-level and 0.49 at M-level), a significant and relatively large improvement in spectral resolution was found (+1.34 ripples per octave on the SMRT). Although the results of the take-home trial were difficult to interpret due to the learning effects found, the beneficial effects of listening with DCF in *that* implementation were certainly decreased. One obvious explanation would be the reduced degree of current focusing that was used in the take-home trial version of the DCF strategy: the σ value at T-level was 0.8 for all subjects, and the average σ at M-level was just 0.17. This reduced σ at T-level was deliberately introduced to lower the power consumption, and thereby improving clinical applicability of the DCF strategy. This was based on the finding that the current (expressed in clinical units, CU) that is required to reach T-level increases with σ value (Figure 4, unpublished data). Although this did result in an increase in battery life, it also resulted in σ values at M-level that are known to have no beneficial effect[34].

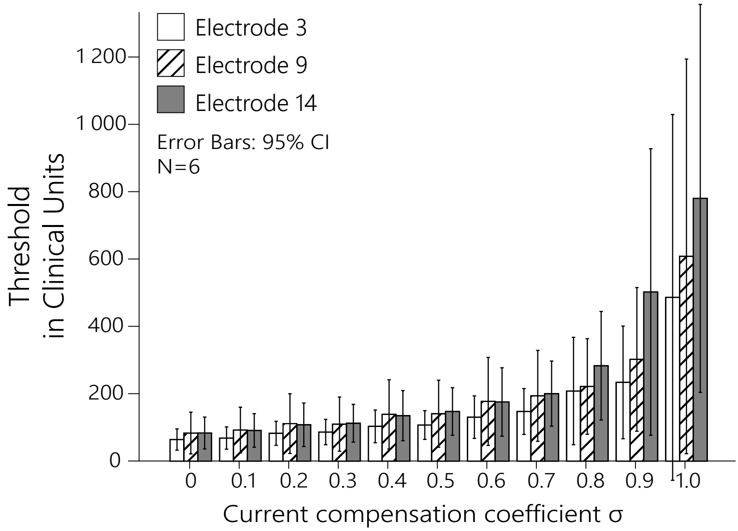


Figure 4. Clinical units at t-level as a function of sigma level (unpublished data). Of six subjects that participated in the study of Chapter 5 threshold levels were measured at multiple sigma values at three electrode contacts. The error bars represent the 95% confidence intervals. The clinical unit level to maintain equal loudness increased with sigma value, which is in line with previous research[40].

Phantom stimulation

Phantom stimulation aims to facilitate the recruitment of apical nerve fibers that are normally not stimulated to create a lower pitch percept. It does this by simultaneous stimulation of two electrode contacts with opposite polarity. Either biphasic or pseudo-monophasic pulse shapes can be used and the ratio between the amplitudes of the compensating and the main electrode contacts (denoted as compensation coefficient σ) can be altered. The pitch shifts following from the phantom configurations displayed in Figure 1 of Chapter 6 were studied in Chapter 6. In line with previous literature[35,36], an average pitch shift of 0.92 (range 0.25-2.0) electrode contact towards the apex was found when the best phantom configuration per tested electrode contact was used. No clear difference between the use of symmetrical biphasic or pseudo-monophasic pulses was found, although the difference between the pitch shift at the apex and base differed considerably for the pseudo-monophasic pulse shape. Overall, the configurations with relatively high σ values of 0.7, 0.75, or 0.8 resulted in the largest pitch shifts. However, the higher the σ value, the more variability between subjects arises, complicating the implementation of a phantom electrode contact in speech coding strategies. Studies that implemented phantom electrodes, therefore, used relatively low

TABLE 2. Overview results acute and take-home trial Dynamic Current Focusing strategy

		Acute study			Take-home trial				
		DCF	Clinical	Difference (DCF - clinical)	DCF	Clinical baseline	Difference (DCF - clinical)	Clinical final	Difference (DCF - clinical)
SMRT (RPO)	45 dB	3.74	2.4	**+1.34 (p=0.016)**	4.66	3.62	+1.04 (NS)	4.49	+0.17 (NS)
	65 dB	4.07	3.27	+0.8 (NS)	4.92	4.05	**+0.87 (p=0.012)**	4.63	+0.29 (NS)
Dutch Matrix (dB SNR)	45 dB				+0.92	+2.38	**-1.46 (p=0.017)**	-0.32	+1.24 (NS)
	65 dB	4.72	4.57	+0.15 (NS)	+1.80	+2.45	-0.65 (NS)	-0.10	**+1.9 (p=0.012)**
MDT (dB re to 100% AM)	65 dB	-8.73	-9.35	+0.62 (NS)	-11.88	-12.3	+0.42 (NS)	-13.45	+1.57 (NS)
Loudness growth (AUC)		32.7	37.8	-5.1 (NS)	31.0	22.1	-8.9 (p=0.007)		

DCF, Dynamic Current Focusing; SMRT, Spectrally Temporally Modulated Ripple Test; RPO, Ripples Per Octave; SNR, Signal-to-Noise Ratio where 50% of the words are repeated correctly; dB re to 100% AM, dB relative to 100% amplitude modulation (lower score is better performance); AUC, area under the curve. All significantly differences in scores are displayed in bold.

7

σ values of maximally 0.625. This, combined with the high variability between subjects may explain the lack of an improved speech understanding[37,38]. Chapter 6 gives arguments why the variability between subjects can be partially averted by performing a pitch ranking or pitch discrimination experiment with multiple phantom configurations before fitting subjects with a phantom strategy. That way the best performing phantom configuration can be used for a specific individual. However, this is time consuming and therefore, less likely to be adopted clinically. Another method to decrease inter-subject variability is to preselect CI users that are more likely to benefit from phantom stimulation. One preselection criterion could be the type of electrode array with which the CI user is implanted. In Chapter 6 it was demonstrated that the CI users who are implanted with a lateral wall electrode array achieve significantly larger pitch shifts than those implanted with mid-scalar or perimodiolar electrode arrays. Unfortunately, the previous studies that implemented a phantom electrode into speech coding strategies did not report on the type of electrode arrays with which the subjects were implanted[37–39].

Effectiveness of electrical field shaping

For both the DCF and phantom strategies the effect sizes on speech perception show great variability between subjects. Of course this occurs in many studies in the field, and unfortunately it seems that there is no speech coding strategy that fits each and every CI user. For that reason, it is of interest to find underlying causes for the presence or absence of a beneficial effect of experimental stimulation strategies. What are the differences between subjects from which we can predict whether a strategy is going to work or not? When looking at the strategies examined in the current thesis, the electrode-neuron interface is of great importance. The following paragraph will discuss multiple factors of the electrode-neuron interface that are likely to influence the effect of both DCF and phantom stimulation.

For effective electrical field shaping two factors are required: (1) a relatively broad current spread and (2) interaction between the center and the adjacent electrode contacts. For example, in tripolar stimulation the broad electrical field can be narrowed down by capturing the flanks of the electrical field before it reaches the auditory nerve. In phantom stimulation the same mechanism occurs, but only at one side of the electrical field so that the center of gravity, and therefore the pitch percept, shifts towards the contralateral side. The more current that is captured because of the opposite stimulation on the flanking electrodes, the more effectively the electrical field is shaped. An inevitable consequence, however, is that more current is required to maintain equal loudness[40,41]. In line with this, we found that higher current levels were required to reach T-level when σ values were increased and thus the excitation patterns were more spatially localized[42] (Figure 4). In reverse, Litvak ***et al***. (2007) used the amount of current that is required to reach T- and M-levels in (partial) tripolar mode to predict the interaction between the center and adjacent electrode contacts, i.e. the effectiveness of electrical field shaping[40]. The following formula is a simplified version of the one presented in the paper of Litvak and colleagues, and describes this interaction coefficient (*K*):

$$K = \frac{1 - TL^{MP} / TL^{0.9}}{0.9} \qquad \text{(eq. 1)},$$

where TL^{MP} is the CU required to reach T-level in monopolar mode, and $TL^{0.9}$ in tripolar mode with a sigma of 0.9. In other words, the more current that is required to reach T-level in tripolar mode relative to monopolar mode, thus more effective interaction, the higher the interaction coefficient *K*. In a computer model, Litvak ***et al***. (2007) found higher *K*s when electrode-to-tissue distances were larger. Also, Kalkman ***et al***. (2014) demonstrated in a computer model of the human cochlea that current focusing techniques are most effective when the electrode-to-tissue distance is relatively large[43]. This is sensible, as a larger distance will lead to a broader current spread, more channel interaction, and therefore more effective current focusing.

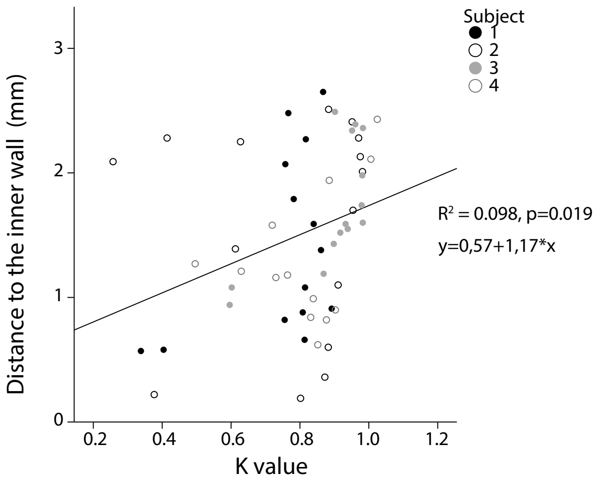


Figure 5. Correlation between K-values and the distance between the electrode contact to the inner wall of the cochlea. Data of four of the subjects that were included in the Dynamic Current Focusing take-home trial were obtained.

To examine this in our own study population, *K*-values were calculated and related to the electrode distance to the inner wall of the cochlea, as measured according to van der Jagt ***et al***. (2016). Unfortunately, only data from four subjects who participated in the DCF take-home trial (14 electrode contacts per subject) were available. Figure 5 shows the weak positive correlation (R2=0.098, p=0.019) between the two variables, indicating that a larger distance between the electrode contacts and the neural tissues indeed leads to a larger *K*. Unfortunately, no such data was available for the subjects who participated in the phantom study (Chapter 6). However, as previously mentioned, we did find a significantly larger effect of phantom stimulation in subjects implanted with a lateral wall electrode array (HiFocus 1J, Advanced Bionics, Sylmar, CA) compared to those implanted with a mid-scalar or perimodiolar electrode array (HiFocus MS or HiFocus 1 with positioner, Advanced Bionics, Sylmar, CA), which is in line with the hypothesis. Moreover, in Chapter 4, where DCF was demonstrated to have a beneficial effect, most subjects were implanted with a lateral wall electrode array, while mostly mid-scalar implanted subjects were studied in Chapter 5, where this beneficial effect was not found. Altogether, this fits the concept that the electrode-neuron distance influences the effectiveness of electrical field shaping. It is also suggested that, when electrodes are placed close to the neural tissue, electrical field shaping can have a detrimental effect. The closer the electrode-to-neuron proximity, the easier the activity of the compensating electrode contact can cause neuronal excitation by itself, creating so-called side lobes[35,40,45,46]. In tripolar mode this means that the region of neural excitation can become broader then it would be in monopolar mode, and in phantom stimulation the center of gravity will be shifted back to the center electrode contact, counteracting the intended pitch shift towards the apex.

Although the electrode-to-neuron distance influences effectiveness of electrical field shaping, the correlation that we found was only weak. An irregularity in the electrode-neuron interface that could also contribute to the inter- and intra-subject variability is the degeneration of spiral ganglion cells, which differs both between and within subjects[47]. If a lot of the spiral ganglion cells are degenerated, all surviving nerve fibers can be excited before sufficient loudness is reached, particularly when the stimulation pattern is narrow due to a current focused stimulation scheme. To increase loudness the excitation profile then has to be broadened by increasing the amplitudes on all three electrode contacts involved, which increases the chance of side lobes[40] stimulating, rather than simply returning current. The broader current spread and the side lobes then reduce the spatial selectivity[45]. Local spiral ganglion degeneration can also have a beneficial effect in phantom stimulation. If the compensating electrode is located in a dead region it may not excite nearby auditory nerve fibers; thus, potential side lobes will not cause neural excitation and will have no detrimental effect, even for high σ values[48].

FUTURE PERSPECTIVES

It appears that we are approaching the limits of CI systems to deliver useful spectro-temporal information to the auditory system. This thesis showed that the currently available new stimulation strategies can add considerable additional value to some CI users, but definitely not to all. The inter- and intra-subject variability is substantial, complicating the development of a one-fits-all speech coding strategy. One aim of future research should be to individualize speech coding strategies. For the DCF strategy this means varying the parameters from channel to channel, depending on the state of the electrode-neuron interface. However, to make the implementation of such an individualized strategy feasible, clinicians have to be able the evaluate the electrode-neuron interface. For example, K-values can be measured for each electrode contact to estimate the effect of the electrical field shaping. If K is low, the beneficial effect would be negligible and therefore this channel should be used in monopolar mode. In phantom stimulation this means that a pitch ranking experiment should precede the fitting, so that the best configuration for that specific CI user can be configured.

Much of the variability across implant listeners is a result of inefficiencies in their use of the information provided by the prosthesis. Therefore, another aim of future research could be to improve the utilization of the input by the brain, as passive adaptation to new speech coding strategies might not be sufficient. Auditory training may be necessary for CI users to access the additional spectral and temporal cues provided by advanced speech processing strategies. The presence of (perceptual) learning effects as found in the current thesis emphasizes the likelihood of achieving benefits in hearing for CI users with this approach.

An area that has been underexposed is loudness encoding. The DCF strategy is a first attempt at changing the way loudness is encoded in speech coding strategies, but obviously many more opportunities can be explored. One example of a loudness encoding strategy, that potentially improves the spatial selectivity for CI users, is a strategy that we call “sequential current steering”. As the outer edges of excitation profiles are responsible for the non-specificity of the stimulation patterns, these areas have to be eliminated. The excitation profile especially broadens when high amplitude pulses are administered. Therefore, it was hypothesized that only using low amplitude pulses can improve the spatial selectivity. The targeted excitation area is then stimulated by a series of low amplitude pulses, that are spatially distributed along the basilar membrane with the use of current steering. Loudness can be achieved by increasing the number of pulses and stimulation sites. The method is based on the assumption that an implant user cannot differentiate between a single large pulse and multiple spatially separated consecutive pulses. Previous research showed that the stimulation of one electrode interleaved with that of an adjacent electrode using short temporal offsets, also known as “sequential dual-electrode stimulation”, leads to intermediate pitch percepts[49,50]. Frijns *et al*. (2009) compared simultaneous and sequential stimulation methods in a computational model and found similar excitation patterns[51]. It is hypothesized that also with multiple pulses the populations of neurons excited by each of the pulsatile stimuli could combine to produce a single region of excitation, provided that the electrodes are sufficiently close together and the time delay between the pulses is small. It is thought that

7

this way of encoding a stimulus leads to a more controllable excitation profile along the basilar membrane and therefore to an improved pitch perception and discrimination. The attentive reader will have noticed that this sequential current steering strategy is also based on shaping the electrical field. These kind of speech coding strategies, like the DCF strategy, are especially promising because multiple CI manufacturers choose atraumatic lateral wall electrode arrays for their new electrode designs. Because the current thesis showed that electrical field shaping is most successful in CI users who are implanted with these electrode designs, future CI users could also benefit from these new speech coding strategies.

REFERENCES

1. Bonnet, R.M., Boermans, P.-P.B., Avenarius, O.F., Briaire, J.J. & Frijns, J.H.M. (2012) Effects of pulse width, pulse rate and paired electrode stimulation on psychophysical measures of dynamic range and speech recognition in cochlear implants. ***Ear Hear. 33,*** 489–496.
2. Gifford, R.H., Noble, J.H., Camarata, S.M., ***et al.*** (2018) The relationship between spectral modulation detection and speech recognition differs for adult versus pediatric cochlear implant recipients. ***Trends Hear. 22,*** 1–14.
3. Limb, C.J. & Roy, A.T. (2014) Technological, biological, and acoustical constraints to music perception in cochlear implant users. ***Hear. Res. 308,*** 13–26.
4. Liu, Z. (1999) Perceptual learning in motion discrimination that generalizes across motion directions. ***Proc. Natl. Acad. Sci. U. S. A. 96,*** 14085–14087.
5. Fine, I. & Jacobs, R.A. (2002) Comparing perceptual learning tasks: a review. ***J. Vis. 2,*** 190–203.
6. Van Wieringen, A. & Wouters, J. (2008) LIST and LINT: sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands. ***Int. J. Audiol. 47,*** 348–355.
7. Houben, R., Koopman, J., Luts, H., ***et al.*** (2014) Development of a Dutch matrix sentence test to assess speech intelligibility in noise. ***Int. J. Audiol. 53,*** 760–763.
8. Drennan, W.R., Won, J.H., Timme, A.O. & Rubinstein, J.T. (2015) Nonlinguistic Outcome Measures in Adult Cochlear Implant Users Over the First Year of Implantation. ***Ear Hear. 37,*** 1–11.
9. Won, J.H., Drennan, W.R. & Rubinstein, J.T. (2007) Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. ***J. Assoc. Res. Otolaryngol. 8,*** 384–392.
10. Drennan, W.R., Longnion, J.K., Ruffin, C. & Rubinstein, J.T. (2008) Discrimination of Schroeder-phase harmonic complexes by normal-hearing and cochlear-implant listeners. ***J. Assoc. Res. Otolaryngol. 9,*** 138–149.
11. Drennan, W., Won, J., Nie, K., Jameyson, E. & Rubinstein, J.T. (2010) Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. ***Hear. Res. 262,*** 1–8.
12. Irvine, D.R.F. (2018) Auditory perceptual learning and changes in the conceptualization of auditory cortex. ***Hear. Res.*** 1–14. doi:10.1016/j.heares.2018.03.011
13. Yund, E.W. & Woods, D.L. (2010) Content and procedural learning in repeated sentence tests of speech perception. ***Ear Hear. 31,*** 769–778.
14. Kollmeier, B., Warzybok, A., Hochmuth, S., ***et al.*** (2015) The multilingual matrix test: Principles, applications, and comparison across languages: A review. ***Int. J. Audiol. 54,*** 3–16.

15. Staller, S., Menapace, C., Domico, E., ***et al.*** (1997) Speech perception abilities of adult and pediatric Nucleus implant recipients using the Spectral Peak (SPEAK) coding strategy. ***Otolaryngol. - Head Neck Surg. 117,*** 236–242.
16. Rouger, J., Lagleyre, S., Fraysse, B., ***et al.*** (2007) Evidence that cochlear-implanted deaf patients are better multisensory integrators. ***Proc. Natl. Acad. Sci. 104,*** 7295–7300.
17. Ruffin, C.V., Tyler, R., Witt, S., ***et al.*** (2007) Long-Term Performance of Clarion 1.0 Cochlear Implant Users. ***Laryngoscope 117,*** 1183–1190.
18. Dorman, M.F. & Loizou, P.C. (1997) Changes in speech intelligibility as a function of time and signal processing strategy for an Ineraid patient fitted with continuous interleaved sampling (CIS) processors. ***Ear and hearing 18,*** 147.
19. Fu, Q.J. & Galvin, J.J. (2008) Maximizing cochlear implant patients' performance with advanced speech training procedures. ***Hear. Res. 242,*** 198–208.
20. Liu, Z. & Weinshall, D. (2000) Mechanisms of generalization in perceptual learning. ***Vision Res. 40,*** 97–109.
21. Moberly, A.C., Bhat, J. & Shahin, A.J. (2015) Acoustic Cue Weighting by Adults with Cochlear Implants : A Mismatch Negativity Study. ***Ear Hear. 37,*** 465–472.
22. Alhanbali, S., Dawes, P., Lloyd, S. & Munro, K.J. (2017) Self-Reported Listening-Related Effort and Fatigue in Hearing-Impaired Adults. ***Ear Hear. 38,*** e39–e48.
23. Rabbitt, P.M.A. (1968) Channel-capacity, intelligibility and immediate memory. ***Q. J. Exp. Psychol. 20,*** 241–248.
24. McCoy, S.L., Tun, P.A., Cox, L.C., ***et al.*** (2005) Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. ***Q. J. Exp. Psychol. Sect. A Hum. Exp. Psychol. 58,*** 22–33.
25. Brant, J.A., Eliades, S.J., Kaufman, H., Chen, J. & Ruckenstein, M.J. (2018) AzBio Speech Understanding Performance in Quiet and Noise in High Performing Cochlear Implant Users. ***Otol. Neurotol. 39,*** 571–575.
26. Hochmair, I., Nopp, P., Jolly, C., ***et al.*** (2006) MED-EL cochlear implants: state of the art and a glimpse into the future. ***Trends Amplif. 10,*** 201–219.
27. Muller, J., Brill, S., Hagen, R., ***et al.*** (2012) Clinical trial results with the med-el fine structure processing coding strategy in experienced cochlear implant users. ***Orl 74,*** 185–198.
28. Vermeire, K., Punte, A.K. & Van De Heyning, P. (2010) Better speech recognition in noise with the fine structure processing coding strategy. ***J. Oto-Rhino-Laryngology, Head Neck Surg. 72,*** 305–311.
29. Riss, D., Arnoldner, C., Reiß, S., Baumgartner, W.D. & Hamzavi, J.S. (2009) 1-year results using the Opus speech processor with the fine structure speech coding strategy. ***Acta Otolaryngol. 129,*** 988–991.
30. Arnoldner, C., Riss, D., Brunner, M., ***et al.*** (2007) Speech and music perception with the new fine structure speech coding strategy: Preliminary results. ***Acta Otolaryngol. 127,*** 1298–1303.
31. Dillon, M., Buss, E., King, E.R., ***et al.*** (2016) Comparison of two cochlear implant coding strategies on speech perception. ***Cochlear Implants Int. 17,*** 263–270.
32. Büchner, A., Brendel, M., Krüger, B., ***et al.*** (2008) Current steering and results from novel speech coding strategies. ***Otol. Neurotol. 29,*** 203–207.
33. Cochran, W.T., Cooley, J.W., Favin, D.L., ***et al.*** (1967) What is the Fast Fourier Transform? ***Trans. Audio Electroacoust. 15,*** 45–55.
34. Landsberger, D.M., Padilla, M. & Srinivasan, A.G. (2012) Reducing current spread using current focusing in cochlear implant users. ***Hear. Res. 284,*** 16–24.
35. Saoji, A. a & Litvak, L.M. (2010) Use of 'phantom electrode' technique to extend the range of pitches available through a cochlear implant. ***Ear Hear. 31,*** 693–701.
36. Macherey, O., Deeks, J.M. & Carlyon, R.P. (2011) Extending the limits of place and temporal

pitch perception in cochlear implant users. ***J. Assoc. Res. Otolaryngol. 12,*** 233–51.

37. Carlyon, R.P., Monstrey, J., Deeks, J.M. & Macherey, O. (2014) Evaluation of a cochlear-implant processing strategy incorporating phantom stimulation and asymmetric pulses. ***Int. J. Audiol. 53,*** 871–879.
38. Nogueira, W., Litvak, L.M., Saoji, A. a. & B??chner, A. (2015) Design and evaluation of a cochlear implant strategy based on a 'Phantom' channel. ***PLoS One 10,*** 1–25.
39. Munjal, T., Roy, A.T., Carver, C., Jiradejvong, P. & Limb, C.J. (2015) Use of the Phantom Electrode strategy to improve bass frequency perception for music listening in cochlear implant users. ***Cochlear Implants Int. 16,*** S121–S128.
40. Litvak, L.M., Spahr, A.J. & Emadi, G. (2007) Loudness growth observed under partially tripolar stimulation: model and data from cochlear implant listeners. ***J. Acoust. Soc. Am. 122,*** 967–981.
41. Vellinga, D., Briaire, J.J., van Meenen, D.M.. & Frijns, J.H.M. (2017) Comparison of multipole stimulus configurations with respect to loudness and spread of excitation. ***Ear Hear. 38,*** 487–496.
42. Bierer, J.A., Bierer, S.M. & Middlebrooks, J.C. (2010) Partial tripolar cochlear implant stimulation: Spread of excitation and forward masking in the inferior colliculus. ***Hear. Res. 270,*** 134–42.
43. Kalkman, R.K., Briaire, J.J. & Frijns, J.H.M. (2014) Current focussing in cochlear implants: An analysis of neural recruitment in a computational model. ***Hear. Res. 322,*** 89–98.
44. van der Jagt, M.A., Briaire, J.J., Verbist, B.M. & Frijns, J.H.M. (2016) Comparison of the HiFocus Mid-Scala and HiFocus 1J electrode array: Angular insertion depths and speech perception outcomes. ***Audiol. Neurotol. 21,*** 316–325.
45. Goldwyn, J.H., Bierer, S.M. & Bierer, J.A. (2010) Modeling the Electrode-Neuron Interface of Cochlear Implants: Effects of Neural Survival, Electrode Placement, and the Partial Tripolar Configuration. ***268,*** 93–104.
46. Frijns, J.H.M., Dekker, D.M.T. & Briaire, J.J. (2011) Neural excitation patterns induced by phased-array stimulation in the implanted human cochlea. ***Acta Otolaryngol. 131,*** 362–70.
47. Nadol, J. jr, Young, Y. & RJ, G. (1989) Survival of spiral ganglion cells in profound sensorineural hearing loss: implications for cochlear implantation. ***Ann Otol Rhinol Laryngol 98,*** 411–416.
48. Moore, B.C.J. (2004) Dead Regions in the Cochlea: Conceptual Foundations, Diagnosis, and Clinical Applications. ***Ear Hear. 25,*** 98–116.
49. Mcdermott, H.J. & Mckay, C.M. (1994) Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea. ***96,*** 155–162.
50. Kwon, B.J. & van den Honert, C. (2006) Dual-electrode pitch discrimination with sequential interleaved stimulation by cochlear implant users. ***J. Acoust. Soc. Am. 120,*** EL1-L6.
51. Frijns, J.H.M., Kalkman, R.K., Vanpoucke, F.J., Bongers, J.S. & Briaire, J.J. (2009) Simultaneous and non-simultaneous dual electrode stimulation in cochlear implants: evidence for two neural response modalities. ***Acta Otolaryngol. 129,*** 433–9.
52. Hole, D.T.R., Van Winden, M.E.C., Alink, M., Ricker, L.J.A.G., Van Groenendael, L., Zwart, V., Wijdicks, S., Goutier, N.Y., Van Der Straaten, T. (2018) Ingredients for an awesome weekend. Het Gat. 4, 1-6

CHAPTER 8

Nederlandse samenvatting

Bij zeer ernstige slechthorendheid en doofheid kan cochleaire implantatie een uitkomst bieden voor zowel volwassenen als kinderen. Met een cochleair implantaat (CI) wordt de gehoorzenuw direct gestimuleerd, waardoor een functionerend buiten- midden- en binnenoor niet nodig zijn om geluid te kunnen waarnemen. In de afgelopen decennia is de techniek van CI's enorm vooruitgegaan. Waar de implantaten in de jaren 80 slechts één frequentiekanaal (d.w.z. één toonhoogte) hadden, hebben de CI's van vandaag 12 tot 22 kanalen. Dit, in combinatie met meer geavanceerde spraakcodering strategieën, heeft ertoe geleid dat patiënten die geïmplanteerd zijn met een CI tot wel 100% kunnen verstaan. Echter, er is nog veel winst te behalen. Met name wanneer de luistercondities moeilijker worden, bijvoorbeeld met achtergrondgeruis, gaat het verstaan van spraak enorm achteruit. Het doel van dit proefschrift was om nieuwe geluidscoderingsstrategieën te ontwikkelen die het luisteren in een lawaaiige omgeving en naar meer complexe stimuli makkelijker maken voor CI-gebruikers.

In **hoofdstuk 1** wordt ingegaan op de verschillende typen gehoorverlies, en wordt het basisprincipe van een CI uitgelegd. Ook de verschillende fases in de ontwikkeling van de huidige CI's komen aan bod en er wordt besproken hoe deze nieuwe technieken geëvalueerd kunnen worden. Naast de bekende spraak-in-ruistests zijn er methoden om de meer basale functies van het gehoor te evalueren. Hiermee wordt gedoeld op de spectrale en temporele resolutie. Spectrale resolutie is het vermogen om de samenstellende componenten (de verschillende toonhoogtes) van een geluid afzonderlijk te horen. Temporele resolutie is het onderscheidingsvermogen naar tijdsverschillen binnen een geluid.

De spectrale resolutie kan onder anderen worden gemeten met de "*spectral-ripple test*", en de temporele resolutie met de "*temporal modulation detection test*". Deze correleren afzonderlijk met het spraakverstaan en laten geen leereffecten zien wanneer ze uitgevoerd worden op één dag. Een ander voordeel is dat de tests taalonafhankelijk zijn, zodat resultaten internationaal vergeleken kunnen worden. Ze worden dan ook regelmatig gebruikt in het wetenschappelijk onderzoek naar CI's. Echter, de studie die wordt gepresenteerd in **hoofdstuk 2** liet zien dat CI-gebruikers beter worden in het maken van de *spectral ripple* en *temporal modulation detection tests* wanneer zij deze herhaaldelijk uitvoeren op verschillende dagen. In dit onderzoek werd de prestatie op deze twee tests zowel vóór als na deelname aan een eerder onderzoek, waarin twee spraakcodering strategieën werden geëvalueerd, gemeten. Tussen de twee meetmomenten bleken de deelnemers significant beter te hebben gescoord. Het is onduidelijk of de scores verbeterden door het herhaaldelijk uitvoeren van de tests of dat het deelnemen aan een wetenschappelijk onderzoek, en dus ervaring opdoen met een ander spraakcoderingstrategie, een effect heeft gehad op de scores. Wel laten deze resultaten zien dat een gerandomiseerd onderzoeksdesign essentieel is voor een correcte interpretatie van de resultaten.

Zoals in hoofdstuk 1 van dit proefschrift is beschreven, maakt het geluid dat wordt opgevangen door een CI meerdere verwerkingsstappen door alvorens het als een stimulatiepatroon aan de gehoorzenuw wordt aangeboden. Vaak worden meerdere stappen aangepast om een nieuwe spraakcoderingsstrategie implementeerbaar te maken. Een voorbeeld is de HiRes Fidelity 120 (HiRes120) spraakcoderingsstrategie. Deze strategie maakt gebruik van "*current steering*" waarbij door het gelijktijdig activeren van twee elektrodecontacten het tussengelegen gebied van de gehoorzenuw gestimuleerd wordt. Door de verhouding tussen de twee elektrodecontacten te variëren, kunnen per elektrodepaar in theorie 8 verschillende toonhoogtes worden gecreëerd, dus zelfs tot 120 verschillende toonhoogtes over de gehele elektrode-array. Om dit mogelijk te maken, moest de analyse van het geluid verbeterd worden en werden de filterbanken vervangen door filters die zijn gebaseerd op Fast Fourier Transformatie (FFT). In **hoofdstuk 3** werd onderzocht of de toepassing van FFT een effect heeft op de spectrale resolutie, de temporele resolutie en daarmee op het spraakverstaan. Drie strategieën werden met elkaar vergeleken; een zonder FFT en zonder *current steering*, een met FFT maar zonder *current steering* en een met zowel FFT als *current steering*. Er werd geen verschil in de scores op de *spectral ripple, temporal modulation detection* en spraak in ruis test gevonden. Dit betekent dat de FFT-filters geen effect hebben op de temporele en spectrale resolutie, of op het spraakverstaan. De filters gebaseerd op FFT kunnen dus zonder problemen worden geïmplementeerd in nieuwe spraakcoderingsstrategieën.

Een valkuil van CI's is dat elk afzonderlijk elektrodecontact een relatief breed gebied van de gehoorzenuw stimuleert. Dit leidt tot een matige spectrale resolutie, wat een negatief effect heeft op het spraakverstaan. Onderzoek heeft aangetoond dat "tripolaire stimulatie" tot een smaller intra-cochleair elektrisch veld leidt en daarmee de spectrale resolutie verhoogt. Dit gaat echter gepaard met verhoogd stroomverbruik, waardoor bij sommige patiënten geen volledige luidheidsgroei kan worden behaald. In **hoofdstuk 4 en 5** worden twee onderzoeken naar een luidheidscoderingstrategie, genaamd *Dynamic Current Focusing* (DCF), beschreven. Deze strategie maakt gebruik van tripolaire stimulatie bij drempelwaarde en verhoogt de luidheid door de mate van tripolaire stimulatie te verlagen. Hierbij blijft het totale stroomverbruik gelijk. De resultaten zoals beschreven in **hoofdstuk 4** lieten zien dat de spectrale resolutie bij lage luidheden significant verbeterde, zelfs wanneer de CI-gebruikers slechts enkele uren de tijd hebben gehad om te wennen aan deze nieuwe strategie. Het was veelbelovend dat het spraakverstaan met de DCF-strategie gelijk was aan die met de klinisch strategie, ondanks de acute test set-up. Omdat spectrale resolutie is gecorreleerd met het spraakverstaan op de lange termijn, werd verwacht dat een langere gewenningsperiode met de DCF-strategie zal leiden tot een verbeterd spraakverstaan. Het batterijverbruik met deze implementatie van de DCF-strategie was aanzienlijk verhoogd. Om die reden is, zoals beschreven in **hoofdstuk 5,** het effect van een energie-zuinige versie van de DCF-strategie na vijf weken gewenning onderzocht. Deze versie van de DCF-strategie zorgde inderdaad voor een langere batterijduur van minimaal 8 uur. Zowel de spectral-ripple als de spraak-in-ruistest lieten een significant leereffect in de tijd zien. Het verschil tussen DCF en de baseline meting was +0.9 ripples per octaaf voor spectrale resolutie (65dB) (p=0.012) en -1.4 dB signaal-ruis-verhouding bij de spraak-in-ruis test (45dB) (p=0.012). Bij vergelijking met de uiteindelijke meting verdween deze significante verbetering en gaf DCF zelfs een significante (p=0.012) achteruitgang van +1.9 dB signaal-ruisverhouding in de spraak-in-ruistest (65dB). Al bemoeilijken de gevonden leereffecten de interpretatie van de resultaten, was het positieve effect van DCF op de lange termijn minder groot dan verwacht op basis van de acute studie.

In **hoofdstuk 6** werd een strategie om meer laagfrequente informatie over te brengen, genaamd fantoomstimulatie, bestudeerd. Door twee elektrodecontacten (het primaire en het compenserende contact) tegelijkertijd te stimuleren met tegengestelde polariteit, wordt het elektrisch veld in de tegengestelde richting van het compenserende elektrodecontact gestuurd. Wanneer dit op het meest apicale elektrodecontact plaatsvindt, kan je in theorie gebieden van de gehoorzenuw stimuleren waar het meest diep geïmplanteerde elektrodecontact (het meest apicaal) anders niet kan komen. Er zijn verschillende configuraties van fantoomstimulatie (zie figuur X in hoofdstuk 6). Zo kan de vorm van de puls symmetrisch zijn, of een pseudo-monofasische vorm hebben waarbij de eerste fase 4 keer zo smal en hoog is als de tweede fase. Daarnaast kan de verhouding in amplitude tussen de twee elektrodecontacten variëren, waarbij een compensatie coëfficiënt van 1.0 betekent dat de amplitudes gelijk zijn en 0.0 betekent dat de amplitude van het compenserende contact 0% van de amplitude op het primaire contact bedraagt. Het onderzoek in hoofdstuk 6 laat met behulp van psychofysische experimenten zien dat, wanneer de beste configuratie voor elke patiënt wordt gekozen, de toonhoogte gemiddeld 0.92 elektrodecontacten opschuift. Wat de beste configuratie is verschilde per persoon maar het vaakst is dat een symmetrische puls met een compensatie coëfficiënt van 0.7 of 0.8. Bij patiënten die geïmplanteerd zijn met een elektrode-array die ontworpen is om dichter bij de gehoorzenuw te liggen, werd een significant minder grote toonhoogteverschuiving gevonden dan wanneer de patiënten geïmplanteerd waren met een electrode-array die verder van de gehoorzenuw af ligt. Het feit dat luidheid geen effect had op de toonhoogteverschuiving vergemakkelijkt de implementatie van fantoomstimulatie.

Hoofdstuk 7 geeft een algemene discussie van de belangrijkste resultaten en conclusies van de studies beschreven in dit proefschrift. Daarnaast worden implicaties voor de klinische praktijk en toekomstig onderzoek besproken.

APPENDIX

Abbreviations
Contributing Authors
List of Publications
About the Author
Dankwoord

ACE	Advanced Combinational Encoder
ANOVA	Analysis of variance
AUC	Area under the curve
BEDCS	Bionic Ear Data Collection System
CA	Compressed Analog
CI	Cochlear Implant
CIS	Continuous Interleaved Sampling
CU	Clinical units
CVC	Consonant-vowel-consonant
DCF	Dynamic Current Focusing
DC-VC	Dynamically Compensated Virtual Channel
FFT	Fast Fourier Transformation
FIR	Finite Impulse Response
FSP	Fine Structure Processing
Hifocus MS	Hifocus Mid Scala
HINT	Hearing in noise test
HiRes	HiResolution
HiRes FFT	HiResolution with FFT-based filters
HiResF120	HiResolution Fidelity 120
HiRes Optima	HiResolution Optima
LIST	Leuven intelligibility sentences test
LUMC	Leiden University Medical Center
MDT	Modulation detection threshold
M-level	Most comfortable level
MP	Monopolar
PS	Pseudo-monophasic pulse shape
PSA_X	Pseudo-monophasic pulse shape with the anodic phase first, with compensation coefficient σ "X"
pTP	partial tripolar
RPO	Ripples per octave
SAS	Simultaneous Analog Stimulation
SB	Symmetric biphasic pulse shape
SBC_X	Symmetric biphasic pulse shape with the cathodic phase first, with compensation coefficient σ "X"
SE	Standard error
SMRT	Spectral-Temporally Modulated Ripple Test
SNR	Signal to noise ratio
SPEAK	Spectral Peak Speech Coding
SPL	Sound pressure level
SRT	Speech reception threshold
SSQ	Spatial and qualities of hearing scale
T-level	Threshold level
TMTF	Temporal Modulation Transfer Function
TP	Tripolar
3AFC	Three-alternative forced choice

Prof. J.H.M. Frijns, MD, PhD
Department of Otorhinolaryngology and Head & Neck Surgery, Leiden University Medical Center, The Netherlands
Leiden Institute for Brain and Cognition, The Netherlands

J.J. Briaire, PhD
Department of Otorhinolaryngology and Head & Neck Surgery Leiden University Medical Center, The Netherlands

J.D. Biesheuvel, MSc
Department of Otorhinolaryngology and Head & Neck Surgery Leiden University Medical Center, The Netherlands

J. Snel-Bongers, MD, PhD
Department of Otorhinolaryngology, Zuyderland Medical Center, Heerlen, The Netherlands

S. Böhringer, PhD
Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands

G.R.F.M Timp, MSc
Department of Otorhinolaryngology and Head & Neck Surgery, University Medical Center Utrecht, Utrecht, The Netherlands

S.F.S. van der Woude, MSc
Department of Biomedical Engineering, University of Technology, Delft, The Netherlands

De Jong, M.A.M., Briaire, J.J., Frijns, J.H.M. (2017). Take-Home Trial Comparing Fast Fourier Transformation-Based and Filter Bank-Based Cochlear Implant Speech Coding Strategies. ***Biomed Res. Int.***, 2017.

De Jong, M.A.M., Briaire, J.J., Frijns, J.H.M. (2017). Learning Effects in Psychophysical Tests of Spectral and Temporal Resolution. ***Ear Hear.,*** 39, 475–481.

De Jong, M.A.M., Briaire, J.J., Frijns, J.H.M. (2019). Dynamic Current Focusing: A Novel Approach to Loudness Coding in Cochlear Implants. ***Ear Hear***., *40(1), 34-44*

De Jong, M.A.M., Briaire, J.J., Van Der Woude, S.F.S., Frijns, J.H.M. (2019). Dynamic Current Focusing for Loudness Encoding in Cochlear Implants: A Take-Home Trial. ***Accepted for publication in International Journal of Audiology***

De Jong, M.A.M., Briaire, J.J., Biesheuvel, J.D., Snel-Bongers, J., Böhringer, S., Timp G.R.F.M., Frijns, J.H.M. Effectiveness of Phantom Stimulation in Shifting the Pitch Percept in Cochlear Implant Users. ***Under review***

De Jong, M.A.M., Smithuis, L.O., Stokroos, R.J. (2014) Eosinofiele Otitis Media. ***Ned Tijdschr Geneeskd, 158.***

Biesheuvel J.D., Briaire J.J., De Jong M.A.M., Böringer S., Frijns J.H.M. (2019) Channel discrimination along all contacts of the cochlear implant electrode array and its relation to speech perception. ***Int J Audiol.*** *20, 1-7*

A

Monique de Jong was born in Gouda, The Netherlands, on 20 December 1988. She completed her secondary education (VWO) at St. Antoniuscollege, Gouda, in 2007. In the same year she started studying Biomedical Sciences at the University of Amsterdam. In 2010 she obtained the BSc degree and was elected to obtain a master degree in Medicine and Clinical Research (Arts-Klinisch Onderzoeker, A-KO) at Maastricht University. She finished this double master program in 2014 with a thesis about a prediction model for congenital sensorineural hearing loss in a high risk population. Subsequently, she started as a PhD candidate at the Department of Otorhinolaryngology at Leiden University Medical Center under the supervision of prof. Johan Frijns, MD PhD, resulting in this thesis. In 2018 she started her residency in Otorhinolaryngology at the same department.

Hij is af! En ik maar denken dat ik het wel alleen kon..
Niet alleen die twee meesterbreinen achter dit promotieboekje, Jeroen en Johan, waren onmisbaar tijdens mijn promotieonderzoek. Naast de inhoudelijke hulp van o.a. Dick, Randy en Leo Litvak was de morele support in J2-55, J2-80 en op H5 onvergetelijk. De wacht werd nogal eens afgewisseld en steeds wanneer ik dacht dat de support dan ook minder zou worden werd ik weer verrast door nieuwe superleuke collega's die eigenlijk ook vrienden zijn geworden. Zonder die karaoke-, Lebkov-, vrimibo-, zeil-, step-, BBQ-, Sinterklaas-, boulder- en ga zo maar door momenten was dit proefschrift ongetwijfeld dikker en indrukwekkender geweest. En ik een stuk minder gelukkig. Eerlijk gezegd, denk ik dat ik de inhoud van dit proefschrift vooral te danken heb aan Advanced Bionics, de mede-auteurs van de artikelen, de CI-gebruikers die hebben deelgenomen aan het onderzoek, en mezelf. Maar zonder mijn ouders, mijn lieve zus en (schoon)broer, de leukste vrienden (waaronder mijn twee paranimfen Marieke en Marleen) en natuurlijk Lukas denk ik niet dat dit me allemaal was gelukt.

Dank jullie allemaal!