Arne Kirkhorn Rødvik

Speech sound confusions in well-performing adults and children with cochlear implants, measured by repetition of mono- and bisyllabic nonsense words

UNIVERSITAS OSLOENSIS
MDCCCXI

Faculty of Educational Sciences 2019

Speech sound confusions in well-performing adults and children with cochlear implants.

A systematic review and meta-analysis, and two experimental studies on the confusion of consonants and vowels in cochlear implant users.

Series of dissertations submitted to the
Faculty of Educational Sciences, University of Oslo
No. 316

ISSN 1501-8962

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Abstract

This thesis describes a three-step quantitative, cross-sectional project, which investigates the confusion of consonants and vowels in well-performing, pre- and postlingually deaf adults and children with cochlear implants (CIs). Consonants and vowels are presented in a nonsense syllable repetition test (NSRT), in the context of monosyllabic consonant-vowel-consonant (CVC) nonsense words and bisyllabic vowel-consonant-vowel (VCV) nonsense words, named nonsense syllables in this thesis. The test is conducted in an open-set design, in which the number of response alternatives is limited only by the size of the participant's mental lexicon.

The first step, reported in Article I, was a systematic review and meta-analysis (Appendix I; Rødvik et al., 2018) aiming to establish a multilingual baseline for consonant and vowel identification scores in pre- and postlingually deaf users of multichannel CIs, tested with CVC and VCV nonsense syllables.

Forty-seven articles covering 50 studies with 647 participants, of whom 581 were postlingually deaf and 66 prelingually deaf, met the inclusion criteria. The mean performance on consonant identification tasks for the postlingually deaf CI users, 58% (n = 44), was higher than for the prelingually deaf CI users, 47% (n = 6), but the difference was not statistically significant. The most common confusions occurred between consonants with the same manner of articulation and the same voicing (/k/ as /t/, /m/ as /n/, and /p/ as /t/). The mean performance on vowel identification tasks for the postlingually deaf CI users was 77% (n = 5), which was higher than the mean performance for the prelingually deaf, 68% (n = 1). A univariate meta-regression model, although not statistically significant, indicated that duration of implant use in postlingually deaf adults predicts a substantial portion of their consonant identification ability.

The second step, reported in Article II, was a study of consonant and vowel confusions in adult CI users measured by an NSRT (Appendix II; Rødvik et al., 2019a). Thirty-nine adults with CIs and a reference group of 20 adults with normal hearing participated. The main objective was to investigate, in detail, the properties of speech sound confusions in adults with CIs, such as the influence of voicing and nasality on perception. The study also aimed to

investigate how a subgroup of users of Med-El's fine structure (FS) stimulation strategies perceived consonant features compared to a subgroup of users of non-FS strategies.

The mean score on the NSRT was significantly lower than the mean score on the real-word monosyllable test, 62% (*SD* = 13%) versus 73% (*SD* = 11%). Hence, the NSRT appeared to reveal more speech sound perception challenges than the real-word monosyllable test.

Other findings:

- The consonant scores were lower than the vowel scores, 57% (*SD* = 14%) versus 72% (*SD* = 17%), and the voiced consonant scores were lower than the unvoiced consonant scores, 53% (*SD* = 15%) versus 63% (*SD* = 16%).
- The stops had a devoicing bias, as voiced stops were often repeated as unvoiced stops, but unvoiced stops were never repeated as voiced stops.
- The nasals were confused with other nasals in one third of the cases and repeated correctly in only one third of the cases.
- [yː] was perceived as [iː] in most of the cases and [iː] was perceived correctly in all cases.
- The perception of nasals versus nonnasals, nasals versus the lateral [l], and stops versus fricatives was significantly higher for a small sample of the non-FS strategy users than for a matched group of the FS strategy users. The perception of voicing was significantly higher for the FS strategy users than for the non-FS strategy users.

The study revealed a general devoicing bias for the stops and a high confusion rate of nasals with other nasals. The subgroup comparison of small samples of users of FS and non-FS stimulation strategies suggests that more research to improve the coding of the low-frequency information in the speech signal is needed.

The third step, reported in Article III, was a study of consonant and vowel confusions in children with CIs measured by an NSRT (Appendix III; Rødvik et al., 2019b) and performed with 36 children with CIs, and two normal-hearing reference groups of 17 six-year-olds and 12 thirteen-year-olds.

The first objective was to measure the confusion of consonants and vowels in well-performing children and adolescents with CIs. The second objective was to investigate how pre- and postlingual deafness influenced the confusions and the perception of speech features.

For the participants with CIs, the mean voiced consonant repetition score was 64% (*SD* = 11%), the mean unvoiced consonant repetition score was 77% (*SD* = 10%), and the mean vowel repetition score was 85% (*SD* = 11%). Subgroup analyses showed no statistically significant differences between the consonant scores for pre- and postlingually deaf participants.

The participants with CIs obtained scores close to ceiling on vowels and real-word monosyllables, but their perception was substantially lower for voiced consonants. This may partly be related to limitations in the CI technology for the transmission of low-frequency sounds.

Taken together, the results show that the mainly prelingually deaf children and adolescents with CIs obtained overall higher scores on the NSRT and on a real-word monosyllable test than the mainly postlingually deaf adult CI users. For both groups, the perception score of vowels was higher than the unvoiced consonant score, which was higher than the voiced consonant score. This confirms the well-known phenomenon that the frequency-place mismatch of the implants is most pronounced for the postlingually deaf.

Although the participating CI users had a 100% correct pronunciation score, none of them obtained scores for voiced consonants above 78%. As their speech is much better than their perception capability would indicate, people they encounter in their everyday life might underestimate the severity of their hearing impairment.

The CI technology has developed substantially since the advent of the commercial multi-channel implants in the early 1980's. However, our results indicate that there still are limitations in today's CI technology for the transmission of low-frequency sounds.

Abbreviations and acronyms

Abbreviation/acronym	Meaning
ANSD	Auditory nerve spectrum disorder
ART	Auditory response telemetry
CI	Cochlear implant
CIS	Continuous interleaved sampling
CNC	Consonant-vowel nucleus-consonant
CVC	Consonant-vowel-consonant
EABR	Electrically evoked auditory brainstem response
ECAP	Electrically evoked compound action potential
ENT	Ear, nose, and throat
ESRT	Electrically evoked stapedius reflex threshold
FS	Fine structure (common feature of Med-El's three stimulation strategies: FSP, FS4, and FS4-p)
HINT	Hearing in Noise Test
IPA	International phonetic alphabet
NAV	Norwegian Labour and Welfare Administration
NRI	Neural response imaging
NRT	Neural response telemetry
NSRS	Nonsense syllable repetition score
NSRT	Nonsense syllable repetition test
OUS	Oslo University Hospital
REC	Regional ethical committee
VCV	Vowel-consonant-vowel

Table of contents

List of articles

Article I: Rødvik, AK, Torkildsen, JvK, Wie, OB, Storaker, MA, and Silvola, JT (2018). Consonant and vowel identification in cochlear implant users measured by nonsense words: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research* 61**,** 1023–1050.

Article II: Rødvik, AK, Torkildsen, JvK, Wie, OB, Tvete, O, Skaug, I, Silvola, JT. (2019a). Consonant and vowel confusions in well-performing adult cochlear implant users, measured by a nonsense syllable repetition test. Submitted to *The Journal of The Acoustical Society of America*.

Article III: Rødvik, AK, Tvete, O, Torkildsen, JvK, Wie, OB, Skaug, I, and Silvola, JT (2019b). Consonant and vowel confusions in well-performing children and adolescents with cochlear implants, measured by a nonsense syllable repetition test. *Frontiers in Psychology* 10**,** 1–17.

Table of figures

Acknowledgements

I would like to express my gratitude to the Department of Special needs education at the University of Oslo, for giving me the opportunity to accomplish this work. Also, thanks to my previous colleagues at the Department. You were all supportive during the four years that I was employed here. As a member of the research group Child Language and Learning (CLL), I was included in a dynamic and vibrant research environment under the experienced and inclusive leadership of Professor Bente Hagtvet. As a member of the National Graduate School in Educational Research (Nated), I learned a lot to help me through the PhD work.

Thanks to former PhD student colleagues for being such good company during these years. Especially thanks to my office room-mate Magnar Ødegaard, for enduring my company for four years, during which he stoically accepted my frequent refurnishing of our office, and to Anne Arnesen and Hanne Næss Hjetland, for nice and stimulating company and interesting conversations during our last summer at the Department.

I wish to express my gratitude to the ENT Department at Oslo University Hospital (OUS), who generously granted me a 4-year-leave to do this project and provided access to clinical data and to the facilities at the hospital during my leave. Having been employed at the ENT Department for 14 years before the start of my PhD-work, I was lucky to know the administrative systems at the hospital and the employees at the Ear Section at the department well in advance. My previous and present colleagues at the Ear section helped me collecting the data I needed, were benevolent discussion partners and were positive to my frequent intrusions at the clinic to test patients in the project.

I want to thank the members of the Cochlear Implant Unit at the ENT Department, who have always done their best to support me. I especially wish to thank Kjell Rasmussen, who came up with the project idea at a playground in Bærum in the beginning of the 2000s, Jens Jørgen Dammerud, who programmed a randomization script for the nonsense syllables in the test, and Marte Myhrum, who would always try her best to answer my intricate questions and help me out with the analyses. Last but not least, I wish to thank Ole Tvete, long time colleague, who has spent a lot of his vacant time to help me through the writing and analysis process. Even in sunny days, during which he would probably rather spend time in the woods, on top of his skis, or on his bicycle, he may sacrifice this to work with me.

Thanks to my supervisors, who are also my co-authors, Juha Tapio Silvola, Janne von Koss Torkildsen, and Ona Bø Wie. You have never counted supervision hours and dismissed my questions because of lack of time. On the contrary, you have always been available and supportive, and responded quickly to my inquiries. You have also provided generous economic support from your research funds when needed in my project. Janne has an enormous research experience and high skills as a researcher, particularly in scientific writing, which I have been lucky enough to draw benefit from, having her as a co-author and supervisor. Ona has been very supportive and helped my project onward with her great scientific experience, broad audiological background, and large international network.

Juha is a nice person and an experienced researcher and ear surgeon, with whom I have also worked as a colleague. He has never insisted that I only call him during working hours. Instead, he has basically been available all waking hours of the day and has done his best to push me, and to pull me through this work.

Thanks to my parents Knut Henrik Rødvik and Magnhild Kirkhorn Rødvik, who have always, and in all possible ways, supported me. My father is a teacher, and he has always encouraged me to study hard and to choose a higher education. He has always been interested in research himself and has inspired me with his scientific curiosity. My mother has loved and supported me unconditionally and limitless. Moreover, my three younger siblings, Torgrim, Trond, and Marit have always been a great support for me.

Thanks to my dear wife and soul mate, Ingebjørg Skaug, who has sacrificed countless holidays, weekends, and evenings to support my project. Ingebjørg is also co-author in two of my articles and proof-reader of my manuscripts. I have been fortunate to draw on her long-time experience with speech training of adults and children with CIs, and her knowledge of articulation and phonetics. Thanks for your infinite love and support. Thanks also to our dear adolescents, Iselin and Arian, who have endured a periodically absent-minded and often absent, father, for all these years.

1 Introduction

Cochlear implants (CIs) establish hearing in prelingually deaf children with severe to profound hearing loss and re-establish hearing in individuals who have lost their hearing. The treatment is standard care in most of the developed countries in the world for profoundly deaf children and adults who cannot benefit from amplification. CIs are also often used as treatment for tinnitus and for *auditory neuropathy spectrum disorder* (ANSD). In 2016 there were 600,000 CI users world-wide (Ear Foundation, 2017), and the multichannel CIs have been commercially available for more than three decades. CIs bypass the outer and middle ear and the approximately 3,000 damaged sensory cells of the inner ear, and directly stimulate residual auditory nerve fibres with patterns of electrical pulses delivered via an electrode array implanted into the cochlea. For severely and profoundly deaf individuals, the CIs are effective and can provide perception of speech and music. However, CIs do not restore normal hearing, outcomes vary among patients, performance is considerably degraded by ambient noise, and music perception is limited (O'Donoghue, 2013).

The CI is composed of an exterior and an interior part (see Figure 1). The exterior part combines a microphone, speech processor, coil, and magnet. The speech processor transforms incoming sound into electrical pulses, which are filtered into frequency bands. The pulses are submitted through intact skin to the electrode array by a transmitting coil. The interior part is an electrode array (a cable) connected with a receiver and transmitter; surgically inserted into the cochlea in the inner ear. The stimulus pulses are distributed to the implant according to a specific stimulation strategy. For instance, for the continuous interleaved sampling (CIS) strategy (Wilson et al., 1991), slow changes of the temporal sound envelope of the speech signal are converted into amplitude-modulated trains of biphasic (having both negative and positive components) pulses at the electrodes. The speech processor sends information from low frequency channels to electrodes in the apex of the cochlea and information from high frequency channels to electrodes in the base. This frequency distribution mimics the tonotopy of a normal cochlea (Dorman and Wilson, 2004).

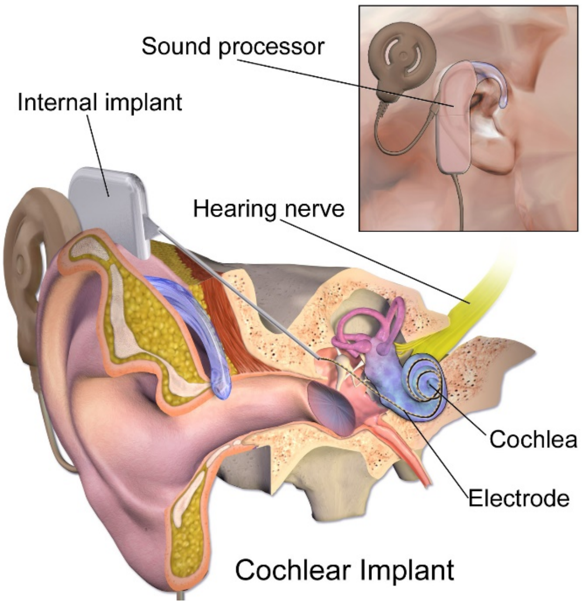


Figure 1. Drawing of a cochlear implant with external and internal parts (Blausen.com staff, 2014).

The implant is threaded into the cochlea and directly stimulates portions of the auditory nerve with the electrical pulses. The auditory nerve carries encoded sound information from the cochlea to the first auditory relay station in the brain, the cochlear nucleus. The signal ascends through multiple nuclei until it reaches the auditory cortex in the temporal lobe of the brain, in which the signal will eventually be perceived as consonants, vowels, or other sounds.

Knowledge of the extent to which CI users identify speech sounds correctly is important to ensure that those who fit and program the CIs find an optimal setting and provide the user with the maximum benefit. There is no overview of the most common confusions of speech sounds for Norwegian-speaking CI users, and studies conducted in other languages cannot automatically be translated into Norwegian.

CI users have different aetiologies and obtain varying benefit from the implants. The benefit may constitute better speech understanding in quiet and noise and sound localization if or when bilateral hearing is established or re-established. However, the implant seldomly provides auditory skills close to normal hearing, and there still is a need for extensive research to exploit its potential, to which this thesis aims to contribute.

1.1 Background

Before starting this PhD project, I worked 14 years as a clinical physicist in the CI clinic at *Oslo University Hospital* (OUS) Rikshospitalet, programming CIs for adults and children. As a quality control of the implant programming, I conducted speech perception testing of adult CI users, employing recorded sentences and monosyllabic words in quiet, and sentences with added noise.

Due to, among other factors, shorter duration of deafness before implantation and improved implants and speech processors, CI users have gradually obtained higher scores on speech perception tests. Today, many CI users reach ceiling level on the sentence and monosyllable tests in quiet. Acknowledging this as a weakness in the speech perception tests employed at the clinic, I started examining a Norwegian adaptation (Teig et al., 1992) of the *Iowa Vowel and Medial Consonant Test* in the Iowa Cochlear Implant test battery (Tyler et al., 1983). I conducted a pilot study in 2002 with five adult CI users, which was presented at the conference of the International Clinical Phonetics and Linguistics Association (ICPLA) in Dubrovnik, Croatia, in 2006, and later published (Rødvik, 2008). In the present PhD project, which was initiated in 2013, I have explored this topic further by collecting and analysing results from speech perception testing, conducted with a new *nonsense syllable repetition test* (NSRT), of a larger number of children and adults with CIs.

In 2006 the Norwegian version of the *Hearing in Noise Test* (HINT; Nilsson et al., 1994; Myhrum and Moen, 2008) was introduced in the CI unit at OUS Rikshospitalet. In this sentence test, speech levels are adjusted depending on whether the subject’s responses are correct or incorrect, while the noise is kept at constant magnitude. The test produces a speech reception threshold, which is defined as the mean signal-to-noise ratio at which the listener can repeat 50% of the sentences correctly. Thus, ceiling level cannot be reached in this test.

The HINT is very useful for assessing top-down inferential speech perception skills in the individual implantee, that is to say how well the implantee perceives spoken language by relying on both language skills and intelligent guessing in addition to auditory input. The test does not, however, provide detailed information about the perception of individual speech sounds per se, and thus I sought to explore the NSRT as a means of obtaining fine-grained information about consonant and vowel perception that could provide analytical information to be used directly in CI programming and auditory training.

1.2 Cochlear implants

1.2.1 Definitions and concepts

Classification of CI users by age at onset of deafness

The term prelingual severe to profound deafness refers to deafness occurring before the acquisition of speech and language. In Studies 2 and 3 (Articles II and III), participants were classified as *prelingually deaf* if congenitally deaf or severely to profoundly deafened before age 12 month (Myhrum et al., 2017). Furthermore, the prelingually deaf CI users can be divided into two groups: those who have had no or minimal access to sound and hence acquired very little oral language before implantation (receiving a CI before age 1), and those who have acquired oral language and benefited from hearing aids (HAs) due to residual hearing (receiving a CI at higher ages).

Those with onset of severe to profound deafness between one and three years of age are often classified as *perilingually deaf*. Participants who become profoundly deaf after three years of age and have acquired some speech and language before onset of deafness, are classified as *postlingually deaf*.

Definitions of unilateral, bilateral, and bimodal cochlear implant users

Unilaterally implanted CI users wear a CI on one ear only, and bilaterally implanted users wear a CI on each ear. Bimodal CI users have one CI and one contralateral HA. In Article II, bimodal CI users were pooled with unilateral and bilateral CI users on the premise that their perception of monosyllables in the implanted ear was more than 40% better than their perception of monosyllables in the ear with an HA (Crew et al., 2015; Yoon et al., 2015).

1.2.2 History

In 1812, the Italian scientist Alessandro Volta experimented on himself and provided the first annotation of electrical stimulation of the hearing nerve by placing two wires at his water-filled outer ear canals and connecting them to an electric circuit. Volta described his experience as a “jolt in the head,” followed by a sound that resembled “a kind of crackling, jerking, or bubbling as if some dough or thick material was boiling” (Eshraghi et al., 2012, p. 1968).

The French scientists Djourno and Eyriès are credited with the first CI (Eisen, 2003). They collaborated in 1957 to place a permanent stimulating electrode in a patient's temporal bone. The electrode was transcutaneously stimulated by an induction coil, and the patient reported hearing some simple sounds.

In 1972, the House Ear Institute in Los Angeles, U.S.A., released a single-channel implant, building on the works of Djourno and Eyriès. The device was updated and became commercially available in 1982 as the House/3M implant. This early device worked for many users as a lip-reading enhancement and an aid for the perception of environmental sounds (Fretz and Fravel, 1985). Some users even achieved open-set word recognition.

Many consider the first successful device for speech recognition to be the multi-channel CI, developed at the Bionic Ear Institute in Melbourne, Australia, which was first implanted in an adult male in August 1978 (Clark et al., 1981). The first commercial multi-channel CI, the Nucleus CI22, was launched in 1982 (Cochlear, 2016). Several large-scale clinical studies compared single-channel and multi-channel CIs (e.g., Tyler et al., 1988; Cohen et al., 1993), and the conclusion was that postlingually deaf adults' performance with multichannel CIs with four channels or more, was better than the performance of postlingually deaf adults with single-channel devices, despite the large variability between the subjects.

In Norway, the first CI surgery was performed in 1981 at Regionsykehuset i Trondheim (Trondheim regional hospital), with a single-channel CI from the company 3M. The first multi-channel CI in Norway was implanted in 1986 in an adult patient at OUS Rikshospitalet. This was a four-channel transcutaneous device, the Ineraid CI from the U.S. company Symbion. The same year, a Nucleus CI22 from the Australian company Cochlear, a 22-channel percutaneous implant without a built-in magnet, was implanted in a patient at OUS Rikshospitalet. Both implants were used simultaneously until 1990, when implantation of the Ineraid was discontinued.

Haukeland University Hospital in Bergen implanted their first multi-channel CI, an Ineraid, in 1988. Regionsykehuset i Trondheim implanted their first multi-channel CI in 2005.

The first postlingually deaf child in Norway received a CI at OUS Rikshospitalet in 1988. The first prelingually deaf child received a CI at OUS Rikshospitalet in 1989.

Today, approximately 2,200 individuals have received a CI at the three hospitals in Norway, in which this surgical procedure is performed. There are also three “CI-satellites” that offer technical support for adult users, located at Sørlandet sykehus (Sørlandet hospital) in Arendal, Universitetssykehuset Nord-Norge (UNN; University Hospital of North Norway) in Tromsø, and Stavanger universitetssykehus (Stavanger University Hospital). More Norwegian CI satellites of this kind are expected to be established shortly.

1.2.3 Candidacy

As a result of the launch of neonatal hearing screening in Norway in 2008, severely and profoundly deaf children are discovered at an earlier age than before. Hence, the age at implantation has been lowered substantially and some infants as young as five months of age currently undergo CI surgery at OUS Rikshospitalet. Early implantation is one of the factors that explains why today’s prelingually deaf implantees are much better performers than those implanted more than 15 years ago (Niparko et al., 2010; Wie, 2010; Geers et al., 2011).

Almost all children in Norway who meet the criteria for receiving CIs are now bilaterally implanted, and this procedure is covered by the public health system. In some cases, for instance for children with ANSD or with residual hearing in one ear, they may be implanted in two unilateral surgical interventions.

Adults in Norway who are medically accepted for cochlear implantation under the public health system receive in general only one implant. At OUS Rikshospitalet, unilaterally implanted adults are usually offered a second implant if there is a probability that they will gain benefit from it for speech understanding. Candidates with additional medical conditions, such as annoying tinnitus, onset of profound deafness due to meningitis, or blindness may also be offered a second implant, even if improved speech understanding is not expected.

1.2.4 Stimulation strategies

The main purpose of CI stimulation strategies is to set up an electrical signal in the hearing nerve that resembles the signal in the normal ear, by means of electrical stimulation patterns in the CI electrode array. These patterns vary somewhat between stimulation strategies and between implant manufacturers, but they all attempt to convey spectral and temporal information of the original signal to the implant (Wouters et al., 2015).

The spectral information of the speech signal (e.g., the first and second formant, F1 and F2) is conveyed by the multichannel organization of the implants, by mimicking the tonotopic (place) organization of the cochlea from low frequencies in the apical part to high frequencies in the basal part. This information is implemented in all stimulation strategies from the main implant manufacturers today. These are in alphabetical order: Advanced Bionics (Stäfa, Switzerland), Cochlear (Sydney, Australia), Med-El (Innsbruck, Austria), and Oticon Medical/Neurelec (Vallauris, France). The manufacturers use mostly spectral information strategies to convey pitch information, in which tonotopic information is transmitted by stimulating a set of predefined electrodes on the implant array in each stimulation cycle.

The temporal information of the speech signal is commonly decomposed into envelope (2–50 Hz), periodicity (50–500 Hz), and temporal fine structure (TFS; 500–10 kHz), described by for instance Wouters et al. (2015). The envelope is the slow variations in the speech signal. Periodicity corresponds with the vibrations of the vocal cords, which conveys fundamental frequency (F0) information. TFS is the fast fluctuations in the signal and contributes to pitch perception, sound localization, and binaural segregation of sound sources.

All stimulation strategies represent high-frequency sounds only by place coding. Moreover, the stimulation rate is constant for all stimulation strategies, varying between 500 and 3,500 pulses per second for each manufacturer and for each implantee. Low-frequency sounds can be represented by both temporal and place coding.

The TFS strategy HiRes120 from Advanced Bionics creates “virtual channels” between the 15 electrode pairs by varying the relative currents in each pair, in effect increasing the spectral resolution compared to the conventional HiRes strategy. For the implant array of 16 electrodes, 120 virtual channels can be created, which may potentially improve pitch perception in the HiRes120 strategy compared to the HiRes strategy. The sound signal is conveyed by place coding and not by temporal coding. Studies have shown some positive effects of HiRes120 over HiRes on measures of music perception and of speech perception in quiet and in noise (e.g., Donaldson et al., 2011).

The advanced combination encoder (ACE) strategy, Cochlear’s most-used stimulation strategy for more than 20 years, conveys low-frequency information by place coding. Cochlear does not currently apply TFS coding, although in recent years the company developed and is testing the Optimized Pitch and Language (OPAL) strategy. OPAL aims to enhance perception of F0 as a cue to pitch in music, voice pitch in speech, and lexical tone in tonal languages (Vandali and van Hoesel, 2012; Vandali et al., 2017; Vandali et al., 2018). Its approach is to enhance the coding of F0 amplitude modulation in the envelope of the stimuli delivered to each channel containing F0 harmonics. Vandali et al. (2018) have reported promising results in an intonation test for a sample of CI users using OPAL compared with users of ACE.

Med-El represents low-frequency sounds with its TFS stimulation strategies (FSP, FS4, and FS4-p), which pick up the oscillation frequency of the vocal cords by phase-locking the hearing nerve in the low frequencies to convey temporal information to the hearing nerve, thus mimicking how a normal ear treats a low-frequency signal. According to Caldwell et al. (2017), these strategies' encoding may be more similar to the natural signal than that of spectral information strategies, such as CIS, in the case of complex stimuli heavily dependent on pitch, such as music.

1.2.5 Programming

Optimal programming of the CI speech processor for the individual CI recipient is crucial for the correct identification of speech sounds. The goal of the programming is to provide access to all speech sounds and to ensure that the sounds are easily perceived at normal stimulation levels and are never uncomfortably loud. Regardless of the implant model, traditionally two basic psychophysical measures need to be obtained on each intracochlear electrode: 1) electrical thresholds (T-levels), defined as the softest level at which a patient is stimulated 100% of the time, and 2) most comfortable loudness levels (C/M levels), defined as the loudest sound a patient can listen to comfortably for a sustained period of time (Shapiro and Bradham, 2011). Stimulation levels should be set so that both environmental sounds and speech are perceived by the implantee. Moreover, the speech processor should be fitted so that the loudness levels match that of persons with normal hearing sensitivity; soft sounds should be soft to a CI user, while loud sounds should also be loud to the user (Wolfe and Schafer, 2015). Basically, the maximum and minimum loudness levels for each electrode channel need to be found and used as a basis for the combination of CI electrodes in a mono- or bipolar way, thus providing sound to the implantee.

CI programming is usually based on behavioural methods with support from objective measurements. Below is an outline of the two main behavioural methods of CI programming outlined, succeeded by a section describing some of the most common objective programming methods.

Behavioural methods

The main method for programming the CIs for adults and adolescents is the *feedback* method, in which the patient determines the maximum and minimum loudness levels for each electrode channel, as well as the frequency allocation table, in close cooperation with the clinician who does the programming. Balancing the loudness between the electrodes is an important part of the programming. The quality of the programming depends on the precision of the feedback from the patient.

The main method of programming CIs for young children or multi-handicapped patients is the *observation* method, in which the implants are programmed without detailed feedback from the patient. The patient's reaction to electrical stimulation is at our clinic at OUS Rikshospitalet observed by a speech-language therapist and a clinical physicist, in two sessions a day for three days. Other clinics may have different procedures. The parents' observations of their child's reactions are invaluable during the CI programming session.

Objective methods

Behavioural methods can be combined with electrical hearing thresholds (T and C/M levels) suggested by objective electrophysiological measurements. These measurements are implemented in the CI programming software by all the implant manufacturers and are described below.

The *electrically evoked compound action potential* (ECAP) method measures the response of the nerve fibres inside the cochlea after electrical stimuli from the implant and is the most widely used objective measurement (Vaerenberg et al., 2014). Several studies have suggested that the ECAPs possess a weak to moderate correlation with T-levels and C/M-levels (e.g., Van Den Abbeele et al., 2012; McKay et al., 2013; Greisiger et al., 2015), and it is therefore useful as a tool to guide the clinician in determining stimulation levels for recipients who cannot provide reliable feedback regarding the loudness of the signals they receive from their implants. Each CI manufacturer has patented its own ECAP version, such as *neural response telemetry* (NRT; Cochlear), *auditory response telemetry* (ART; Med-El), *neural response imaging* (NRI; Advanced Bionics), and *neuro electrically evoked compound action potentials*

(Neuro ECAP 2.0; Oticon Medical). The basic principles behind these algorithms are the same.

The *electrically evoked stapedius reflex threshold* (ESRT) measurement is often carried out during CI surgery. Single electrodes get stimulated while the surgeon visually observes the reflexes of the stapes muscle. The threshold can be determined by lowering the current or charge delivered to the electrodes and thus to the hearing nerve fibres and observing at which current the reflex disappears. Studies have shown that the correlation between the ESRT and the T- and C/M-levels is poor, but that the subjective C/M-levels rarely exceeds the ESRT levels (Lorens et al., 2004; Caner et al., 2007; Walkowiak et al., 2011; Greisiger et al., 2015).

In addition, the *electrically evoked auditory brainstem response* (EABR) is often measured during CI surgery to verify the coupling of the CI electrodes to the nerve fibres. Also, this might prove a valuable measurement for ANSD patients, which is characterized by dyssynchrony of their nerve fibres. Greisiger (2016) has shown a significant relationship between observed intra-operative EABR measures and post-operative speech recognition. EABR is rarely used in CI programming (Vaerenberg et al., 2014).

Nonphysiological measurements are also useful in the programming session (Hughes, 2012). Faulty electrode arrays may be discovered by impedance measurements, and incorrect combinations of implant parameters may be discovered by examining the voltage compliance levels. CI imaging, by magnetic resonance (MR), x-ray, and computer tomography, may also contribute to the individual programming of the implant, for instance by confirming correct electrode placement or discovering displacement into *scala vestibuli* or perforation of the *basilar membrane*.

There are large variations in the correlation between objective and behavioural threshold levels. The objective methods are useful as a guide to the programming but should always be supplemented with individually directed programming or observation.

1.2.6 Transmission of speech sounds

The transmission of consonants and vowels in CIs is designed to reproduce a speech signal that closely resembles the original by means of electrical stimulation patterns in the electrode array of the CI. Failure to resemble the original signal is always explained from two viewpoints: (1) cognitive and physiological limitations in the implant user's auditory system and (2) technical limitations in the CI system.

Limiting factors in the auditory system might be reduced neural plasticity in the brain due to high age at implantation, perhaps combined with no auditory stimulation prior to implantation. Reduced neural plasticity can also be experienced in case of long duration of deafness before implantation for patients. Other limitations may be congenital malformations in the hearing system, such as missing or damaged auditory nerve, *mondini* deformity of the cochlea, cochlear atresia, or profound deafness due to the [1]CHARGE syndrome (Pagon et al., 1981).

Technical limitations in the CI system are also likely to affect the perception of consonants and vowels. In a CI user with optimal conditions for the perception of speech, the degree of success in implantation surgery, such as placement and insertion depth of the electrode array in the cochlea, will be important for perception. Moreover, the speech coding, the input dynamic range and frequency range of the speech signal, and the implant electrode array properties such as length, hugging/nonhugging, and soft/stiff tip may also influence perception.

As vowels are characterized by long duration and high intensity compared with most consonants, they are usually easily perceived by the implantees, although they may be confused with other vowels with formants close in frequency. Furthermore, as the two primary vowel formants in Norwegian can be found in the frequency range between 200 and 2,500 Hz and the input frequency range of the implant usually includes frequencies as low as 100 Hz and as high as 8,000 Hz, all vowels should be recognizable to the CI user.

The high-frequency parts of the consonants are easily picked up by the CI speech processors. However, the transmission of low-frequency sounds in the implants, specifically F0, has its limitations. Perception of voicing depends on how F0 is processed by the CIs. As the tonotopy of the cochlea is organized with the low frequency sounds in the apex and the high frequency sounds in the base, the more apical part of the cochlea that is stimulated, the lower the pitch perceived by the implantee. As the insertion depth of the electrode array usually is

---

[1] Previously used name of syndrome affecting newborn children with the congenital features of coloboma of the eye, heart defects, atresia of the nasal choanae, retardation of growth and/or development, genital and/or urinary abnormalities, and ear abnormalities and deafness. These features are no longer used in making a diagnosis of CHARGE syndrome, but the name remains.

quite shallow compared with the length of the cochlea, the apical stimulation may induce a perceived frequency transposition of the pitch to a higher frequency than the input stimuli (frequency-place mismatch; Venail et al., 2015). Users of the implants with the longest electrodes should therefore generally be expected to obtain the most correct pitch perception. However, this is not always the case. Other reasons for erroneous pitch perception may be the limited bandwidth of the input frequencies, inaccurate stimulation due to spread of excitation, imprecise coding of temporal cues, malfunctioning auditory nerve, and lack of hearing experience prior to implantation. Many studies have confirmed that CI users do not perceive pitch as well as normal-hearing (NH) listeners (e.g., Sucher and McDermott, 2007; Wang et al., 2011; Tao et al., 2015).

1.2.7 Auditory training

In general, all listening will improve the benefit of the CIs compared with non-use. At our clinic at OUS Rikshospitalet, the CI recipients are recommended to use their CIs all day, and to lower the volume on their CIs rather than take them off if they get tired of the sound. Several studies have shown that auditory training improves the auditory skills of the implantees (e.g., Stacey et al., 2010; Plant et al., 2015; Schumann et al., 2015). The training should focus on speech sounds that are challenging for the individual implantee, employing minimal pairs and triplets as well as normal, connected speech. In other words, a mix of top-down and bottom-up training. Analogously, auditory training with music generally improves the ability to perceive music (e.g., Petersen et al., 2012).

Auditory training for implanted children may vary somewhat for the pre- and postlingually deaf children. Postlingually deaf implanted children may resemble postlingually deaf implanted adults more than prelingually deaf implanted children (Niparko, 2004), as they use their CIs to map new and different sound impressions onto an existing linguistic code. Prelingually deaf implanted children, on the other hand, must use the information from the implant to develop completely novel linguistic codes.

Important issues to keep in mind when dealing with the post-implantation habilitation of children with CIs, are described in numerous guides written by the implant manufacturers and others (e.g., Cochlear, 2006; Med-El, 2008; Wolfe, 2015, pp. 258–260). All emphasize that habilitation must be based on early intervention and aim to equip parents with the skills to maximize their deaf child's speech and language development by stimulating auditory brain development and enabling deaf children with CIs to make sense of the sound relayed by their devices. Parents often profit from observing the speech therapist's communication strategies with the child and adopting the same approach in their own everyday life. Professional, auditory training should always be conducted in consideration of the CI user's age; the younger the child, the more play-based exercises. Closer to school-age, it is important to help children achieve both precise articulation and precise perception of all speech sounds to be well-prepared for learning new words by ear and for learning to write. The best possible reading and writing skills are of course very important in the life of a person who is hard of hearing, both for compensatory information seeking and for communication in general. Advanced vocabulary, which is necessary for higher education, is also mostly acquired through written texts (e.g., Duff et al., 2015). At home, kindergarten, and in the classroom at school, the acoustics should be optimal for speech perception, with a good signal-to-noise ratio. Digital transmission of the voice of all speakers directly into the child's CIs is probably the most efficient way to provide this.

1.3 Aims

The overarching aim of this thesis is to investigate how well pre- and postlingually deaf adults and children with CIs can identify speech sounds, presented in the contexts of monosyllabic CVC and bisyllabic VCV nonsense words.

1.3.1 Article I

The aim of the systematic review and meta-analysis is to examine, pool, and synthesize previous research to investigate how well users of multichannel CIs identify consonants and vowels in tests using monosyllabic and bisyllabic nonsense words as stimuli. The included studies were pooled in a meta-analysis, empirical findings and measurements were aggregated to increase the statistical strength, and a baseline of consonant and vowel perception scores in previous research was established.

The research questions are:

1. What are the typical vowel and consonant identification scores in CI users when measured by nonsense syllables, and how do the typical vowel and consonant identification scores differ between prelingually and postlingually deaf implantees?

2. Which consonants and vowels are most frequently confused by CI users, and which consonants and vowels are most frequently identified correctly?
3. To what extent are age at implantation, duration of implant use, and real-word monosyllable score associated with variations in consonant and vowel identification performance in nonsense syllable tasks for prelingually and postlingually deaf CI users?

1.3.2 Article II

The study's first objective was to identify the most common vowel and consonant confusions and the most common confusions of the phonetic features voicing, nasality, stopping, frication, and the lateral [l] with an NSRT in an open-set design, in a sample of well-performing adult CI users.

The second objective was to investigate how a subgroup of users of Med-El's fine structure (FS) stimulation strategies perceive consonant features compared to a matched subgroup of users of non-FS strategies from Advanced Bionics, Cochlear, and Med-El.

1.3.3 Article III

The overall objective was to measure the perception of speech sounds in well-performing children and adolescents with CIs with an NSRT in an open-set design.

The two sub-objectives were as follows:

Objective 1: To identify the most common vowel and consonant confusions and the most common confusions of the phonetic features voicing, frication, stopping, nasality, and laterality.

Objective 2: To investigate how age at onset of severe to profound (pre-, peri-, and postlingual) deafness influences confusion of speech sounds and features.

1.4 Outline of the thesis

This thesis consists of two main parts: a) the extended abstract and b) three papers (Articles I, II, and III), each of which is written in collaboration with different co-authors.

The studies build on each other as follows: The systematic review and meta-analysis (Article I) provided both an overview of previous research and a baseline of consonant and vowel scores, calculated from the included studies. Based on this, we conducted an experimental study with adult participants with CIs and an NH reference group (Article II) and an experimental study with children with CIs and two NH reference groups (Article III).

Article I summarizes the empirical evidence on cross-linguistic identification of consonants and vowels by CI users, measured by nonsense syllable identification tests. Articles II and III examine the outcomes of adults and children with CIs tested with a Norwegian NSRT. References to the PhD project include all these studies.

2 Empirical and theoretical foundations

2.1 Empirical foundation

Many previous studies have been based on the assumption that repeating nonsense syllables measures the participants' actual auditory skills rather than inferential skills and vocabulary, which participants naturally heavily rely on in real-word tests (e.g., Mulder et al., 1992; Välimaa et al., 2002a; 2002b; Munson et al., 2003). The following discusses the origins of this assumption.

2.1.1 Previous research

One of the first articles that describe the use of nonsense syllables for measuring consonant and vowel perception in CI users, reported on the implantation of multichannel CIs in two profoundly deaf persons (Clark et al., 1981). The stimuli were presented live, as VCV nonsense syllables, visually, auditorily, and visually and auditorily combined. The expectations for open speech understanding without lip-reading were very low as the implanted CIs were of an early version.

Several decades earlier, Miller and Nicely (1955) used nonsense syllables with added noise for measuring the consonant confusions of five NH adults. In their classical study, a novel method of measuring the transmission of the five speech features voicing, duration, nasality, affrication, and place of articulation was applied. Confusion matrices (CMs) were collapsed with regard to voicing, nasality, affrication, duration, and place, and the percentage of information transmitted for each sub-matrix was calculated.

Aside from the studies described in this dissertation, and the study conducted by Teig et al. (1992), there are only a few investigations of the confusions of Norwegian speech sounds (Ormestad, 1955; Ottem, 1972; Tetzchner, 1975), none with CI users and none published internationally.

In the systematic review and meta-analysis conducted by the author and colleagues (Study 1), an exhaustive selection of previous research was examined, pooled, and synthesized to investigate how well users of multichannel CIs identify consonants and vowels in tests using monosyllabic and bisyllabic nonsense words as stimuli. Included studies had participants with multi-channel CIs. Both consonant and vowel scores were reported. The tests were presented with auditory stimuli only, and scores were reported numerically, with both means and *SD*s. The 47 included articles spanned 27 years, which is a rather low number of articles in such a long period. The main exclusion criteria were: stimuli presented live, consonant and vowel scores not measured with nonsense syllables, and scores not reported numerically with means and *SD*s.

This study provided a baseline for Study 2 in terms of consonant and vowel scores for the pre- and postlingually deaf. It also contained a meta-CM, which was constructed by 17 consonant CMs from the included articles and provided a cross-lingual overview of the most common consonant confusions.

2.1.2 Different approaches for measuring speech perception in CI users

Speech perception is the process by which a person hears, interprets, and understands the sounds of language. Speech perception research explores how listeners recognize speech sounds and use this information to understand spoken language. There are multiple relevant theoretical models, and several academic disciplines involved in the research. Audiology, phonetics, linguistics, electronics, and psychology, all contribute to explaining the phenomenon.

Residual hearing on one or both ears may influence speech perception, especially when using amplification such as a hybrid CI speech processor (one with a built-in HA for stimulation of the residual hearing in the low frequencies). CI users with no residual hearing in either ear will, when tested in quiet, not necessarily perceive speech better with two CIs than with one. However, if tested with added noise, they will typically obtain higher speech perception score with two implants than with one.

The most common audiological test, pure tone audiometry, measures hearing loss by stimulation with sine tones. The test indicates whether, and at what threshold level, the test subject can perceive single frequencies, and provides little information about perception of speech and of speech sounds.

Repetition tests of sentences and real-word monosyllables are widely used in *ear, nose, and throat* (ENT) clinics to measure speech perception in CI users and in other hearing-impaired individuals. It has been shown that monosyllable repetition tests correlate with audiometric thresholds (Stach, 2009, p. 296). Scores on the monosyllable and sentence tests are calculated by counting the numbers of correctly repeated target words and dividing it by the total number of presented words.

The scores on sentence tests are usually higher than the scores on monosyllable tests, and there is more often a ceiling effect on the scores, as a greater range of language skills will influence scores on the sentence repetition tests than scores on the monosyllable repetition tests. Vickers et al. (2009) constructed a conversion table of scores between Bamford, Kowal and Bench (BKB) sentences and Arthur Boothroyd (AB) words in quiet, the two most commonly used standardized speech tests in the United Kingdom for the assessment of CI users. This conversion table showed that the monosyllable word score equivalent for 50% correct on the BKB sentences was 18.5% on the AB test, and 34.5% when the phoneme score was calculated.

Tests of sentences with adaptively added noise, for instance the HINT, have increasingly been applied in clinics of late. A big advantage with these tests is that there is no ceiling effect on the results, as the outcome is a signal-to-noise ratio and not a score. This ratio has been validated with NH individuals. Previously used sentence-in-noise tests had a fixed noise level throughout the test.

Consonant and vowel scores can be calculated by counting the numbers of correctly repeated phonemes in a monosyllable test. The *consonant-vowel nucleus-consonant* (CNC) test developed by Peterson and Lehiste (1962), is widely used for this purpose in English-speaking countries. Ling’s 6-Sound Test (Ling, 1976), which checks the perception and production of the three consonants and three vowels, [s, ʃ, m, ɑː, iː, uː], is language-independent and in use in clinics all around the world. Tests of consonant and vowel perception can also be measured by an NSRT, as in our study. There are many different NSRTs, and they are usually composed by VC and CV combinations. A brief overview of modern clinically used speech perception tests can be found in Article I (Rødvik et al., 2018, p. 1024).

Open- and closed-set test design

Open-set and closed-set test designs are different in many aspects: In a closed-set test design, the responses are measured by a forced-choice task, in which the number of response alternatives is limited by the experimenter. The perception score is adjusted with regard to chance performance. In an open-set test design, the number of response alternatives is limited by the size of the mental lexicon (Clopper et al., 2006). When studies of word recognition tasks began to be common in science in the 1940s and 1950s, open-set testing with word or syllable recognition was usually conducted (documented by Miller et al., 1946). After some years, many scientists started doing closed-set testing, and according to Black (1957), this may be explained by closed-set design being less time consuming and more easily administered and scored than open-set design. Today, speech perception tests with real words and sentences are usually conducted open-set, and consonant and vowel tests with nonsense syllables are usually conducted closed-set (e.g., Rødvik et al., 2018).

Articles I and III elaborate further on the differences between, and on the advantages and disadvantages of open- and closed-set testing of CI users with nonsense syllables (Rødvik et al., 2018; Rødvik et al., 2019a). Today, the most frequently applied method of assessing speech sound perception is closed-set testing, and contemporary studies using open-set designs are difficult to find. An open-set approach was used in a study by Eisenberg et al. (2002), in which two trained audiologists instantly transcribed the speech of the participating children as they repeated nonsense syllables. An open-set approach was also used in studies of the speech sound perception in Finnish CI users (e.g., Välimaa et al., 2002a; 2002b), and in a pilot study by Rødvik (2008).

We chose an open-set design to minimize the opportunity for using inferential and top-down skills, and to create a test situation similar to real life, aiming at optimizing the ecological validity of the test. The participants' speech sound perception was assessed by analysing their repetitions of nonsense syllables, after verifying in advance that they all could spontaneously pronounce the tested speech sounds 100% correctly. In real-life, listeners may experience challenging situations similar to NSRTs when they try to catch an unfamiliar name or are confronted with new vocabulary, and new and difficult words are perceived as nonsense syllables until they become internalized as meaningful units.

2.1.3 Rationale for using nonsense syllables as stimuli

The most common way of measuring consonant and vowel perception is by counting the numbers of correct phonemes in real-word monosyllables. The drawback of this method is that the participants do not have to perceive all the sounds in the word to identify them, as they must when responding to nonsense syllables. In nonsense syllable repetition tasks, inference based on vocabulary, language proficiency, and inferential skills will not be possible.

As long as all the syllables in an NSRT are phonotactically legal and indigenous in the language of the listener, studies using nonsense syllables as stimuli might be conducted with the same test in different languages and compared cross-lingually. A few studies have been conducted to investigate this (e.g., Tyler and Moore, 1992; Pelizzone et al., 1999).

The learning effect in multiple experiments with the same nonsense syllables is very small compared with tests using real-word stimuli (Dubno and Dirks, 1982). It is thus possible to use the same NSRT for repeated examination of speech perception in the same individual to check for progress in auditory skills.

Experiments using nonsense syllables have been shown to evoke fewer associations in the participants and thus reduce between-participants variability in test results compared with experiments using real words (Glaze, 1928).

2.2 Theoretical foundation

2.2.1 Theories of phonological development

Perhaps the most widely known theory that endeavours to explain the speech sound development in NH children from infant to adolescent is by Jakobson (1941). He claimed that children first acquire distinctive features, such as voicing, nasality, and manner of articulation, rather than specific phonemes, and do so in a particular order. When a feature is acquired, it will be reflected in all the phonemes in the child’s phoneme inventory that possess this feature. The acquisition order of the phonemes builds on the principle of maximum contrast (the features that are most prominent are learned first). Figure 2 shows the acquisition of distinctive features according to Jakobson’s hierarchy (Singh and Frank, 1972). The VC contrast comes first, followed by the stop-nasal contrast. Thereafter the distinction between labials and dentals emerges; at this stage, the child can produce the CV syllables /pa/ and /ma/. The next step is being able to distinguish between different vowels.

The basis for Jakobson's theory is that children's speech sound inventory gradually develops to become more and more similar to an adult's and that all children's speech sound development follows the same main steps. Jakobson proposed a theory to which all future researchers in child language development have related. Still, his theory has been widely criticized, especially for its inflexibility, as it does not allow for individual variations.

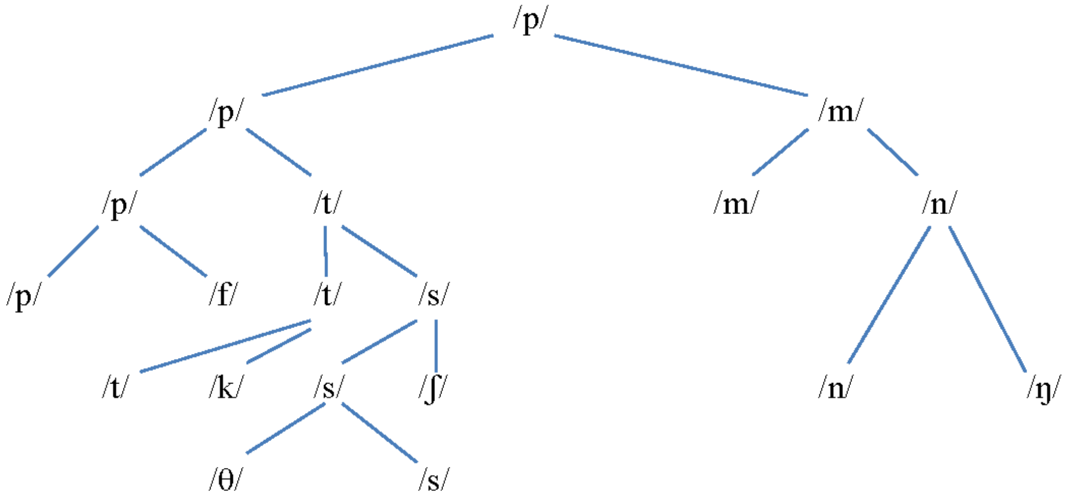


Figure 2. Acquisition of distinctive features, according to Jakobson's hierarchy for distinctive phonological features (Singh and Frank, 1972).

In an alternative model for the speech sound development of children, Waterson (1971; 1976) proposed a prosody-based theoretical model, which builds on the prosodic aspects of speech sound acquisition, such as stress, tone, and intonation, and not merely on the segmental development of speech. In another model, speech sound acquisition is described as a result of simplification processes (Stampe, 1969; Smith, 1973; Ingram, 1974). According to this model, the rules of simplification are gradually reduced as the child's speech becomes more and more similar to adult speech. The rules of simplification are divided into *paradigmatic processes*, in which phonemes are substituted, and *syntagmatic processes*, for instance assimilation, in which pronunciation is influenced by the context of a speech sound.

In a more recent model by Vihman (1993), the *articulatory filter hypothesis* was proposed. This theory suggests that input speech forms that are a rough match to a child's own vocalizations become especially salient to the child. The model was proposed as a response to a paradox reported by Ferguson and Farwell (1975); children seem to be phonologically selective, as they their first words seem to be relatively accurate and related to their existing babble pattern. Vihman (2002) argues that this motor skill is triggered neurologically by "mirror neurons." The construct of an articulatory filter has been supported by clinical evidence in the last few years (e.g., Depaolis et al., 2013; Majorano et al., 2013).

2.2.2 Speech production and perception theories

Many theories, the two most common are *auditory* and *motor* theories, have attempted to explain the mechanisms behind speech production and perception. Auditory theories address the process of speech perception as primarily auditory, with the same hearing mechanism and perceptual processing for any type of sound. In such models the perception of speech is explained on the basis of acoustic cues. The listener simply identifies acoustic patterns and/or features and matches them directly to the learned and stored acoustic-phonetic features of the language.

According to the auditory "*acoustic invariance theory*", listeners abstract the essential acoustic features of an incoming sound to make a decision about its identity. A considerable number of studies have contributed to the elaboration of this theory. Most of them have focused on invariant acoustic properties that can be used to classify stop consonants according to place of articulation (Fant, 1958; Stevens and Blumstein, 1978; Sussman et al., 1991).

The Swedish pioneer in speech research, Gunnar Fant (1958), modelled speech perception as primarily *auditory/sensory*, or "*non-motor*." He maintained that the perceptual and production mechanisms share a common pool of distinctive features but that the listener needs not refer to production to perceive speech. He acknowledged that it would be hard to provide conclusive evidence for either of the two groups of theories. The model in Figure 3 (Fant, 1967) shows his proposal for a model of the connection between speech perception and production that includes linguistic and auditory or acoustic processes. Fant claimed that the motor and sensory functions become more involved from the peripheral to the central stages of the model.

Figure 4 presents another model of the speech production and speech perception processes in humans, and the machine counterparts to the processes appear next to each physiologic process (Rabiner and Juang, 1993, p. 12). The model emphasizes the auditory and acoustic processes in speech perception and production and has been included to show that speech processing can be described in different ways.

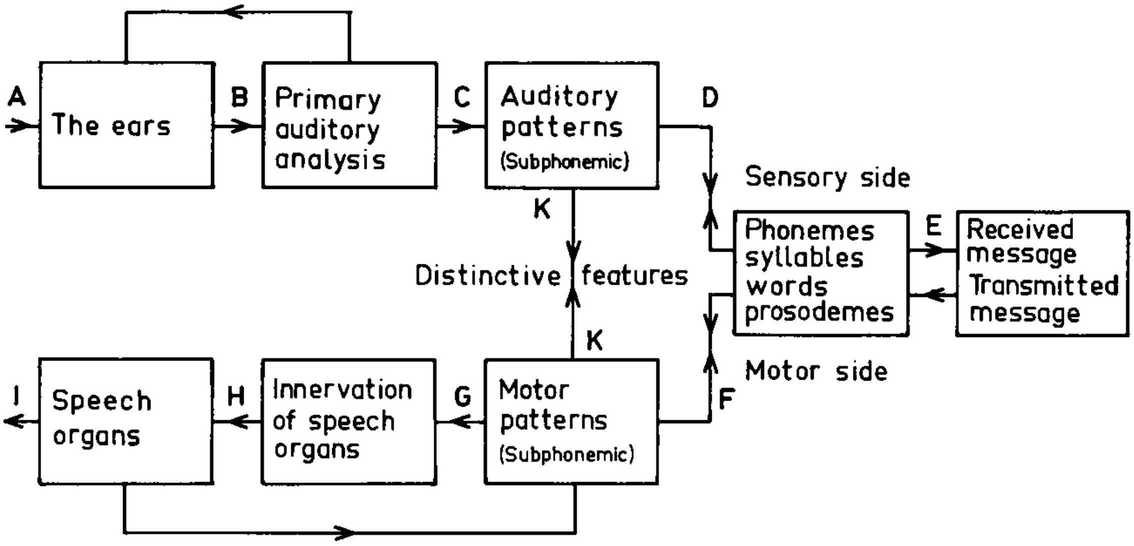


Figure 3. Hypothetical model of brain functions in speech perception and production (Fant, 1967).

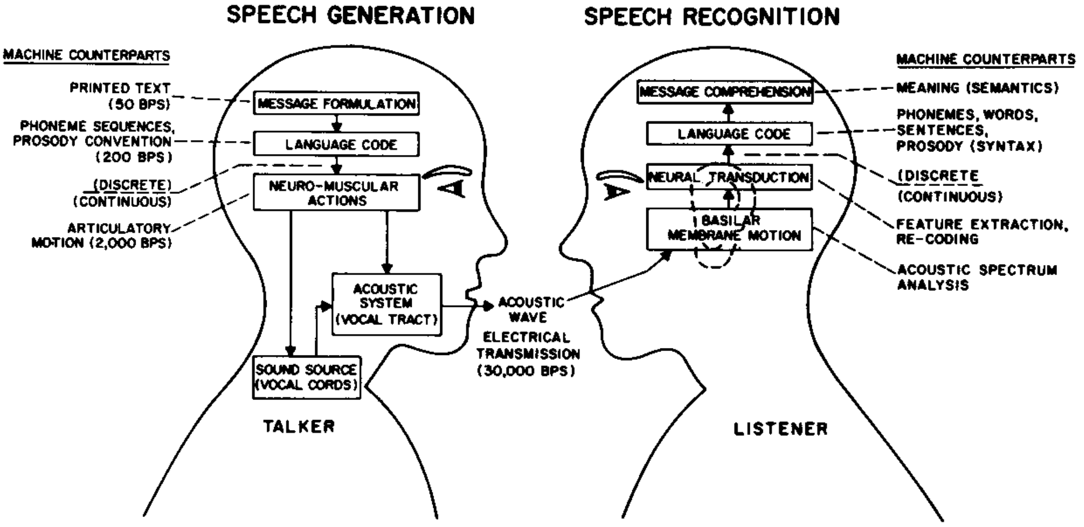


Figure 4. Schematic diagram of processes in speech perception and production (Rabiner and Juang, 1993).

Motor theories is the second of the two large groups of speech perception and production theories. They consist of mechanisms referencing to articulation in the perception process and thus link speech perception to production (Borden et al., 1994, p. 221). One of the most influential motor theories is the "*motor theory of speech perception*," which essentially claims that an individual perceives speech because he or she produces speech; that is, the individual identifies the vocal tract gestures with which speech is pronounced rather than the sound patterns that speech generates (Liberman et al., 1967; Liberman and Mattingly, 1985).

Theories for speech perception and production can also be categorized as *bottom-up* and *top-down* theories (Samuel, 2011). Bottom-up theories presume that all the information necessary for the recognition of speech sounds is contained within the acoustic signal, and therefore the listener does not need to involve linguistic and cognitive processes in decoding sounds. By contrast, top-down theories emphasize higher level linguistic and cognitive operations as crucial to the identification and analysis of sounds. Most theories are neither completely top-down nor bottom-up, but place more or less weight on acoustic versus linguistic-cognitive contributions to speech perception.

A final mention of relevant theories of speech production and perception is the *phonological loop* in the model of *working memory* proposed by Baddeley and Hitch (1974). The loop comprises a phonological store (or acoustic or articulatory store) within which memory traces fade after 2 s unless an articulatory control process refreshes them by *subvocal rehearsal*. The phonological store thus acts as an "inner ear" by remembering speech sounds in their temporal order, while the articulatory control process acts as an "inner voice" by repeating the series of words on a loop to prevent them from decay. For example, if one tries to remember a telephone number in the few moments before dialling by repeating it over and over, this would take place in the phonological loop.

2.2.3 Rationale for the design of the NSRT by idealization of a theoretical model

In general, the fit between a theoretical model and the real world is based on an evaluation of a theoretical hypothesis as true or false (Giere, 1991). The design of a test for assessing the perception of speech sounds by the repetition of nonsense syllables must maximize the degree to which the test measures the participants' actual auditory skills. By eliminating, or "controlling away," as many as possible of the other processes involved in speech production and perception, such as language proficiency, inferential skills, and vocabulary, the auditory skills are expected to reveal its capacity (Cartwright, 2009).

Since the focus of our study is auditory skills in CI users, we decided that the auditory model of Fant (1967) would be most suitable. This model emphasizes auditory and acoustic processes in speech perception and production, as opposed to the motor theories. However, recent brain scanning experiments have revealed that speech motor functions are to some extent activated simultaneously with the auditory cognitive process (Liberman and Mattingly, 1985). Thus, a more correct, but more complicated, model of speech perception and production would be a mixture of the two.

Idealization

An idealization is a deliberate simplification of something complicated with the objective of making it more tractable (Frigg and Hartmann, 2012). The idealization of a real-world phenomenon may be achieved by formulating a theory that includes only a few parameters from the phenomenon. The unselected parameters usually do have some influence, so the theory does not characterize the actual phenomenon but rather the contribution to it by the selected abstract parameters. Galilean idealization is the gold standard in natural sciences, as it falsifies nature by simplification (e.g., vacuum or a frictionless plane), and was applied in the design of the NSRT.

Some negative consequences of idealization are less accuracy, fewer details, and possible over-simplification. The gains from idealization outweigh the drawbacks and include a better overview of the phenomenon, easier manipulability, and easier access to complicated phenomena.

In the present project, the model was idealized to measure the auditory skills as exclusively as possible and should explain whether a participant who chooses to repeat a nonsense syllable in an ideal test environment does so correctly or incorrectly. The model was tentatively based on the following theoretical hypothesis: “Under idealized test conditions, the repetition of one- or two-syllabic nonsense words solely measures the auditory skills of the participants.”

If this was a causal hypothesis, presenting a nonsense syllable could be framed as a positive causal factor for the production of that syllable. However, the participant might decide to produce something quite different from that nonsense syllable. Thus, there cannot be a causal relationship between perception and production, as production is controlled by will rather than cause.

In the following, the model is sketched before and after idealization. All the processes involved in the theoretical model are assumed to be static. A real-life model (if such a model could be imagined) would include a mixture of causal and static processes.

Figure 5 shows the theoretical model before idealization. Even in this first model, considerable idealization from the real-world phenomenon has already been done. The model specifies a group of cognitive functions, a group of physiological factors, two environmental factors (test environment and social environment), and personal well-being. These factors are selected to represent the main mechanisms involved in speech perception and speech production and represent a subjective transformation from the real world to a theoretical model.

The idealization process took place in four steps, described below. Figure 6 shows the model after idealization.

Step 1: Eliminating model variables by the inclusion and exclusion criteria

- Participants with mental challenges were excluded.
- Participants with impaired phonation or articulation were excluded.
- Only speech sounds common to all Norwegian dialects were included, to avoid difficulties in repeating speech sounds unknown to the participants.

Step 2: Eliminating model variables by the test design

- Test words with only one or two syllables were chosen to eliminate the influence of possible low short-term memory capacity (Baddeley et al., 1975; Gathercole et al., 1994).
- Nonsense syllable stimuli were used instead of real-word stimuli to ensure that inferential skills, vocabulary, and general language proficiency played a minor role. NB: No syllables in the NSRT have any lexical meaning in Norwegian.
- The test words were presented auditorily with no visual cues, to avoid perception by vision to influence the results.

Step 3: Creating an ideal test environment

- Degraded speech signal due to poor acoustics was avoided, as the testing was performed in an anechoic chamber.

- The effect of low attention span and exhaustion was minimized, as the participants took frequent breaks during testing.
- Poor quality of the speech signal was avoided, as high-quality recordings of both the test and the participants' repetitions were secured.

Step 4: Simplifying the responses in the analysis phase

- The repetitions were categorized as either a nonsense syllable included in the test or an "unclassified" speech sound. The unclassified speech sounds were not further analysed.

As shown in Figure 6, test design could only partly eliminate the influence of cognitive functions. The theoretical model is constructed for hearing-impaired adults and children. In a model for NH adults and children, the box that lists hearing would be removed.

Hearing impairment is an obstacle to the perception of acoustic cues. Depending on the nature of the hearing loss, many acoustic cues will not be perceived, and speech perception will be poorer. It was expected that participants with CIs would achieve lower scores than NH listeners, which could be explained mainly by their impaired hearing, as other factors assumed to influence performance in this task were minimized. We expected ceiling effects in the NH listeners' scores, as their hearing is the norm for the test.

An idealized model for hearing-impaired participants could be imagined, in which the influence of all the cognitive processes had been eliminated. However, this would imply that the only process involved in the repetition of nonsense syllables was their auditory skills, which would violate all existing theories on speech perception. Thus, it can already be stated that the theoretical hypothesis is deficient, or over-simplified, because the effect of cognitive functions cannot be eliminated completely (but can be minimized). An improved formulation of the hypothesis is therefore: Under idealized test conditions, repetition of nonsense syllables by hearing-impaired participants measures their auditory skills as exclusively as possible, with minimal influence of cognitive functions.

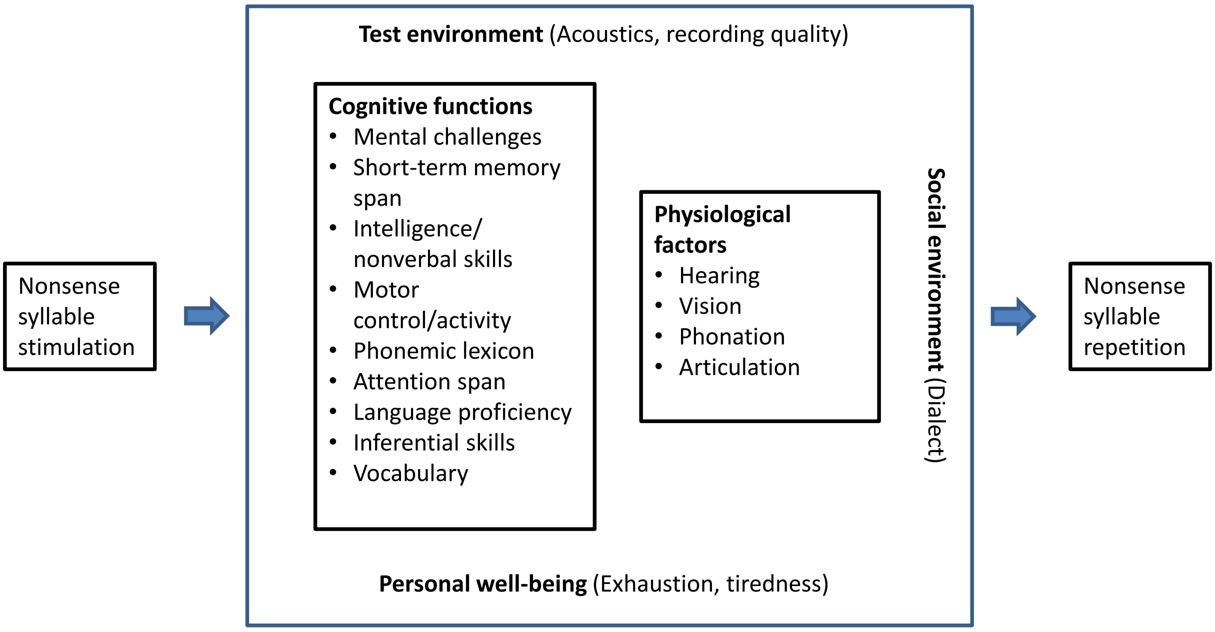


Figure 5. Theoretical model before idealization showing factors involved in speech perception and production in hearing-impaired adults and children.

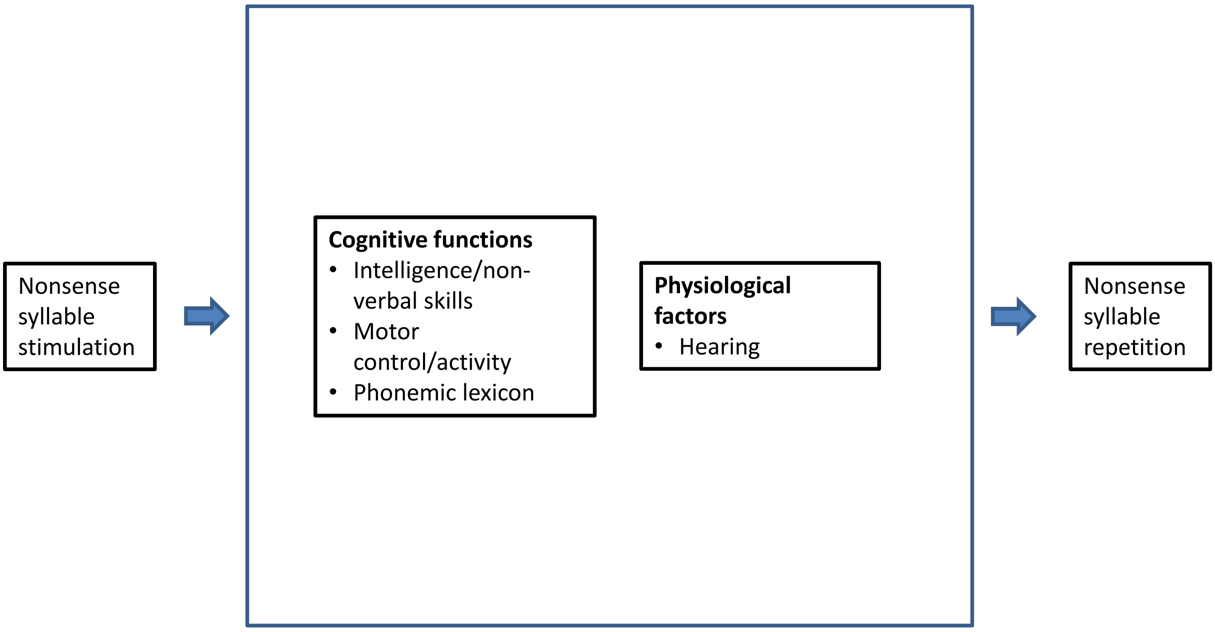


Figure 6. Theoretical model after idealization; showing factors involved in speech perception and production in hearing impaired adults and children.

3 Methodological reflections

3.1 Outline of the connection between the articles

This thesis describes a three-step PhD project composed of a systematic review and meta-analysis (Study 1), a cross-sectional study including adults with CIs and a reference group of NH adults (Study 2), and a cross-sectional study including children and adolescents with CIs and two reference groups of NH 6- and 13-year-olds (Study 3).

In Study 1, an exhaustive systematic review and meta-analysis on previous research was conducted. The motivation behind this study was the absence of previous systematic reviews in studies reporting consonant and vowel perception in CI users measured with nonsense syllable identification and the desire to obtain a general baseline of consonant and vowel scores in CI users assessed by mono- and bisyllabic nonsense syllable identification tests.

The outcomes of Study 1 and from a pilot study (Rødvik, 2008) was used as a basis for the development of the NSRT that was used in Studies 2 and 3 for measuring consonant and vowel repetition scores in Norwegian-speaking adults and children with CIs, for study design, analysis, and discussion.

Article I reports results from Study 1, Article II reports results from Study 2, and Article III reports results from Study 3. In Article II, the outcomes of adult CI users are reported and discussed for the whole sample and for subgroups of FS and non-FS strategy users. In Article III, the outcomes of children and adolescents with CIs are reported and discussed for the whole sample and for subgroups of pre- and postlingually deaf children. Article II was written before Article III, and Article III was published in August 2019.

An overview of the materials and methods used in the three articles is shown in Table 1.

Table 1

Overview of the materials and methods used in the articles in the dissertation

	Article I	Article II	Article III
Title	Consonant and vowel identification in cochlear implant users, measured by nonsense words: A systematic review and meta-analysis	Consonant and vowel confusions in well-performing adult cochlear implant users, measured by a nonsense syllable repetition test	Consonant and vowel confusions in well-performing children and adolescents with cochlear implants, measured by a nonsense syllable repetition test
Design	Systematic review and meta-analysis	Experimental and cross-sectional	Experimental and cross-sectional
Sample size	50 studies reported in 47 articles, encompassing 647 participants with CIs	39 adult CI users 20 NH adult listeners	36 children and adolescents with CIs 12 NH 13-year-old listeners 17 NH 6-years-old listeners
Test instruments	Mono- and bisyllabic VCV and CVC nonsense words	Mono- and bisyllabic VCV and CVC nonsense words	Mono- and bisyllabic VCV and CVC nonsense words
Procedure	Repetition of recorded nonsense words	Verbal repetition of recorded nonsense words	Verbal repetition of recorded nonsense words
Status	Published April 2018 in Journal of Speech, Language, and Hearing Research	Submitted	Published August 2019 in Frontiers in Psychology

3.2 Test instruments

3.2.1 The NSRT

The testing in Studies 2 and 3 (Articles II and III) was conducted with an NSRT that contains the 16 consonants common for all Norwegian dialects [p, t, k, s, ʃ, f, h, b, d, ɡ, ʝ, v, m, n, ŋ, l], presented in a 2-syllable VCV context with the three main cardinal vowels in Norwegian, /ɑː, iː, uː/. In addition, the test contains 11 consonants that are used in some local Norwegian dialects. The latter were not included in the analyses. However, the inclusion of these dialectal speech sounds made the test very perceptually open, as the participants had no prior knowledge about which or how many speech sounds would be presented to them. The NSRT makes use of a Standard East Norwegian tone 2 (Kristoffersen, 2000, p. 242) throughout the test of consonants. The test design collects information about consonant confusions in three different vowel contexts, and about how formant transitions influence perception in each context.

The NSRT also contains the nine Norwegian long vowels, [ɑː, eː, iː, uː, ʉː, yː, æː, øː, ɔː], presented in a CVC context with /b/ as the chosen consonant. None of the CVC or VCV combinations presented in the test have any lexical meaning in Norwegian. Table 2 shows the included consonants placed in the *International Phonetic Alphabet* (IPA) chart, showing manner and place of articulation, and voicing, and Table 3 presents a complete list of the included nonsense words in an IPA notation. Figure 7 displays a simplified vowel chart with all the nine long Norwegian vowels used in the NSRT, plotted according to the two lowest formant frequencies, F1 and F2 (Kristoffersen, 2000, p. 17, modified).

Vowel-consonant-vowel (VCV) and consonant-vowel-consonant (CVC) nonsense syllables have frequently been used in the last century, in working memory tests and learning experiments for measuring consonant and vowel perception. The NSRT-syllables comply with the phonotactic rules of Norwegian. Limiting the number of syllables to one or two ensures that the working memory capacity will not be strained (Gathercole et al., 1994). Consonant scores of an NSRT are rarely at ceiling for CI users, as vowel scores may be (e.g., Rødvik et al., 2018). By using verbal repetitions of recorded nonsense syllables, detailed information regarding speech perception and listening capacity for certain acoustic properties will be provided.

Table 2

Simplified IPA chart displaying the speech sounds used in NSRT-C

	Place of articulation													
	Bilabial		**Labiodental**		**Dental**		**Postalveolar**		**Palatal**		**Velar**		**Glottal**	
Manner of articulation	**U**	**V**	**U**	**V**	**U**	**V**	**U**	**V**	**U**	**V**	**U**	**V**	**U**	**V**
Stops	[p]	[b]			[t]	[d]					[k]	[ɡ]		
Fricatives			[f]		[s]		[ʃ]			[j]			[h]	
Nasals		[m]				[n]						[ŋ]		
Lateral						[l]								

Note. U = unvoiced, V = voiced.

3.2.2 Inclusion of consonants and vowels in the NSRT

Consonants and vowels can be classified by perceptual units such as the formants in vowels and voiced consonants, formant transitions between consonants and vowels, aspiration, duration, and nasality. Their acoustic properties can be characterized by time-varying spectral patterns called acoustic cues. The cues are resolved when a speech wave propagates on the basilar membrane. The relationship between acoustic cues and perceptual units is in many cases difficult to find and is a common problem in speech perception research. The inclusion of phonetic features of consonants and vowels to be examined in Study 2 is motivated below.

Consonants

Consonants are a heterogeneous group of speech sounds. They are classified by manner of articulation (e.g., stopping, frication, nasality, and laterality), place of articulation (e.g., labial, dental, velar, alveolar, and glottal), and voicing (see Table 2). Earlier research has shown that acoustically similar consonants are most likely to be confused (Fant, 1973), and in general, implant users more frequently confuse consonants that are pronounced in the same manner but with a constriction in different places in the mouth cavity, than consonants that are pronounced in different manners with the tongue in the same place.

Table 3

VCV and CVC nonsense syllables included in the NSRT

No.	aCa Words	iCi Words	uCu Words	bVb Words
1	['ɑːbɑ]	['iːbi]	['uːbu]	[bɑːb]
2	['ɑːdɑ]	['iːdi]	['uːdu]	[beːb]
3	['ɑːfɑ]	['iːfi]	['uːfu]	[biːb]
4	['ɑːɡɑ]	['iːɡi]	['uːɡu]	[buːb]
5	['ɑːhɑ]	['iːhi]	['uːhu]	[bʉːb]
6	['ɑːjɑ]	['iːji]	['uːju]	[byːb]
7	['ɑːkɑ]	['iːki]	['uːku]	[bæːb]
8	['ɑːlɑ]	['iːli]	['uːlu]	[bøːb]
9	['ɑːmɑ]	['iːmi]	['uːmu]	[bɔːb]
10	['ɑːnɑ]	['iːni]	['uːnu]	—
11	['ɑːpɑ]	['iːpi]	['uːpu]	—
12	['ɑːsɑ]	['iːsi]	['uːsu]	—
13	['ɑːtɑ]	['iːti]	['uːtu]	—
14	['ɑːvɑ]	['iːvi]	['uːvu]	—
15	['ɑːʃɑ]	['iːʃi]	['uːʃu]	—
16	['ɑːŋɑ]	['iːŋi]	['uːŋu]	—

Note. aCa = a-consonant-a; iCi = i-consonant-i; uCu = u-consonant-u; bVb = b-vowel-b; CVC = consonant-vowel-consonant; NSRT = nonsense syllable repetition test; VCV = vowel-consonant-vowel. All the speech sounds in the test are prevalent in all Norwegian dialects.

All consonants can be classified as either voiced or unvoiced and hence, voicing was contrasted with nonvoicing in the analyses. Each voiced consonant differs substantially acoustically from its unvoiced counterpart by the presence of the fundamental frequency (F0), which is generated by the vibrations of the vocal cords. Voiced consonants have higher intensity than their unvoiced counterparts.

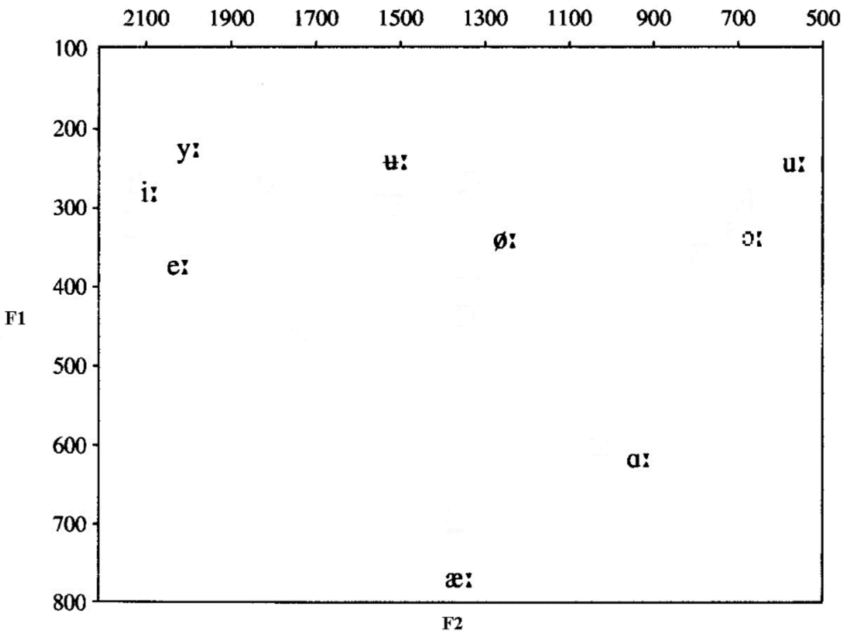


Figure 7. The Norwegian vowel system. All the Norwegian long vowels are plotted according to their two first formant frequencies, F1 and F2 (Kristoffersen, 2000, p. 17, modified).

Stops and fricatives have often been shown to be mutually confused (e.g., van Wieringen and Wouters, 1999; Munson et al., 2003). Stops were therefore contrasted with fricatives in the analyses. Stops are produced by blocking the vocal tract to cease all airflow. The acoustic output of the stop is the sudden release of the blocking. The subgroups of voiced and unvoiced stops can be distinguished by the presence of a silent gap in the unvoiced stops (Lisker, 1981). Norwegian unvoiced stops are characterized by a strong aspiration; a distinct final auditory breathy pause, which is created by closing the vocal cords from a maximally spread position leading up to the explosion of the plosive, lasting longer than the occluded phase of the stop articulation (Kristoffersen, 2000). Stops can be difficult to identify, as they are very short. Unvoiced stops have little acoustic energy. In identifying stops, CI users also rely considerably on the spectral properties of the surrounding vowels, such as locus and length of the formant transitions, spectral height and steepness, and the time between air release and vocal-cord vibration; voice onset time (Välimaa et al., 2002b). For fricatives, the occlusion is partial and the airflow in the vocal tract is constricted, not blocked.

Nasality is an important speech feature in most languages. The Norwegian nasals are voiced and produced with the air flowing through the nose and with the oral cavity blocked in different places. The overall intensity level of nasals is lower than that of vowels. They can similarly be characterized by their formant frequencies; F1 is low in frequency for all nasals and F2 is in general at the frequency level of F3 for vowels. In addition, nasals can be characterized by a *nasal murmur*, the acoustic output associated with nasal radiation of sound energy, having a spectrum dominated by the prominence of low frequencies around 250 Hz, but also with resonances of higher frequencies (between 800 and 2000 Hz). The F2 transitions are important in distinguishing nasals of different places of articulation. In the analyses, the three nasal consonants, [m, n, ŋ], were contrasted with the nonnasals.

The lateral [l] is produced by an airstream passing along the sides of the tongue, and it is blocked by the tip of the tongue from the middle of the mouth. Acoustically, the formant pattern of [l] is similar to the vowels. /l/ is the consonant phoneme in our sample with the largest acoustic space, as F2 can vary substantially with the tongue position more or less retracted. The F1s of [l] and [n] are similar, around 250 Hz, but the F2s are located around 1,200 Hz and 2,500 Hz, respectively. Moreover, [n] has more energy in the low frequencies than [l]. The lateral [l] was contrasted with the nasals in the analyses.

Vowels

Vowels normally constitute the core of the syllable in spoken language and are usually more easily perceived than the consonants (e.g., Rødvik et al., 2018), due to their combination of high intensity and long duration. Vowels are only defined by place of articulation, and in Norwegian, the F1 and F2 relationship indicates tongue placement, mouth opening, and lip rounding (see the vowel chart in Figure 7). Front vowels have the body of the tongue pushed forward in the mouth and somewhat arched. Back vowels are produced with the back part of the tongue raised toward velum. High vowels are pronounced with the tongue arched toward the roof of the mouth. Low vowels are produced with the tongue relatively flat and low in the mouth and with the mouth open a little wider than for high vowels. Midvowels have a tongue position between that of high and low vowels. High and front vowels usually have low F1 and high F2. Low and back vowels usually have high F1 and low F2 (see Figure 7). In the present study, vowels have been regarded as one group, and contrasted with consonants.

3.2.3 Basic features of the Norwegian language

The NSRT in our study is based on Norwegian, which is a Northern Germanic language, belonging to the Scandinavian language group. There is no official common Norwegian pronunciation norm, as oral Norwegian is a collection of dialects, and Norwegians normally speak the dialect of their native region. This makes it challenging to plan, execute, and evaluate speech sound repetition tests for CI users, as the implantees are normally and randomly distributed across all dialect divides. Most of the adult, postlingually deaf CI users have been exposed to speech sounds of other dialects before their deafness. Nevertheless, recognizing them, identifying them, and even repeating them will, supposedly, present more of a challenge to them than to NH listeners.

Norwegian has two lexical tones (except for certain few dialects), which span across two syllables and are used as a distinguishing lexical factor. The tones' melodies are indigenous to each dialect and are recognized as a dominant and typical prosodic element of the dialect, distinguishing it from other dialects.

Norwegian has a semi-transparent orthography, meaning that there is not a consistent one-to-one correspondence between letters and speech sounds, like for instance in Finnish, but a much more consistent relationship between letters and speech sounds than in English (Elley et al., 1992).

3.3 Article I: Consonant and vowel identification in cochlear implant users measured by nonsense words: A systematic review and meta-analysis

3.3.1 Abstract

In Article I (Rødvik et al., 2018), studies that measured consonant and vowel identification in CI users by nonsense syllable stimulation were pooled.

The aims of this study are:

1) to establish a baseline of the vowel and consonant identification scores in prelingually and postlingually deaf users of multichannel CIs tested with CVC and VCV nonsense syllables

and to study how the typical vowel and consonant identification scores differ between prelingually and postlingually deaf implantees,

2) to investigate which consonants and vowels are most frequently confused by CI users, and which consonants and vowels are most frequently identified correctly,

3) to investigate to what extent age at implantation, duration of implant use, and real-word monosyllable score are associated with variations in consonant and vowel identification performance in nonsense syllable tasks for prelingually and postlingually deaf CI users.

Forty-seven articles covering 50 studies with 647 participants, of which 581 were postlingually deaf and 66 prelingually deaf, met the inclusion criteria. The mean performance on vowel identification tasks for the postlingually deaf CI users was 77% ($n = 5$), which was higher than the mean performance for the prelingually deaf CI users (68%; $n = 1$). The mean performance on consonant identification tasks for the postlingually deaf CI users was higher (58%; $n = 44$) than for the prelingually deaf CI users (47%; $n = 6$). The most common consonant confusions occurred between those with the same voicing and manner of articulation (/k/ as /t/, /m/ as /n/, and /p/ as /t/).

There were no statistically significant differences between the mean performance scores on consonant identification tasks for pre- and postlingually deaf CI users. The consonants that were incorrectly identified were typically confused with other consonants with the same acoustic properties: voicing, duration, nasality, and silent gaps. A univariate meta-regression model, although not statistically significant, indicated that duration of implant use in postlingually deaf adults predicts a substantial portion of their consonant identification ability.

3.3.2 Validity and reliability

In Study 1, the findings of other researchers' primary studies were pooled, quantified, and analysed in a systematic review and meta-analysis. The primary studies generally addressed issues concerning validity and ethics, leaving heterogeneity, quality, and publication bias to be discussed in Article I. The research questions in Study 1 were constructed independently, in accordance with the 27-item checklist in the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) statement (Moher et al., 2009).

At initiation, Study 1 was registered in PROSPERO, an international prospective register of systematic reviews:
(http://www.crd.york.ac.uk/prospero/display_record.asp?ID=CRD42014015141).

The data from the included articles were coded broadly, but not all coded data were discussed or analysed (e.g., aetiology, stimulation level, number of speech sounds included in the primary studies, and implant models). The impact of implant model and of stimulation strategies on the consonant and vowel scores was initially part of the factors we wanted to look into in the systematic review. However, due to missing reports of implant models, implant model for each participant, and stimulation strategies in the primary studies, there were insufficient data for inclusion in the analyses. The reports on sentence repetition tests were also too sparse to be included in the analyses. Several tests of different degrees of difficulty were used, making it challenging to pool the data.

We showed in the systematic review article that there was a low and not significant correlation between year of publication and consonant score in the included primary studies. This indicated that it was reasonable to include articles as far back as 1989. The effect of the variability on the scores between the participants was thus probably larger than improvements in implant technology over the years, duration of deafness prior to implantation (which has probably decreased over the years), age at onset of deafness (which has been lowered over the years), and age at implantation (which has decreased over the years for the prelingually deaf children). The large variability between CI users has been shown in e.g., Valimaa and Sorri (2000); Dowell et al. (2002); Rotteveel et al. (2010).

3.4 Article II: Consonant and vowel confusions in well-performing adult cochlear implant users, measured by a nonsense syllable repetition test

Article II reports and discusses results from Study 2. The outcomes of Study 1 and of the pilot study (Rødvik, 2008) were used as a basis for the development of study design, analysis, and discussion.

3.4.1 Pilot study

The pilot study had five adult, postlingually deaf CI users. Speech sound perception was investigated by using the nonsense syllables a-consonant-a (aCa), i-consonant-i (iCi), and b-vowel-b (bVb) as stimuli. Results from the pilot study include the following:

- Voiced consonants were rarely confused with unvoiced consonants.
- The consonants were more frequently confused in the iCi context than in the aCa context.
- For the consonants, manner of articulation was rarely confused, and place of articulation was often confused.
- The mean correct consonant repetition score was 48% (*SD* = 32%).
- The mean correct vowel repetition score was 71% (*SD* = 25%).

Based on the pilot study (Rødvik, 2008), the statistical power for Studies 2 and 3 were calculated and an appropriate sample size for the large-scale studies was predicted. Optimization of the design of Studies 2 and 3 was then performed regarding recruitment of participants, optimization of test environment, sound recordings, and inclusion and exclusion criteria.

3.4.2 Abstract

In Article II (Rødvik et al., 2019a), consonant and vowel perception in Norwegian-speaking adults was measured by an NSRT. The participants consisted of a convenience sample of 39 adults with CIs and a reference group of 20 NH adults. Verbal repetitions of consonants and vowels in mono- and bisyllabic contexts were recorded. The recordings were transcribed by two experienced phoneticians and confusions of vowels and consonants were registered.

Due to expectations of large differences between the scores on the outcomes of the groups of CI users and of NH listeners, we had no research questions regarding differences between the scores of the two groups, and statistical comparisons were performed only descriptively.

The objective of this article was to use an NSRT to investigate, in detail, the properties of speech sound confusions in adult CI users, such as the influence of voicing and nasality on perception, and to investigate how the users of Med-El's fine structure (FS) stimulation strategies perceive consonant features compared to users of non-FS stimulation strategies from Cochlear and Med-El.

For the CI users, the mean score on the NSRT was significantly lower than the mean score on the real-word monosyllable test (62% [*SD* = 13%] versus 73% [*SD* = 11%]). Hence, the NSRT appeared to reveal more speech sound misperceptions than the real-word monosyllable test did.

Important results:

- Voiced stops were often repeated as unvoiced stops, whereas unvoiced stops were never repeated as voiced stops.
- Consonants were confused more often than vowels (57% [*SD* = 14%] versus 72% [*SD* = 17%] correct), and voiced consonants were confused more often than unvoiced consonants (53% [*SD* = 15%] versus 63% [*SD* = 16%] correct).
- The nasals were confused with other nasals in one third of the cases and repeated correctly in only one third of the cases.
- The subgroup comparison showed that the perception scores of nasals versus nonnasals, nasals versus the lateral [l], and stops versus fricatives were significantly lower for the five CI users employing FS strategies than for a matched sample of 5 CI users employing non-FS strategies. The perception of voicing was significantly lower for these non-FS strategy users than for the FS strategy users.

The study revealed a general devoicing bias for the stops and a high confusion rate of nasals with other nasals. The subgroup comparison of small samples of users of FS and non-FS stimulation strategies suggests that more research is needed in order to improve the coding of the low-frequency information in the speech signal.

3.4.3 Validity and reliability

Validity is a property of inferences, and the relevance of various aspects of validity depends on the types of inferences drawn, not on the kinds of data used as a basis for the inferences (Kleven, 2008). Shadish et al. (2002) use the term to refer to the approximate truth of an inference. In their validity system, which often serves as a methodological frame of reference for the evaluation of a scientific study in modern quantitative methodology (Lund, 2002), there are four types of validity: construct, statistical conclusion, internal, and external. The validity of the inferences of the present study will be evaluated according to these four terms in the following paragraphs.

Construct validity

Construct validity measures to what extent the constructs of theoretical interest are successfully operationalized in the research. The two major groups of threats to construct validity are traditionally called *systematic* and *random* measurement errors. Systematic measurement errors, which include construct underrepresentation and construct irrelevance, require the most attention. Random measurement errors tend to even out over the long term (Kleven, 2008). The larger the study sample, the smaller the measurement error.

In the present study, the following minimized the measurement errors using experiences from the pilot study (Rødvik, 2008):

- Participants' pronunciation was checked before testing, and only those with 100% correct pronunciation of the tested speech sounds were included.
- The recorded nonsense syllables were randomized before testing, so that the transcribers would not learn their order.
- The speech processor condition was examined before testing. The speech processor settings preferred by each individual participant were used.
- In participants with an HA on the contralateral ear, we required that his/her speech perception with the HA alone be 40% poorer than the speech perception with the CI alone. This was measured by testing each ear seperately with a real-word monosyllable test.
- Only participants with a real-word monosyllable score above 50% were included in the study.
- Reference groups of NH individuals were included to verify our assumption that they would score at ceiling.

The following factors may have compromised construct validity:

1. The participants' ability to maintain concentration during testing varied, as the test material was quite extensive. The participants were supposed to repeat 90 nonsense syllables played sequentially. They probably repeated the nonsense syllables less accurately near the end of the lists than in the beginning. However, due to the randomization of the lists, the inaccuracy was not consistently related to certain syllables. Therefore, this represents random error, which is present in all similar tests.

2. The quality of the recorded nonsense syllable repetitions may have influenced the transcriptions and thereby the scoring reliability, although the background noise was minimal in the anechoic chamber. The sound level of the nonsense syllables was balanced according to the sound level of natural speech. If the participants were restless and frequently moved their heads during testing, the sound quality would vary, which would introduce a random error.

3. Background variables, such as age, sex, type of implant, number of implants (one or two), choice of ear, number of active channels in the CI, residual hearing, duration of implant use, or time since last sound programming session, may all have been relevant issues. We partially investigated the impact of these variables on the results, and age at implantation had a moderate and statistically significant correlation with the *nonsense syllable repetition score* (NSRS). The correlations between the NSRS and the other variables were low and non-significant. Ideally, the group should be matched for all these variables, which would minimize their possible impact on the results. However, this was not feasible in our study, as the pool of eligible participants was insufficient in our convenience sample.

Statistical validity

We performed an analysis of statistical power based on the consonant and vowel scores from the pilot study, to estimate how many participants were needed to decide with statistical significance whether some consonant confusions were more prevalent than others. The total number of participants was shown to be sufficient for statistically significant conclusions.

External validity

The important question regarding external validity is the issue of generalization, or transferability. External validity is strengthened by having as many participants as possible and choosing them to be as representative as possible with regard to the target population. All factors that make it difficult or impossible to draw such generalizations are threats to external validity.

For the results to be generalizable, the participants should represent a random sample of the population of adult CI users. Article II states that the participants were selected in order of consecutive appearance at the clinic and by specific inclusion criteria. The convenience sample selection is thus a threat to external validity.

The study was statistically valid, securing external validity in this respect. Moreover, as only speech sounds that exist in all Norwegian dialects, were included, the results could be generalized to all CI users with Norwegian as their native language.

The demographic factors may also have been a threat to external validity, as they were not evenly distributed in the groups of study participants. These included sex, age, implant type, number of implants (one or two), modality (one or two implanted ears, or HA on one ear and CI on the other), speech processor, speech processing strategy, number of active channels, and duration of implant use.

Reliability

Reliability is often discussed together with validity. Validity is a property of inferences in a study, and reliability indicates whether the tools measure the same if the test is performed several times. The measurement tools in a study must be reliable for inferences to be valid. If the reliability of a test or measurement is poor, this will be a threat to statistical power, and thus to statistical validity, and also to construct validity (Lund, 2002). In the validity system by Shadish et al. (2002), poor test reliability is defined as a threat to statistical validity only.

Inter-rater reliability was measured by the *point by point agreement*, which is a proportion defined by: percentage agreement/(percentage agreement + percentage disagreement). The point by point agreement between the two transcribers was 89.8% (*SD* = 7.3%; range: 68.4–100%), indicating acceptable reliability (Shriberg et al., 2010).

The internal consistency in the NSRT was calculated by Cronbach's alpha (α), which was 0.824 for the adult CI users. This value is commonly regarded as "moderate to high level" (Murphy and Davidshofer, 2001, p. 142).

The sound quality may have varied slightly from word to word due to movements of the participants' heads relative to the microphone, introducing problems with the reliability of the recordings.

Part of the NSRT was applied in a pilot study with five adult CI users (Rødvik, 2008), and consonant and vowel scores were calculated as 49% and 71%, respectively. In the systematic review and meta-analysis, mean consonant and vowel scores were calculated from 50 pooled studies (Rødvik et al., 2018). The actual outcomes of the NSRT for consonants and vowels (57% and 72%, respectively) were almost similar to the results of this meta-analysis (56% and 72%, respectively) and also close to what was found in the pilot study. Thus, the NSRT seems to possess high reliability.

In Study 2, the reliability was also checked using groups of NH listeners, presenting all the consonants in three vowel contexts, and having two independent transcribers for the recorded repetitions. The NH listeners scored close to ceiling on the NSRT, as expected.

3.5 Article III: Consonant and vowel confusions in well-performing children and adolescents with CIs, measured by a nonsense syllable repetition test

Article III reports and discusses results from Study 3. The outcomes of Study 1 and of the pilot study (Rødvik, 2008) were used as a basis for the development of study design, analysis, and discussion.

3.5.1 Abstract

In Article III an NSRT was chosen to measure the consonant and vowel perception scores of Norwegian-speaking children and adolescents with CIs. The participants consisted of 36 children with CIs and reference groups of 17 NH six-year-olds and 12 NH 13-year-olds. Sound recordings were obtained of the repetitions of 16 consonants in three vowel contexts, and nine vowels in the /b/-context. Confusions of vowels and consonants, and speech sound features were registered for the groups of participants.

The main aim of this study was to assess the effectiveness of CIs by obtaining a measure of the speech sound confusions in children with CIs, using nonsense syllables. The study also aimed to investigate how perception of speech features was influenced by pre- and postlingual deafness.

Unvoiced consonants were repeated correctly in 77% (*SD* = 10%) of the cases and voiced consonants were repeated correctly in 64% (*SD* = 11%) of the cases. The difference between the means was statistically significant (13%, $p < 0.001$). The mean vowel repetition score was

85% (*SD* = 11%). There were no statistically significant differences between the pre- and postlingually deaf neither for the distinction of voicing and nonvoicing, nor the distinction of nasality and non-nasality, nor the distinction of stops and fricatives.

The children and adolescents with CIs obtained scores close to ceiling on vowels and real-word monosyllables, but their perception was substantially lower for voiced consonants. This may partly be related to limitations in the CI technology for the transmission of low-frequency sounds, such as insertion depth and stimulation mode.

3.5.2 Validity and reliability

Validity

The validity of the inferences in the present study is similar to Study 2. This was evaluated above (section 3.4) according to construct validity, statistical validity, internal and external validity.

Reliability

Inter-rater reliability of the test was measured by the point by point agreement, which was explained in the previous section. The point by point agreement between the transcribers was 85.4% (*SD* = 7.6%), indicating good, though not excellent, reliability.

The internal consistency in the NSRT was calculated by Cronbach's alpha (α), which was 0.550 for the children with CIs. This value is commonly regarded as "poor," possibly due to the children having varying degree of concentration during the test.

3.6 Ethics

Studies 2 and 3 received approval from the *Regional Committee for Research Ethics* (REC; project no. 2013/1580) and the *Personal security department* [Personvernombodet] at OUS Rikshospitalet (project no. 2013/12632).

The project did not raise difficult ethical issues. The participants' anonymity was preserved, and they did not have to reveal intimate and personal information. The strain inflicted on the participants during testing was minor, and the testing did not imply any physical or mental risk for the participants. The sound level that the participants were exposed to during testing was below 65 dB(A). This level corresponds to the sound pressure level of speech at a 1 m distance and is not regarded as harmful.

The testing was conducted as part of the follow-up program at the clinic, a program in which all CI recipients are advised to participate. Thus, the participants were not asked to come to the hospital solely to be tested in the study.

In Study 3, special concerns were made to facilitate the participation of children, by making sure that the children felt safe in the test situation, and they were allowed to bring a parent/guardian into the test room if they were anxious. They were rewarded by small gifts and particular attention was given to their consent. According to *Helseforskningsloven* (Health Research Act), there are clearly defined rules concerning the children's consent for participation in research. These issues were addressed in four different consent forms: for the children aged 7-12 years, for the children aged 12-16 years, for those older than 16 years, and for their parents/guardians.

Advantages and disadvantages

The participants appreciated becoming more conscious of the benefit of their CIs in perceiving speech sounds. They learned which speech sounds were most difficult to recognize for them personally and which speech sound confusions were most common for CI users in general. All NH participants benefited from a hearing check and middle-ear status examination by an ENT specialist in connection with this project.

There were no serious disadvantages for the participants. They might have perceived the testing as exhausting, as it occurred after an ordinary CI follow-up appointment and possibly following testing in another research project on the same day. Participants had to allocate enough time to accomplish the tests, which may have required time away from employment and other obligations.

Consent

All participants provided informed consent. CI users who had cognitive or physical challenges were excluded. Cognitive issues may make it difficult to understand the purpose of the project or to complete the test. The latter might also be the case with physical issues.

There were no restrictions regarding publication of the study results. The consent form signed by the participants states that their anonymized test results can be published in at least two articles in international, peer-reviewed scientific journals and presented in national and international congresses.

The participants could withdraw from the project or postpone their participation at any time with no consequences for their further follow-up at the clinic.

Financing of the project

I received a scholarship from the University of Oslo (UiO) during the four years of the project. My main supervisor received some compensation from UiO. Research funds from the Department of Special Needs Education paid two phoneticians to transcribe the sound recordings.

The participants with CIs were not paid, but the clinic covered their costs for traveling and accommodation, according to current welfare laws and regulations (Pasient- og brukerrettighetsloven, 1999, §2–6), as for every follow-up appointment. The parent or guardian accompanying his or her own child to the clinic also had his or her transportation expenses covered by the *Norwegian Labour and Welfare Administration* (NAV).

The NH listeners in the reference group did not have any expenses in connection with the project.

Insurance

Personal insurance for the participants was regarded as unnecessary, as no risk was expected in the project. Governmental rules for the coverage of travel reimbursement, food, and insurance applied for all participants.

Conflict of interests

There were no conflicts of interest for the author or co-authors in this project.

4 Discussion

In a systematic review and meta-analysis, we have investigated in detail the properties of speech sound confusions in studies of adult CI users with an NSRT; for instance, the influence of voicing and nasality on perception. We have also obtained a measure of the speech sound perception in well-performing children and adults with CIs. In a sub-group analysis with small samples of children and adolescents, we have investigated how prelingually deaf perceive consonant features compared to postlingually deaf. In another sub-group analysis, we have investigated how a small sample of adult users of Med-El's FS stimulation strategies perceive consonant features compared to a matched sample of users of non-FS stimulation strategies from Cochlear and Med-El.

Our results show that unvoiced consonants were significantly less confused than voiced consonants for both groups of participants with CIs. Moreover, there was a devoicing bias for the stops for both groups: Unvoiced stops were seldom confused with voiced stops, and voiced stops were often confused unvoiced stops. For the adults, unvoiced stops were never perceived as voiced stops. No significant difference was found between the perception score of consonant features for pre- and postlingually deaf children with CIs. The small sample of adult non-FS strategy users discriminated nasals versus nonnasals, nasals versus the lateral [l], and stops versus fricatives better than a small sample of FS strategy users, who, on the other hand, perceived voicing better than the non-FS strategy users.

4.1 Comparison of vowel and consonant scores of children and adults

In Article I (Rødvik et al., 2018), a systematic review reporting results of a majority of adult participants, the consonant score was 58% (*n* = 44) for the postlingually deaf and 47% (*n* = 6) for the prelingually deaf. The consonant score for the postlingually deaf adults in Article II was 56% (*n* = 34) and for the prelingually deaf 63% (*n* = 5). We notice that the consonant score for the postlingually deaf participants is almost similar in the systematic review as in Article II, and that the consonant score for the prelingually deaf is substantially lower in the systematic review than in Article II. The reason is probably that the systematic review included several older studies, in which the prelingually deaf participants received their implant at higher ages than what is common today. In clinics in our part of the world, congenitally deaf children are normally implanted before the age of 12 months.

The vowel score in the systematic review was 77% (*n* = 5) for postlingually deaf CI users and 68% (*n* = 1) for the prelingually deaf CI users. The vowel score for the postlingually deaf adults in Article II was 73% (*n* = 34) and for the prelingually deaf adults 62% (*n* = 5). We notice that the scores in the systematic review and in Article II are closer for the vowels than for the consonants. The vowels are in general more easily perceived, due to their long duration and high intensity, and it appears that age at implantation is not as important for the perception of vowels as for consonants.

The consonant and vowel scores were substantially lower for the mainly postlingually deaf adults in Article II than for the mainly prelingually deaf children and adolescents in Article III. This may have several reasons: Firstly, many of the postlingually deaf adults have had a progressive hearing loss and been hard-of-hearing for many years before implantation. This may have led to a degradation of their auditory pathways (Peelle and Wingfield, 2016). Secondly, a majority of the children in the study received a CI before the age of 12 months, which is close to what is often regarded as the optimal age for implantation (Busa et al., 2007). In particular, the perception score of voiced consonants was substantially higher for the prelingually deaf children than for the postlingually deaf adults. For the prelingually deaf children, the limitations in CI technology to convey low-frequency sounds seem less of a challenge than for the postlingually deaf adults, for whom the frequency distribution of the apical part of the implant is skewed compared to the natural tonotopy of their pre-implant hearing experience. This skewness is not present for the prelingually deaf and early implanted children who adapt to the tonotopy of the cochlea if implanted early.

4.2 Impact of stimulation strategy on speech sound perception

Stimulation strategy is one of many factors that affect speech sound perception in CI users. Other relevant factors are duration of deafness prior to implantation (a long duration appears to have a negative effect on auditory performance), age of onset of deafness (younger age is associated with better outcome), age at implantation (earlier implantation is associated with better outcome for prelingually deaf subjects), duration of hearing aid use before implantation, and duration of CI use (longer duration of CI experience is associated with better outcome). These factors have been described in a review article by Loizou (1999). Furthermore, aetiology of hearing loss, cognitive abilities, number of surviving spiral ganglion cells, electrode placement and insertion depth, electro-neural interface (the distance between the electrode and the neurons affects the stimulation thresholds), electrical dynamic range of the CI, and signal processing strategy are all important factors (Spencer, 2004; Wie et al., 2007; Rotteveel et al., 2010; Blamey et al., 2013; Long et al., 2014; Blamey et al., 2015). A strict research design is crucial for investigating the impact of stimulation strategies on speech sound discrimination, and the results of our comparison should be interpreted against this backdrop.

Due to the low number of participants in our comparison of the abilities of FS and non-FS strategy users to perceive different speech features, reported in Article II, the following was done to increase statistical power: Both groups had equal number of participants, both had the same mean consonant score, both included users of CIs only (not bimodal users), all participants were well-performing with a monosyllable score above 50%, and the participants in both groups had 100% correct pronunciation score on a test of Norwegian speech sounds. Due to the small sample, however, our results should be regarded only as a trend that should be investigated further.

4.3 Impact of pre- and postlingual deafness on speech sound discrimination

The comparison of pre- and postlingually deaf children and adolescents' perception of speech features, reported in Article III, showed no statistically significant differences. The uneven size of the two groups, in addition to the low sample size of prelingually deaf children and adolescents may have lowered the statistical validity. Moreover, the group of prelingually deaf children was heterogeneous, including both congenitally deaf and children who were born with hearing (normal or impaired) but experienced rapid hearing loss before the age of one year. For the participants in Article III, brain plasticity at implantation measured by cortical responses (e.g., the review article by Sharma et al., 2015) may be a more relevant prognostic factor for the development of speech perception skills than age at onset of deafness, because of the large individual variations in age at implantation and HA use before implantation.

The mechanisms of brain plasticity and the consequences of age at onset of deafness (pre-, peri-, and postlingual) have been shown to be important factors for both auditory and linguistic development. Buckley and Tobey (2011) found that the influence of cross-modal plasticity on speech perception is greatly influenced by age at onset of severe to profound (pre- or postlingual) deafness rather than by the duration of auditory deprivation before cochlear implantation.

4.4 Speech sound confusions visualized by acoustic cues in spectrograms

Although this dissertation is based on subjective assessments of speech recordings, the assessments are obviously based on objective measures such as acoustic cues in the speech signal. Some of the acoustic cues that contribute to perception are formants, formant transitions (locus and length), voicing, silent gaps, voice onset time, peak amplitude, and length of the bursts. For instance, Donaldson and Kreft (2006) pointed out that the stimulus level cue in particular may have dramatic effects on consonant recognition, primarily because of changes in the audibility of low-level cues such as those related to place of articulation.

In the following, we have constructed sound waves and spectrograms of some of the target words in the NSRT with the computer program Praat (Boersma and Weenink, 2019), to visualize some relevant acoustic cues used for speech recognition. We have included the first two formants, which are important for vowel and consonant recognition in Norwegian. A frequency span of 0–3,000 Hz has been chosen.

4.4.1 Confusions of [iː] and [yː]

In Norwegian, all vowels are distinguishable from one another by F1 and F2 alone. In Figure 8 below, we notice in the spectrograms of the recorded stimuli, that the F1s of the two vowels [iː] and [yː] (the first red line from the bottom of the figure) are close in frequencies (F1([iː]) = 310 Hz and F1([yː]) = 279 Hz). The same goes for F2 of [iː] and [yː] (F2([iː]) = 2693 Hz and F2([yː] = 2563 Hz). This may explain the CI users' frequent confusions of [iː] and [yː], both for the adults and for the children and adolescents. The small mean differences in F1 and F2 between the two vowels (31 Hz and 130 Hz, respectively) are probably poorly rendered by the implants, since this systematic error was not found for the NH groups. This is probably due to frequency-to-place mismatch, which will affect F1 more than F2.

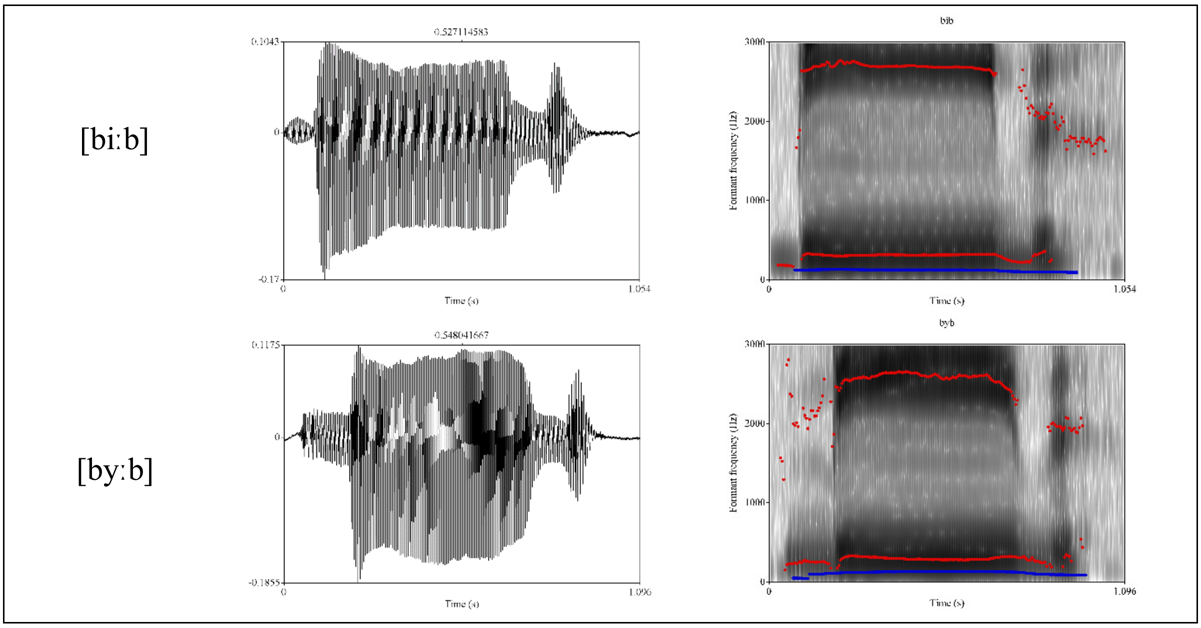


Figure 8. Transcription, sound wave, and spectrogram (shown from left to right) for the two nonsense words [biːb] and [byːb]. In the spectrogram, F0 is shown in blue and F1 and F2 are shown in red.

4.4.2 Devoicing bias in the perceptions of stops

The devoicing bias in the perception of stops in Articles II and III, illustrated by the spectrograms and waveforms of the minimal pair [ˈɑːpɑ] and [ˈɑːbɑ] in Figure 9, shows clearly the voicing in the rendering of [b], as opposed to the unsystematic acoustic noise in the rendering of [p]. Our results indicated that the adult implantees confuse voicing with nonvoicing but never nonvoicing with voicing. This could be explained by the speech coding of the implants being poorer in the lower frequencies. For the children, this devoicing bias was less pronounced.

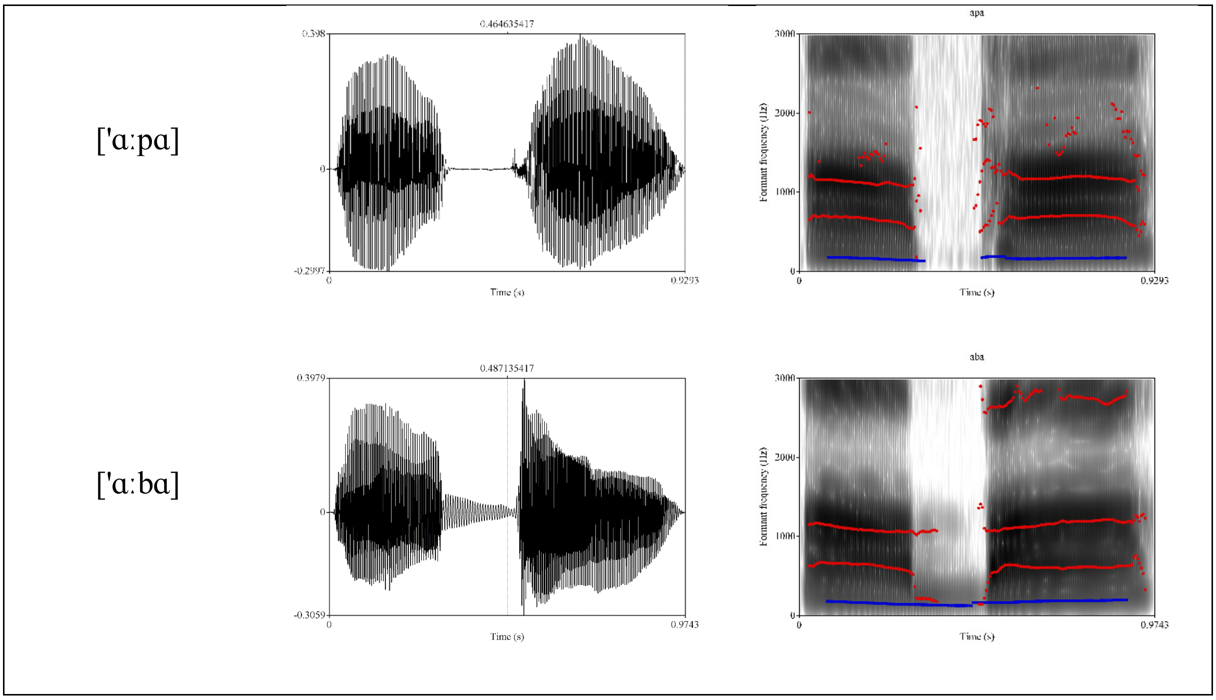


Figure 9. Transcription, sound wave, and spectrogram (shown from left to right) for the two nonsense words [ˈɑːpɑ] and [ˈɑːbɑ]. In the spectrogram, F0 is shown in blue and F1 and F2 are shown in red.

4.4.3 Confusion of nasals

Our results show that the nasals [m, n, ŋ] had the lowest recognition score of all the speech sounds, as they were repeated correctly in only one third of the cases and repeated as other nasals in one third of the cases. In Figure 10, the spectrograms and waveforms of ['ɑːmɑ], ['ɑːnɑ], and ['ɑːŋɑ] are shown. As for the other voiced consonants, the implants probably have difficulties rendering the nasals, which have a lot of energy in the lower frequencies.

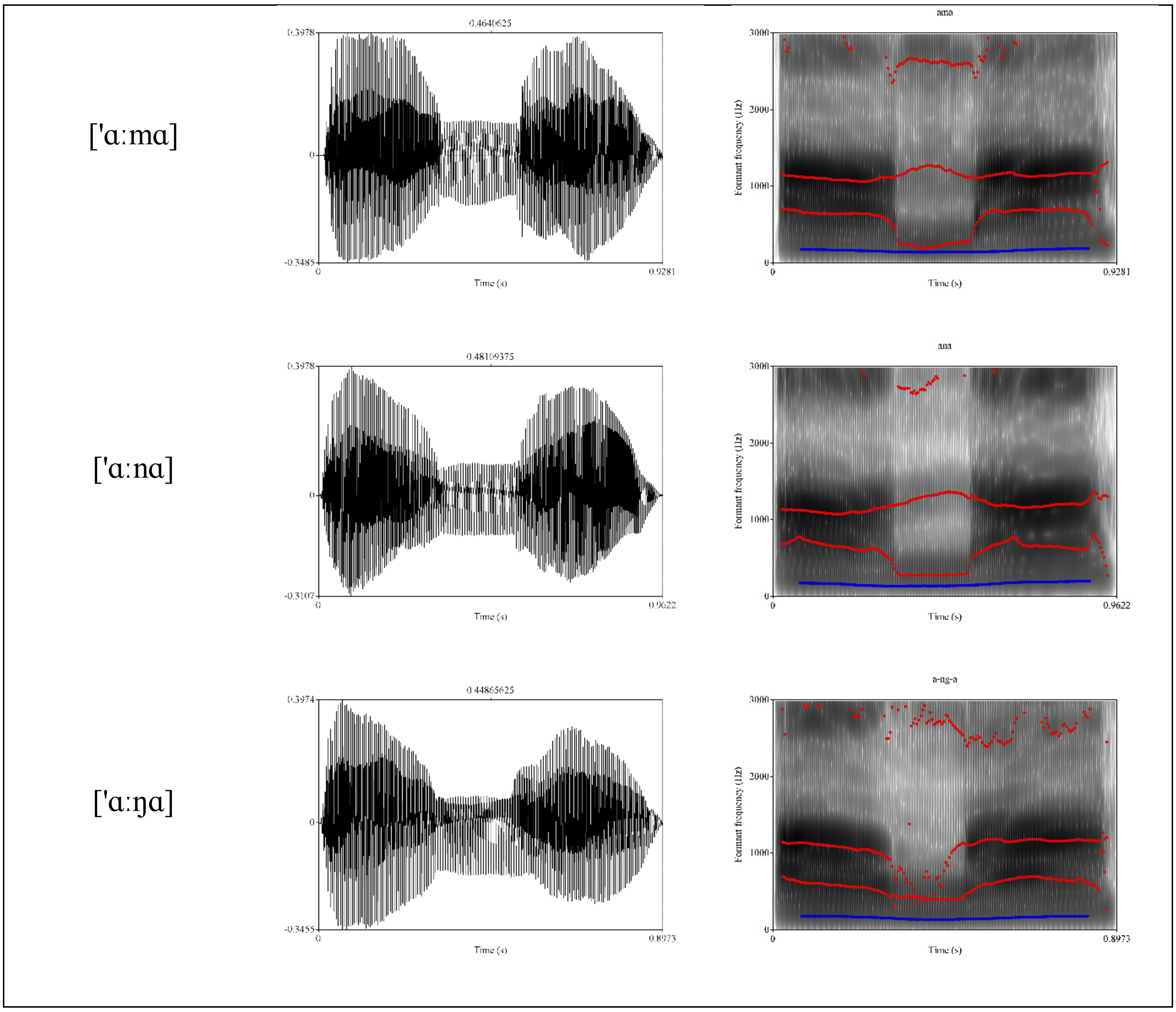


Figure 10. Transcription, sound wave, and spectrogram (shown from left to right) for the three nonsense words ['ɑːmɑ], ['ɑːnɑ], and ['ɑːŋɑ]. In the spectrogram, F0 is shown in blue and F1 and F2 are shown in red.

4.4.4 Higher consonant recognition score in the /a/ context than in the /i/ and /u/ contexts

The results in Study 2 show that the adult CI users' recognition score of consonants in the /a/ context was higher than in the /i/ and /u/ contexts. Presentation of the consonants in three vowel contexts was useful for the analyses, as it increased the statistical power and allowed for averaging the influence from formant transitions on the results. In Figure 11, the consonant [k] is shown in the three vowel contexts.

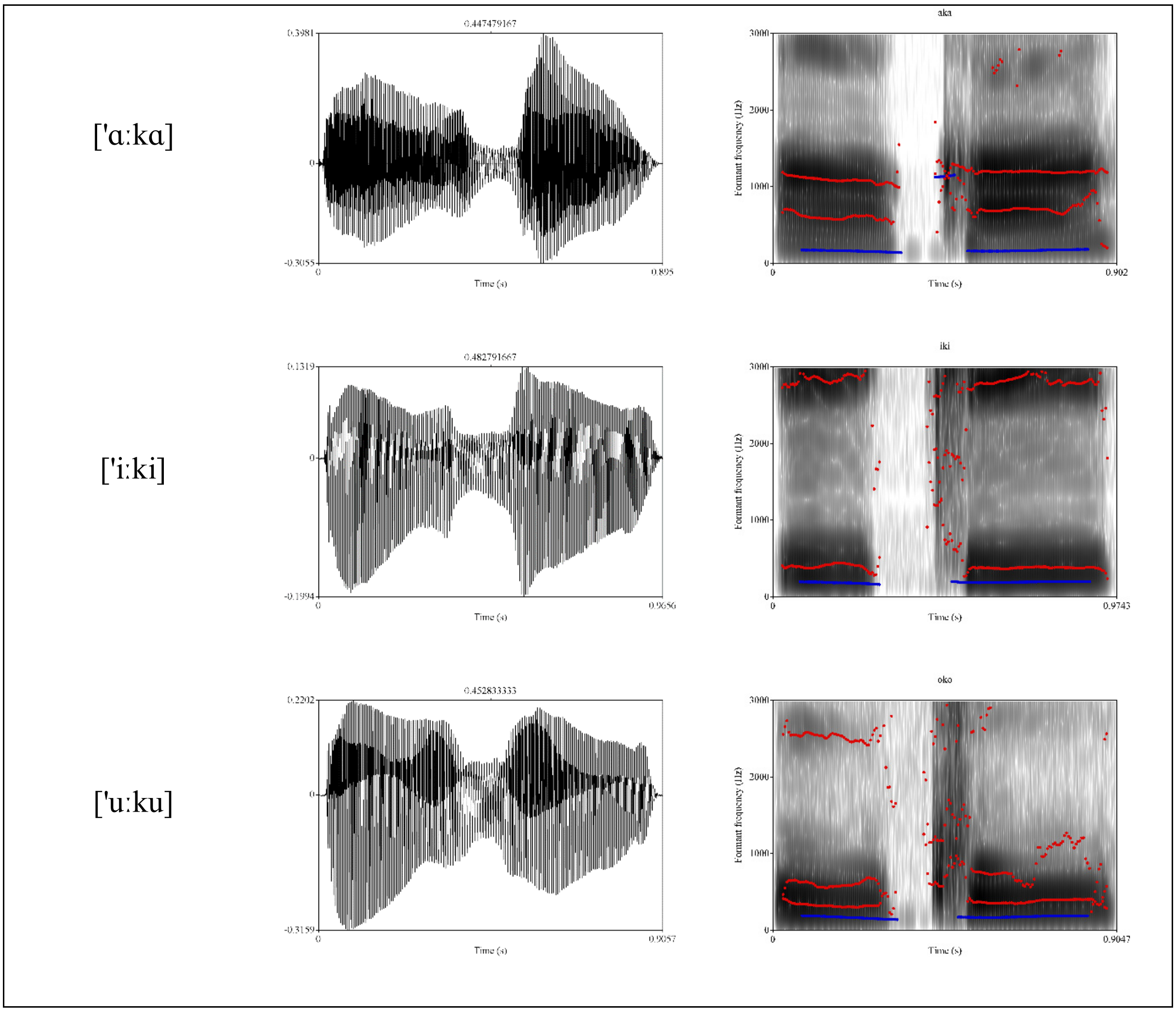


Figure 11. Transcription, sound wave, and spectrogram (shown from left to right) for the three nonsense words ['ɑːkɑ], ['iːki], and ['uːku]. In the spectrogram, F0 is shown in blue and F1 and F2 are shown in red.

4.5 Selection of participant groups

Study 2 was designed with a convenience sample of adult CI users and a group of NH listeners, both above 16 years of age. Study 3 was designed with a convenience sample of children and adolescents with CIs below the age of 16, and two NH reference groups, one with 6-year-olds and one with 13-year-olds. The reason for having two reference groups with different ages in Study 3, was to disentangle the effect of language development from the analyses. Although the NH 6-year-olds in the study had 100% correct pronunciation of all the included speech sounds in the test, their higher language skills were probably not as fully developed as the 13-year-olds, and this may have contributed to lower consonant and vowel repetition scores for them than for the 13-year-olds, who, in fact, obtained scores similar to those of the NH adults. As the 6-year-olds were the better match in hearing age with the children with CIs (i.e., the age at which their CIs were activated), we chose to employ only the 6-year-olds in our comparative analyses.

The statistical comparisons between the two samples of NH listeners for descriptive purposes, suggest that the auditory systems of 6-year-olds are indeed less developed than the auditory systems of 13-year-olds. Litovsky (2015, p. 58) constructed a table based on experiments performed in her lab, that gives an idea of the ages at which different auditory mechanisms reach adult-level maturity. According to the table, NH children do not develop full low-frequency detection and intensity discrimination until early school age, and auditory development matures during the teenage years.

An improved design of the study might be obtained by constructing subgroups of CI users and NH listeners of the same sample size and with same hearing age. The recruitment of the CI users by convenience samples made this impossible, as groups matched from these samples would be too small to achieve adequate statistical power.

4.6 Overarching strengths and limitations

4.6.1 Strengths

Articles I–III support our claim that the NSRT is a useful tool for assessing the perception of speech sounds in Norwegian CI users. The results of the systematic review indicate that consonant and vowel scores measured by monosyllabic CVC and bisyllabic VCV nonsense words are stable in large populations and less dependent on language and implant technology than on the variability between the CI users. The findings in Studies 2 and 3 may thus be relevant for studies in other languages, provided that the speech perception tests are modified to only include speech sounds existing in the particular language.

We regard the use of nonsense syllables for measuring actual consonant and vowel perception as the major strength of our studies, as the stimuli in the NSRT are without lexical meaning and listeners cannot rely on linguistic context or use top-down mechanisms to recognize the words. The Norwegian NSRT fulfil these criteria, as it does not have a single test entity with any lexical meaning in Norwegian. Ceiling effects on the scores are also more likely avoided with nonsense syllable stimuli than with real words.

Our inclusion requirement of a 100% pronunciation score of all the included speech sounds and a score above 50% on a real-word monosyllable test is a strength of the study, as it excluded any misinterpretation of data due to mispronunciation.

The use of an open-set test design with verbal repetitions is a strength for our purpose, as the test scores will be influenced by neither the test subjects' reading or writing ability nor their computer skills, and thus more detailed information about speech perception and listening capacity for acoustic properties is provided. Furthermore, open-set tests have relatively small learning effects compared to closed-set tests and can therefore be performed reliably at desired intervals (Drullman, 2005, p. 8). Finally, the information that is given by the size and contents of the unclassified category in the CMs is not present in a closed-set test design.

The consonants were presented in three vowel contexts, and as such the statistical power for the consonant score was higher than for the vowel score, which was calculated from only one repetition of each vowel. The design of the consonant test also facilitated collection of useful information about how formant transitions influence perception in different vowel contexts (see Figure 11).

4.6.2 Limitations and future directions

The NSRS-V is based on one single repetition of each vowel. The vowels /ɑ, i, u/, which constitute the vowel context of the consonants in the NSRT and were presented in initial and final position, were not included in the calculations of NSRS-V, as we focused on medial vowels. This provided less information to draw conclusions about inter-individual differences than the consonant test, in which there are three repetitions of each consonant. Still, the results provided useful descriptive information about group means and *SD*s, and they compared very well with the mean vowel and consonant repetition scores in the systematic review and meta-analysis. Moreover, in our study, only group means and *SD*s were used in the calculations, not individual data.

Since the test lists of the NSRT counted as many as 90 CVC and VCV nonsense words, fatigue and diverted attention of the participants may have influenced the scores of the NSRT, especially for the younger children. The participants' response accuracy probably decreased toward the end of the test session, which for a majority lasted around 5 minutes. However, as the test lists were randomized, the same word was prevented from always appearing at the end of the test list and thus systematic errors were avoided. The participants' concentration therefore probably did not influence the repetition of specific speech sounds, only the total scores.

The participants were a convenience sample, tested during their regular visits to the hospital for follow-ups during a data collection period spanning 1.5 years. Matching of the participants with factors such as demographics, implant model, speech processor settings, age at implantation, duration of implant use, and anatomy would have been interesting, but was not feasible due to the time constraints of the study. The statistical validity and the external validity were probably somewhat diminished because of this choice of study design. On the other hand, the participants in Study III represent a completely random sample of Norwegian-speaking children with CIs, since all implanted children in Norway have received their CI at OUS Rikshospitalet.

It would, in future research, be interesting to recruit suitable participants who could be matched with regard to the abovementioned factors in a multi-centre study, as having a larger patient pool would enable a more time-efficient data collection phase.

Another interesting research project, would be to investigate how the variables that were eliminated in the construction of the NSRT (e.g., working memory, language skills, and attention), would predict the variability of the NSRT scores. This might also be compared with the extent to which the same variables predict the results on top-down speech perception tests such as the HINT.

4.7 Clinical implications

This study might be used as a basis for the development, validation, and norming of a simplified version of the NSRT to be included in the standard test battery in audiology clinics. Children with CIs tested regularly with the NSRT would be provided with individual feedback on what needs to be targeted in the programming of their CIs and in their listening therapy sessions. Individually directed pre- and posttesting with the NSRT can be used as a quality control tool of the programming of the CI. A clinical NSRT would also meet the increasing challenge of assessing speech perception in patients with different language backgrounds, as it can be adjusted for different languages by modifying it to only include speech sounds existing in a particular language.

The NSRT helps determine which consonants and vowels and which consonant features are most frequently confused by adult CI users. Knowledge of the implantee's speech sound confusions is useful in the CI programming session. For instance, poor perception of unvoiced speech sounds might indicate that programming of the basal electrodes should be considered. Low scores on the perception of voiced speech sounds, which was revealed for the adult participants in Article II, might indicate that reprogramming of the apical electrodes should be considered, as well as auditory training focusing on voiced consonants.

A close examination of the CMs for the individual CI user may be useful when choosing between reprogramming and listening therapy. Speech sounds within the same manner-group are more acoustically similar than speech sounds in different manner groups. Hence, a rule-of-thumb may be that in case of confusions within the same manner-group, start with listening therapy, and in case of confusions between two manner-groups, reprogramming of the implant may prove to be the better option.

5 Concluding remarks

5.1 Conclusions

In Studies 2 and 3, participants with CIs achieved the highest scores for vowel repetitions, the second highest for unvoiced consonant repetitions, and the lowest for voiced consonant repetitions. A devoicing bias was found for the stops and a confusion bias was found for [yː] and [iː], in favour of [iː]. This bias was most pronounced for the adults.

In general, consonants were mostly confused with consonants with the same voicing and manner. Voiced consonants proved more difficult to perceive than unvoiced consonants. Vowels were confused with other vowels, in which both first and second formants were close in frequency.

A comparison of small samples of adult users of FS and non-FS stimulation strategies in Study 2 indicated that the distinction of nasals and nonnasals, nasals and the lateral [l], and stops and fricatives was higher for the non-FS strategy users than for the FS strategy users. The distinction of voicing and nonvoicing was indicated to be higher for the FS strategy users than for the non-FS strategy users.

For the children and adolescents with CIs in Study 3, subgroup analyses showed no statistically significant differences between the consonant scores and phonetic features for pre- and postlingually deaf participants.

Although the participating CI users had a 100% correct pronunciation score, none of them obtained scores for voiced consonants above 78%. As their speech is much better than their perception capability would indicate, people they encounter in their everyday life might underestimate the severity of their hearing impairment.

For the CI users in both studies, the low-frequency transmission of the implants appeared to function more poorly than the high-frequency transmission, and this issue seemed to be most pronounced for the perception of voiced consonants. Furthermore, the frequency-place mismatch in the apical region of the cochlea seemed to be most pronounced for the postlingually deaf. This indicates that there are still limitations in the CI technology for the transmission of low-frequency sounds.

5.2 Further perspectives

In the endeavour to reach the implant's potential, the NSRT is a useful tool for revealing the challenging speech sounds to establish an individual baseline from which to plan auditory training. This, combined with the best possible programming of the CIs, will maximise their effect.

The discrepancy of 16–17 percentage points between the real-word monosyllable scores and the NSRS-C, confirms that the former to a great extent also tests the CI users' ability to guess, with reference to their own established vocabulary. This discrepancy might be seen in light of the mirror neuron theory for the acquisition of spoken language, as the CI users seem to more easily perceive a word that they have internalized as a sound unit and most likely pronounce correctly, than a word that is unfamiliar to them. The somewhat unusual VCV words of the NSRT do not seem to be included among these established sound units. Considering that all new words are perceived as nonsense words before they are internalized, one should put more focus on bottom-up processes at the syllable-level in listening therapy, and thus endeavour to establish a toolkit for catching unfamiliar words more easily. This is especially pertinent for the prelingually deaf children with CIs, who will have to develop all vocabulary from scratch.

Consequently, routines for expansion of the vocabulary in children with CIs should be developed and their correct pronunciation verified. The more words children with CIs have internalized and can pronounce correctly, the more spoken words they will be able to perceive in their everyday surroundings.

References

Baddeley, AD, and Hitch, G (1974). Working memory. *Psychology of Learning and Motivation - Advances in Research and Theory* 8**,** 47–89.

Baddeley, AD, Thomson, N, and Buchanan, M (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior* 14**,** 575–589.

Black, JW (1957). Multiple-choice intelligibility tests. *The Journal of speech and hearing disorders* 22**,** 213–235.

Blamey, PJ, Artieres, F, Başkent, D, Bergeron, F, et al. (2013). Factors Affecting Auditory Performance of Postlinguistically Deaf Adults Using Cochlear Implants: An Update with 2251 Patients. *Audiology and Neurotology* 18**,** 36–47.

Blamey, PJ, Maat, B, Başkent, D, Mawman, D, Burke, E, Dillier, N, et al. (2015). A Retrospective Multicenter Study Comparing Speech Perception Outcomes for Bilateral Implantation and Bimodal Rehabilitation. *Ear and hearing* 36**,** 408–416.

Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". WikiJournal of Medicine 1 (2). DOI:10.15347/wjm/2014.010. ISSN 2002–4436.

Boersma, P., and Weenink, D. (2019). Praat: Doing phonetics by computer [Computer program]. Version 6.0.56. Available at: http://www.praat.org (accessed September 1, 2019).

Borden, GJ, Harris, KS, and Raphael, LJ (1994). *Speech science primer: physiology, acoustics, and perception of speech.* Philadelphia: Lippincott Williams & Wilkins.

Busa, J, Harrison, J, Chappell, J, Yoshinaga-Itano, C, Grimes, A, Brookhouser, P, et al. (2007). Year 2007 position statement: Principles and guidelines for early hearing detection and intervention programs. *Pediatrics* 120**,** 898–921.

Caldwell, MT, Jiam, NT, and Limb, CJ (2017). Assessment and improvement of sound quality in cochlear implant users. *Laryngoscope Investigative Otolaryngology* 2**,** 119–124.

Caner, G, Olgun, L, Gültekin, G, and Balaban, M (2007). Optimizing Fitting in Children Using Objective Measures Such as Neural Response Imaging and Electrically Evoked Stapedius Reflex Threshold. *Otology & Neurotology* 28**,** 637–640.

Cartwright, N (2009). If no capacities then no credible worlds. But can models reveal capacities? *An International Journal of Analytic Philosophy* 70**,** 45–58.

Clark, GM, Tong, YC, Martin, LF, Busby, PA, Dowell, RC, Seligman, PM, et al. (1981). A multiple-channel cochlear implant: An evaluation using nonsense syllables. *Annals of Otology, Rhinology and Laryngology* 90**,** 227–230.

Clopper, CG, Pisoni, DB, and Tierney, AT (2006). Effects of open-set and closed-set task demands on spoken word recognition. *Journal of the American Academy of Audiology* 17**,** 331–349.

Cochlear (2006). “HOPE Note. The Newly Implanted Preschooler.” (Cochlear Americas, Englewood,CO).

Cochlear (2016). “Cochlear celebrates 30 years of hearing revolution.” Retrieved from http://www.cochlear.com/wps/wcm/connect/au/about/cochlear-30-anniversary/

Cohen, NL, Waltzman, SB, and Fisher, SG (1993). A prospective, randomized study of cochlear implants. *New England Journal of Medicine* 328**,** 233–237.

Crew, J, Galvin, J, and Landsberger, D (2015). Contributions of Electric and Acoustic Hearing to Bimodal Speech and Music Perception. *PLoS One* 10**,** 1–18.

Depaolis, RA, Vihman, MM, and Nakai, S (2013). The influence of babbling patterns on the processing of speech. *Infant Behavior and Development* 36**,** 642–649.

Donaldson, GS, Dawson, PK, and Borden, LZ (2011). Within-subjects comparison of the HiRes and Fidelity120 speech processing strategies: speech perception and its relation to place-pitch sensitivity. *Ear and Hearing* 32**,** 238–250.

Donaldson, GS, and Kreft, HA (2006). Effects of vowel context on the recognition of initial and medial consonants by cochlear implant users. *Ear and Hearing* 27**,** 658–677.

Dorman, MF, and Wilson, BS (2004). The Design and Function of Cochlear Implants: Fusing medicine, neural science and engineering, these devices transform human speech into an electrical code that deafened ears can understand. *American Scientist* 92**,** 436–445.

Dowell, RC, Dettman, SJ, Blamey, PJ, Barker, EJ, and Clark, GM (2002). Speech perception in children using cochlear implants: prediction of long-term outcomes. *Cochlear Implants International* 3**,** 1–18.

Drullman, R (2005). "D-7-1: Speech recognition tests for different languages", in: *FP6–004171 HEARCOM Hearing in the Communication Society.* (Germany: Fraunhofer Institute for Applied Information Technology).

Dubno, JR, and Dirks, DD (1982). Evaluation of Hearing-Impaired Listeners Using a Nonsense-Syllable Test, I: Test Reliability. *Journal of Speech and Hearing Research* 25**,** 135–141.

Duff, D, Tomblin, JB, and Catts, H (2015). The influence of reading on vocabulary growth: a case for a Matthew effect. *Journal of Speech, Language, and Hearing Research* 58**,** 853–864.

Ear Foundation (2017, December 17). Cochlear Implant Information Sheet. Retrieved from http://www.earfoundation.org.uk/hearing-technologies/cochlear-implants/cochlear-implant-information-sheet.

Eisen, MD (2003). Djourno, Eyries, and the first implanted electrical neural stimulator to restore hearing. *Otology & Neurotology* 24**,** 500–506.

Eisenberg, LS, Martinez, AS, Holowecky, SR, and Pogorelsky, S (2002). Recognition of lexically controlled words and sentences by children with normal hearing and children with cochlear implants. *Ear and Hearing* 23**,** 450–462.

Elley, WB, The Intenational Association for the Evaluation of Educational, A, and Literacy, IEaSOR (1992). *How in the world do students read: IEA study of reading literacy.* The Hague: The International Association for the Evaluation of Educational Achievement (IEA).

Eshraghi, AA, Nazarian, R, Telischi, FF, Rajguru, SM, Truy, E, Gupta, C, et al. (2012). The cochlear implant: Historical aspects and future prospects. *Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology* 295**,** 1967–1980.

Fant, G (1958). *Acoustic theory of speech production.* Stockholm: Royal Institute of Technology.

Fant, G (1967). "Auditory patterns of speech", in *Models for the perception of speech and visual form*, ed. W. Wathen-Dunn (Cambridge, Massachusetts, and London, England: MIT Press), 111–125.

Fant, G (1973). *Speech sounds and features.* Cambridge, Mass: MIT Press.

Ferguson, CA, and Farwell, CB (1975). Words and sounds in early language acquisition. *Language* 51**,** 419–439.

Fretz, RJ, and Fravel, RP (1985). Design and function: a physical and electrical description of the 3M House cochlear implant system. *Ear and hearing* 6**,** 14S–19S.

Frigg, R, and Hartmann, S (2012). *Models in science* [Online]. Available: http://plato.stanford.edu/archives/fall2012/entries/models-science [Accessed 31.05.2014].

Gathercole, SE, Willis, CS, Baddeley, AD, and Emslie, H (1994). The children's test of nonword repetition: a test of phonological working memory. *Memory* 2**,** 103–127.

Geers, AE, Brenner, CA, and Tobey, EA (2011). Long-term outcomes of cochlear implantation in early childhood: sample characteristics and data collection methods. *Ear and hearing* 32**,** 1–27.

Giere, RN (1991). *Understanding scientific reasoning.* Fort Worth: Holt, Rinehart, and Winston.

Glaze, JA (1928). The Association Value of Non-Sense Syllables. *The Pedagogical Seminary and Journal of Genetic Psychology* 35**,** 255–269.

Greisiger, R (2016). *Objective Measurements and Cochlear Implants Imaging.* Doctoral dissertation, University of Oslo.

Greisiger, R, Shallop, JK, Hol, PK, Elle, OJ, and Jablonski, GE (2015). Cochlear implantees: Analysis of behavioral and objective measures for a clinical population of various age groups. *Cochlear Implants International* 16**,** 1–19.

Hughes, ML (2012). *Objective measures in cochlear implants.* San Diego: Plural.

Ingram, D (1974). Phonological rules in young children. *Journal of Child Language* 1**,** 49–64.

Jakobson, R (1941). *Kindersprache, Aphasie und allgemeine Lautgesetze [Child language, aphasia, and general phonetics].* Uppsala: Almqvist & Wiksell.

Kleven, TA (2008). Validity and validation in qualitative and quantitative research. *Nordic Studies in Education* 03**,** 219–233.

Kristoffersen, G (2000). *The phonology of Norwegian.* Oxford: Oxford University Press.

Liberman, AM, Cooper, FS, Shankweiler, DP, Studdert-Kennedy, M, and Cofer, CN (1967). Perception of the speech code. *Psychological Review* 74**,** 431–461.

Liberman, AM, and Mattingly, IG (1985). The motor theory of speech perception revised. *Cognition* 21**,** 1–36.

Ling, D (1976). *Speech and the hearing-impaired child: Theory and practice.* Washington DC: The Alexander Graham Bell Association for the Deaf.

Lisker, L (1981). On Generalizing the Rapid-Rabid Distinction Based on Silent Gap Duration. *Haskins Laboratories Status Report on Speech Research* 65**,** 251–259.

Litovsky, RY (2015). "The human auditory system: fundamental organization and clinical disorders," in *Handbook of Clinical Neurology,* ed. G.G. Celesia. (Amsterdam: Elsevier Science).

Loizou, PC (1999). Introduction to cochlear implants. *Engineering in Medicine and Biology Magazine, IEEE* 18**,** 32–42.

Long, CJ, Holden, TA, Mcclelland, GH, Parkinson, WS, Shelton, C, Kelsall, DC, et al. (2014). Examining the Electro-Neural Interface of Cochlear Implant Users Using

Psychophysics, CT Scans, and Speech Understanding. *Jaro-Journal of the Association for Research in Otolaryngology* 15**,** 293–304.

Lorens, A, Walkowiak, A, Piotrowska, A, Skarzynski, H, and Anderson, I (2004). ESRT and MCL correlations in experienced paediatric cochlear implant users. *Cochlear Implants International* 5**,** 28–37.

Lund, T (2002). "Metodologiske prinsipper og referanserammer. [Methodological principles and frames of reference]," in *Innføring i forskningsmetodologi. [Introduction to research methodology].* (Oslo: Unipub).

Majorano, M, Vihman, MM, and Depaolis, RA (2013). The Relationship Between Infants' Production Experience and Their Processing of Speech. *Language Learning and Development* 10**,** 1–26.

McKay, C, Chandan, K, Akhoun, I, Siciliano, C, and Kluk, K (2013). Can ECAP Measures Be Used for Totally Objective Programming of Cochlear Implants? *Journal of the Association for Research in Otolaryngology* 14**,** 879–890.

Med-El (2008). "Handbook for educators. Teaching children who listen with a cochlear implant," v. 2.0. In Med-El's White Papers (Med-El, Innsbruck).

Miller, GA, and Nicely, PE (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* 27**,** 338–352.

Miller, GA, Wiener, FM, and Stevens, SS (1946). *Transmission and reception of sounds under combat conditions.* Washington D. C.: Office of Scientific Research and Development.

Moher, D, Liberati, A, Tetzlaff, J, and Altman, DG (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *PLoS Medicine* 6**,** 1–6.

Mulder, HE, Van Olphen, AF, Bosman, A, and Smoorenburg, GF (1992). Phoneme recognition by deaf individuals using the multichannel Nucleus cochlear implant. *Acta Oto-Laryngologica* 112**,** 946–955.

Munson, B, Donaldson, GS, Allen, SL, Collison, EA, and Nelson, DA (2003). Patterns of phoneme perception errors by listeners with cochlear implants as a function of overall speech perception ability. *Journal of the Acoustical Society of America* 113**,** 925–935.

Murphy, KR, and Davidshofer, CO (2001). *Psychological testing: principles and applications.* Upper Saddle River, NJ: Prentice Hall.

Myhrum, M, and Moen, I (2008). The Norwegian Hearing in Noise Test. *International Journal of Audiology* 47**,** 377–378.

Myhrum, M, Strøm-Roum, H, Heldahl, MG, Rødvik, AK, Eksveen, B, Landsvik, B, et al. (2017). Sequential bilateral cochlear implantation in children: outcome of the second implant and long-term use. *Ear and Hearing* 38**,** 301–313.

Nilsson, M, Soli, SD, and Sullivan, JA (1994). Development of the Hearing in Noise Test for the Measurement of Speech Reception Thresholds in Quiet and in Noise. *Journal of the Acoustical Society of America* 95**,** 1085–1099.

Niparko, JK (2004). "Cochlear Implants: Clinical Applications," in *Cochlear Implants: Auditory Prostheses and Electric Hearing,* eds. F.-G. Zeng, A.N. Popper & R.R. Fay. (New York, NY: Springer).

Niparko, JK, Tobey, EA, Thal, DJ, Eisenberg, LS, Wang, N-Y, Quittner, AL, et al. (2010). Spoken language development in children following cochlear implantation. *Journal of the American Medical Association* 303**,** 1498–1506.

O'Donoghue, G (2013). Cochlear Implants — Science, Serendipity, and Success. *The New England Journal of Medicine* 369**,** 1190–1193.

Ormestad, H (1955). *En akustisk undersøkelse av høreskarphet og taletydelighet og forståelsen av norske språklyder [An acoustic investigation of hearing sharpness and speech intelligibility, and of the understanding of Norwegian speech sounds].* Doctoral dissertation, University of Oslo.

Ottem, E (1972). *En sammenligning mellom bokstavforvekslinger hos barn med antatt utviklingsmessig dysleksi og bokstavforvekslinger hos normalt lesende barn [A comparison between letter confusions in children with assumed developmental dyslexia and letter confusions in normal-reading children].* Master dissertation, University of Oslo.

Pagon, RA, Graham, JM, Zonana, J, and Yong, S-L (1981). Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association. *The Journal of Pediatrics* 99**,** 223–227.

Pasient- og brukerrettighetsloven (1999, July 2). [Act on patient rights]. No 63.

Peelle, JE, and Wingfield, A (2016). The Neural Consequences of Age-Related Hearing Loss. *Trends in Neurosciences* 39**,** 486–497.

Pelizzone, M, Cosendai, G, and Tinembart, J (1999). Within-patient longitudinal speech reception measures with continuous interleaved sampling processors for Ineraid implanted subjects. *Ear and Hearing* 20**,** 228–237.

Petersen, B, Mortensen, MV, Hansen, M, and Vuust, P (2012). Singing in the Key of Life: A Study on Effects of Musical Ear Training After Cochlear Implantation. *Psychomusicology: Music, Mind & Brain December* 22**,** 134–151.

Peterson, GE, and Lehiste, I (1962). Revised CNC lists for auditory tests. *The Journal of speech and hearing disorders* 27**,** 62–70.

Plant, G, Bernstein, C, and Levitt, H (2015). Optimizing Performance in Adult Cochlear Implant Users through Clinician Directed Auditory Training. 36**,** 296–310.

Rabiner, L, and Juang, B-H (1993). *Fundamentals of speech recognition.* Englewood Cliffs, NJ: PTR Prentice Hall.

Rotteveel, LJC, Snik, AFM, Cooper, H, Mawman, DJ, van Olphen, AF, and Mylanus, EAM (2010). Speech Perception after Cochlear Implantation in 53 Patients with Otosclerosis: Multicentre Results. *Audiology and Neuro-Otology* 15**,** 128–136.

Rødvik, AK (2008). Perception and confusion of speech sounds by adults with a cochlear implant. *Clinical Linguistics and Phonetics* 22**,** 371–378.

Rødvik, AK, Torkildsen, JvK, Wie, OB, Storaker, MA, and Silvola, JT (2018). Consonant and vowel identification in cochlear implant users measured by nonsense words: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research* 61**,** 1023–1050.

Rødvik, AK, Torkildsen, JvK, Wie, OB, Tvete, O, Skaug, I, Silvola, JT. (2019a). Consonant and vowel confusions in well-performing adult cochlear implant users, measured by a nonsense syllable repetition test. Submitted to *The Journal of The Acoustical Society of America*.

Rødvik, AK, Tvete, O, Torkildsen, JvK, Wie, OB, Skaug, I, and Silvola, JT (2019b). Consonant and vowel confusions in well-performing children and adolescents with cochlear implants, measured by a nonsense syllable repetition test. *Frontiers in Psychology* 10**,** 1–17.

Samuel, AG (2011). Speech perception. *Annual Review of Psychology* 62**,** 49–72.

Schumann, A, Serman, M, Gefeller, O, and Hoppe, U (2015). Computer-based auditory phoneme discrimination training improves speech recognition in noise in experienced adult cochlear implant listeners. *International Journal of Audiology* 54**,** 190–198.

Shadish, WR, Cook, TD, and Campbell, DT (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

Shapiro, WD, and Bradham, TS (2011). Cochlear Implant Programming. *Otolaryngologic Clinics of North America* 45, 111–127.

Sharma, A, Glick, H, Deeves, E, Duncan, E, and Sharma, A (2015). The P1 biomarker for assessing cortical maturation in pediatric hearing loss: a review. Otorinolaringologia 65, 103–114.

Shriberg, LD, Fourakis, M, Hall, SD, Karlsson, HB, Lohmeier, HL, Mcsweeny, JL, et al. (2010). Perceptual and acoustic reliability estimates for the Speech Disorders Classification System (SDCS). *Clinical Linguistics & Phonetics* 24**,** 825-846.

Singh, S, and Frank, DC (1972). A distinctive feature analysis of the consonantal substitution pattern. Language and Speech 15, 209–218.

Smith, N (1973). *The Acquisition of Phonology: a Case Study.* Cambridge: University Press.

Spencer, PE (2004). Individual differences in language performance after cochlear implantation at one to three years of age: Child, family, and linguistic factors. *Journal of Deaf Studies and Deaf Education* 9**,** 395–412.

Stacey, PC, Raine, CH, O'Donoghue, GM, Tapper, L, Twomey, T, and Summerfield, AQ (2010). Effectiveness of computer-based auditory training for adult users of cochlear implants. *International Journal of Audiology* 49**,** 347–356.

Stach, BA (2009). *Clinical audiology: an introduction.* New York: Delmar - Cengage Learning.

Stampe, D (1969). *The Acquisition of Phonetic Representation.* Chicago: University of Chicago.

Stevens, KN, and Blumstein, SE (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64**,** 1358–1368.

Sucher, CM, and Mcdermott, HJ (2007). Pitch ranking of complex tones by normally hearing subjects and cochlear implant users. *Hearing Research* 230**,** 80–87.

Sussman, HM, Mccaffrey, HA, and Matthews, SA (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America* 90**,** 1309–1325.

Tao, JD, Deng, JR, Jiang, JY, Galvin, JJ, Fu, JQ-J, and Chen, JB (2015). Melodic pitch perception and lexical tone perception in mandarin-speaking cochlear implant users. *Ear and Hearing* 36**,** 102–110.

Teig, E, Lindeman, HH, Flottorp, G, Tvete, O, Hanche-Olsen, S, and Arntsen, O (1992). Patient Performance with Two Types of Multiple Electrode Intracochlear Implant. *Scandinavian Audiology* 21**,** 93–99.

Tetzchner, SV (1975). *En sammenligning mellom bokstavforvekslinger hos døve og normalt hørende barn [A comparison of letter confusions in deaf and normally hearing children].* Master dissertation, University of Oslo.

Tyler, RS, Abbas, P, Tye-Murray, N, Gantz, BJ, Knutson, JF, Mccabe, BF, et al. (1988). Evaluation of five different cochlear implant designs: Audiologic assessment and predictors of performance. *Laryngoscope* 98**,** 1100–1106.

Tyler, RS, and Moore, BCJ (1992). Consonant recognition by some of the better cochlear-implant patients. *Journal of the Acoustical Society of America* 92**,** 3068–3077.

Tyler, RS, Preece, JP, and Lowder, M (1983). The Iowa cochlear implant tests. *University of Iowa, Iowa City*.

Vaerenberg, B, Smits, C, De Ceulaer, G, Zir, E, Harman, S, Jaspers, N, et al. (2014). Cochlear Implant Programming: A Global Survey on the State of the Art. *The Scientific World Journal* 2014**,** 1–12.

Valimaa, TT, and Sorri, MJ (2000). Speech perception after multichannel cochlear implantation in Finnish-speaking postlingually deafened adults. *Scandinavian Audiology* 29**,** 276–283.

Van Den Abbeele, T, Noël-Petroff, N, Akin, I, Caner, G, Olgun, L, Guiraud, J, et al. (2012). Multicentre investigation on electrically evoked compound action potential and stapedius reflex: how do these objective measures relate to implant programming parameters? *Cochlear Implants International* 13**,** 26–34.

van Wieringen, A, and Wouters, J (1999). Natural vowel and consonant recognition by Laura cochlear implantees. *Ear and Hearing* 20**,** 89–103.

Vandali, AE, Dawson, PW, and Arora, K (2017). Results using the OPAL strategy in Mandarin speaking cochlear implant recipients. *International Journal of Audiology* 56**,** 74–85.

Vandali, AE, Dawson, PW, Au, A, Yu, Y, Brown, M, Goorevich, M, et al. (2018). Evaluation of the optimized pitch and language strategy in cochlear implant recipients. *Ear and Hearing* 40**,** 555–567.

Vandali, AE, and Van Hoesel, RJM (2012). Enhancement of temporal cues to pitch in cochlear implants: effects on pitch ranking. *Journal of the Acoustical Society of America* 132**,** 392–402.

Venail, F, Mathiolon, C, Menjot De Champfleur, S, Piron, JP, Sicard, M, Villemus, F, et al. (2015). Effects of electrode array length on frequency-place mismatch and speech perception with cochlear implants. *Audiology and Neurotology* 20**,** 102–111.

Vickers, D, Eyles, J, Brinton, J, Glasberg, B, and Graham, J (2009). Conversion of scores between Bamford, Kowal and Bench (BKB) sentences and Arthur Boothroyd (AB) words in quiet for cochlear implant patients. Cochlear Implants International 10, 142–149.

Vihman, M (1993). Variable Paths to Early Word Production. *Journal of Phonetics* 21**,** 61–82.

Vihman, M (2002). The Role of Mirror Neurons in the Ontogeny of Speech. In *Mirror Neurons and the Evolution of Brain and Language,* eds. M. I. Stamenov, V. Gallese. (Amsterdam/Philadelphia: John Benjamins Publishing Company).

Välimaa, TT, Määttä, TK, Löppönen, HJ, and Sorri, MJ (2002a). Phoneme recognition and confusions with multichannel cochlear implants: Vowels. *Journal of Speech, Language, and Hearing Research* 45**,** 1039–1054.

Välimaa, TT, Määttä, TK, Löppönen, HJ, and Sorri, MJ (2002b). Phoneme recognition and confusions with multichannel cochlear implants: Consonants. *Journal of Speech, Language, and Hearing Research* 45**,** 1055–1069.

Walkowiak, A, Lorens, A, Polak, M, Kostek, B, Skarzynski, H, Szkielkowska, A, et al. (2011). Evoked Stapedius Reflex and Compound Action Potential Thresholds versus Most Comfortable Loudness Level: Assessment of Their Relation for Charge-Based Fitting Strategies in Implant Users. *ORL* 73**,** 189–195.

Wang, W, Zhou, N, and Xu, L (2011). Musical pitch and lexical tone perception with cochlear implants. *International Journal of Audiology* 50**,** 270–278.

Waterson, N (1971). Child phonology: a prosodic view. *Journal of Linguistics* 7**,** 179–211.

Waterson, N (1976). "Perception and production in the acquisition of phonology," in *Baby Talk and Infant Speech,* eds. W.V. Raffler-Engel & Y. Lebrun. (Amsterdam: Swets and Zeitlinger).

Wie, OB (2010). Language development in children after receiving bilateral cochlear implants between 5 and 18 months. *International Journal of Pediatric Otorhinolaryngology* 74**,** 1258–1266.

Wie, OB, Falkenberg, E-S, Tvete, O, and Tomblin, B (2007). Children with a cochlear implant: Characteristics and determinants of speech recognition, speech-recognition growth rate, and speech production. *International Journal of Audiology* 46**,** 232–243.

Wilson, B, S. , Finley, CC, Lawson, D, T. , Wolford, RD, Eddington, DK, and Rabinowitz, WM (1991). Better speech recognition with cochlear implants. *Nature* 352**,** 236–238.

Wolfe, J, and Schafer, EC (2015). *Programming cochlear implants.* San Diego, CA: Plural Publishing.

Wouters, J, Francart, T, and Mcdermott, HJ (2015). Sound coding in cochlear implants: from electric pulses to hearing. *IEEE Signal Processing Magazine* 32**,** 67–80.

Yoon, YS, Shin, YR, Gho, JS, and Fu, QJ (2015). Bimodal benefit depends on the performance difference between a cochlear implant and a hearing aid. *Cochlear implants international* 16**,** 159–167.

Appendix 1—Article I

JSLHR

Review Article

Consonant and Vowel Identification in Cochlear Implant Users Measured by Nonsense Words: A Systematic Review and Meta-Analysis

Arne Kirkhorn Rødvik,[a] Janne von Koss Torkildsen,[a] Ona Bø Wie,[a,b]
Marit Aarvaag Storaker,[c,d] and Juha Tapio Silvola[b,c,e]

Purpose: The purpose of this systematic review and meta-analysis was to establish a baseline of the vowel and consonant identification scores in prelingually and postlingually deaf users of multichannel cochlear implants (CIs) tested with consonant–vowel–consonant and vowel–consonant–vowel nonsense syllables.
Method: Six electronic databases were searched for peer-reviewed articles reporting consonant and vowel identification scores in CI users measured by nonsense words. Relevant studies were independently assessed and screened by 2 reviewers. Consonant and vowel identification scores were presented in forest plots and compared between studies in a meta-analysis.
Results: Forty-seven articles with 50 studies, including 647 participants, thereof 581 postlingually deaf and 66 prelingually deaf, met the inclusion criteria of this study. The mean performance on vowel identification tasks for the postlingually deaf CI users was 76.8% (*N* = 5), which was higher than the mean performance for the prelingually deaf CI users (67.7%; *N* = 1). The mean performance on consonant identification tasks for the postlingually deaf CI users was higher (58.4%; *N* = 44) than for the prelingually deaf CI users (46.7%; *N* = 6). The most common consonant confusions were found between those with same manner of articulation (/k/ as /t/, /m/ as /n/, and /p/ as /t/).
Conclusions: The mean performance on consonant identification tasks for the prelingually and postlingually deaf CI users was found. There were no statistically significant differences between the scores for prelingually and postlingually deaf CI users. The consonants that were incorrectly identified were typically confused with other consonants with the same acoustic properties, namely, voicing, duration, nasality, and silent gaps. A univariate metaregression model, although not statistically significant, indicated that duration of implant use in postlingually deaf adults predict a substantial portion of their consonant identification ability.

As there is no ceiling effect, a nonsense syllable identification test may be a useful addition to the standard test battery in audiology clinics when assessing the speech perception of CI users.

The offering of multichannel cochlear implants (CIs) to profoundly deaf and hard-of-hearing adults and children is a well-established medical procedure today, and there are more than 600,000 CI users in the world (The Ear Foundation, 2017). The CI is offered to patients with a large variety of causes for their hearing loss and leads to a considerable improvement in hearing for the majority of users. There is, however, large variability in speech perception outcomes after cochlear implantation (Dowell, Dettman, Blamey, Barker, & Clark, 2002; Rotteveel et al., 2010; Välimaa & Sorri, 2000). Thus, it is critical to have precise measures of how well CI users can perceive different speech sounds. Such measures are important for the fitting of CIs and testing of new implant technology but also for planning and assessing the effects of listening training and speech therapy. In recent years, traditional speech perception tests using sentences and words as stimuli have increasingly produced ceiling or near-ceiling effects in CI users (Blamey et al., 2013). This may be due to a number of factors, such as shorter time of deafness before implantation,

---

[a] Department of Special Needs Education, Faculty of Educational Sciences, University of Oslo, Norway
[b] Oslo University Hospital, Norway
[c] Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Norway
[d] Lillehammer Hospital, Norway
[e] Akershus University Hospital, Lørenskog, Norway
Correspondence to Arne Kirkhorn Rødvik: a.k.rodvik@isp.uio.no

Editor-in-Chief: Frederick (Erick) Gallun
Editor: Matthew Fitzgerald

Received January 9, 2017
Revision received September 4, 2017
Accepted December 18, 2017
https://doi.org/10.1044/2018_JSLHR-H-16-0463

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

increased residual hearing of the implant candidates, and better hearing preservation in CI surgery. There is therefore an increasing need for more difficult tests, which provide fine-grained information on perception of consonants and vowels. Speech perception tests with nonsense words, which are more difficult than real-word tests and less reliant on prior experience with a specific language, appear to be a valuable alternative for future clinical practice and research. However, in order for nonsense word tests to be maximally useful, it is necessary to establish a baseline of the typical level of consonant and vowel perception that CI users achieve on these tests. Additionally, it is important to determine how this baseline relates to performance on other speech perception tests for both prelingually and postlingually deaf CI users. The present systematic review and meta-analysis investigates the typical performance of CI users in nonsense word tests and the influence of some clinically relevant background factors on performance in these tests.

Testing of Speech Perception in CI Users

In the first years after the advent of the CI, speech perception in CI users was assessed more thoroughly and frequently than today, as the CI technology was new and regarded as experimental by many. In these assessments, the CI users were asked to repeat monosyllabic and bisyllabic words to assess their word perception and their consonant and vowel perception and to repeat sentences with and without audiovisual support. Later, with improved implant technology, modified indications for implantation and, thus, improved hearing in the implantees, the test batteries were supplemented with sentences-in-noise tests.

The test batteries for clinical assessment of the quality of hearing in adults and children with CIs today typically consist of monosyllabic words and sentences presented in quiet and with added noise in free field, sometimes also with pure-tone audiometry in free field (Berrettini et al., 2011; Faulkner & Pisoni, 2013; Lorens et al., 2016). Usually, these tests are conducted without the possibility of lipreading, except for the poorest performers.

Testing of the speech perception of CI users is normally done with test lists of real-word monosyllables and sentences in the implantees' native language. Because 80% of the included articles in our meta-analysis are done with English-speaking participants, we will focus on tests with English words in the following paragraph. Speech perception tests in other languages follow the same principles as the tests in English.

A common monosyllabic test is the consonant–vowel nucleus–consonant test created by Peterson and Lehiste (1962). This test is a special case of the consonant–vowel–consonant (CVC) test, which both tests the perception of real words and of speech sounds. The consonant–vowel nucleus–consonant word lists are a set of 10 lists of 50 phonemically balanced words. The test has been controlled for text-based lexical frequency across lists. The Northwestern University Auditory Test No. 6 (NU-6) monosyllable test is another test of word and speech sound recognition with monosyllables in the CVC format, consisting of 50 words and 150 speech sounds (Tillman & Carhart, 1966). Yet, another commonly used test is the Phonetically Balanced Kindergarten Word Test (Haskins, 1949). The test contains four 50-word lists and is still extensively used for assessing speech perception of children who have hearing impairment. All these three tests are commonly used in English-speaking countries and have been adapted to many other languages.

Real-word monosyllable recognition scores have been shown to have a high correlation with audiometric thresholds. In a study by Dubno, Lee, Klein, Matthews, and Lam (1995), a confidence limit for maximum word recognition scores of the NU-6 was obtained from 407 ears in a large group of young and aged subjects with confirmed cochlear hearing losses. The relationship between the pure-tone averages and the maximum word recognition scores on the basis of this study is displayed in a table by Stach (2009, p. 296).

As part of the development of implant technology, the implant companies run clinical studies regularly to test the benefits of new implants, speech processors, or speech-processing strategies. New technology is also tested in CI clinics, wherein company-supported or independent studies are conducted. Standard speech perception tests are used in testing, typically repetition of words or sentences, but also more sophisticated tests involving, for instance, consonant and vowel identification or discrimination (Carlyon, Monstrey, Deeks, & Macherey, 2014; Frijns, Briaire, De Laat, & Grote, 2002; McKay, McDermott, Vandali, & Clark, 1992). A common test design for deciding which one of two or more speech-processing strategies gives the best speech perception for the CI user is to measure the consonant and vowel identification with each of the strategies and, then, compare the scores.

Open- or Closed-Set Tests

Speech perception is usually measured in either open- or closed-set/forced-choice test conditions, depending on what kind of information the clinician is seeking. Open-set tests provide a collection of detailed information about speech perception, listening capacity, and acoustic properties but require a substantial effort from the test leader for posttest analysis. Open-set tests have relatively small learning effects for the patient and can therefore be performed reliably at desirable intervals.

Closed-set tests are quickly performed and easily administered but give limited information about perception of individual speech sounds. The person being tested responds by pushing a button or touching a screen, and the results are interpreted automatically and instantly by a computer. However, the learning effect is considerably larger than in open-set tests because of the limited number of possible answers (Drullman, 2005). In closed-set tests, all participants should perform significantly above chance level.

Tests of monosyllabic and bisyllabic words and sentences have traditionally been performed in open-set conditions, whereas vowel and consonant identification tests have been performed in closed-set conditions. Some commonly used closed-set tests of consonant and vowel identification are those by Hillenbrand, Getty, Clark, and Wheeler (1995); Shannon, Jensvold, Padilla, Robert, and Wang (1999); Tyler, Preece, and Tye Murray (1987); and Van Tasell, Greenfield, Logemann, and Nelson (1992). An open-set test of phoneme recognition and confusion in Finnish is described by Välimaa, Määttä, Löppönen, and Sorri (2002a, 2002b).

Consonant and Vowel Identification

Consonants are part of a heterogeneous group of speech sounds characterized by voicing, duration, manner, and place of articulation. Phonetically, consonants are speech sounds with the air stream passing one or more constrictions on its way from the lungs through the vocal tract.

Vowels are characterized by the tongue position in the mouth cavity and by the lip-rounding. Tongue position can be high, low, back, or front. Normally, vowels are voiced, and the air stream passes frictionless along the middle of the mouth cavity while the tongue is in a static position. The vowel is the nucleus of a syllable, and a syllable can be one vowel alone or a vowel with surrounding consonants. Consonants carry more varied types of phonetic information than vowels, but many of them have lower duration and less acoustic energy. Because of this, vowel sounds are often easier to perceive than consonants, and it is widely accepted that vowels carry most of the intelligibility information in sentences (e.g., Kewley-Port, Burkle, & Lee, 2007).

Previous research has confirmed that CI users have more difficulties identifying consonants and vowels than persons with normal hearing, who typically achieve a score of 95%–100% on consonant and vowel identification tests (Kirk, Tye-Murray, & Hurtig, 1992; Sagi, Kaiser, Meyer, & Svirsky, 2009). In addition, consonant identification scores have usually been measured to be lower than vowel scores. For instance, in two Finnish studies of CI users, it was shown that 24 months after switch-on of the CIs, the average vowel recognition score was 80% and the average consonant recognition score was 71% (Välimaa et al., 2002a, 2002b).

Postlingually deaf CI users often have substantial problems identifying vowels, despite their long duration and high acoustic energy. The reason might be that the first and second formants (F1 and F2) are altered by the implant compared with what the users once used to hear. The same problem applies to the voiced consonants. Therefore, the failure rate in vowel identification by CI users may be as large as, or even larger than, the failure rate for voiced consonant identification.

Consonant and vowel identification tests provide more detailed information about the hearing of CI users than word or sentence tests. Identification of consonants and vowels can be measured both with real-word or nonsense syllable identification tests, and the scoring can be done by counting the number of correctly identified speech sounds. Other commonly used consonant and vowel identification tests have vowel–consonant–vowel (VCV) or consonant–vowel (CV) nonsense syllables as stimuli, and the consonants are typically presented in an [ɑ, i] or [u] context with the target consonant in medial or initial position.

Different vowel contexts give somewhat different test results for the identification of consonants because the formant transitions of the first and second formants differ in the vowel–consonant or consonant–vowel transition phase for the different vowels and consonants. The advantages and disadvantages of the different vowel contexts have been thoroughly evaluated by Donaldson and Kreft (2006), who concluded that the choice of vowel context has small but significant effects on consonant-recognition scores for the average CI listener, with the back vowels /ɑ/ and /u/ producing better performance than the front vowel /i/.

In typical vowel identification tests, vowels are presented in CVC or CV contexts, for example, in hVd, bVd, wVb, or bVb context, or alone. The hVd vowel-test (Hillenbrand et al., 1995; Tyler, Preece, & Lowder, 1983) has been widely used with English-speaking CI users, although vowels in hVd context form real words in English (Munson, Donaldson, Allen, Collison, & Nelson, 2003).

Although a large number of studies have been published on the subject of speech perception in CI users, there is no international consensus or standard on how to measure the identification of vowels and consonants. Several countries use nationally standardized tests for speech perception measurements. An overview of different speech perception tests (sentence identification, CVC words, and number triplets) in Danish, Dutch, (British) English, French, German, Polish, and Swedish is given in a report from the European HEARCOM project (Drullman, 2005). However, this document only reports the use of meaningful CVC words (i.e., not nonsense words) for consonant and vowel identifications.

Consonant and Vowel Confusions

Since the early 1980s, it has been common to carry out investigations of consonant and vowel confusions to assess the benefits of CIs in speech perception (e.g., Clark et al., 1981). Acoustic similarity has usually been identified as the most important variable to explain confusions of speech sounds (Fant, 1973). Consonant and vowel confusion studies have been conducted in several languages, among them English (Baskent & Shannon, 2004; Bhattacharya & Zeng, 2007), Flemish (Van Wieringen & Wouters, 1999; Wouters & van den Berghe, 2001), and Finnish (Välimaa et al., 2002a, 2002b; Välimaa, Sorri, Laitakari, Sivonen, & Muhli, 2011).

In vowel and consonant recognition studies of postlingually deaf adult CI users, some predominant confusions have been identified. Van Wieringen and Wouters (1999) tested vowel and consonant recognition in Flemish-speaking CI users and found that /y/ was often confused with /e/ and that /ɪ/ is often confused with /ə/, showing that vowel length was recognized correctly. The consonant /t/ was often confused with /k/, and /ɤ/ was often confused with /z/, indicating that voicing and manner of articulation were recognized correctly. Munson et al. (2003) found that English-speaking CI users often confused /ɛ/ with /ɪ/ and /ɪ/ with /ɛ/, concluding that they recognized vowel length. Moreover, /d/ was confused with /g/ and /θ/ with /f/, concluding that they recognized voicing and manner of articulation. Välimaa et al. (2011) presented longitudinal data of vowel recognition and confusion patterns in Finnish informants from before CI surgery until 4 years post implantation. They also studied the effect of duration of profound hearing impairment before implantation and the effect of the use of different implant devices after implantation. After 4 years, the most frequent confusions were /ø/ perceived as /æ/ and /e/ perceived as /ø/ or /æ/, which led to the conclusion that the Finnish front vowels were the most difficult to distinguish. This is in agreement with previous studies showing that vowels with smaller spectral differences are often the most difficult to identify (Munson et al., 2003; Skinner, Fourakis, Holden, Holden, & Demorest, 1996; Van Wieringen & Wouters, 1999).

A widely used method for evaluation of the transmission of speech features is described in an article by Miller and Nicely (1955). Their method of classifying the consonant confusions by arranging them into confusion matrices (CMs) and calculating the information transmission of the linguistic features voicing, nasality, affrication, duration, and place of articulation is still in use.

Nonsense Syllable Test Words

Nonsense syllables have no meaning but are typically phonotactically legal in the language of the listener. The primary advantage of using nonsense syllables instead of real words to measure vowel and consonant identification is that the informant cannot guess which word is presented but has to rely on his or her hearing alone. Thus, the influence of other cognitive factors, such as vocabulary and inferential skills, is reduced compared with when conducting the test with real words. Consequently, nonsense syllable tests tend to be more difficult than real-word tests, as the stimuli ideally do not match any existing representation in the user's mental lexicon.

Another advantage of nonsense syllable tests is that learning effects in multiple experiments with the same stimuli are very small compared with tests using real-word stimuli (Dubno & Dirks, 1982). Thus, it is possible to use the same nonsense syllable test for repeated examination of speech perception in the same individual to check for progress in listening ability.

Nonsense syllables are convenient to use in experiments measuring speech perception. In his classical article, Glaze (1928) showed that experiments using nonsense syllables evoke fewer associations in the participants and thus reduce between-participants variability in test results compared with experiments using real words.

Studies using nonsense syllables as stimuli can be compared across languages as long as the included speech sounds in the tests exist in both languages and a few such studies have been conducted (e.g., Pelizzone, Cosendai, & Tinembart, 1999; Tyler & Moore, 1992).

Nonsense words used in studies of speech perception usually contain only one or, at most, two syllables to avoid the influence of possibly poor phonological working memory span on performance. However, some studies have used tests, such as the Children's Test of Nonword Repetition (Gathercole, Willis, Baddeley, & Emslie, 1994) and other nonsense word tests primarily constructed to assess children's working memory span and cognitive abilities, to study speech perception (Burkholder-Juhasz, Levi, Dillon, & Pisoni, 2007; Casserly & Pisoni, 2013; Nakeva Von Mentzer et al., 2015). The nonsense word test battery of Gathercole et al. (1994) contains nonsense words with two, three, four, and five syllables, but even the bisyllabic nonsense words are poorly suited to measure vowel and consonant identification, as the same vowel or consonant can be found several times in the same word in different positions and several times in the same test sequence. This makes it more complicated to measure the prevalence of consonant or vowel confusions.

Milestones in the Development of CI Technology

A significant advance in the CI technology was the transformation from single-channel to multichannel implants in the beginning of the 1980s. The single-channel implants provided limited spectral information and very rarely gave open speech understanding, as only one site in the cochlea was stimulated. Multichannel implants with four channels and more, however, provide electrical stimulation at multiple sites in the cochlea with an electrode array and can also convey frequencies covering most of the frequency range of the speech sounds. All multichannel strategies are spectral resolution strategies, as they convey spectral information to the implantees.

The stimulation strategies of the early multichannel implants were either analog or pulsatile. The main difference between the two groups of strategies is that the first employs simultaneous stimulation, whereas the latter employs sequential stimulation. A major disadvantage with the analog stimulation strategy is channel interaction, an effect that obstructs speech perception by sound distortion. This problem is less prevalent in pulsatile, nonsimultaneous stimulation. All the stimulation strategies currently used are pulsatile.

The discontinued implants from Ineraid/Symbion and from University of California, San Francisco/Storz employed the compressed analog (CA) stimulation strategy. The CA strategy was also employed by Advanced Bionics in their previous implants. Some years later, Advanced Bionics released simultaneous analog stimulation, which is a modified CA strategy. This strategy was applied until the mid-2000s. Several clinical studies have demonstrated open speech understanding with analog stimulation strategies (e.g., Dorman, Hannley, Dankowski, Smith, & McCandless, 1989), and several studies have also compared implants running pulsatile and analog stimulation (Tyler et al., 1996; Tyler, Lowder, Parkinson, Woodworth, & Gantz, 1995; Xu, Zwolan, Thompson, & Pfingst, 2005). The results have pointed toward better speech perception with pulsatile stimulation than with analog, although there has been large variability in the outcomes. Analog strategies are not used in CI processors today.

Variables Influencing Speech Perception in CI Users

It has been shown in many studies that there is a large variability in speech recognition performance of CI users (Dowell et al., 2002; Rotteveel et al., 2010; Välimaa & Sorri, 2000). For a given type of implant, auditory performance may vary from 0% to 100% correct, and thus, the individual differences between CI users appear to be vastly larger than the effect of implant manufacturer. Auditory performance is here understood as the ability to discriminate, detect, identify, or recognize speech. A typical measure of auditory performance is the percentage correct score on open-set speech recognition tests. The review article by Loizou (1999) lists the following factors that have been found to affect auditory performance: the duration of deafness prior to implantation (a long duration appears to have a negative effect on auditory performance), age of onset of deafness (younger age is associated with better outcome), age at implantation (earlier implantation is associated with better outcome for prelingually deaf subjects), and duration of CI use (longer duration of CI experience is associated with better outcome). Other factors that may affect auditory performance include etiology of hearing loss, number of surviving spiral ganglion cells, electrode placement and insertion depth, electrical dynamic range of the CI, cognitive abilities, duration of hearing aid use before implantation, and signal processing strategy (Blamey et al., 2013, 2015; Rotteveel et al., 2010; Spencer, 2004; Wie, Falkenberg, Tvete, & Tomblin, 2007).

It is critical to be aware of the influence of these factors when assessing and evaluating speech perception outcomes in CI users. Furthermore, it should be kept in mind that the influence of these and other factors on speech perception may be different for prelingually and postlingually implanted children and adults.

Some studies have even found that age at implantation is not a significant predictor of speech perception outcome for prelingually deaf children (e.g., Geers, Brenner, & Davidson, 2003; Wie et al., 2007). Wie et al. (2007) found that the variations in performance on speech perception tasks could be explained by daily user time, nonverbal intelligence, duration of CI use, educational placement, and communication mode (use of sign language or spoken language). The authors explained this result by the relatively high age at implantation for the participants in the study, as only one participating child was implanted before 24 months of age.

For a group of 65 postlingually implanted adults, Plant, McDermott, van Hoesel, Dawson, and Cowan (2016) showed different factors which predicted word recognition scores for unilaterally and bilaterally implanted CI users. For the unilaterally implanted group, predictors included a shorter duration of severe-to-profound hearing loss in the implanted ear and poorer pure-tone-averaged thresholds in the contralateral ear. For the bilateral group, shorter duration of severe-to-profound hearing loss before implantation, lower age at implantation, and better contralateral hearing thresholds were associated with higher bilateral word recognition in quiet and speech reception threshold in noise.

Transmission of Consonants and Vowels in an Implant

The transmission of consonants and vowels in CIs is designed to reproduce a speech signal that closely resembles the original by means of electrical stimulation patterns in the CI electrode. Failure to resemble the original signal is always explained from two viewpoints: limitations in the hearing system of the implant user caused by different variables (cf. previous section) and technical limitations in the CI system. In a CI user with optimal conditions for the reception of speech, some important factors for the transmission of speech are the speech coding, the length and insertion depth of the implant, the input dynamic and input frequency range of the speech processor, and implant electrode properties.

Vowels are characterized by long duration and high energy compared with consonants, and as such, they are easily perceived by the implantees. Furthermore, vowels are characterized mainly by F1 and F2, the first two formants, which can be found in the frequency range between 200 Hz and 2500 Hz. Thus, provided the input frequency range of the implant includes frequencies as low as 200 Hz, all vowels should be possible to recognize.

For the perception of pitch, the insertion depth of the implant plays an important role. The tonotopy of the cochlea is organized with the low frequency sounds in the apical region and the high frequency sounds in the basal region. When the more apical part of cochlea is stimulated, darker pitch is received by the implantee. Thus, one should expect that users of the implants with the longest electrodes, like Med-El's, would obtain best pitch perception. However, this is not always the case.

Some stimulation strategies are supposed to be better for the perceptions of voiced sounds than others. For example, the FSP/FS4/FS4-p strategies from Med-El will code the fundamental frequencies on the most apical electrodes in addition to running ordinary continuous interleaved sampling (CIS) stimulation. The HiRes120 strategy from Advanced Bionics is marketed as being supposed to improve the spatial precision of stimulus delivery and be more suitable for the perception of pitch and music than spectral envelope strategies like CIS or Advanced Combination Encoder (Wouters, Francart, & McDermott, 2015).

The microphone sensitivity in the speech processors plays an important role in the perception of soft sounds, and the higher the microphone sensitivity is, the better these speech sounds are picked up. None of the implants have problems with picking up soft speech sounds, as long as the sounds are within the input frequency range of the speech processor.

Consonants are a more heterogeneous group of speech sounds than the vowels. They can be characterized, for example, by long or short duration, by voicing or nonvoicing, or by being nasal or nonnasal. Many of the consonants, especially the unvoiced stops and fricatives, have high frequency parts, which are easily picked up by the CI speech processors. Earlier research has shown that acoustic similarity of the consonants is the most important reason for confusion (Fant, 1973), as implant users most frequently confuse consonants that are pronounced in the same manner but with a constriction in different places in the mouth cavity. Consonants that are pronounced with different manner in the same place are seldom confused. Furthermore, CI users have more trouble distinguishing between voiced consonants than between unvoiced and have the most trouble distinguishing between nasals and laterals.

Cognitive explanatory factors obviously play an important role in the perception of consonants and vowels but are outside the scope of this discussion.

Aim and Research Questions

The aim of this systematic review and meta-analysis was to examine previous research in order to investigate how well users of multichannel CIs identify consonants and vowels in tests using monosyllabic and bisyllabic nonsense words as stimuli. We wanted to ascertain the baseline of consonant and vowel perception in previous nonsense word research, use aggregated empirical findings and measurements to increase the statistical strength, and pool these studies in a meta-analysis. Specifically, we aimed to investigate the following three research questions:

1. What are the typical vowel and consonant identification scores in CI users when measured by nonsense syllables, and how do the typical vowel and consonant identification scores differ between prelingually and postlingually deaf implantees?
2. Which consonants and vowels are most frequently confused by CI users, and which consonants and vowels are most frequently identified correctly?
3. (a) To what extent are age at implantation, duration of implant use, and real-word monosyllable score associated with variations in consonant and vowel identification performance in nonsense syllable tasks for prelingually deaf CI users?
   (b) To what extent are duration of implant use and real-word monosyllable score associated with variations in consonant and vowel identification performance in nonsense syllable tasks for postlingually deaf CI users?

To our knowledge, a systematic review and meta-analysis of studies on consonant and vowel identification in CI users tested by nonsense syllables has not been published before.

Method

This systematic review was conducted in accordance with the 27-item checklist in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (Moher, Liberati, Tetzlaff, & Altman, 2009).

Details of the systematic review protocol were registered with PROSPERO, the international prospective register of systematic reviews, on December 15, 2014. The protocol is available online at: http://www.crd.york.ac.uk/prospero/display_record.asp?ID=CRD42014015141.

The systematic review was performed in the following steps:

- Literature search.
- Screening of articles for inclusion and exclusion.
- Extraction of information from the articles (coding).
- Pooling of data for statistical analysis.

A flow diagram displaying the process from searching, via screening and eligibility to the final number of included articles, is shown in Figure 1. The diagram is based on a template designed by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Moher et al., 2009).

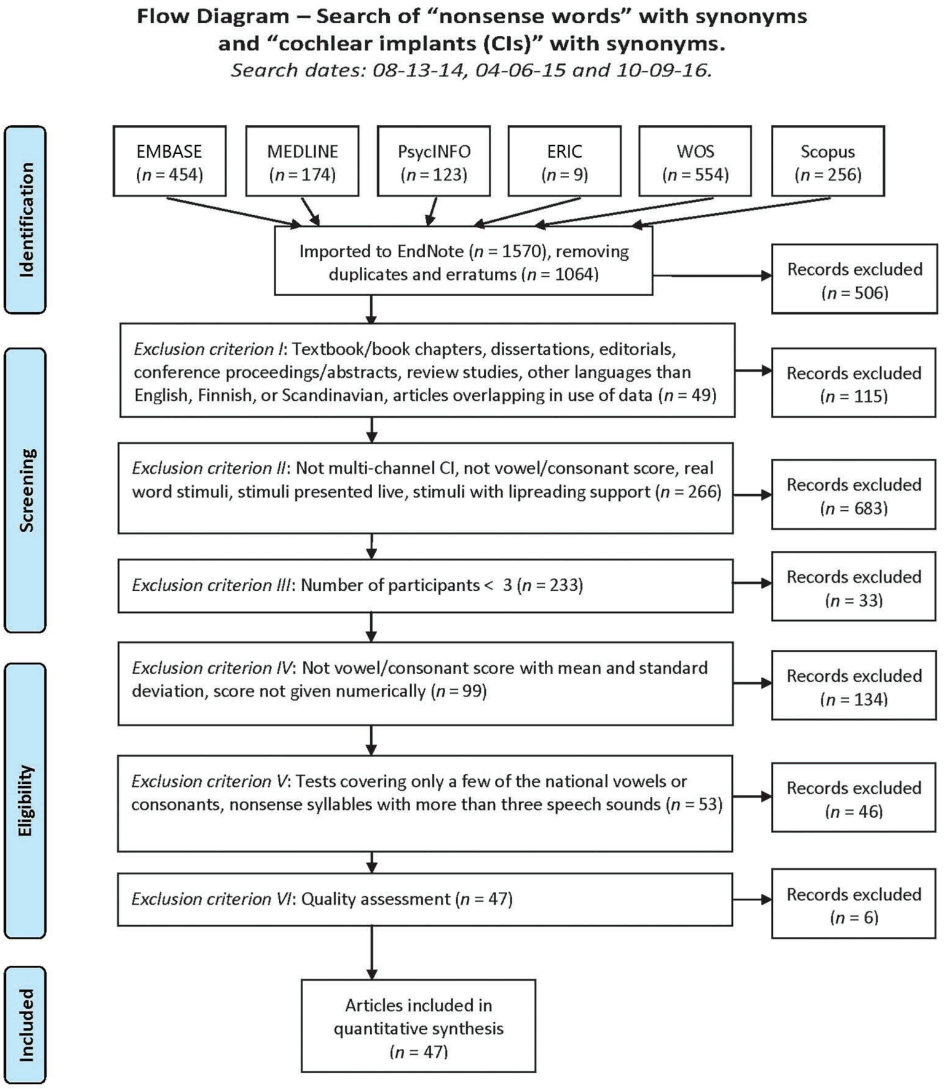
The forest plots displayed in Figures 2, 3, and 4 were generated by means of the software Comprehensive Meta-Analysis (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2014).

Literature Searches

Detailed searches for primary and retrospective studies were performed in the following six databases: EMBASE, MEDLINE, PsycINFO, ERIC, Web of Science/Web of Knowledge, and Scopus. Initially, the databases Cochrane Library, Speech Bite, Svemed, Pubpsych, Proquest, Norart, Researchgate.com, and Academia.edu were also searched by the review team, but these searches returned no results.

The searches were run three times on August 13, 2014, April 6, 2015, and October 9, 2016 and were limited to peer-reviewed journal articles written in English, in Scandinavian languages (Norwegian, Swedish, and Danish), and in Finnish. The search strings consisted of two elements: (a) various terms referring to nonsense words and speech discrimination and (b) terms referring to CIs. All the search elements were truncated in order for the searches to include all conjugations of the nouns. Truncation was represented by an asterisk (*).

Figure 1. Flow diagram, searches for “nonsense words” with synonyms and “cochlear implants” with synonyms. CI = cochlear implant; EMBASE = Excerpta Medica Database; MEDLINE = Medical Literature Analysis and Retrieval System Online; PsycINFO = Psychological Information Database; ERIC = Education Resources Information Center; WOS = Web of Science. Copyright © 2009 Moher et al. (Creative Commons Attribution License).



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

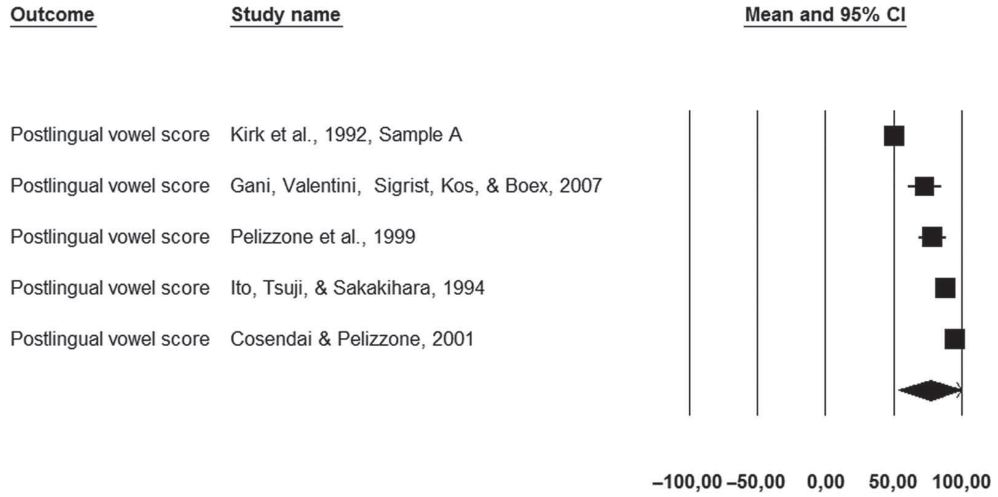
For more information, visit www.prisma-statement.org.

(a) Nonsense word repetition with the synonyms nonword*, NW*, nonsense word*, pseudo word*, nonsense syllable*, nonword syllable*, pseudo syllable*, CV* word*, VC* word*, speech sound repetition*, speech sound recognition*, speech sound confusion*, speech sound identification*, speech sound discrimination*, speech sound perception*, phoneme repetition*, phoneme recognition*, phoneme confusion*, phoneme identification*, and phoneme discrimination*.

(b) Cochlear implants with the synonyms CI, cochlear prosthes*, hearing aid*, sensory aid*, hearing instrument*, and hearing device*.

Because “cochlear implant” is an unambiguous concept, unlike “nonsense word repetition,” the number of search terms in (b) turned out to be considerably lower than in (a). The complete search syntaxes for the four Ovid databases EMBASE, MEDLINE, PsycINFO, and ERIC, as well as for Web of Science and Scopus, are listed in the Appendix.

Figure 2. Forest plot of vowel identification scores for postlingually deaf cochlear implant users. The primary studies are represented by boxes, which are bounded by the confidence interval (CI) for the effect sizes in each study. The effect sizes are measured in percent.



Screening of Abstracts and Review of Full-Text Articles

The search results were imported into EndNote, v. X7.7.1 (Thompson Reuters), for removal of duplicates, books and book chapters, dissertations, editorials, systematic reviews, and articles in languages other than Danish, English, Finnish, Norwegian, and Swedish. Thereafter, the references were imported into the web-based systematic review software DistillerSR (EvidencePartners), which was used for the screening process.

Assessment of articles was performed in two phases: (a) screening of abstracts and titles and (b) full-text review of the remaining articles, as described in Figure 1. In Phase (a), two researchers (the first author, AKR, and the fourth author, MAS) independently evaluated all the identified titles and abstracts and excluded the studies missing one or both of the search terms *cochlear implants* and *nonsense words* with synonyms. Disagreements were solved by discussion or by reading the full text of the articles. Further on, the abstracts were screened by AKR for number of participants, and studies with less than three participants

Figure 3. Forest plot of consonant identification scores for prelingually deaf cochlear implant users. The primary studies are represented by boxes, which are bounded by the confidence interval (CI) for the effect sizes in each study. The effect sizes are measured in percent.

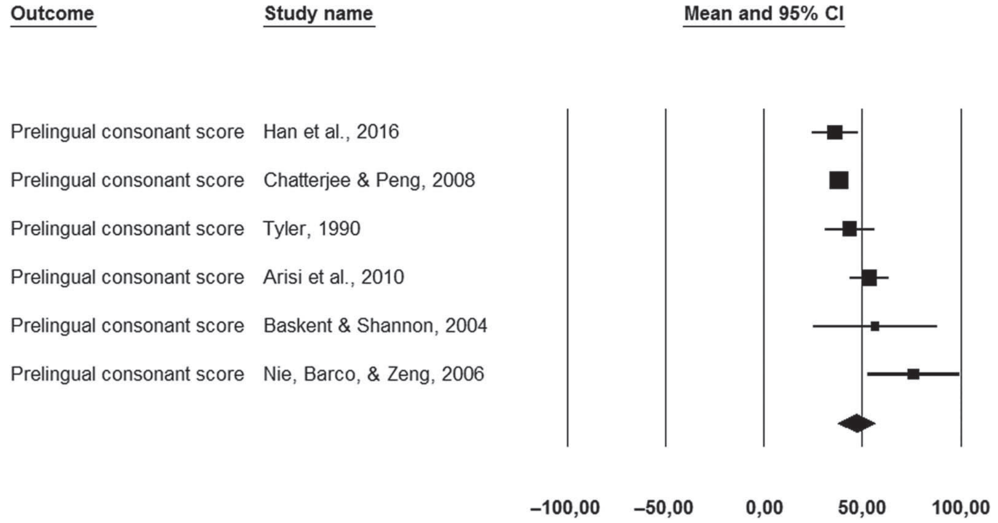
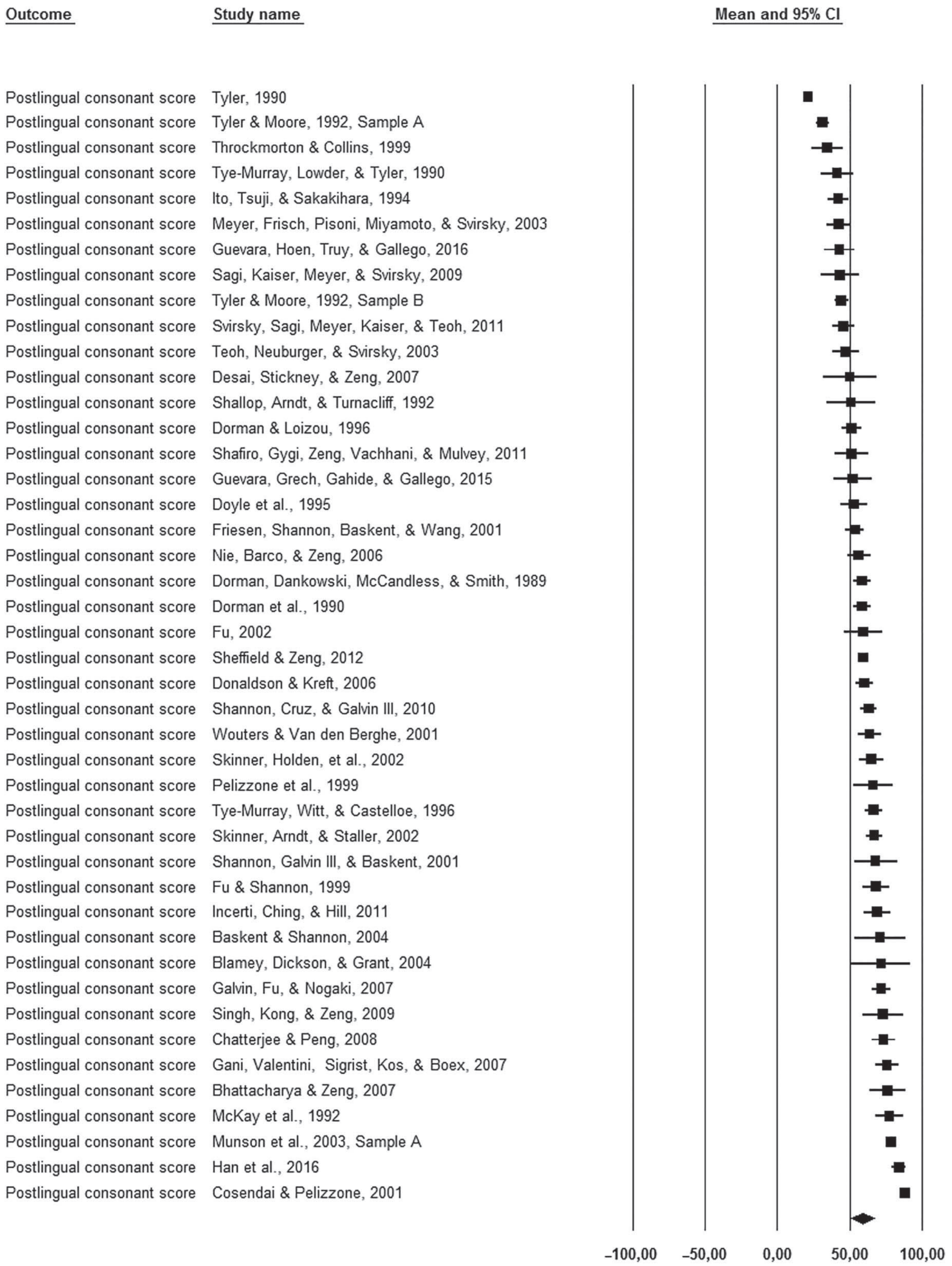


Figure 4. Forest plot of consonant identification scores for postlingually deaf cochlear implant users. The primary studies are represented by boxes, which are bounded by the confidence interval (CI) for the effect sizes in each study. The effect sizes are measured in percent.



were excluded, as case studies with one or two participants did not fit into the methodology of the systematic review.

In Phase (b), full-text articles were reviewed according to exclusion Criteria IV and V in Figure 1. During this phase, some of the articles were also excluded according to Criterion I, II, or III when this applied. Further details on the inclusion and exclusion criteria are found in the subsequent paragraphs.

Inclusion Criteria

Inclusion criteria were based on the Participants, Intervention, Control, Outcomes, and Study designs strategy (Santos, Pimenta, & Nobre, 2007; see Table 1).

The included articles described studies with three participants or more. We focused on the outcome of consonant and vowel identification tests measured by nonsense words in free field 6 months or more after implantation. If use of repeated measures in longitudinal studies was reported in the article, we registered the most recent nonsense word scores. If different nonsense word tests for the same groups of participants were used, for example, in Kirk et al. (1992), we included the test that provided results with the highest score. If the article referred to other articles by the same authors for more details about the tests, we extracted the necessary information from these.

Exclusion Criteria

- Studies on participants with single-channel CIs were excluded. This was based on research showing that implants need at least four channels to provide adequate speech perception in quiet (Cohen, Waltzman, & Fisher, 1993; Tyler et al., 1988).
- Studies measuring consonant or vowel score by real-word stimuli and not by nonsense syllables were excluded.
- Studies measuring consonant or vowel score by nonsense words with three or more syllables were excluded, as it is difficult to disentangle effects of working memory span from hearing when interpreting these results. In addition, the same target consonants or vowels are often presented more than once in such multisyllable test words.
- Studies assessing the identification of less than about 50% of the national inventory of vowels and consonants were excluded, as these studies presented vowel and consonant identification scores on the basis of too few consonants and vowels to represent the phoneme inventory of this language. For instance, there are 20–24 consonants in English, depending on the dialect, and for the study to be included, at least half of these had to be used to calculate a consonant identification score.
- Studies in which means and standard deviations of the consonant and vowel identification score were not reported, only reported graphically in diagrams, or could not be calculated from confidence intervals or standard errors were excluded. For those excluded studies published less than 10 years ago, we wrote to the corresponding author to ask for the raw data from the study. Studies from which the raw data were received were included in the meta-analysis.
- Studies in which nonsense words were presented live instead of recorded were excluded because of less expected consistency in the test results than in recorded materials (Mendel & Owen, 2011).
- Studies in which the stimuli were presented with lip-reading support were excluded.
- Studies using synthesized or electronically generated test stimuli were excluded.
- Studies displaying speech sound scores not separated into a vowel and a consonant score were excluded.
- Studies in which the identification score for consonants was only reported as categories according to consonant properties like place, manner, or voicing (e.g., Nelson, Van Tasell, Schroder, Soli, & Levine, 1995) were excluded.
- In those cases where different articles were based on the same study participants and/or the same data, all but one of these articles were excluded. The article that included the highest number of participants was selected for further analysis.
- Studies including participants with a contralateral hearing aid in addition to an implant were excluded unless it was clearly stated in the article that the

Table 1. PICOS criteria for inclusion in the systematic review and meta-analysis.

Acronym	Definition	Application of the criteria on the present study
P	Participants	Adults and/or children with one or two multichannel CIs
I	Intervention	None
C	Control	Studies included both with and without control group
O	Outcomes	Consonant and/or vowel identification scores, measured by nonsense words
S	Study designs	Cross-sectional studies, longitudinal studies, case studies ($N \geq 3$)

Note. PICOS = Participants, Intervention, Control, Outcomes, and Study designs (adapted from Santos, Pimenta, & Nobre, 2007); CI = cochlear implant.

benefit of the implant was better than the benefit of the hearing aid.

Risk of Publication Bias

Risk of publication bias was commented on qualitatively and by inspection of funnel plots generated in CMA. A symmetrical funnel plot could indicate the absence of publication bias. However, an asymmetrical funnel plot could indicate several conditions, for instance, heterogeneity, publication bias, or chance, and the interpretation of the asymmetry with regard to publication bias has been highly disputed in previous research (Lau, Ioannidis, Terrin, Schmid, & Olkin, 2006; Sterne et al., 2011). Although it is common in meta-analyses to correct the asymmetry in funnel plots by the "Trim-and-fill" method, we chose not to make use of this technique in our study, as there are substantial methodological problems related to it (Lau et al., 2006). Effect sizes may be underestimated when publication bias does not exist and overestimated when publication bias does exist, and thus, it can be argued that the method is inadequate as a corrective technique (Simonsohn, Nelson, & Simmons, 2014). Therefore, we chose not to draw definite conclusions about publication bias in the case of asymmetry.

Quality Assessment

Publications considered to be of weak overall quality by the review team were excluded from the systematic review. These quality criteria were

- inconsistent presentation of results;
- errors in the analyses; and
- lack of transparency, for example, missing description of the study methods.

Selection and Coding of Data

A pilot coding was performed on 11 articles by MAS, to test the strength of the categories in the coding form. After this, an evaluation of the pilot coding was performed by the review team to develop the final coding form, in which the selection of coding parameters was done based on our research questions. The following data were extracted from the articles: author, title of article, publication year, journal, aim, language, and study design, and absence or presence of a control group. For studies including participants with an implant, the following measures were coded: number of participants; number of postlingually/prelingually implanted participants; number of participants with auditory neuropathy spectrum disorder; implant type; speech-processing strategy; age at testing; age at implantation; duration of implant use; duration of deafness before implantation; age at onset of deafness; stimulation level; number of unilaterally or bilaterally, sequentially, or bimodally implanted participants; identification score for vowels; most confused vowel; identification score for consonants; most confused consonant; monosyllable real-word identification score; and score from postoperative audiometric measurements. For participants with normal hearing serving as control groups, the following measures were coded: number of participants, identification score for vowels, most confused vowel, identification score for consonants, most confused consonant, and monosyllable real-word identification score. The data were extracted to the form by AKR.

Strategy for Data Synthesis

Both aggregate and individual participant data were used. We used quantitative methodology on the included studies, which were sufficiently homogeneous. Vowel and consonant identification scores and vowel and consonant confusions were compared between studies and between languages, despite cross-linguistic differences (Tyler & Moore, 1992).

Analysis

Our meta-analysis included studies reporting means and standard deviations. A random effects model was chosen over a fixed effects model to average the effect sizes across studies, as this does not assume a shared common true effect (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Research Question 1, "What are the typical vowel and consonant identification scores in cochlear implanted participants when measured with nonsense syllables, and how do the typical vowel and consonant identification scores differ between prelingually and postlingually deaf implantees?" was answered statistically by pooling of the studies in CMA. Individual consonant and vowel identification scores were weighted by the random effects model, averaged across studies and presented as forest plots in Figures 2, 3, and 4.

To answer Research Question 2, "Which consonants and vowels are most frequently confused by CI users, and which consonants and vowels are most frequently identified correctly?" we constructed meta CMs to display the three most common vowel and consonant confusions, from the 11 studies in which this information was available. In some articles, this information was given qualitatively, and in these cases, our presentation of the results was also given qualitatively.

To answer Research Question 3, only users with postlingual deafness were included in the analysis, as very few studies reported consonant and vowel scores for the prelingually deaf group. We performed a univariate regression analysis with the weighted mean consonant identification score against duration of CI use. Real-word monosyllable score and vowel identification score were omitted as independent and dependent variables in the analyses because this was only reported in 17 studies and 6 studies, respectively. We obtained beta regression coefficients to characterize the univariate relationship and explained the percentage of between-studies variance by using R^2, which quantifies the proportion of variance explained by the covariates (Borenstein et al., 2009).

Results

Study Characteristics

The results are based on analyses of the 50 studies reported in the 47 included articles, and the study characteristics are summarized in Table 2 and below. The articles that met our inclusion criteria were published between 1989 and 2016. Three of these articles were treated as two independent studies each in the meta-analysis, with different participants in each study (Kirk et al., 1992; Munson et al., 2003; Tyler & Moore, 1992). In 38 of the studies, the participants were speaking English, and 32 of these studies had participants with American English as their mother tongue. In eight of the remaining nine studies, the participants spoke either Flemish, French, German, Italian, or Japanese. In the final study, the participants reportedly spoke one out of seven mother tongues, namely, Albanian, French, German, Italian, Russian, Spanish, and Swahili (Pelizzone et al., 1999). The large majority of participants (581 of 647) were reported as postlingually deaf and the rest (66) as prelingually deaf. As the criteria for prelingual and postlingual deafness differed between studies, and often were not reported, we used the studies' own report of prelingual and postlingual deafness in our statistics.

Six hundred thirteen participants were unilaterally implanted, 10 bilaterally and 24 bimodally. The number of participants per study varied between three and 56. Three articles described CI users with a hearing aid on the contralateral ear (bimodal users; Gani, Valentini, Sigrist, Kos, & Boex, 2007; Incerti, Ching, & Hill, 2011; Sheffield & Zeng, 2012). From these articles, we included in our meta-analysis only the results obtained without a hearing aid. In one of the articles, the participants' vowel perception was tested both with wVb and with bVb words (Kirk et al., 1992). According to our inclusion criteria stating that the participants should not be represented in the material more than once, we chose to use the bVb words in our analyses, as these gave the highest mean score of vowel perception.

The participants used implants from the CI manufacturers Advanced Bionics, Cochlear, Digisonic/Neurelec, Ineraid/Symbion, Laura, and Med-El. Many studies reported results from participants with implants from more than one manufacturer and results from studies in which one implant used several stimulation strategies, thus it was not always possible to pool results per implant model or per stimulation strategy.

The mean age at onset of deafness was 31.6 years (*SD* = 18.0 years, range = 2.6–52.4 years), reported in 28 studies, and the duration of profound deafness before CI was 14.8 years (*SD* = 8.1 years, range = 2.7–38.9 years), reported in 29 studies.

Only two of the included studies had children or adolescents as participants (Arisi et al., 2010; Tyler, 1990). In a study by Tyler (1990), the five children who participated had a mean age of 8.5 years (*SD* = 1.6 years, range = 6.8–10.3 years) and obtained a consonant identification score of 30% (*SD* = 13.2%, range = 19%–50%). In a study by Arisi et al. (2010), 45 adolescent participants had a mean age of 13.4 years (*SD* = 2.6 years, range = 11–18 years) and obtained a consonant identification score of 53.5%.

Research Question 1: What are the Typical Vowel and Consonant Identification Scores in CI Users When Measured by Nonsense Syllables, and How Do the Typical Vowel and Consonant Identification Scores Differ in Prelingually and Postlingually Deaf Implantees?

Table 3 shows the vowel and consonant identification scores for the studies with prelingually deaf participants, the studies with postlingually deaf participants, and for the whole sample of 50 studies. All scores are weighted by the random effects model (Borenstein et al., 2009). Only five studies reported scores on vowel identification for the postlingually deaf (Cosendai & Pelizzone, 2001; Gani et al., 2007; Ito, Tsuji, & Sakakihara, 1994; Kirk et al., 1992; Pelizzone et al., 1999). Four of these studies (including 30 participants) reported both consonant and vowel identification scores. For the prelingually deaf, a vowel score for one CI user was reported in only one article, which also reported a consonant score for the same user (Gani et al., 2007). Another article reported the consonant score of one prelingually deaf CI user (Bhattacharya & Zeng, 2007). These scores could not be included in the analyses because of an *SD* of 0. Finally, vowel identification scores for the normal-hearing group were only calculated in one study, and a mean score of 98.3% (*SD* = 1.0%) was reported (Kirk et al., 1992).

Consonant identification scores were reported in 46 articles (48 studies). Four of these articles had to be excluded because the consonant scores could not be split into one score for the prelingually deaf and one for the postlingually deaf (Kirk et al., 1992; Munson et al., 2003; Stacey et al., 2010; Van Wieringen & Wouters, 1999). Consonant identification scores were not reported for any of the normal-hearing control groups, which were included in 13 of the studies. In many of these studies, the control group was used for calibrating the consonant and vowel identification test in the local dialect. This was done by requiring a score of 95% or higher on the test by the control group, before the test could be used for testing cochlear-implanted participants. If the score for the control group turned out to be lower than the limit set in the study, the consonant identification test was modified to get the score above the limit, for instance, by removing nonsense syllables with high failure rates from the test, for example, certain test words pronounced in a dialect little known to the participants.

In Figures 2, 3, and 4, the vowel and consonant identification scores are presented as forest plots, showing the weighted mean and the 95% confidence interval for each study, arranged in ascending order. Ceiling effects were observed in the individual scores of the included studies, especially in the vowel scores.

Only five studies reported consonant identification scores for both the prelingually and postlingually deaf CI

Table 2. Study characteristics, task characteristics, and results for the 50 included studies.

Authors	*N*		Language of participants	Stimulus context	No. of consonants and vowels in the test	Age (years) of implantation, *M* (*SD*)	Duration (years) of implant use, *M* (*SD*)	Speech sound score (%)		Monosyllables	
	CI	NH						Vowels, *M* (*SD*)	Consonants, *M* (*SD*)	Score (%), *M* (*SD*)	Name of test
Arisi et al., 2010	45	0	Italian	VCV	—	13.4 (2.6)	> 3	—	53.5 (33.6)	—	—
Baskent & Shannon, 2004	6	0	English (USA)	aCa	20	38.3 (13.6)	3.0 (1.5)	—	63.6 (21.7)	—	—
Bhattacharya & Zeng, 2007	7	6	English (USA)	aCa	20	63.3 (10.7)	3.4 (1.8)	—	68.4 (23.6)	—	—
Blamey et al., 2004	4	0	English (Australia)	aCa	16	54.0 (21.6)	3.3 (2.6)	—	71.1 (20.9)	—	—
Chatterjee & Peng, 2008	10	4	English (USA)	VCV	20	53.0 (16.5)	7.0 (5.0)	—	66.0 (17.0)	—	—
Cosendai & Pelizzone, 2001	3	0	French (Switzerland)	aCa, V	14 (VCV), 7 (V)	32.0 (12.8)	10.0 (4.4)	95 (2.7)	88.0 (1.0)	—	—
Desai et al., 2008	8	14	English (USA)	VCV	20	62.5 (13.9)	3.8 (3.5)	—	49.6 (26.7)	—	—
Donaldson & Kreft, 2006	20	0	English (USA)	aCa, iCi, uCu, Ca, Ci, Cu	19	55.2 (11.4)	3.3 (3.6)	—	59.8 (13.9)	—	—
Dorman, Dankowski, et al., 1989	10	0	English (USA)	aCa	16	Adult	—	—	58.1 (9.8)	—	—
Dorman & Loizou, 1996	7	0	English (USA)	aCa	16	Adult	> 4	—	51 (9)	—	—
Dorman et al., 1990	10	0	English (USA)	aCa	16	Adult	—	—	58.2 (9.7)	41.3 (17.5)	NU-6
Doyle et al., 1995	14	0	English (USA)	iCi	14	54.3 (15.9)	1.0 (0.6)	—	52.7 (17.2)	—	—
Friesen et al., 2001	19	5	English (USA)	aCa	14	59.0 (13.3)	2.6 (2.2)	—	53.2 (14.4)	44.5 (20.2)	CNC
Fu, 2002	9	0	English (USA)	aCa	16	46.9 (9.3)	8.6 (2.5)	—	59.1 (20.3)	—	—
Fu & Shannon, 2000	6	0	English (USA)	aCa	16	53.0 (12.4)	5.3 (2.3)	—	67.8 (11.7)	—	—
Galvin et al., 2007	11	9[a]	English (USA)	aCa	20	49.0 (14.9)	7.3 (4.8)	—	49 (14.9)	—	—
Gani et al., 2007	4	0	French (Switzerland)	aCa, V	14 (aCa), 7 (V)	46.8 (15.3)	1.4 (0.6)	67.5 (12.1)	64.2 (18.8)	—	—
Guevara et al., 2015	8	0	French (France)	VCV	16	49.3 (8.7)	2.5 (1.7)	—	51.8 (19.0)	—	—
Guevara et al., 2016	16	0	French (France)	VCV	16	48.8 (14.2)	5.1 (3.7)	—	42.5 (21.6)	—	—
Han et al., 2016	10	11[a]	English (USA)	aCa	16	45.1 (18.2)	6.1 (4.0)	—	74 (21.1)	—	—
Incerti et al., 2011	15	0	English (Australia)	aCa	24	Adult	4.2 (–)	—	68.4 (18.7)	—	—
Ito et al., 1994	10	0	Japanese	V, aCa	5 (V), 13 (aCa)	Adult	—	87.7 (10.5)	41.7 (11.2)	—	—
Kirk et al., 1992, Sample A	10	12	English (USA)	wVb, bVb	8	47.3 (9.7)	1.7 (1.3)	50.5 (4.8)	—	22.8 (17.0)	NU-6
Kirk et al., 1992, Sample B	11	12	English (USA)	wVb, bVb	8	53.4 (17.3)	1.8 (1.6)	52 (4.0)	—	14.2 (16.1)	NU-6
McKay et al., 1992	4	0	English (Australia)	aCa	12	45.0 (16.9)	2.7 (1.7)	—	77 (9.7)	57.6 (26.6)	NU-6
Meyer et al., 2003	26	0	English (Australia)	aCa	24	Adult	> 11 months	—	42.3 (22.2)	38.0 (24.5)	CNC
Munson et al., 2003, Sample A	14	0	English (USA)	aCa	19	41.4 (10.3)	3.6 (3.5)	—	78.5 (6.1)	66.6 (17.7)	NU-6
Munson et al., 2003, Sample B	16	0	English (USA)	aCa	19	55.1 (12.2)	4.3 (4.1)	—	46.9 (13.0)	26.8 (18.5)	NU-6
Nie et al., 2006	5	0	English (USA)	aCa	20	35.8 (12.8)	4.6 (1.1)	—	64 (14.7)	—	—
Pelizzone et al., 1999	12	0	Albanian, French, German, Italian, Russian, Spanish, Swahili	aCa, V	14 (aCa), 7 (V)	43.3 (16.5)	5.1 (3.0)	78.3 (17.2)	65.6 (24.1)	—	—

(*table continues*)

Table 2. *(Continued)*.

Authors	*N*		Language of participants	Stimulus context	No. of consonants and vowels in the test	Age (years) of implantation, *M (SD)*	Duration (years) of implant use, *M (SD)*	Speech sound score (%)		Monosyllables	
	CI	NH						Vowels, *M (SD)*	Consonants, *M (SD)*	Score (%), *M (SD)*	Name of test
Sagi et al., 2009	11	16	English (USA)	aCa	16	50.4 (14.5)	3.6 (1.9)	—	43.1 (22.6)	38.6 (21.1)	CNC
Shafiro et al., 2011	17	0	English (USA)	aCa	20	55.0 (11.2)	3.2 (2.1)	—	51 (24)	51.8 (30.5)	CNC
Shallop et al., 1992	7	0	English (USA)	aCa	14	58.4 (14.0)	1 (0)	—	50.6 (19.0)	21.2 (12.9)	NU-6
Shannon et al., 2011	7	0	English (USA)	aCa	20	48.3 (9.1)	1.0 (0.3)	—	62.7 (7.3)	66.4 (17.9)	CNC
Shannon et al., 2002	6	6[b]	English (USA)	aCa	14	52.2 (11.1)	5.0 (2.9)	—	67.6 (18.5)	44 (22.5)	NU-6
Sheffield & Zeng, 2012	8	8	English (USA)	aCa	20	58.6 (11.4)	4.4 (2.0)	—	59.1 (5.2)	—	—
Singh et al., 2009	5	0	English (USA)	aCa	20	68.8 (7.6)	3.0 (2.6)	—	72.6 (16.1)	—	—
Skinner, Arndt, et al., 2002	56	0	English (USA)	aCa	14	54.4 (17.1)	—	—	66.6 (20.7)	43.7 (22.7)	CNC
Skinner, Holden, et al., 2002	12	0	English (USA and Australia)	aCa	14	50.9 (22.6)	Short	—	64.6 (14.6)	—	—
Stacey et al., 2010	11	0	English (UK)	aCa	20	49.1 (19.5)	5.7 (2.7)	—	39.8 (18.8)	—	—
Svirsky et al., 2011	28	0	English (USA)	aCa	24	49.7 (15.9)	3.3 (2.5)	—	45.1 (20.7)	—	—
Teoh et al., 2003	15	0	English (USA)	aCa	24	52.9 (12.7)	3.6 (1.8)	—	46.7 (18.1)	—	—
Throckmorton & Collins, 1999	7	0	English (USA)	aCa	14	54.6 (11.6)	6.9 (2.8)	—	34 (14.2)	14.9 (10.7)	NU-6
Tye-Murray et al., 1990	5	0	English (USA)	iCi	14	Adult	> 10 months	—	41 (13.1)	27.6 (13.2)	NU-6
Tye-Murray et al., 1996	40	0	English (USA)	aCa	13	51.5 (–)	3.6 (2.2)	—	66 (20)	—	—
Tyler, 1990	5	0	English (Australia)	iCi	13	7.4 (1.9)	1.1 (0.6)	—	30 (13.2)	16 (20.3)	PBK
Tyler & Moore, 1992, Sample A	10	6	German	iCi	13	43.8 (9.5)	2.0 (0.6)	—	31.1 (7.0)	—	—
Tyler & Moore, 1992, Sample B	19	2	English (USA)	iCi	13	37.5 (13.5)	2.4 (2.8)	—	43.9 (10.6)	—	—
Van Wieringen & Wouters, 1999	25	20	Flemish	aCa	16	43.4 (14.2)	2.1 (1.4)	—	33 (13)	—	—
Wouters & van den Berghe, 2001	4	0	Flemish	aCa	16	39.8 (10.4)	2.0 (1.4)	—	63.3 (7.9)	—	—

Note. The means and standard deviations are given with one decimal, except when the included articles reported these values without decimals. Em dashes indicate data not obtained. CI = cochlear implant; NH = normally hearing; VCV = vowel–consonant–vowel; aCa = a–Consonant–a; V = Vowel; iCi = i–Consonant–i; uCu = u–Consonant–u; Ca = Consonant–a; Ci = Consonant–i; Cu = Consonant–u; NU–6 = The Northwestern University Auditory Test No. 6; CNC = the consonant–vowel nucleus–consonant test; wVb = w–Vowel–b; bVb = b–Vowel–b; PBK = The Phonetically Balanced Kindergarten Word Test.

[a] Not tested with the consonant test. [b] Tests performed using a CI simulator and the test results therefore not included.

Table 3. Means, standard deviations, and ranges of the study variables for the prelingually and postlingually deaf CI users.

	Postlingually deaf			Prelingually deaf			Total		
Study variables	***M* (*SD*) (%)**	***N***	**Range (%)**	***M* (*SD*) (%)**	***N***	**Range (%)**	***M* (*SD*) (%)**	***N***	**Range (%)**
Consonant score	58.4 (26.3)	44	18.7–91.6	46.7 (11.5)	6	36.0–76.0	56.3 (23.1)	48	30.0–88.0
Vowel score	76.8 (26.5)	5	50.5–95.0	67.7 (0.0)	1	–	72.4 (23.1)	6	50.5–95.0
Real-word monosyllable score	40.1 (16.6)	14	14.9–66.6	–	–	–	36.9 (16.8)	17	14.2–66.6

Note. In three of the studies, the real-word monosyllable scores could not be separated into separate scores for the groups of prelingually and postlingually deaf CI users. Em dashes indicate data not obtained. CI = cochlear implant.

users, and no studies reported vowel identification scores for both groups. Consonant identification scores for the postlingually deaf users were on average 10.9% better than for the prelingually deaf users ($SD = 39.7\%$, range = −22.5%–47.5%, $z[5] = 0.61$). This difference in scores was not statistically significant ($p = .54$, $df = 4$). Hence, it is unclear whether there is a difference in consonant perception between prelingually and postlingually deaf CI users.

Research Question 2: Which Consonants and Vowels are Most Frequently Confused by CI Users, and Which Consonants and Vowels are Most Frequently Identified Correctly?

Vowel Confusions

Details on individual vowel confusions were reported in only one of the included articles (containing two studies; Kirk et al., 1992) but were based on quantitative data from 27 CMs. This article reports results from participants with normal hearing and two groups of CI users: Ineraid and Nucleus-users. Vowel stimuli were given both in bVb context and in wVb context. Identifications and misidentifications were reported qualitatively, and for the subjects with normal hearing, only a few errors were made. In the bVb context, mean vowel identification was 50.5% ($SD = 4.8\%$, range = 30.0%–77.7%) for cochlear CI users and 52.0% ($SD = 4.0\%$, range = 32.5%–82.5%) for Ineraid CI users. In the wVb context, the vowel identification scores were somewhat lower than in the bVb context for both implants. In summary, the long vowels /iː, æː, ɑː/, and /uː/ were seldom misidentified, but the short vowels /ɪ, ɛ, ʌ/, and /ʊ/ were often confused with other short vowels. /ʊ/ was sometimes, however, also confused with /ɑː/ in wVb context. Additionally, a higher number of short vowels were confused in the wVb context than in the bVb context.

Consonant Confusions

Details about consonant confusions were reported in 13 of the included articles (15 studies; Donaldson & Kreft, 2006; Dorman & Loizou, 1996; Dorman et al., 1990; Doyle et al., 1995; Incerti et al., 2011; McKay et al., 1992; Munson et al., 2003; Pelizzone et al., 1999; Sagi et al., 2009; Teoh, Neuburger, & Svirsky, 2003; Tyler, 1990; Tyler & Moore, 1992; Van Wieringen & Wouters, 1999). In 11 of these articles, the consonant confusions were reported in CMs. Table 4 gives an overview of these 11 articles. Detailed results of the three most frequently correctly identified consonants from the 11 articles are shown in Table 5, and details about the most common consonant confusions from the 11 articles are presented in a meta-CM in Table 6. Because of the low number of articles presenting CMs (11), we chose to base our study's matrices on the nine consonants that were used in all the 15 studies, /b, d, p, t, k, n, m, s/, and /z/. We also chose to pool articles reporting studies conducted in different languages (Australian English, American English, and Flemish) and to pool those with different kinds of stimuli, Cɑ, Ci, Cu, ɑCɑ, iCi, and uCu. We also pooled the only article, which included children as participants (Tyler, 1990) with the remaining articles.

In two studies (Dorman et al., 1990; Munson et al., 2003), the participants were divided into poor and better performers; in one study, the participants were divided into poor, intermediate, and better performers (Van Wieringen & Wouters, 1999); and in two studies, the participants were divided into three groups according to type of implant (Doyle et al., 1995) or according to native language of participants (Tyler & Moore, 1992). In each of these studies, the data from the CM of each group were plotted into the table and the meta-CM. Thus, a total of 17 CMs were pooled into Table 5 and the meta-CM in Table 6.

In three of the articles, several consonant identification tests were given to the same participants. We chose the better of the two outcomes when two speech processors were compared (Dorman & Loizou, 1996; McKay et al., 1992). We chose the outcomes on the basis of use of CI alone if one CM was made based on the CI alone and one on CI + hearing aid (Incerti et al., 2011). In one article (Donaldson & Kreft, 2006), the consonant identification tests were performed in six contexts, Cɑ, Ci, Cu, ɑCɑ, iCi, and uCu, and averaged over all conditions. We included the pooled data in our analyses. When several CMs were presented, obtained with and without background noise and with and without lipreading (Incerti et al., 2011), testing in quiet and auditory-only condition was chosen.

As Table 5 shows, the consonants that were most frequently identified correctly were the unvoiced stops /t/ and /k/.

The meta-CM in Table 6 shows that the most frequent confusions were /k/ confused with /t/ and /m/ confused with /n/.

Table 4. Description of the articles presenting consonant confusions in matrices.

Authors	Consonant context	No. of consonants	No. of participants	Language
Donaldson & Kreft, 2006	Cɑ, Ci, Cu, ɑCɑ, iCi and uCu, both female and male reader in each condition	19	20	English (USA)
Dorman & Loizou, 1996	ɑCɑ	16	7	English (USA)
Dorman et al., 1990	ɑCɑ	16	10	English (USA)
Doyle et al., 1995	iCi	14	14	English (USA)
Incerti et al., 2011	ɑCɑ	24	15	English (Australia)
McKay et al., 1992	ɑCɑ	12	4	English (Australia)
Munson et al., 2003	ɑCɑ	19	30	English (USA)
Teoh et al., 2003	ɑCɑ	24	14	English (USA)
Tyler, 1990	iCi	13	4	English (Australia)
Tyler & Moore, 1992	iCi	13	28	English (USA)
Van Wieringen & Wouters, 1999	ɑCɑ	16	25	Flemish

Note. Ca = Consonant–a; Ci = Consonant–i; Cu = Consonant–u; aCa = a–Consonant–a; iCi = i–Consonant–i; uCu = u–Consonant–u.

Research Question 3: (a) To What Extent are Age at Implantation, Duration of Implant Use, and Real-Word Monosyllable Score Associated With Variations in Consonant and Vowel Identification Performance in Nonsense Syllable Tasks for Prelingually Deaf CI Users? (b) To What Extent are Duration of Implant Use and Real-Word Monosyllable Score Associated With Variations in Consonant and Vowel Identification Performance in Nonsense Syllable Tasks for Postlingually Deaf CI Users?

(a) The weighted scores of age at implantation and duration of implant use for the prelingually and postlingually deaf CI users are reported in Table 7. The monosyllable scores are reported in Table 3. Because only six studies report results for prelingually deaf CI users, a bivariate metaregression was not carried out, and Research Question 3 (a) could not be answered.

Table 5. Overview of the three most frequently correctly identified consonants in the included studies.

Stimulus	Index of correct identifications (%)
/t/	18.1
/k/	17.7
/m/	14.8
/n/	10.6
/p/	9.7
/z/	8.9
/d/	7.8
/b/	7.0
/s/	5.3

Note. The three most frequently correctly identified consonants in each study were picked, assigned to an index weighed by the number of participants in the study, added together with the results from the other studies, and included in this table. The percentages in the second column were calculated by dividing the number of correct identifications of each consonant by the total number of correct responses. The consonant with the highest percentage was the most frequently, correctly identified of the nine consonants. The consonants are arranged in descending order according to percentage of correct identification.

(b) Only five studies reported a vowel identification score for the group of postlingually deaf. This is too few to provide an adequate representation of the included studies, and further analyses were therefore not performed on this group. The vowel identification scores can be examined in Table 3.

We decided to omit monosyllable scores from the multiple regression model with postlingually deaf CI users due to a small number of studies ($N = 14$). A univariate regression model was then constructed with the moderator variable duration of implant use and the independent variable consonant identification score. The results of the univariate regression were $\beta = 2.6$, $SE = 1.4$, 95% confidence interval = [−0.22, 5.3], $z[36] = 1.81$, and not significant ($p = .071$). The proportion of total between-studies variance explained by the model was $R^2 = .59$, $N = 36$.

Publication Bias

In order to optimize the quality of our included study sample, we have only included peer-reviewed, published studies written in English, Finnish, and in Scandinavian languages. Although we performed searches in a number of grey material databases in the beginning of our systematic review process, without finding any relevant studies, some unpublished and even published research may still be missing from our searches. Also, relevant studies may have experienced delayed publishing for various reasons. Thus, there might be some publication bias in our systematic review.

By visual inspection of the funnel plot for the consonant identification scores of the postlingually deaf, we noticed that the studies were slightly scattered to the left of the mean of the funnel plot. The asymmetry in the funnel plot may be a sign of publication bias, heterogeneity, or chance.

Discussion

The purpose of this systematic review and meta-analysis is to establish a baseline of the vowel and consonant identification scores in prelingually and postlingually

Table 6. Confusion matrix of the three most frequently confused consonants pooled across 13 studies.

	Response (%)									
Stimulus	**/p/**	**/t/**	**/k/**	**/b/**	**/d/**	**/m/**	**/n/**	**/s/**	**/z/**	**Sum (%)**
/p/		11.4	6.7					0.8		18.9
/t/	7.5		10.1							17.6
/k/	3.7	18.5								22.2
/b/					3.0		4.3			7.3
/d/				5.6						5.6
/m/							18.1			18.1
/n/						7.5				7.5
/s/										0.0
/z/								2.8		2.8

Note. The three most frequently confused consonants in each confusion matrix were picked, assigned to an index equal to the number of participants in the study, added together with the results from the other matrices, and included in this table. The percentages in the table cells were calculated by dividing the number of confusions in each cell by the total number of confusions. The cell with the highest percentage shows the most frequent consonant confusion of the 13 studies.

deaf users of multichannel CIs tested with CVC and VCV nonsense syllables.

The mean consonant and vowel identification scores for the prelingually and postlingually deaf CI users show that performance was well below ceiling for both groups and that there were higher scores for vowels than for consonants. The mean differences between the consonant identification scores for the prelingually and postlingually deaf CI users were not statistically significant.

Details of the vowel confusions were given qualitatively and in only one article. Details of the consonant confusions were given in CMs in 11 articles. Our meta-CM showed that the most frequently confused consonants were /k/ confused with /t/ and /m/ confused with /n/.

In a univariate regression model between duration of implant use and consonant identification score for postlingually deaf CI users, duration of implant use explained 59% of the variance in effect sizes. The model was not statistically significant (p = .071).

Research Question 1: Typical Vowel and Consonant Identification Scores

We could not draw definite conclusions about differences in consonant identification between prelingually and postlingually deaf CI users because of the large difference in sample size between the groups (six studies with prelingually deaf and 44 studies with postlingually deaf). The same reason applies to why Research Question 1 could not be answered with regard to vowel identification score, as only one article with one participant reported a vowel score of prelingually deaf CI users and five articles reported vowels scores of postlingually deaf CI users.

Visual inspection of Table 3 shows that the vowel identification scores were substantially higher than the consonant identification scores for both prelingually and postlingually deaf CI users and that the total vowel score was approximately 16% higher than the total consonant score. This can be explained by the known fact that vowels have more acoustic energy than most consonants. The vowels in the CVC test words also have longer duration than the consonants in the VCV test words and may therefore be easier to perceive, as the participants have more time to listen to them.

The consonant score for the prelingually deaf implant users was below 50% and more than 10% lower than the consonant score for the postlingually deaf (see Table 3). When we examined the six included studies with prelingually deaf participants, we noticed that they all had participants with a high age at implantation (range = 6.8–31.5 years) and, thus, long duration of deafness before implantation. Many studies have shown that prelingually deaf individuals younger than 2 years of age at implantation are more likely to obtain higher benefit from the implant for open speech perception than prelingually deaf implanted at a higher age (May-Mederake, 2012; Quittner, Cejas, Wang, Niparko, & Barker, 2016; Tobey et al., 2013). Studies conducted with prelingually deaf children implanted earlier than at 1 year of age show even that their speech perception measures are superior to the corresponding measures for postlingually deaf CI users, for prelingually deaf, later-implanted children, and for CI users with a progressive hearing loss before implantation (Colletti, Mandalà, & Colletti, 2012; Dettman et al., 2016; Holman et al., 2013).

Research Question 2: Vowels and Consonants Most Frequently Confused and Most Frequently, Correctly Identified

In 11 of the included articles in our meta-analysis, consonant confusions were presented in CMs, making

Table 7. Means, standard deviations, and ranges for the moderator variables for the prelingually and postlingually deaf CI users.

	Postlingually deaf			Prelingually deaf			Total		
Moderator variable	***M (SD)*** **(years)**	***N***	**Range (years)**	***M (SD)*** **(years)**	***N***	**Range (years)**	***M (SD)*** **(years)**	***N***	**Range (years)**
Age at implantation	49.7 (18.3)	37	7.9–68.8	39.9 (6.8)	6	6.8–31.5	48.0 (22.5)	42	7.4–68.8
Duration of implant use	3.4 (1.6)	35	1.0–10.0	6.0 (4.6)	5	0.8–11.5	3.4 (1.6)	39	1.0–10.0

Note. CI = cochlear implant.

the results easy to quantify. In the spirit of meta-analytic approach, the CMs from the 11 articles were merged into one meta-CM displaying the three most frequently confused consonants from each CM.

It is a well-known phenomenon in phonetic and audiologic research that confusions between speech sounds most frequently happen within a group of sounds with different place of articulation but similar manner of articulation. Fant (1973) showed that the acoustic similarities of consonants grouped according to manner of articulation, for instance, stops, fricatives, and nasals, are significant for speech sound perception. The most frequently confused consonants in this study had the same manner of articulation and were thus acoustically similar and differed only in place of articulation. /t/ is an unvoiced dental/alveolar stop, and /k/ is an unvoiced velar stop. /m/ and /n/ are voiced nasals. In both confusions, different places of articulation were confused within the same category of manner of articulation.

The relatively high percentage of correct identification scores for the unvoiced stop consonants /t/ and /k/ in VCV context displayed in Table 5 can be explained by the fact that CI users listen to formant transitions in the adjacent vowels for identification. Consonants with the same manner but different place of articulation would be difficult to identify if formant transitions were not available. Moreover, the quality of the aspiration of the unvoiced stops also makes them easier to recognize than the voiced stops. There is a distinct audible difference between the aspiration following the pronunciation of /p/, /t/, and /k/, resembling the sound of the corresponding fricatives produced in the same place.

/k/ and /t/ were found to be the most frequently, correctly identified consonants, but /k/ was also the consonant most frequently confused, namely, with /t/. This may seem contradictory, but the explanation is most likely that the other consonants in the CMs of the included studies, /b, d, n, m, s, z/, are confused more broadly and more frequently with a number of other speech sounds, and also with those not included in our study, whereas the three unvoiced stops are almost exclusively confused among themselves. Apparently, CI users perceive the unvoiced stops as the most audibly distinct group among the consonants included in this study.

Research Question 3: The Association Between Age at Implantation, Duration of Implant Use, and Real-Word Monosyllable Score on Vowel and Consonant Identification Scores in Prelingually and Postlingually Deaf CI Users

Due to the low number of included studies reporting consonant or vowel identification score for the prelingually deaf, a statistical analysis of the associations with the moderators could not be performed for this group. However, many previous studies have investigated this, and it is well known that age at implantation plays an important role for the outcome of speech perception tests for prelingually deaf CI users (Holman et al., 2013; Tobey et al., 2013). Presumably, this is also the case for vowel and consonant tests measured by nonsense words.

For the postlingually deaf CI users, we constructed a univariate regression model in which duration of implant use could explain 59% of the variance in consonant score. After implantation, the CI users need a period of adaptation to the implant sound, which, in most cases, can vary from 3 months to 1 year. Thus, until stability of the fitting parameters is reached, the implantees will experience a gradual improvement of the benefit of the implants. Schmidt and Griesser (1997) showed that this stability was reached after about 1 year.

Earlier studies have shown that there is a close relationship between consonant and vowel identification scores and real-word monosyllable scores (e.g., Rødvik, 2008). Due to the low number of studies that reported real-word monosyllable scores in quiet for the postlingually deaf implantees ($N = 14$), we could not confirm this relationship in the meta-analysis. It also needs to be pointed out that, in the included studies, three different real-word monosyllable tests were used, and the consistency of the pooled means may therefore not be satisfactory.

Limitations

Exclusion of Studies Reporting Vowel Identification Scores Measured by Real Words

Our set criterion of only including studies which measured vowel and consonant scores by nonsense words demanded the exclusion of studies in which real words were used. The hVd nine-vowel test by Tyler et al. (1983) and the hVd 12-vowel test by Hillenbrand et al. (1995) were used to calculate vowel identification scores in 28 of the included studies, in which consonant identification scores were also measured. The test scores were excluded from this meta-analysis, as all of the hVd-combinations produced real English words, and also included diphthongs. Among the six included studies in which vowel scores were measured using nonsense words, three described Swiss participants (French-speaking; Cosendai & Pelizzone, 2001; Gani et al., 2007; Pelizzone et al., 1999), one described Japanese (Ito et al., 1994), and two described English-speaking participants from the United States (Kirk et al., 1992). It appears that many of the studies conducted with English-speaking participants use tests with real words in vowel identification testing, but tests with nonsense syllables in consonant identification testing. Studies conducted with participants with other native languages more often use nonsense syllables for obtaining vowel identification score as well. The reason might be lack of a validated nonsense syllable vowel test in English or that other languages do not have as many minimal pairs or triplets as the English language.

The consequences of excluding studies in which real words were used to measure consonant and vowel identification scores can be considered both positive and negative. On the positive side, consonant and vowel scores are collected from a homogenous material and can be compared cross-linguistically. On the negative side, the collected material is smaller than it would have been if consonant and vowel scores measured by real words were included, and thus, the statistical power is lower.

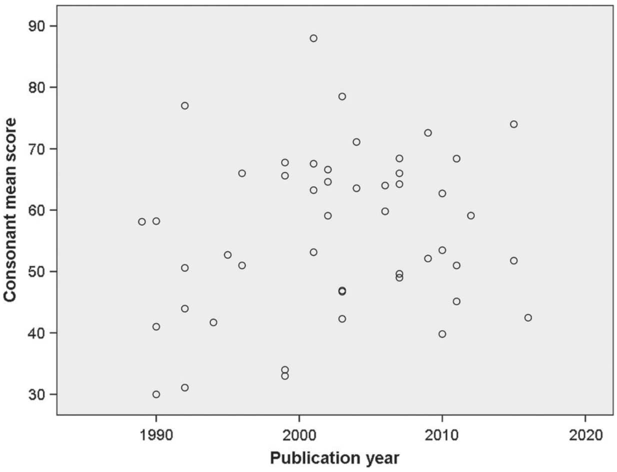
Use of Nonsense Syllable Tests to Avoid Ceiling Effects in Speech Perception Testing

When the outcomes of speech perception tests approach the ceiling effect, the tests should be replaced with more difficult tests. This is usually done in two different ways, either by adding noise to test words and sentences or by exchanging the real-word tests with nonsense syllable tests. These are two very different approaches of increasing the levels of difficulty, and both have advantages and disadvantages. A speech-in-noise test is most frequently preferred in clinics, and one reason may be that such tests allow for the assessment of speech perception in everyday situations, which often involve a degree of environmental noise. Although the nonsense syllable identification test does not correspond closely to everyday speech perception situations, it has a major advantage in its relative independence of cognitive and contextual factors, such as language abilities, language experience, inferential skills, working memory capacity, and use of sentence context for comprehension. Such a test is valuable in research and in clinics, as it provides information about minute details of the speech sound perception of the implantees, details that cannot easily be obtained with other tests. This is useful for the fitting of implants and for the planning of individual listening therapy.

Choice of Time Frame for the Inclusion of Articles

The articles included in the meta-analysis range in publication year from 1989 to 2016 and report test results on CI users with multichannel implants of four channels or more. The validity of our choice is confirmed by Figure 5, which shows that the correlation between publication year and consonant score in the included articles is low and not statistically significant (.187; $p = .202$). Hence, other factors than implant technology would probably explain the consonant score or dominate in a regression model with consonant score as the dependent variable.

Figure 5. Scatter plot of consonant mean scores versus publication years in the 48 included studies reporting consonant scores. The consonant mean scores are measured in percent. The cases are weighted by number of participants.



Since 1989, there has been a transition from analog strategies in Symbion/Ineraid and feature extraction strategies in previous Cochlear devices (F0F2 and F0F1F2), to n-of-m and derivate of CIS stimulation strategies. More recently, there has been a transition to the fine structure stimulation strategies from Med-El. These strategies convey the fundamental frequency in the coding algorithm. All these modern strategies are spectral resolution strategies and, thus, can deliver pitch information to the inner ear, unlike the previous single-channel implants. The spectral resolution strategies are mainly pulsatile strategies, except for the analog strategies, and thus, the information is delivered to the electrodes using a set of narrow pulses in a nonsimultaneous fashion. Some of the recent stimulation strategies from Advanced Bionics even employed combined pulsatile and simultaneous (analog) stimulation strategies.

There has been a development in the microphone technology since the early years of CI. The input frequency range has increased, and the overall microphone quality has improved. However, the microphone sensitivity and the internal noise of the microphones have not improved noteworthily, although the availability of good microphones has increased. The benefit of increased frequency range in the speech processors for the postlingually deaf can also be discussed because the perceived pitch depends on where the implant is located in the cochlea rather than on the input frequency range of the microphone. Thus, the improvements in speech processor technology may not be of great importance in a clinical test situation with a good signal-to-noise ratio.

The largest improvements and developments of the implant technology since 1989 have followed the advances in conventional hearing aids by integrating a large amount of technology from the hearing aid industry. For instance, refined and further developed automatic gain controls with new noise reduction and compression algorithms have been implemented in the speech processors from all implant manufacturers. Also, there has been a trend toward smaller processors and toward controlling the speech processors by remote controls or by “apps” on the users’ smartphones. All this may have had substantial impact on the speech perception in daily life but probably only minor impact on speech perception in a clinical environment.

Conclusions

This systematic review and meta-analysis included peer-reviewed studies using nonsense syllables to measure the consonant and vowel identification scores of CI users, both with and without control groups.

The mean performance on consonant identification tasks for the postlingually deaf CI users from 44 studies was higher than for the prelingually deaf users, reported in six studies. No statistically significant difference between the scores for prelingually and postlingually deaf CI users was found.

The consonants that were not correctly identified were typically confused with other consonants with the same acoustic properties, namely, voicing, duration, nasality, and silent gaps.

A univariate metaregression model with consonant score against duration of implant use for postlingually deaf adults indicated that duration of implant use predicts a substantial portion of their consonant identification ability. No statistical significance was found using this model.

Tests with monosyllabic and bisyllabic nonsense syllables have been employed in research studies on CI users' speech perception for several decades. These kinds of studies expose information about the hearing of cochlear-implanted patients, which the standard test batteries in most audiologic clinics do not reveal, information that is very useful for the mapping of CIs and for the planning of habilitation and rehabilitation therapy. Such tests may also give valuable information for further development of CI technology. We therefore propose that nonsense syllable tests be used as part of the standard test battery in audiology clinics when assessing the speech perception of CI users.

Acknowledgments

The authors did not receive any external funding for this project, and none of the authors have potential conflicts of interest to be disclosed. The authors wish to express their gratitude to chief librarian Ingerid Straume at the University of Oslo for invaluable assistance in the search for relevant literature and Tamar Kalandadze-Tabakhmelashvili for valuable guidance in the use of the CMA meta-analysis program.

References

Studies marked with an asterisk have been included in the meta-analysis.

*__Arisi, E., Forti, S., Pagani, D., Todini, L., Torretta, S., Ambrosetti, U., & Pignataro, L.__ (2010). Cochlear implantation in adolescents with prelinguistic deafness. *Otolaryngology—Head & Neck Surgery, 142*(6), 804–808.

*__Baskent, D., & Shannon, R. V.__ (2004). Frequency-place compression and expansion in cochlear implant listeners. *The Journal of the Acoustical Society of America, 116*(5), 3130–3140. https://doi.org/10.1121/1.1804627

Berrettini, S., Baggiani, A., Bruschini, L., Cassandro, E., Cuda, D., Filipo, R., . . . Forli, F. (2011). Systematic review of the literature on the clinical effectiveness of the cochlear implant procedure in adult patients. *Acta Otorhinolaryngologica Italica, 31*(5), 299–310.

*__Bhattacharya, A., & Zeng, F. G.__ (2007). Companding to improve cochlear-implant speech recognition in speech-shaped noise. *The Journal of the Acoustical Society of America, 122*(2), 1079–1089.

Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., . . . Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiology and Neuro-Otology, 18*(1), 36–47. https://doi.org/10.1159/000343189

*__Blamey, P. J., Dickson, B., & Grant, L. M.__ (2004). An incremental excitation scale for cochlear implants. *Acoustics Research Letters Online-Arlo, 5*(2), 50–55. https://doi.org/10.1121/1.1687471

Blamey, P. J., Maat, B., Başkent, D., Mawman, D., Burke, E., Dillier, N., . . . Lazard, D. S. (2015). A retrospective multicenter study comparing speech perception outcomes for bilateral implantation and bimodal rehabilitation. *Ear and Hearing, 36*(4), 408–416. https://doi.org/10.1097/AUD.0000000000000150

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2014). Comprehensive meta-analysis (Version 3) [Software]. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Hoboken, NJ: Wiley.

Burkholder-Juhasz, R. A., Levi, S. V., Dillon, C. M., & Pisoni, D. B. (2007). Nonword repetition with spectrally reduced speech: Some developmental and clinical findings from pediatric cochlear implantation. *Journal of Deaf Studies and Deaf Education, 12*(4), 472–485.

Carlyon, R. P., Monstrey, J., Deeks, J. M., & Macherey, O. (2014). Evaluation of a cochlear-implant processing strategy incorporating phantom stimulation and asymmetric pulses. *International Journal of Audiology, 53*(12), 871–879. https://doi.org/10.3109/14992027.2014.932024

Casserly, E. D., & Pisoni, D. B. (2013). Nonword repetition as a predictor of long-term speech and language skills in children with cochlear implants. *Otology & Neurotology, 34*(3), 460–470.

*__Chatterjee, M., & Peng, S. C.__ (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hearing Research, 235*(1–2), 143–156. https://doi.org/10.1016/j.heares.2007.11.004

Clark, G. M., Tong, Y. C., Martin, L. F., Busby, P. A., Dowell, R. C., Seligman, P. M., & Patrick, J. F. (1981). A multiple-channel cochlear implant: An evaluation using nonsense syllables. *Annals of Otology, Rhinology & Laryngology, 90*(3), 227–230.

Cohen, N. L., Waltzman, S. B., & Fisher, S. G. (1993). A prospective, randomized study of cochlear implants. *The New England Journal of Medicine, 328*(4), 233–237. https://doi.org/10.1056/NEJM199301283280403

Colletti, L., Mandalà, M., & Colletti, V. (2012). Cochlear implants in children younger than 6 months. *Otolaryngology—Head & Neck Surgery, 147,* 139–146.

*__Cosendai, G., & Pelizzone, M.__ (2001). Effects of the acoustical dynamic range on speech recognition with cochlear implants. *Audiology, 40*(5), 272–281.

*__Desai, S., Stickney, G., & Zeng, F. G.__ (2008). Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. *The Journal of the Acoustical Society of America, 123*(1), 428–440. https://doi.org/10.1121/1.2816573

Dettman, S. J., Dowell, R. C., Choo, D., Arnott, W., Abrahams, Y., Davis, A., . . . Briggs, R. J. (2016). Long-term communication outcomes for children receiving cochlear implants younger than 12 months: A multicenter study. *Otology and Neurotology, 37*(2), e82–e95. https://doi.org/10.1097/MAO.0000000000000915

*__Donaldson, G. S., & Kreft, H. A.__ (2006). Effects of vowel context on the recognition of initial and medial consonants by cochlear implant users. *Ear and Hearing, 27*(6), 658–677. https://doi.org/10.1097/01.aud.0000240543.31567.54

*__Dorman, M. F., Dankowski, K., McCandless, G., & Smith, L.__ (1989). Consonant recognition as a function of the number of channels of stimulation by patients who use the Symbion cochlear implant. *Ear and Hearing, 10*(5), 288–291.

Dorman, M. F., Hannley, M. T., Dankowski, K., Smith, L., & McCandless, G. (1989). Word recognition by 50 patients fitted with the Symbion multichannel cochlear implant. *Ear and Hearing, 10*(1), 44–49. https://doi.org/10.1097/00003446-198902000-00008

*__Dorman, M. F., & Loizou, P. C.__ (1996). Improving consonant intelligibility for Ineraid patients fit with continuous interleaved sampling (CIS) processors by enhancing contrast among channel outputs. *Ear and Hearing, 17*(4), 308–313.

*__Dorman, M. F., Soli, S., Dankowski, K., Smith, L. M., Parkin, J., & McCandless, G.__ (1990). Acoustic cues for consonant identification by patients who use the Ineraid cochlear implant. *The Journal of the Acoustical Society of America, 88*(5), 2074–2079. https://doi.org/10.1121/1.400104

Dowell, R. C., Dettman, S. J., Blamey, P. J., Barker, E. J., & Clark, G. M. (2002). Speech perception in children using cochlear implants: Prediction of long-term outcomes. *Cochlear Implants International, 3*(1), 1–18. https://doi.org/10.1179/cim.2002.3.1.1

*__Doyle, K. J., Mills, D., Larky, J., Kessler, D., Luxford, W. M., & Schindler, R. A.__ (1995). Consonant perception by users of Nucleus and Clarion multichannel cochlear implants. *American Journal of Otology, 16*(5), 676–681.

Drullman, R. E. (2005). Speech recognition tests for different languages. *FP6–004171 HEARCOM Hearing in the Communication Society Report no. D-7-1*. Germany: Fraunhofer Institute for Applied Information Technology.

Dubno, J. R., & Dirks, D. D. (1982). Evaluation of hearing-impaired listeners using a Nonsense-Syllable Test. I: Test reliability. *Journal of Speech and Hearing Research, 25*(1), 135–141. https://doi.org/10.1044/jshr.2501.135

Dubno, J. R., Lee, F. S., Klein, A. J., Matthews, L. J., & Lam, C. F. (1995). Confidence limits for maximum word-recognition scores. *Journal of Speech and Hearing Research, 38*(2), 490–502.

The Ear Foundation. (2017, December 17). *Cochlear Implant Information Sheet*. Retrieved from http://www.earfoundation.org.uk/hearing-technologies/cochlear-implants/cochlear-implant-information-sheet

Fant, G. (1973). *Speech sounds and features* (Vol. 4). Cambridge, MA: MIT Press.

Faulkner, K. F., & Pisoni, D. B. (2013). Some observations about cochlear implants: Challenges and future directions. *Neuroscience Discovery, 1*(1), 9. https://doi.org/10.7243/2052-6946-1-9

*__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. *The Journal of the Acoustical Society of America, 110*(2), 1150–1163.

Frijns, J. H. M., Briaire, J. J., De Laat, J. A. P. M., & Grote, J. J. (2002). Initial evaluation of the Clarion CII cochlear implant: Speech perception and neural response imaging. *Ear and Hearing, 23*(3), 184–197.

*__Fu, Q. J.__ (2002). Temporal processing and speech recognition in cochlear implant users. *NeuroReport, 13*(13), 1635–1639.

*__Fu, Q. J., & Shannon, R. V.__ (2000). Effect of stimulation rate on phoneme recognition by Nucleus-22 cochlear implant listeners. *The Journal of the Acoustical Society of America, 107*(1), 589–597. https://doi.org/10.1121/1.428325

*__Galvin, I. J. J., Fu, Q. J., & Nogaki, G.__ (2007). Melodic contour identification by cochlear implant listeners. *Ear and Hearing, 28*(3), 302–319.

*__Gani, M., Valentini, G., Sigrist, A., Kos, M. I., & Boex, C.__ (2007). Implications of deep electrode insertion on cochlear implant fitting. *Journal of the Association for Research in Otolaryngology, 8*(1), 69–83.

Gathercole, S., Willis, C., Baddeley, A., & Emslie, H. (1994). The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory, 2*(2), 103–127. https://doi.org/10.1080/09658219408258940

Geers, A., Brenner, C., & Davidson, L. (2003). Factors associated with development of speech perception skills in children implanted by age five. *Ear and Hearing, 24*(1), 24S–35S.

Glaze, J. A. (1928). The association value of non-sense syllables. *The Pedagogical Seminary and Journal of Genetic Psychology, 35*(2), 255–269. https://doi.org/10.1080/08856559.1928.10532156

*__Guevara, N., Grech, C., Gahide, I., & Gallego, S.__ (2015). Assessment of the contralateral routing of signal system in unilateral cochlear implantation. *Clinical Otolaryngology, 40*(6), 535–544. https://doi.org/10.1111/coa.12404

*__Guevara, N., Hoen, M., Truy, E., & Gallego, S.__ (2016). A cochlear implant performance prognostic test based on electrical field interactions evaluated by eABR (electrical auditory brainstem responses). *PLoS ONE, 11*(5), e0155008. https://doi.org/10.1371/journal.pone.0155008

*__Han, J.-H., Zhang, F., Kadis, D. S., Houston, L. M., Samy, R. N., Smith, M. L., & Dimitrijevic, A.__ (2016). Auditory cortical activity to different voice onset times in cochlear implant users. *Clinical Neurophysiology, 127*(2), 1603–1617. https://doi.org/10.1016/j.clinph.2015.10.049

Haskins, H. (1949). *A phonetically balanced test of speech discrimination for children*. Evanston, IL: Northwestern University.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America, 97*(5, Pt. 1), 3099–3111. https://doi.org/10.1121/1.411872

Holman, M. A., Carlson, M. L., Driscoll, C. L. W., Grim, K. J., Petersson, R. S., Sladen, D. P., & Flick, R. P. (2013). Cochlear implantation in children 12 months of age and younger. *Otology & Neurotology, 34*(2), 251–258.

*__Incerti, P. V., Ching, T. Y. C., & Hill, A.__ (2011). Consonant perception by adults with bimodal fitting. *Seminars in Hearing, 32*(1), 90–102.

*__Ito, J., Tsuji, J., & Sakakihara, J.__ (1994). Reliability of the promontory stimulation test for the preoperative evaluation of cochlear implants: A comparison with the round window stimulation test. *Auris Nasus Larynx, 21*(1), 13–16.

Kewley-Port, D., Burkle, T. Z., & Lee, J. H. (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America, 122*(4), 2365–2375. https://doi.org/10.1121/1.2773986

*__Kirk, K. I., Tye-Murray, N., & Hurtig, R. R.__ (1992). The use of static and dynamic vowel cues by multichannel cochlear implant users. *The Journal of the Acoustical Society of America, 91*(6), 3487–3498. https://doi.org/10.1121/1.402838

Lau, J., Ioannidis, J., Terrin, N., Schmid, C., & Olkin, I. (2006). Evidence based medicine: The case of the misleading funnel plot. *British Medical Journal, 333*(7568), 597–600.

Loizou, P. C. (1999). Introduction to cochlear implants. *IEEE Engineering in Medicine and Biology Magazine, 18*(1), 32–42. https://doi.org/10.1109/51.740962

Lorens, A., Skarzynski, H., Rivas, A., Rivas, J. A., Zimmermann, K., Parnes, L., . . . Pulibalathingal, S. (2016). Patient management for cochlear implant recipients in audiology departments: A practice review. *Cochlear Implants International, 17*(3), 123–128. https://doi.org/10.1080/14670100.2015.1115188

May-Mederake, B. (2012). Early intervention and assessment of speech and language development in young children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology, 76*(7), 939–946. https://doi.org/10.1016/j.ijporl.2012.02.051

*__McKay, C. M., McDermott, H. J., Vandali, A. E., & Clark, G. M.__ (1992). A comparison of speech-perception of cochlear implantees using the spectral maxima sound processor (SMSP) and the MSP (MULTIPEAK) processor. *Acta Oto-Laryngologica, 112*(5), 752–761. https://doi.org/10.3109/00016489209137470

Mendel, L. L., & Owen, S. R. (2011). A study of recorded versus live voice word recognition. *International Journal of Audiology, 50*(10), 688–693. https://doi.org/10.3109/14992027.2011.588964

*__Meyer, T. A., Frisch, S. A., Pisoni, D. B., Miyamoto, R. T., & Svirsky, M. A.__ (2003). Modeling open-set spoken word recognition in postlingually deafened adults after cochlear implantation: Some preliminary results with the neighborhood activation model. *Otology and Neurotology, 24*(4), 612–620.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America, 27*(2), 338–352. https://doi.org/10.1121/1.1907526

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

*__Munson, B., Donaldson, G. S., Allen, S. L., Collison, E. A., & Nelson, D. A.__ (2003). Patterns of phoneme perception errors by listeners with cochlear implants as a function of overall speech perception ability. *The Journal of the Acoustical Society of America, 113*(2), 925–935.

Nakeva Von Mentzer, C., Lyxell, B., Sahlén, B., Dahlström, Ö., Lindgren, M., Ors, M., . . . Uhlén, I. (2015). Segmental and suprasegmental properties in nonword repetition—An explorative study of the associations with nonword decoding in children with normal hearing and children with bilateral cochlear implants. *Clinical Linguistics & Phonetics, 29*(3), 216–235. https://doi.org/10.3109/02699206.2014.987926

Nelson, D. A., Van Tasell, D. J., Schroder, A. C., Soli, S., & Levine, S. (1995). Electrode ranking of place pitch and speech recognition in electrical hearing. *The Journal of the Acoustical Society of America, 98*(4), 1987–1999. https://doi.org/10.1121/1.413317

*__Nie, K., Barco, A., & Zeng, F. G.__ (2006). Spectral and temporal cues in cochlear implant speech perception. *Ear and Hearing, 27*(2), 208–217.

*__Pelizzone, M., Cosendai, G., & Tinembart, J.__ (1999). Within-patient longitudinal speech reception measures with continuous interleaved sampling processors for Ineraid implanted subjects. *Ear and Hearing, 20*(3), 228–237.

Peterson, G. E., & Lehiste, I. (1962). Revised CNC lists for auditory tests. *Journal of Speech and Hearing Disorders, 27,* 62–70.

Plant, K., McDermott, H., van Hoesel, R., Dawson, P., & Cowan, R. (2016). Factors predicting postoperative unilateral and bilateral speech recognition in adult cochlear implant recipients with acoustic hearing. *Ear and Hearing, 37*(2), 153–163. https://doi.org/10.1097/aud.0000000000000233

Quittner, A. L., Cejas, I., Wang, N.-Y., Niparko, J. K., & Barker, D. H. (2016). Symbolic play and novel noun learning in deaf and hearing children: Longitudinal effects of access to sound on early precursors of language. *PLoS ONE, 11*(5), e0155964. https://doi.org/10.1371/journal.pone.0155964

Rødvik, A. K. (2008). Perception and confusion of speech sounds by adults with a cochlear implant. *Clinical Linguistics & Phonetics, 22*(4–5), 371–378.

Rotteveel, L. J. C., Snik, A. F. M., Cooper, H., Mawman, D. J., van Olphen, A. F., & Mylanus, E. A. M. (2010). Speech perception after cochlear implantation in 53 patients with otosclerosis: Multicentre results. *Audiology and Neuro-Otology, 15*(2), 128–136. https://doi.org/10.1159/000235578

*__Sagi, E., Kaiser, A. R., Meyer, T. A., & Svirsky, M. A.__ (2009). The effect of temporal gap identification on speech perception by users of cochlear implants. *Journal of Speech, Language, and Hearing Research, 52*(2), 385–395. https://doi.org/10.1044/1092-4388(2008/07-0219)

Santos, C. M. C., Pimenta, C. A. M., & Nobre, M. R. C. (2007). A estratégia PICO para a construção da pergunta de pesquisa e busca de evidências [The PICO strategy for the research question construction and evidence search]. *Revista Latino-Americana de Enfermagem, 15*(3), 508–511. https://doi.org/10.1590/S0104-11692007000300023

Schmidt, M., & Griesser, A. (1997). Long-term stability of fitting parameters with the COMBI 40. *The American Journal of Otology, 18*(6, Suppl.), S109–S110.

*__Shafiro, V., Gygi, B., Cheng, M. Y., Vachhani, J., & Mulvey, M.__ (2011). Perception of environmental sounds by experienced cochlear implant patients. *Ear and Hearing, 32*(4), 511–523. https://doi.org/10.1097/AUD.0b013e3182064a87

*__Shallop, J. K., Arndt, P. L., & Turnacliff, K. A.__ (1992). Expanded indications for cochlear implantation: Perceptual results in seven adults with residual hearing. *Journal of Speech-Language Pathology and Audiology, 16*(2), 141–148.

*__Shannon, R. V., Cruz, R. J., & Galvin, J. J., III.__ (2011). Effect of stimulation rate on cochlear implant users' phoneme, word and sentence recognition in quiet and in noise. *Audiology and Neuro-Otology, 16*(2), 113–123.

*__Shannon, R. V., Galvin, J. J., III, & Baskent, D.__ (2002). Holes in hearing. *Journal of the Association for Research in Otolaryngology, 3*(2), 185–199. https://doi.org/10.1007/s101620020021

Shannon, R. V., Jensvold, A., Padilla, M., Robert, M. E., & Wang, X. (1999). Consonant recordings for speech testing. *The Journal of the Acoustical Society of America, 106*(6), L71–L74.

*__Sheffield, B. M., & Zeng, F. G.__ (2012). The relative phonetic contributions of a cochlear implant and residual acoustic hearing to bimodal speech perception. *The Journal of the Acoustical Society of America, 131*(1), 518–530.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General, 143*(2), 534–547.

*__Singh, S., Kong, Y. Y., & Zeng, F. G.__ (2009). Cochlear implant melody recognition as a function of melody frequency range, harmonicity, and number of electrodes. *Ear and Hearing, 30*(2), 160–168.

*__Skinner, M. W., Arndt, P. L., & Staller, S. J.__ (2002). Nucleus ((R)) 24 advanced encoder conversion study: Performance versus preference. *Ear and Hearing, 23*(1), 2S–17S. https://doi.org/10.1097/00003446-200202001-00002

Skinner, M. W., Fourakis, M. S., Holden, T. A., Holden, L. K., & Demorest, M. E. (1996). Identification of speech by cochlear implant recipients with the multipeak (MPEAK) and spectral peak (SPEAK) speech coding strategies. I. Vowels. *Ear and Hearing, 17*(3), 182–197.

*__Skinner, M. W., Holden, L. K., Whitford, L. A., Plant, K. L., Psarros, C., & Holden, T. A.__ (2002). Speech recognition with the Nucleus 24 SPEAK, ACE, and CIS speech coding strategies in newly implanted adults. *Ear and Hearing, 23*(3), 207–223.

Spencer, P. E. (2004). Individual differences in language performance after cochlear implantation at one to three years of age: Child, family, and linguistic factors. *Journal of Deaf Studies and Deaf Education, 9*(4), 395–412.

*__Stacey, P. C., Raine, C. H., O'Donoghue, G. M., Tapper, L., Twomey, T., & Summerfield, A. Q.__ (2010). Effectiveness of computer-based auditory training for adult users of cochlear implants. *International Journal of Audiology, 49*(5), 347–356.

Stach, B. A. (2009). *Clinical audiology: An introduction* (2nd ed.). New York, NY: Delmar–Cengage Learning.

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., . . . Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal, 343,* d4002. https://doi.org/10.1136/bmj.d4002

*__Svirsky, M. A., Sagi, E., Meyer, T. A., Kaiser, A. R., & Teoh, S. W.__ (2011). A mathematical model of medial consonant identification by cochlear implant users. *The Journal of the Acoustical Society of America, 129*(4), 2191–2200.

*__Teoh, S. W., Neuburger, H. S., & Svirsky, M. A.__ (2003). Acoustic and electrical pattern analysis of consonant perceptual cues used by cochlear implant users. *Audiology and Neuro-Otology, 8*(5), 269–285.

*__Throckmorton, C. S., & Collins, L. M.__ (1999). Investigation of the effects of temporal and spatial interactions on speech-recognition skills in cochlear-implant subjects. *The Journal of the Acoustical Society of America, 105*(2, Pt. 1), 861–873.

Tillman, T. W., & Carhart, R. (1966). *An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6* [Technical report]. San Antonio, TX: USAF School of Aerospace Medicine.

Tobey, E. A., Thal, D., Niparko, J. K., Eisenberg, L. S., Quittner, A. L., & Wang, N.-Y. (2013). Influence of implantation age on school-age language performance in pediatric cochlear implant users. *International Journal of Audiology, 52*(4), 219–229. https://doi.org/10.3109/14992027.2012.759666

*__Tye-Murray, N., Lowder, M., & Tyler, R. S.__ (1990). Comparison of the F0F2 and F0F1F2 processing strategies for the cochlear corporation cochlear implant. *Ear and Hearing, 11*(3), 195–200.

*__Tye-Murray, N., Witt, S., & Castelloe, J.__ (1996). Initial evaluation of an interactive test of sentence gist recognition. *Journal of the American Academy of Audiology, 7*(6), 396–405.

*__Tyler, R. S.__ (1990). Speech perception with the nucleus cochlear implant in children trained with the auditory/verbal approach. *American Journal of Otology, 11*(2), 99–107.

Tyler, R. S., Abbas, P., Tye-Murray, N., Gantz, B. J., Knutson, J. F., McCabe, B. F., . . . Kuk, F. (1988). Evaluation of five different cochlear implant designs: Audiologic assessment and predictors of performance. *Laryngoscope, 98*(10), 1100–1106. https://doi.org/10.1288/00005537-198810000-00013

Tyler, R. S., Gantz, B. J., Woodworth, G. G., Parkinson, A. J., Lowder, M. W., & Schum, L. K. (1996). Initial independent results with the Clarion cochlear implant. *Ear and Hearing, 17*(6), 528–536. https://doi.org/10.1097/00003446-199612000-00008

Tyler, R. S., Lowder, M. W., Parkinson, A. J., Woodworth, G. G., & Gantz, B. J. (1995). Performance of adult Ineraid and Nucleus cochlear implant patients after 3.5 years of use. *Audiology, 34*(3), 135–144.

*__Tyler, R. S., & Moore, B. C. J.__ (1992). Consonant recognition by some of the better cochlear-implant patients. *The Journal of the Acoustical Society of America, 92*(6), 3068–3077. https://doi.org/10.1121/1.404203

Tyler, R. S., Preece, J. P., & Lowder, M. (1983). *The Iowa Cochlear Implant Tests*. Iowa City, IA: University of Iowa.

Tyler, R. S., Preece, J. P., & Tye Murray, N. (1987). *Iowa audiovisual speech perception laser videodisc*. Iowa City, IA: University of Iowa, Department of Otolaryngology Head and Neck Surgery.

Välimaa, T. T., Määttä, T. K., Löppönen, H. J., & Sorri, M. J. (2002a). Phoneme recognition and confusions with multichannel cochlear implants: Vowels. *Journal of Speech, Language, and Hearing Research, 45*(5), 1039–1054.

Välimaa, T. T., Määttä, T. K., Löppönen, H. J., & Sorri, M. J. (2002b). Phoneme recognition and confusions with multichannel cochlear implants: Consonants. *Journal of Speech, Language, and Hearing Research, 45*(5), 1055–1069.

Välimaa, T. T., & Sorri, M. J. (2000). Speech perception after multichannel cochlear implantation in Finnish-speaking postlingually deafened adults. *Scandinavian Audiology, 29*(4), 276–283. https://doi.org/10.1080/010503900750022916

Välimaa, T. T., Sorri, M. J., Laitakari, J., Sivonen, V., & Muhli, A. (2011). Vowel confusion patterns in adults during initial 4 years of implant use. *Clinical Linguistics & Phonetics, 25*(2), 121–144. https://doi.org/10.3109/02699206.2010.514692

Van Tasell, D. J., Greenfield, D. G., Logemann, J. J., & Nelson, D. A. (1992). Temporal cues for consonant recognition: Training, talker generalization, and use in evaluation of cochlear implants. *The Journal of the Acoustical Society of America, 92*(3), 1247–1257. https://doi.org/10.1121/1.403920

*__Van Wieringen, A., & Wouters, J.__ (1999). Natural vowel and consonant recognition by Laura cochlear implantees. *Ear and Hearing, 20*(2), 89–103.

Wie, O. B., Falkenberg, E.-S., Tvete, O., & Tomblin, B. (2007). Children with a cochlear implant: Characteristics and determinants of speech recognition, speech-recognition growth rate, and speech production. *International Journal of Audiology, 46*(5), 232–243.

Wouters, J., Francart, T., & McDermott, H. J. (2015). Sound coding in cochlear implants: From electric pulses to hearing. *IEEE Signal Processing Magazine, 32*(2), 67–80. https://doi.org/10.1109/MSP.2014.2371671

*__Wouters, J., & van den Berghe, J.__ (2001). Speech recognition in noise for cochlear implantees with a two-microphone monaural adaptive noise reduction system. *Ear and Hearing, 22*(5), 420–430. https://doi.org/10.1097/00003446-200110000-00006

Xu, L., Zwolan, T. A., Thompson, C. S., & Pfingst, B. E. (2005). Efficacy of a cochlear implant simultaneous analog stimulation strategy coupled with a monopolar electrode configuration. *Annals of Otology, Rhinology & Laryngology, 114*(11), 886–893.

Appendix (p. 1 of 5)

Search Syntax

Database: EMBASE Classic + EMBASE <1947 to 2014 July 02>

1. (speech sound adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (164)
2. (phoneme adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (534)
3. 1 or 2 (693)
4. (nonsense word* or nonword* or pseudo word*).mp. (2167)
5. (nonword* syllable* or nonsense syllable* or pseudo syllable*).mp. (475)
6. Cochlear Implants/ (9918)
7. Cochlear Implantation/ (64151)
8. [or/6–9,19,25] (0)
9. (speech sound adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (164)
10. (phoneme adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (534)
11. 9 or 10 (693)
12. (nonsense word* or nonword* or pseudo word*).mp. (2167)
13. (nonword* syllable* or nonsense syllable* or pseudo syllable*).mp. (475)
14. Cochlear Implants/ (9918)
15. Cochlear Implantation/ (64151)
16. ((cochlear or auditive or auditory or hearing) adj2 (implant* or prosthes*)).mp. (11677)
17. “prostheses and orthoses”/ (12910)
18. sensory aid/ (40)
19. hearing aid/ (11172)
20. exp hearing disorder/th [Therapy] (6721)
21. exp hearing impairment/rh, th [Rehabilitation, Therapy] (7593)
22. (implant* or prosthes*).mp. (561843)
23. 17 or 18 or 19 or 20 or 21 or 22 (575414)
24. cochlea/ (17468)
25. cochlea*.mp. [mp = title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] (49355)
26. 24 or 25 (49355)
27. 23 and 26 (13633)
28. (implant* or prosthes*).mp. (561843)
29. hearing aid/ (11172)
30. exp hearing impairment/rh, th [Rehabilitation, Therapy] (7593)
31. exp hearing disorder/th [Therapy] (6721)
32. 29 or 30 or 31 (18475)
33. 28 and 32 (4939)
34. or/14–17,27,33 (87916)
35. 11 or 12 or 13 (3261)
36. 34 and 35 (145)

Appendix (p. 2 of 5)

Search Syntax

Database: Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE® <1946 to Present>

1. (speech sound adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (117)
2. (phoneme adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (360)
3. 1 or 2 (473)
4. (nonsense word* or nonword* or pseudo word*).mp. (2020)
5. (nonword* syllable* or nonsense syllable* or pseudo syllable*).mp. (389)
6. Cochlear Implants/ (6699)
7. Cochlear Implantation/ (3664)
8. 6 or 7 (8764)
9. ((cochlear or auditive or auditory or hearing) adj2 (implant* or prosthes*)).mp. (10940)
10. “Prostheses and Implants”/ (36221)
11. Sensory Aids/ (987)
12. Hearing Aids/ (6699)
13. exp Hearing Loss/rh, th [Rehabilitation, Therapy] (9705)
14. exp Persons With Hearing Impairments/rh [Rehabilitation] (488)
15. exp Hearing Disorders/th [Therapy] (5609)
16. (implant* or prosthes*).mp. (452036)
17. or/10–16 (462905)
18. cochlea*.mp. (39651)
19. Cochlea/ (15557)
20. 18 or 19 (39651)
21. 17 and 20 (11732)
22. (implant* or prosthes*).mp. (452036)
23. Hearing Aids/ (6699)
24. exp Hearing Loss/rh, th [Rehabilitation, Therapy] (9705)
25. exp Persons With Hearing Impairments/rh [Rehabilitation] (488)
26. exp Hearing Disorders/th [Therapy] (5609)
27. 23 or 24 or 25 or 26 (15328)
28. 22 and 27 (5255)
29. or/6–9,21,28 (12981)
30. 3 or 4 or 5 (2829)
31. 29 and 30 (144)

Database: PsycINFO <1806 to June Week 4 2014>

1. (speech sound adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (168)
2. (phoneme adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (484)
3. 1 or 2 (648)
4. (nonsense word* or nonword* or pseudo word*).mp. (4252)
5. (nonword* syllable* or nonsense syllable* or pseudo syllable*).mp. (1587)
6. exp Cochlear Implants/ (1620)

Appendix (p. 3 of 5)

Search Syntax

7. ((cochlear or auditive or auditory or hearing) adj2 (implant* or prosthes*)).mp. (2136)
8. exp Prostheses/ or "prostheses and implants".mp. (2236)
9. exp Partially Hearing Impaired/ or exp Hearing Disorders/ or exp Hearing Aids/ or sensory aids.mp. (17231)
10. (implant* or prosthes*).mp. (12952)
11. 8 or 9 or 10 (28173)
12. cochlea*.mp. or cochlea/ [mp = title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (5518)
13. 11 and 12 (2795)
14. (implant* or prosthes*).mp. (12952)
15. exp Hearing Aids/ (2776)
16. exp Deaf/ or exp Partially Hearing Impaired/ or exp Hearing Disorders/ (16099)
17. 15 or 16 (17202)
18. 14 and 17 (2009)
19. or/6–7,13,18 (2854)
20. 3 or 4 or 5 (6354)
21. 19 and 20 (56)

Database: ERIC <1965 to June 2014>
NB: Because of difficulties in adapting the search strategy in part b), we ran two searches; one adapted search with ERIC subject headings (18 hits) and the EMBASE search, which produced additionally five articles.

1. (speech sound adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (53)
2. (phoneme adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (109)
3. 1 or 2 (160)
4. (nonsense word* or nonword* or pseudo word*).mp. [mp = abstract, title, heading word, identifiers] (1107)
5. (nonword* syllable* or nonsense syllable* or pseudo syllable*).mp. [mp = abstract, title, heading word, identifiers] (100)
6. Cochlear implants/ or Cochlear implantation/ (1846)
7. ((cochlear or auditive or auditory or hearing) adj2 (implant* or prosthes*)).mp. (499)
8. "prostheses and implants"/ or sensory aids/ or hearing aids/ or exp hearing loss/th, rh or hearing impaired persons/rh or hearing disorders/th or (implant* or prosthes*).mp. (2727)
9. Cochlea*.mp. or Cochlea/ (2065)
10. 8 and 9 (2032)
11. (implant* or prosthes*).mp. (706)
12. hearing aids/ or exp hearing loss/th, rh or hearing impaired persons/rh or hearing disorders/th (1846)
13. 11 and 12 (324)
14. or/6–7,10,13 (2035)
15. 3 or 4 or 5 (1349)
16. 14 and 15 (23)
17. (speech sound adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (53)
18. (phoneme adj2 (repetition or recognition or confusion or identification or discrimination or perception)).mp. (109)
19. 17 or 18 (160)
20. (nonsense word* or nonword* or pseudo word*).mp. (1107)
21. (nonword* syllable* or nonsense syllable* or pseudo syllable*).mp. (100)

Appendix (p. 4 of 5)

Search Syntax

22. “cochlear implant*”.mp. [mp = abstract, title, heading word, identifiers] (492)
23. ((cochlear or auditive or auditory or hearing) adj2 (implant* or prosthes*)).mp. (499)
24. (prostheses and implants).mp. [mp = abstract, title, heading word, identifiers] (0)
25. exp Sensory Aids/ (565)
26. hearing aids.mp. (332)
27. hearing impairments/ (6689)
28. exp Deafness/ (6685)
29. (implant* or prosthes*).mp. (706)
30. 25 or 26 or 27 or 28 or 29 (12097)
31. cochlea*.mp. [mp = abstract, title, heading word, identifiers] (530)
32. 30 and 31 (524)
33. (implant* or prosthes*).mp. (706)
34. hearing aids.mp. (332)
35. exp Hearing Impairments/ (11489)
36. 34 or 35 (11513)
37. 33 and 36 (449)
38. or/22–23,32,37 (532)
39. 19 or 20 or 21 (1349)
40. 38 and 39 (18)
41. 16 or 40 (23)
42. 41 not 40 (5)

Database: Web of Science/Web of Knowledge

1. TS = (“speech sound” NEAR/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score))
2. TS = (phoneme NEAR/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score))
3. TS = (consonant NEAR/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score))
4. TS = (vowel NEAR/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score))
5. TS = (“nonsense word*” or “nonword*” or “pseudo word*”)
6. TS = (« nonword* syllable* » or « nonsense syllable* » or « pseudo syllable* »)
7. #6 OR #5 OR #4 OR #3 OR #2 OR #1
8. TS = (“Cochlear implants” or “Cochlear implantation*”)
9. TS = ((cochlear or auditive or auditory or hearing) near/2 (implant* or prosthes*))
10. TS = (“prostheses and implants” or “sensory aids” or “hearing aids” or “hearing loss” or “hearing disorders” or (implant* or prosthes*))
11. TS = (Cochlea*)
12. #11 AND #10
13. TS = (implant* or prosthes*)
14. TS = (“hearing aids” or “hearing loss” or “hearing disorders” or “hearing impair*”)

Appendix (p. 5 of 5)

Search Syntax

15. #14 AND #13
16. #15 OR #12 OR #9 OR #8
17. #16 AND #7

Database: Scopus (Elsevier)
((TITLE-ABS-KEY(“Cochlear implant*”)) OR
(TITLE-ABS-KEY((cochlear or auditive or auditory or hearing) PRE/2 (implant* or prosthes*))) or ((TITLE-ABS-KEY(“prostheses and implants” or “sensory aids” or “hearing aids” or “hearing loss” or “hearing impaired persons” or “hearing disorders” or (implant* or prosthes*))) and
(TITLE-ABS-KEY(cochlea*))) or
((TITLE-ABS-KEY(implant* or prosthes*)) and
(TITLE-ABS-KEY(“hearing aids” or “hearing loss” or “hearing impaired persons” or “hearing disorders”)))) and
(((TITLE-ABS-KEY(“speech sound” PRE/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score)) OR
TITLE-ABS-KEY(phoneme PRE/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score)) OR
TITLE-ABS-KEY(consonant PRE/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score))OR
TITLE-ABS-KEY(vowel PRE/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score)))) or
((TITLE-ABS-KEY(“nonsense word*” or nonword* or “pseudo word*”) OR TITLE-ABS-KEY(“nonword* syllable*” or “nonsense syllable*” or “pseudo syllable*”))) or (((TITLE-ABS-KEY(“speech sound” PRE/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score)) OR TITLE-ABS-KEY(phoneme PRE/2 (repetition or recognition or confusion or identification or discrimination or perception or test or score)))) or ((TITLE-ABS-KEY(“nonsense word*” or nonword* or “pseudo word*”) OR TITLE-ABS-KEY(“nonword* syllable*” or “nonsense syllable*” or “pseudo syllable*”)))))

Appendix 3—Article III

frontiers in Psychology

ORIGINAL RESEARCH
published: 14 August 2019
doi: 10.3389/fpsyg.2019.01813

Consonant and Vowel Confusions in Well-Performing Children and Adolescents With Cochlear Implants, Measured by a Nonsense Syllable Repetition Test

Arne Kirkhorn Rødvik[1,2]*, Ole Tvete[2], Janne von Koss Torkildsen[1], Ona Bø Wie[1,2], Ingebjørg Skaug[3] and Juha Tapio Silvola[1,2,4]

[1] Department of Special Needs Education, Institute of Educational Sciences, University of Oslo, Oslo, Norway, [2] Cochlear Implant Unit, Department of Otorhinolaryngology, Division of Surgery and Clinical Neuroscience, Oslo University Hospital, Oslo, Norway, [3] Cochletten Foundation, Oslo, Norway, [4] Ear, Nose, and Throat Department, Division of Surgery, Akershus University Hospital, Lørenskog, Norway

OPEN ACCESS

Edited by:
Viveka Lyberg Åhlander,
Åbo Akademi University, Finland

Reviewed by:
Anu Sharma,
University of Colorado Boulder,
United States
Ignacio Moreno-Torres,
University of Málaga, Spain
Etienne Gaudrain,
INSERM U1028 Centre de Recherche
en Neurosciences de Lyon, France

***Correspondence:**
Arne Kirkhorn Rødvik
a.k.rodvik@isp.uio.no

Specialty section:
This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Psychology

Received: *07 April 2019*
Accepted: *22 July 2019*
Published: *14 August 2019*

Citation:
Rødvik AK, Tvete O,
Torkildsen JvK, Wie OB, Skaug I and
Silvola JT (2019) Consonant
and Vowel Confusions
in Well-Performing Children
and Adolescents With Cochlear
Implants, Measured by a Nonsense
Syllable Repetition Test.
Front. Psychol. 10:1813.
doi: 10.3389/fpsyg.2019.01813

Although the majority of early implanted, profoundly deaf children with cochlear implants (CIs), will develop correct pronunciation if they receive adequate oral language stimulation, many of them have difficulties with perceiving minute details of speech. The main aim of this study is to measure the confusion of consonants and vowels in well-performing children and adolescents with CIs. The study also aims to investigate how age at onset of severe to profound deafness influences perception. The participants are 36 children and adolescents with CIs (18 girls), with a mean (*SD*) age of 11.6 (3.0) years (range: 5.9–16.0 years). Twenty-nine of them are prelingually deaf and seven are postlingually deaf. Two reference groups of normal-hearing (NH) 6- and 13-year-olds are included. Consonant and vowel perception is measured by repetition of 16 bisyllabic vowel-consonant-vowel nonsense words and nine monosyllabic consonant-vowel-consonant nonsense words in an open-set design. For the participants with CIs, consonants were mostly confused with consonants with the same voicing and manner, and the mean (*SD*) voiced consonant repetition score, 63.9 (10.6)%, was considerably lower than the mean (*SD*) unvoiced consonant score, 76.9 (9.3)%. There was a devoicing bias for the stops; unvoiced stops were confused with other unvoiced stops and not with voiced stops, and voiced stops were confused with both unvoiced stops and other voiced stops. The mean (*SD*) vowel repetition score was 85.2 (10.6)% and there was a bias in the confusions of [iː] and [yː]; [yː] was perceived as [iː] twice as often as [yː] was repeated correctly. Subgroup analyses showed no statistically significant differences between the consonant scores for pre- and postlingually deaf participants. For the NH participants, the consonant repetition scores were substantially higher and the difference between voiced and unvoiced consonant repetition scores considerably lower than for the participants with CIs. The participants with CIs obtained scores close to ceiling on vowels and real-word monosyllables, but their perception was substantially lower for voiced consonants. This may partly be related to limitations in the CI technology for the transmission of low-frequency sounds, such as insertion depth of the electrode and ability to convey temporal information.

Keywords: cochlear implants, speech perception, speech sound confusions, consonants, vowels, hearing

INTRODUCTION

Provided with adequate access to environments in which speech is the common mode of communication, the majority of profoundly deaf children implanted in their sensitive period (before age 3.5–4.0 years) will develop intelligible speech and functional hearing for oral language (Kral and Sharma, 2012; Leigh et al., 2013; Dettman et al., 2016). Early implanted children follow similar development in speech and language as normal-hearing (NH) children do (e.g., the systematic review by Bruijnzeel et al., 2016). However, early implanted children with good speech perception ability do not discriminate minute details of speech, such as voicing, frication, and nasality, as well as their NH peers, even in quiet surroundings (Tye-Murray et al., 1995; Geers et al., 2003).

The present study aims to reveal possible systematic misperceptions of speech sounds in detail for children and adolescents with cochlear implants (CIs) and to investigate how age at onset of severe to profound (pre-, peri-, and postlingual) deafness influences their confusion of speech sounds and features. In the following, we will outline the maturation of the auditory system and the fundamentals of speech processing in CIs, before presenting the rationale for our test design and giving a brief introduction to the Norwegian language.

The human cochlea is fully developed at birth, but the brain's auditory pathways and centers, from the brain stem to the auditory cortex, continue to develop. Conditions for the acquisition of language are optimal in a sensitive period, which can be estimated by measuring the cortical P1 latency response as an index of maturation of the auditory pathway in populations with abnormal auditory experience, such as congenital profound deafness. Sharma et al. (2002a,b,c) found that the optimal sensitive period for cochlear implantation in profoundly deaf children lasts until approximately 3.5–4 years of age, and it is important that children receive auditory stimulation within this critical period. These children can still benefit from CIs until the eventual end of the overall sensitive period, at approximately 6.5–7.0 years of age (Kral and Sharma, 2012). However, later implantation in congenitally deaf children normally results in difficulties with acquiring oral speech and language skills.

As normal maturation of the auditory system depends on adequate auditory input in very early childhood, detection of hearing loss by otoacoustic emissions and/or auditory brainstem responses right after birth is crucial. Immediate programming of hearing aids (HAs) for infants with discovered mild to moderate hearing loss, or of CIs for the profoundly deaf among them, will facilitate stimulation of the brain's auditory pathways in the sensitive period. Clinical findings indisputably show that children with hearing impairments who receive appropriate and early intervention achieve much better hearing and better oral language performance than those who start the process later (Wilson and Dorman, 2008; Niparko et al., 2010; Wie, 2010).

The gradual development and maturation of the auditory system can be seen in outcomes of auditory tests into the late teenage years, with individual variability within a given age (Maxon and Hochberg, 1982; Fischer and Hartnegg, 2004). Children's peripheral hearing is established before their speech. However, the development of the ability to discriminate speech sounds, as well as vocabulary and language, takes many years.

Auditory sensitivity in audiometric tests, in absence of noise or other masking stimuli, is known to improve between infancy and early school age (Olsho et al., 1988; Trehub et al., 1988). Litovsky (2015) suggests that the reason for this improvement is that the tasks used to measure perception of pure-tones do not separate the effects of cognitive ability, motivation, memory, and variability in neural representation of the stimuli. For real-word tests, top-down processing allows for decoding based on context and is facilitated by the lexical content present in real-word stimulus materials or by the intrinsic language proficiency. To diminish the influence of these factors in the present study, auditory skills are measured by a nonsense syllable repetition test (NSRT), which is idealized to measure the perception of speech sounds with only minor influence from top-down processing and with minimal stress on working memory. This test should therefore establish a more correct expression of the true auditory perception skills of a child with CIs.

CI users are often classified into pre-, peri-, and postlingually deaf. In the present study, prelingual deafness is defined as congenital, profound deafness or onset of severe to profound deafness before the age of 12 months. According to the widely used definition by the World Health Organization [WHO] (2019), severe hearing loss is characterized by a pure-tone average (PTA)[1] between a 60 and 80 dB hearing level (HL), and profound hearing loss is characterized by a PTA above 80 dB HL. In prelingually deaf children, the auditory system is immature when hearing is initiated by a CI, whose stimulus signal is different from the signal generated by the inner hair cells in a normal cochlea. The earlier the age at implantation, the faster the adaptation to the novel signal, and the better the speech perception outcomes (Niparko et al., 2010; Tobey et al., 2013; Liu et al., 2015). Furthermore, prelingually deaf children with CIs can be divided into two groups: those who have had no or minimal access to sound and hence acquired very little oral language before implantation (these children are often congenitally deaf

---

[1] PTA is defined as average hearing loss on the frequencies 1,000, 2,000, 3,000, and 4,000 Hz, according to the National Institute for Occupational Safety, and Health [NIOSH] (1996).

and receive a CI before age 1), and those who have acquired oral language and benefited from HAs due to residual hearing, receiving a CI at a higher age.

The children with onset of severe to profound deafness between 1 and 3 years of age are classified as perilingually deaf. postlingual deafness is defined as progressive or sudden hearing loss and onset of severe to profound deafness after age 3 years, with a benefit from HAs and acquired oral language before onset of deafness (Myhrum et al., 2017).

Although language acquisition is a gradual process, the breakpoint of age 1 year for distinguishing between pre- and perilingual deafness is precisely defined for practical reasons. This age corresponds to when infants usually start saying their first words (Darley and Winitz, 1961; Locke, 1983, p. 8). In postlingually deaf adults and children, the neural pathways in the brain have been shaped by acoustic sound perception before onset of deafness. The degree of success with a CI is dependent on how the brain compares the new signal with what was heard previously.

For both the pre-, peri-, and postlingually deaf, auditory deprivation will occur after a period of lack of sensory input. This process entails a degeneration of the auditory system, both peripherally and centrally (Feng et al., 2018), including a degradation of neural spiral ganglion cells (Leake and Hradek, 1988). If profound deafness occurs in the sensitive period before 3.5–4.0 years of age, it arrests the normal tonotopic organization of the primary auditory cortex. This arrest can, however, be reversed after reactivation of afferent input by a CI (Kral, 2013).

The hearing-impaired participants in this study are aided by CIs, which consist of a speech processor on the ear and a surgically implanted electrode array in the cochlea with up to 22 electrical contacts. A speech signal input is received by the built-in speech processor microphone and translated into sequences of electrical pulses in the implant by a stimulation strategy. The main purpose of every such strategy is to set up an electrical signal in the auditory nerve using electrical stimulation patterns in the electrode array to mimic the signal in a normal ear. These patterns vary somewhat between stimulation strategies and implant manufacturers, but they all attempt to convey spectral (frequency-related) and temporal information of the original signal through the implant (Wouters et al., 2015).

The spectral information of the speech signal (e.g., the first and second formant, F1 and F2) is conveyed by the multichannel organization of the implants, by mimicking the tonotopic (place) organization of the cochlea from low frequencies in the apex to high frequencies in the base. This information is implemented in all stimulation strategies from the main (in terms of market share) implant manufacturers today, listed in alphabetical order: Advanced Bionics (Stäfa, Switzerland), Cochlear (Sydney, NSW, Australia), Med-El (Innsbruck, Austria), and Oticon Medical/Neurelec (Vallauris, France).

The temporal information of the speech signal is commonly decomposed into envelope (2–50 Hz), periodicity (50–500 Hz), and temporal fine structure (TFS; 500–10,000 Hz), for instance described by Wouters et al. (2015). The envelope is the slow variations in the speech signal. Periodicity corresponds with the vibrations of the vocal cords, which conveys fundamental frequency (F0) information. TFS is the fast fluctuations in the signal, and contributes to pitch perception, sound localization, and binaural segregation of sound sources.

All stimulation strategies represent high-frequency sounds only by place coding. Moreover, the stimulation rate in every implant is constant, varying between 500 and 3,500 pulses per second for the different manufacturers. Low-frequency sounds can be represented by both temporal and place coding.

In the present study, the consonant and vowel repetition scores and confusions were measured using an NSRT with recorded monosyllabic consonant-vowel-consonant (CVC) and bisyllabic vowel-consonant-vowel (VCV) nonsense words, named nonsense syllables in this article, in an open-set design. By open-set design, we mean that the responses are not made through a forced choice of alternatives, but rather by repetition of what is perceived. The nonsense syllables follow the phonotactic rules of the participants' native language, which in our case is Norwegian (e.g., Coady and Aslin, 2004). To avoid straining the working memory, each stimulus unit was limited to 1 or 2 syllables (Gathercole et al., 1994). In the following, the rationale for the test design is presented.

Speech perception tests for children with CIs are traditionally performed with live or recorded real words or sentences in quiet or in noise (e.g., Harrison et al., 2005; Zeitler et al., 2012; Ching et al., 2018). Such tests indisputably measure the children's language skills in addition to their auditory skills.

There are two methods of making speech perception tests more difficult in order for the test subjects not to perform at ceiling. One is to degrade the speech signal by altering its temporal and spectral information, for instance by adding background noise to the test words or applying high- or low-pass filtering. Perception of speech in background noise is more difficult than in quiet due to factors such as diminished temporal coding (Henry and Heinz, 2012). The other method is to use more challenging test units, such as words without lexical meaning, and assess details in the perception of individual speech sounds under optimal listening conditions. The use of an NSRT in quiet allows for directly studying feature information transmission as opposed to tests relying on a degraded speech signal. In real life, listeners are faced with challenging situations similar to NSRTs when they try to catch an unfamiliar name or are confronted with new vocabulary. New and difficult words are perceived as nonsense syllables until they become internalized as meaningful units.

The measurement of consonant and vowel scores in children with CI's via recorded nonsense syllables has rarely been reported in scientific literature. A systematic review and meta-analysis by Rødvik et al. (2018), found only two studies of this kind (Tyler, 1990; Arisi et al., 2010). Tyler (1990) included five children who were asked to choose between several written alternatives when they identified each nonsense syllable. Their mean (*SD*) age at testing was 8.5 (1.6) years, and they obtained a mean (*SD*) consonant identification score of 30% (13%) (range: 19–50%). The reason for this relatively low score was probably the high age at implantation for these prelingually (N = 2) and postlingually (N = 3) deaf children [mean (*SD*) = 7.4 (1.9) years]. Arisi et al. (2010) included 45 adolescents with a mean (*SD*) age of 13.4 (2.6) years, who obtained a mean (*SD*) consonant identification score of 53.5 (33.6)%. All participants marked their choices with a pen on printed text.

We chose a test with verbal repetition of the test words, to ensure that the test scores would neither be influenced by the test subjects' reading or writing ability nor their computer skills, and that they were not required to relate to anything other than their own hearing and speech as well as their own established phoneme inventory. This design provided detailed information about speech perception and listening capacity for acoustic properties.

Furthermore, an open-set test design was chosen, in which the participants did not know which or how many test units would be presented to them. The participants were thus not limited in their responses and would find no external clues when interpreting what they heard. Previous studies have reported robust effects of competition between items in the mental lexicon and of speaker variability in open-set but not in closed-set tests (e.g., Sommers et al., 1997; Clopper et al., 2006). Moreover, open-set test designs have relatively small learning effects compared to closed-set test designs and can therefore be performed reliably at desired intervals (Drullman, 2005, p. 8).

Open-set test designs also have some disadvantages. For example, they often result in lower overall performance than closed-set test designs and may be challenging to use with low-performing adults and young children. Moreover, they require a substantial effort in post-test analysis if each response is to be transcribed phonetically. Alternatively, responses may be scored simply as correct or incorrect for routine-testing in a clinical practice.

Norwegian is a Northern Germanic language, belonging to the Scandinavian language group. There is no official common Norwegian pronunciation norm, as oral Norwegian is a collection of dialects, and Norwegians normally speak the dialect of their native region. Norwegian has two lexical tones (except for certain dialects), which span across bisyllabic words and are used as a distinguishing, lexical factor. The tones' melodies are indigenous to each dialect and are recognized as a dominant and typical prosodic element of the dialect, distinguishing it from other dialects. Norwegian has a semi-transparent orthography, meaning that there is not a consistent one-to-one correspondence between letters and phonemes, like for instance in Finnish, but a much more transparent relation between phonemes and letters than in English (Elley, 1992). In the present study, only speech sounds common for all Norwegian dialects are included; see **Table 1** and **Figure 1** for an overview.

The overall objective of the present study was to measure the perception of speech sounds in well-performing children and adolescents with CIs with an NSRT.

The two sub-objectives were as follows:

Objective 1: To identify the most common vowel and consonant confusions and the most common confusions of the phonetic features voicing, frication, stopping, nasality, and laterality in a sample of well-performing children and adolescents with CIs.

Objective 2: To investigate how age at onset of severe to profound (pre-, peri-, and postlingual) deafness in children and adolescents with CIs influences their confusion of speech sounds and features.

MATERIALS AND METHODS

Abbreviations and acronyms are presented in **Table 2**.

Participants

Informed written consent was obtained from all participants and their legal guardians, according to the guidelines in the Helsinki declaration (World Medical Association [WMA], 2017). The project was approved by the ethical committee of the regional health authority in Norway (REC South East) and by the data protection officer at Oslo University Hospital.

Participants With CIs

Thirty-six children and adolescents with CIs (18 girls) participated in this study. Their age range was 5.9–16.0 years [mean (*SD*) = 11.6 (3.0) years]. Oral language was the main communication mode for all participants. The study sample included 29 prelingually and 7 postlingually deaf participants using the CI stimulation strategies FS4 ($N = 4$), FSP ($N = 7$), and CIS + ($N = 2$) from Med-El and ACE ($N = 23$) from Cochlear (abbreviations are explained in **Table 2**).

The following inclusion criteria were met for all of these participants: minimum 6 months of implant use, more than 3 months since the activation of the second CI (if they had one), and unchanged processor settings for at least the last 2 months. Furthermore, the participants were required to obtain a score of more than 50% on the HIST monosyllable test in free-field (Øygarden, 2009) and to spontaneously pronounce 100% of all the Norwegian speech sounds correctly. Subjects with a contralateral HA were excluded.

All the included participants were enrolled in the CI program at Oslo University Hospital and were recruited for the present study as part of their ordinary follow-up appointments. Individual demographic information is shown in **Supplementary Table S1**, and individual test results are listed in **Supplementary Table S2**.

Reference Groups

The two reference groups of NH participants were: seventeen 6-year-olds (7 girls; [mean (*SD*) age = 5.9 (0.3) years; range: 5.3–6.3 years]), and twelve 13-year-olds (7 girls; [mean (*SD*) age = 13.0 (0.3) years; range: 12.5–13.3 years]). Six years was an appropriate lower age limit in the reference group, as the majority of children of this age were able to pronounce all the speech sounds correctly in their own dialect. The NH 6-year-olds were mainly recruited from kindergartens near the hospital, and the 13-year-olds were recruited from a primary school nearby.

Normal hearing was confirmed by pure-tone audiometry showing audiometric thresholds at 20 dB (HL) or better on frequencies between 125 and 8,000 Hz. We chose a level of uncertainty of 5 dB, according to the *SDs* of measured audiometric thresholds in a large group of NH listeners in a study by Engdahl et al. (2005). Thus, also children and

TABLE 1 | Simplified IPA chart displaying the speech sounds used in the NSRT.

	Place of articulation													
	Bilabial		**Labiodental**		**Dental**		**Post-alveolar**		**Palatal**		**Velar**		**Glottal**	
Manner of articulation	**U**	**V**	**U**	**V**	**U**	**V**	**U**	**V**	**U**	**V**	**U**	**V**	**U**	**V**
Stops	[p]	[b]			[t]	[d]					[k]	[g]		
Fricatives			[f]		[s]		[ʃ]			[j]			[h]	
Nasals		[m]				[n]						[ŋ]		
Lateral						[l]								

U = unvoiced; V = voiced.

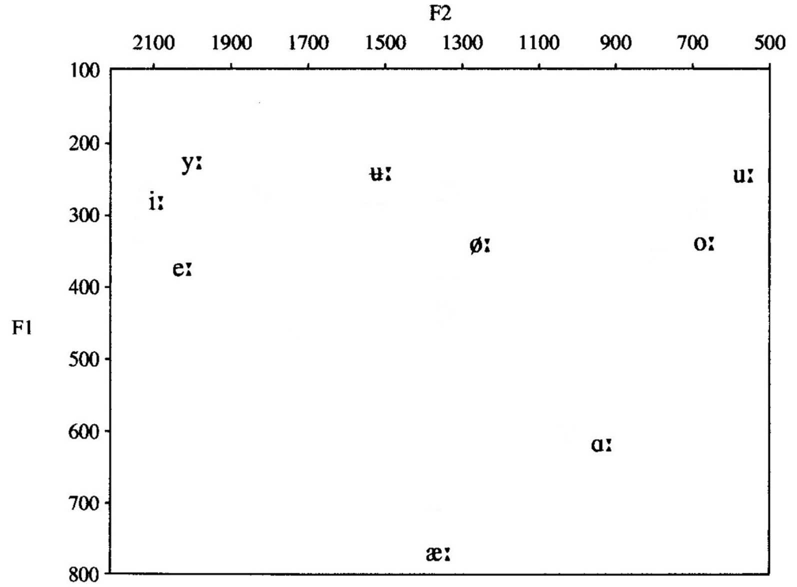


FIGURE 1 | Simplified vowel chart displaying the long Norwegian vowels used in the NSRT, plotted according to the two first formant frequencies, F1 and F2 [modified after Kristoffersen, 2000 (2000, p. 17)].

adolescents with hearing thresholds at 25 dB were included. The middle-ear status of the reference groups was checked with tympanometry and otomicroscopy by an ear, nose, and throat specialist before audiometry.

Inclusion Criteria for All Groups

All participants were required to have Norwegian as their native language and to obtain a 100% score on a pronunciation test of all the target speech sounds in the NSRT.

Test Descriptions

The Nonsense Syllable Repetition Test

The NSRT contains the 16 consonant sounds that are common for all Norwegian dialects, [p, t, k, s, ʃ, f, h, b, d, ɡ, ʝ, v, n, m, ŋ, l], and 11 additional consonant sounds that are used in local Norwegian dialects. To avoid dialect background as a confounding factor in our study, only the first-mentioned 16 consonants were included in the analyses, as they were familiar to all participants. The consonants were placed in a bisyllabic VCV context with the three main cardinal vowels in Norwegian, /ɑː, iː, uː/ (see **Supplementary Table S3**). **Table 1** presents a simplified IPA chart of the included consonants, classified by manner and place of articulation, and by voicing/non-voicing.

The NSRT also contains the nine Norwegian long vowels, [ɑː, eː, iː, uː, ʉː, yː, æː, øː, ɔː], presented in a monosyllabic CVC context with /b/ as the chosen consonant (see the vowel chart in **Figure 1** and an overview of the nonsense syllables in **Supplementary Table S3**).

None of the CVC or VCV combinations presented in the test had lexical meaning in Norwegian. Recording and preparation of the test was mainly done with the computer program Praat (Boersma and Weenink, 2018) and is described in **Supplementary Data Sheet S1** and Introduction provides the rationale for using a repetition test with nonsense syllables in an open-set design.

Real-Word Monosyllable Test

The perception of real-word monosyllables was measured by the HIST monosyllable test in free-field, a test with 50 Norwegian phonetically balanced words, which produces a percent score

TABLE 2 | List of acronyms and abbreviations.

Number	Abbreviation/ acronym	Meaning
1	ACE	Advanced combination encoder (stimulation strategy from Cochlear)
2	CI	Cochlear implant
3	CIS	Continued interleaved sampling (generic stimulation strategy)
4	CM	Confusion matrix
5	CVC	Consonant-vowel-consonant
6	F0, F1, F2	Fundamental frequency, first formant, and second formant
7	FSP/FS4/FS4-p	Fine structure processing (stimulation strategies from Med-El)
8	HA	Hearing aid
9	HIST	Høgskolen i Sør-Trøndelag (real-word monosyllable test)
10	HL	Hearing level
11	NH	Normal-hearing
12	NSRS	Nonsense syllable repetition score
13	NSRS-C	Nonsense syllable repetition score – consonants
14	NSRS-C_{voi}	Nonsense syllable repetition score – voiced consonants
15	NSRS-C_{unvoi}	Nonsense syllable repetition score – unvoiced consonants
16	NSRS-C_{aCa}	Nonsense syllable repetition score – consonants in the aCa context
17	NSRS-C_{iCi}	Nonsense syllable repetition score – consonants in the iCi context
18	NSRS-C_{uCu}	Nonsense syllable repetition score – consonants in the uCu context
19	NSRS-C_{pre}	Nonsense syllable repetition score – consonants repeated by prelingually deaf
20	NSRS-C_{post}	Nonsense syllable repetition score – consonants repeated by postlingually deaf
21	NSRS-V	Nonsense syllable repetition score – vowels
22	NSRS-V_{pre}	Nonsense syllable repetition score – vowels repeated by prelingually deaf
23	NSRS-V_{post}	Nonsense syllable repetition score – vowels repeated by postlingually deaf
24	NSRT	Nonsense syllable repetition test
25	PTA	Pure-tone average
26	REC	Regional ethical committee
27	T, T_{max}, T_{rel}	Speech transmission index (absolute, maximum, and relative)
28	TFS	Temporal fine structure
29	VCV	Vowel-consonant-vowel
30	VOT	Voice onset time

(Øygarden, 2009). The test words were presented at 65 dB(A), and 1 out of 12 lists was chosen.

Pronunciation Test

A sample of "Norsk fonemtest" (Norwegian test of phonemes; Tingleff, 2002) with 28 of its 104 pictures, was used to assess the participants' ability to pronounce all Norwegian consonants and vowels correctly. The selected test items presented the target phoneme in the medial position to match their position in the NSRT. Only those who obtained a 100% score on this test were included in the study.

Procedure and Design

The test words were presented from a SEAS 11F-LGWD 4.5" loudspeaker (Moss, Norway), in an anechoic chamber via the computer program SpchUtil, v. 5 (Freed, 2001). The hard disk recorder Zoom H4n (Hauppauge, NY, United States) was used to record the repeated test words and the naming of the pictures. The distance between the loudspeaker and the participants was 1.5 m, and the equivalent sound level in listening position was 65 dB(A).

Testing of Children and Adolescents With CIs

The NSRT was conducted by playing the recorded CVC and VCV nonsense syllables in randomized order and recording participants' verbal repetitions. The participants were exposed to auditory stimuli only and could not rely on lipreading. They were informed that words with no meaning would be presented to them, but they were not given any further details about how many, which words, and in which consonant or vowel context the speech sounds would be presented.

The participants were instructed to repeat what they heard and to guess if they were unsure, in order to achieve a 100% response rate. Each speech stimulus was presented only once, and the participants were not allowed to practice before being tested or provided with feedback during the testing.

The ecological validity of the testing was optimized by having the participants use the everyday settings of their speech processors instead of switching off front-end sound processing, which has been done in similar studies (e.g., Wolfe et al., 2011). The speech processors were quality checked before testing, and new programming was not performed prior to the testing.

Unaided pure-tone audiometry was performed to check for residual hearing, if these results were not present in the patient's file. Otomicroscopy was performed by an ear, nose, and throat specialist if the participant had residual hearing in one ear or if middle-ear problems were suspected.

Fifty HIST monosyllabic test words in free-field were conducted with all the participants with CIs.

Testing of Normal-Hearing Children and Adolescents

The test setup for the NH reference groups corresponded to that for the participants with CIs, except that the HIST monosyllable test was not conducted, because listeners with normal hearing typically perform at the ceiling level on this test.

Phonetic Transcription and Scoring

The recordings of the participants' repetitions were transcribed by two independent, trained phoneticians, who were blind to the purpose of the study and to what kind of participant groups they transcribed. The transcribers performed a broad phonetic transcription of the nonsense syllables in the test, including primary and secondary stress, and lexical tone, but not suprasegmentals.

The transcriptions of the two phoneticians were compared, and in the case of disagreement between the transcribers, the first author listened to the recordings and picked the transcription that he judged to be correct. The mean (*SD*; range) exact percent agreement between the two transcribers was 82.8 (6.6; 66.7–98.2)% for the participants with CIs and 89.2 (7.5; 68.4–100)% for the NH reference groups.

The repetitions of each target speech sound were scored as either correct (1) or incorrect (0). The total scores were calculated by dividing the number of correctly repeated responses by the total number of stimuli, for the consonants, averaged for the three vowel contexts (NSRS-C), for the vowels (NSRS-V), for the consonants in aCa, iCi, and uCu contexts (NSRS-C_{iCi}, NSRS-C_{aCa}, and NSRS-C_{uCu}), and for the voiced and unvoiced consonants averaged for the three vowel contexts (NSRS-C_{voi} and NSRS-C_{unvoi}). The consonant and vowel scores for the subgroups of prelingually and postlingually deaf were calculated by dividing the number of correctly repeated responses by the total number of stimuli for each subgroup (NSRS-C_{pre}, NSRS-C_{post}, NSRS-V_{pre}, and NSRS-V_{post}). The nonsense syllable repetition score (NSRS) was produced by calculating a weighted mean of NSRS-V and NSRS-C, in which the weights were determined by the number of different vowels (9) and consonants (16) in the test [NSRS = (NSRS-V × 9 + NSRS-C × 16)/25].

Analysis

The 12 variables mentioned in the previous section (#12–23 in **Table 2**) were constructed to score the performance on the NSRT for the three groups of participants, and means, medians, and standard deviations were calculated for all variables. The consonant speech features voicing, stopping, frication, nasality, and laterality were examined separately in the analyses. Assumptions of a normal distribution were violated due to checking of the data with the Shapiro–Wilk test, possibly due to a ceiling effect in some of the variables. Therefore, scores from the participants with CIs were compared by the non-parametric Wilcoxon signed rank z test for related samples, for the following variables:

- Voiced and unvoiced consonant scores (NSRS-C_{voi} and NSRS-C_{unvoi}).
- The HIST real-word monosyllable score and the NSRS.
- NSRS-C_{aCa}, NSRS-C_{iCi}, and NSRS-C_{uCu}.
- The consonant and vowel scores (NSRS-C and NSRS-V).
- Consonant and vowel scores for the pre- and postlingually deaf (NSRS-C_{pre}, NSRS-C_{post}, NSRS-V_{pre}, and NSRS-V_{post}).

Comparisons of NSRS-C and NSRS-V, and NSRS-C_{voi} and NSRS-C_{unvoi}, were also performed for the NH 6- and 13-year olds. Correlations were calculated with Spearman's rho (ρ).

Scores on all variables were compared between the CI users and the NH 6-year-olds, and between the NH 6-year-olds and the NH 13-year-olds, with the Mann–Whitney *U* test for independent samples. To determine statistical significance, an alpha (α) level of 0.05 was chosen for all tests.

Box-and-whiskers were used to display the score distribution for HIST monosyllables, NSRS-V, NSRS-C_{unvoi}, and NSRS-C_{voi} for the three participant groups (see **Figure 2**). All statistical analyses were performed by SPSS v. 24.0 (SPSS Inc., Chicago, IL, United States). A Holm-Bonferroni correction was used to correct for multiple comparisons in all the statistical tests.

Information Transmission for Subgroup Comparisons of Speech Sound Features

The speech sound confusions were organized into confusion matrices (CMs). The CM for the consonant confusions was submitted to an information transfer analysis. This method was introduced by Miller and Nicely (1955) and is an application of the information measure by Shannon (1948) to obtain data from a speech repetition task and measure the covariance of input and output in a stimulus-response system. The method produces a measure of mean logarithmic probability. The logarithm is taken to the base 2, and the measure can thus be called the average number of binary decisions needed to specify the input, or the number of bits of information per stimulus. The method has been used in a large number of studies of the speech sound perception of implantees (e.g., Tye-Murray et al., 1990; Tyler and Moore, 1992; Doyle et al., 1995; Sheffield and Zeng, 2012; Yoon et al., 2012).

The advantage of using this unit instead of recognition scores of correct and incorrect repetitions that are measured binarily is that the repetition errors within the same category of speech sounds obtain higher scores than repetition errors between different categories.

The information transmission (T) was calculated with the formula below:

$$T = -\sum_i \sum_j \frac{n_{ij}}{n} \log_2 \frac{\frac{n_i}{n}\frac{n_j}{n}}{\frac{n_{ij}}{n}}$$

Here, i and j are the stimulus number and response number (the column and row numbers of the CM, respectively), n_{ij} is the cell value, n_i is the row sum, n_j is the column sum, and n is the total sum.

The relative transmission, T_{rel}, is given by $T_{\mathrm{rel}} = T/T_{\max}$, in which $T_{\max}$ is the maximum transmission of information. $T_{\max}$ describes the transmission if all the speech sounds were repeated correctly and no stimulus/response pairs were missing, and T is the absolute transmission. T_{rel} was calculated for the speech sound feature contrasts voicing versus non-voicing, nasality versus non-nasality, frication versus stopping, and nasality versus the lateral [l] for the subgroups of the prelingually (N = 29) and postlingually (N = 7) deaf.

The information transmissions for the subgroups were compared by collapsing the CMs in **Table 6** and analyzing them by χ^2 statistics. Fisher's exact test was applied if the number in one of the quadrants in the 2 × 2 tables was lower than 5. Our null hypothesis was that the information transmission was equally large for both pre- and postlingually deaf participants. A histogram was constructed to visualize the transmission of speech sound features for the two groups (**Figure 3**).

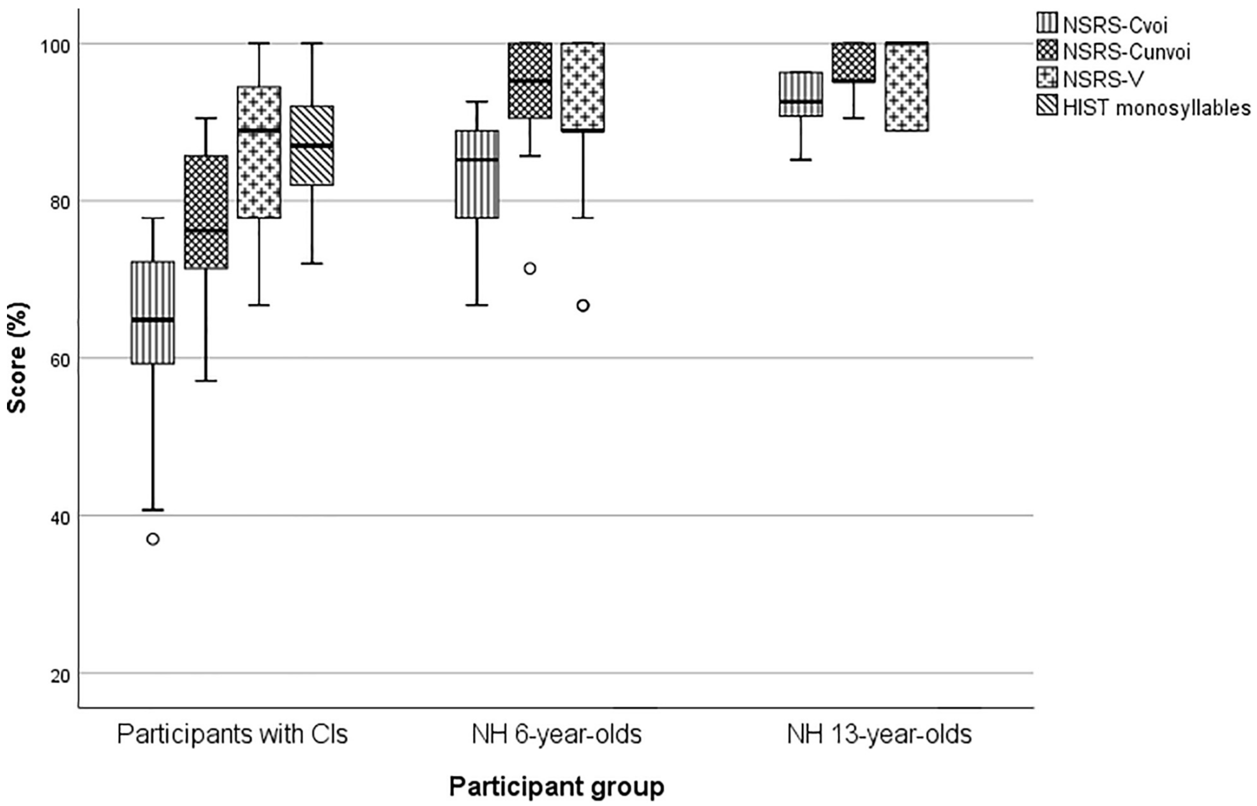


FIGURE 2 | Unvoiced and voiced consonant scores, vowel scores, and monosyllable scores for the three participant groups. The small circles are outliers that represent scores larger than 1.5 times the interquartile range of the box.

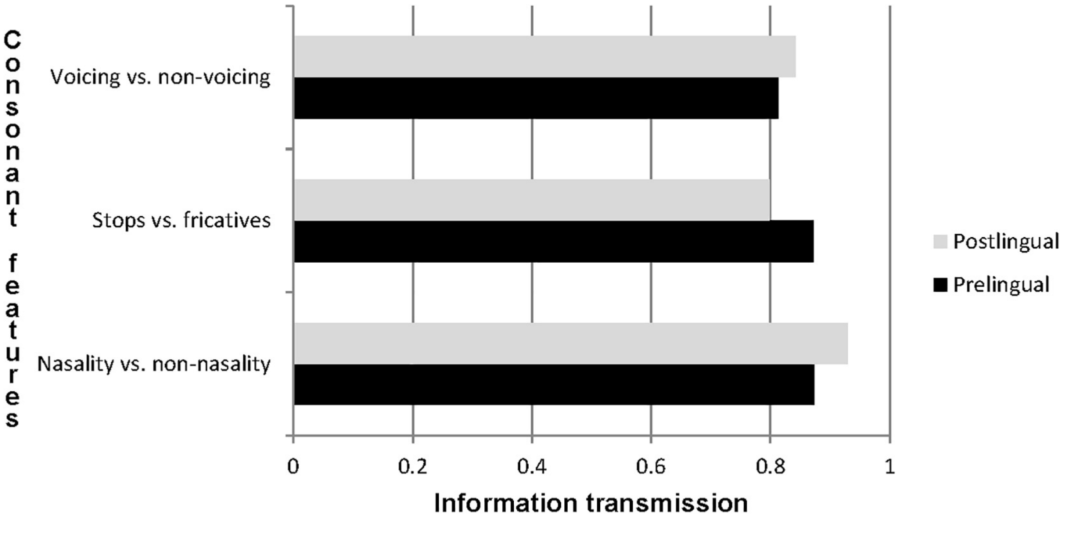


FIGURE 3 | Relative transmission of speech features for pre- and postlingually deaf participants with CIs.

RESULTS

Study Characteristics

The medians of the three groups of participants are displayed in **Table 3**, and comparisons of the participants with CIs and the NH 6-year-olds, and of the NH 6-year-olds and the NH 13-year-olds with independent sample Mann–Whitney tests, are displayed in **Table 4**. The results show, as expected, that the NH 6-year-olds had significantly higher scores than the participants with CIs on all variables, except on the NSRS-V. The comparisons

TABLE 3 | *M, Md*, and *SD* of the study variables for the participants with CIs, the NH 6-year-olds, and the NH 13-year-olds.

	CI users (*N* = 36)			**NH 6-year-olds (*N* = 17)**			**NH 13-year-olds (*N* = 12)**		
Variable (%)	***M* (*SD*)**	***Md***	**Range**	***M* (*SD*)**	***Md***	**Range**	***M* (*SD*)**	***Md***	**Range**
NSRS	75.2 (8.0)	77.3	56.0–89.3	87.6 (5.8)	88.0	72.0–94.7	94.8 (2.0)	94.7	90.7–97.3
NSRS-C	69.6 (8.0)	70.8	50.0–83.3	86.9 (6.1)	87.5	72.9–93.8	94.4 (2.7)	95.8	89.6–97.9
NSRS-C_{aCa}	78.0 (8.6)	81.3	56.3–93.8	90.1 (7.3)	87.5	75.0–100	97.9 (3.1)	100	93.8–100
NSRS-C_{iCi}	69.3 (12.3)	71.9	25.0–87.5	89.3 (4.8)	87.5	81.3–100	96.4 (5.0)	100	87.5–100
NSRS-C_{uCu}	61.5 (13.1)	62.5	31.3–93.8	81.3 (12.1)	87.5	56.3–100	89.1 (3.9)	87.5	81.3–93.8
NSRS-C_{voi}	63.9 (10.6)	64.9	37.0–77.8	82.6 (7.5)	85.2	66.7–92.6	92.6 (3.5)	92.6	85.2–96.3
NSRS-C_{unvoi}	76.9 (9.3)	76.2	57.1–90.5	92.4 (7.5)	95.2	71.4–100	96.8 (3.1)	95.2	90.5–100
NSRS-C_{pre}	69.1 (7.8)	70.8	50.0–81.3	–	–	–	–	–	–
NSRS-C_{post}	71.4 (9.0)	70.8	56.3–83.3	–	–	–	–	–	–
NSRS-V	85.2 (10.9)	88.9	66.7–100	88.9 (11.1)	88.9	66.7–100	95.4 (5.7)	100	88.9–100
NSRS-V_{pre}	86.2 (10.1)	88.9	66.7–100	–	–	–	–	–	–
NSRS-V_{post}	81.0 (13.9)	88.9	66.7–100	–	–	–	–	–	–
HIST monosyllable score	86.9 (6.7)	87.0	72.0–100	–	–	–	–	–	–

of the medians of the NH 6- and 13-year-olds show a significantly higher score for the 13-year-olds for all variables except NSRS-C_{uCu}, NSRS-C_{unvoi}, and NSRS-V.

In **Table 5** the medians for the three groups of participants were compared with Wilcoxon's signed rank test and Mann-Whitney's U-test, and furthermore, correlations between the HIST score and NSRS-C_{voi}, NSRS-C_{unvoi}, and NSRS-V were shown. For the children with CIs, statistically significant differences were found for NSRS-V versus NSRS-C, NSRS-C_{unvoi} versus NSRS-C_{voi}, NSRS-C_{aCa} versus NSRS-C_{iCi}, and NSRS-C_{aCa} versus NoSRS-C_{uCu}. No statistically significant differences were found for NSRS-C_{iCi} versus NSRS-C_{uCu}, NSRS-C_{pre} versus NSRS-C_{post}, and NSRS-V_{pre} versus NSRS-V_{post}. For the NH participants, no statistically significant difference was found, except for the comparison of NSRS-C_{unvoi} and NSRS-C_{voi} for the NH 6-year-olds.

Consonant Confusions

Tables 6, **7** show the CMs for the 16 consonants in aCa, iCi, and uCu contexts for the 36 participants with CIs. The consonants are grouped primarily as voiced and unvoiced and secondarily according to manner of articulation. Of the consonant stimuli, 223 (12.9%) were repeated as consonant clusters or as consonants other than the ones listed in the CM and were excluded from the analyses. These are listed in the unclassified category of the CM.

The consonant CM in **Table 6** shows a devoicing bias for the stops. Unvoiced consonants are in general most frequently confused with other unvoiced consonants and voiced consonants are most frequently confused with other voiced consonants, except for the voiced stops, which are frequently repeated as unvoiced stops. Furthermore, there are highly populated clusters of correct repetitions around voiced and unvoiced stops, voiced and unvoiced fricatives, and nasals.

Table 7 shows that the highest proportion of correct repetitions was within the manner-groups of unvoiced fricatives; 90.5% of these were repeated as the same, or as another unvoiced fricative, and of unvoiced stops; 85.8% were repeated as the same, or as another unvoiced stop. Among the nasals, 81.2% were repeated as the same, or as another nasal, among the voiced fricatives, 79.2% were repeated as the same, or as another voiced fricative, and among the voiced stops, 79.3% were repeated as the same, or as another voiced stop. The highest proportion of consonant confusions was found for the lateral [l], with a correct score of only 61.1%.

The correct repetition scores of the categories of speech features in **Figure 4** ranged from 60% to 80%, except for the nasals, which had a score slightly below 50%. The most common confusions were between consonants with the same manner and same voicing (Type 1 confusions). The least common confusions were between consonants with a different manner and opposite voicing (Type 3 confusions). The number of unclassified confusions, which includes consonant clusters and consonant sounds other than the stimuli, was also substantial, particularly for the lateral [l].

TABLE 4 | Comparisons of the study variables for the participants with CIs, the NH 6-year-olds, and the NH 13-year-olds.

	CI users vs. NH 6-year-olds*				**NH 6-year-olds vs. NH 13-year-olds****			
Variable (%)	***U***	***z***	***p***	***r***	***U***	***z***	***p***	***r***
NSRS	47.0	−4.94	<0.001	0.68	20.5	−0.64	<0.001	0.12
NSRS-C	23.0	−5.41	<0.001	0.74	19.5	−3.73	<0.001	0.69
NSRS-C_{aCa}	84.5	−4.30	<0.001	0.59	34.0	−3.17	0.002	0.59
NSRS-C_{iCi}	22.5	−5.47	<0.001	0.75	35.0	−3.12	0.002	0.58
NSRS-C_{uCu}	85.5	−4.25	<0.001	0.58	60.0	−1.96	0.050***	0.36
NSRS-C_{voi}	40.0	−5.10	<0.001	0.70	18.5	−3.76	<0.001	0.70
NSRS-C_{unvoi}	61.0	−4.70	<0.001	0.65	65.5	−1.69	0.091	0.31
NSRS-V	264.5	−0.83	0.404	0.11	68.5	−1.62	0.105	0.30

The columns show the results of comparisons of means with the Mann–Whitney independent samples U-test between participants with CIs and NH 6-year olds. **The columns show the results of comparisons of means with the Mann–Whitney independent samples U-tests between NH 6- and 13-year-olds. ***The comparison was non-significant after adjusting for multiple testing. The medians and sample sizes that were used in the analyses can be found in* ***Table 3**.*

TABLE 5 | Comparisons of the study variables for the participants with CIs, the NH 6-year-olds, and the NH 13-year-olds.

Comparison	Participant group	Statistical test	ρ	*U*	*z*	*p*	*r*
HIST vs. NSRS-C$_{\text{unvoi}}$	CI	S	0.26	–	–	0.13	–
HIST vs. NSRS-C$_{\text{voi}}$	CI	S	0.41	–	–	0.013*	–
HIST vs. NSRS-V	CI	S	0.18	–	–	0.31	–
HIST vs. NSRS	CI	W	–	–	−4.90	< 0.001	0.82
NSRS-V vs. NSRS-C	CI	W	–	–	−5.12	< 0.001	0.85
	NH6	W	–	–	−0.78	0.43	0.19
	NH13	W	–	–	−0.32	0.75	0.09
NSRS-C$_{\text{unvoi}}$ vs. NSRS-C$_{\text{voi}}$	CI	W	–	–	−4.46	< 0.001	0.74
	NH6	W	–	–	−3.15	0.002	0.76
	NH13	W	–	–	−2.60	0.009*	0.75
NSRS-C$_{\text{aCa}}$ vs. NSRS-C$_{\text{iCi}}$	CI	W	–	–	−3.96	< 0.001	0.66
	NH6	W	–	–	−0.18	0.86	0.04
	NH13	W	–	–	−0.97	0.33	0.27
NSRS-C$_{\text{aCa}}$ vs. NSRS-C$_{\text{uCu}}$	CI	W	–	–	−4.75	< 0.001	0.79
	NH6	W	–	–	−2.64	0.008*	0.64
	NH13	W	–	–	−2.99	0.003*	0.86
NSRS-C$_{\text{iCi}}$ vs. NSRS-C$_{\text{uCu}}$	CI	W	–	–	−2.76	0.006*	0.46
	NH6	W	–	–	−2.51	0.012*	0.61
	NH13	W	–	–	−2.72	0.006*	0.79
NSRS-C$_{\text{pre}}$ vs. NSRS-C$_{\text{post}}$	CI	M-W U	–	85.00	−0.66	0.51	0.11
NSRS-V$_{\text{pre}}$ vs. NSRS-V$_{\text{post}}$	CI	M–W *U*	–	80.00	−0.91	0.36	0.15

*CI = cochlear implant; NH6 = NH 6-year-olds; NH13 = NH 13-year-olds; S = Spearman's correlation test; W = Wilcoxon's signed rank test; M–W U = Mann–Whitney's U-test for independent samples. *Not significant after adjusting for multiple testing. The medians and sample sizes that were used in the analyses can be found in* ***Table 3****.*

TABLE 6 | Confusion matrix for 36 participants with CIs; consonants in the aCa, iCi, and uCu contexts added together.

			Response																	
			Unvoiced							Voiced										
Stimulus			S			F				S			F		Na			L		
			/p/	/t/	/k/	/s/	/ʃ/	/f/	/h/	/b/	/d/	/g/	/j/	/v/	/n/	/m/	/ŋ/	/l/	U	N
Unvoiced	S	/p/	86	6				10		3	1								2	108
		/t/		84	4						2								18	108
		/k/	5	4	89							1							9	108
	F	/s/				93	5	4											6	108
		/ʃ/				13	75												20	108
		/f/	1			14	13	73	4										3	108
		/h/					3	13	81		2		1						8	108
Voiced	S	/b/	13	1	1				1	66	11			4					11	108
		/d/		6							85	3							14	108
		/g/			9				1		2	90		2					4	108
	F	/j/							2			2	88						16	108
		/v/							1		1	1		83		1			21	108
	N	/n/													77	9		2	20	108
		/m/													29	66	2	1	10	108
		/ŋ/											1		43	21	16	4	23	108
	L	/l/										1		1	2			66	38	108
																			Total sum	1,728

S = stops; F = fricatives; Na = nasals; L = lateral [l]; U = unclassified speech sounds and consonant clusters.

TABLE 7 | Confusion matrix of consonant repetitions for participants with CIs, collapsed with regard to manner and place of articulation (percentage of correctly repeated stimulus features in each cell).

			Response (%)																		
			Unvoiced							**Voiced**											
			S			**F**				**S**			**F**		**Na**			**L**			
Stimulus			**/p/**	**/t/**	**/k/**	**/s/**	**/ʃ/**	**/f/**	**/h/**	**/b/**	**/d/**	**/g/**	**/j/**	**/v/**	**/n/**	**/m/**	**/ŋ/**	**/l/**	**U**	**Sum (%)**	**N**
Unvoiced	S	/p/																			
		/t/		85.8				3.1			2.2								9.0	100	324
		/k/																			
		/s/																			
	F	/ʃ/		0.2				90.5			0.5			0.2					8.6	100	432
		/f/																			
		/h/																			
Voiced	S	/b/																			
		/d/		9.3				0.6			79.3			1.9					9.0	100	324
		/g/																			
		/j/																			
	F	/v/						1.4			1.9			79.2		0.5			17.1	100	216
		/n/																			
	N	/m/												0.3		81.2		2.2	16.4	100	324
		/ŋ/																			
	L	/l/									0.9			0.9		1.9		61.1	35.2	100	108
																				Total sum	1, 728

S = stops; F = fricatives; Na = nasals; L = lateral [l]; U = unclassified speech sounds and consonant clusters.

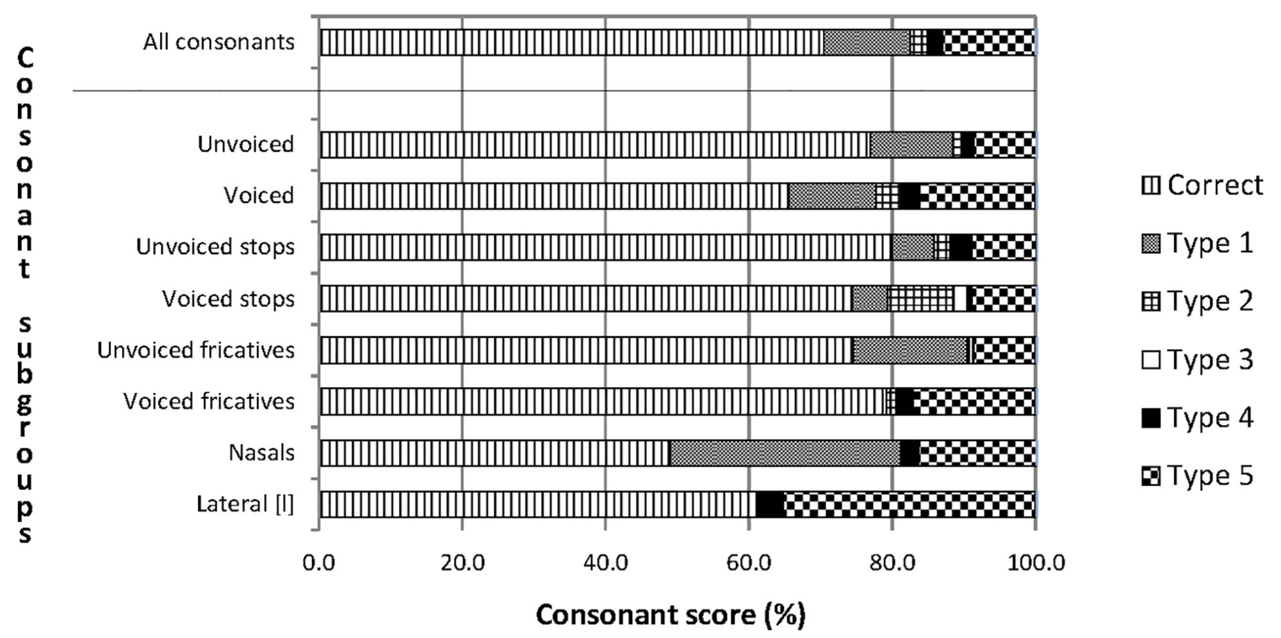


FIGURE 4 | Percentages of correct consonant repetitions and of five types of consonant confusions for participants with CIs. The upper bar describes the complete material of consonant confusions and the eight bars below the horizontal line describe subsets of the material. The units on the horizontal axis are the percentage scores of correct and incorrect repetitions. The bars with a horizontal pattern visualize correct repetitions. Type 1 is confusion between consonants with the same manner and the same voicing. Type 2 is confusion between consonants with the same manner and the opposite voicing. Type 3 is confusion between consonants with a different manner and opposite voicing. Type 4 is confusion between consonants with a different manner and the same voicing. Type 5 is unclassified confusions.

The NH participants repeated almost all the consonants correctly, as shown in **Supplementary Tables S4, S5, S7**, and **S8**. However, we observed an important exception for the 6-year-olds: 10 (19.6%) of the /ŋ/ stimuli were confused with /m/. The 13-year-olds also had an unexpectedly high number of misperceptions of /ŋ/ (7; 19.4%).

Vowel Confusions

Only two cases of unclassified vowels were found among the nine vowels in the bVb context for the 36 participants with CIs (**Table 8**). An [i:]-[y:] perception bias was revealed; [y:] was more frequently repeated as [i:] (67%) than as [y:] (31%).

The CMs for the NH children and adolescents (**Supplementary Tables S6, S9**) show that almost all vowels were repeated correctly. The vowel CM for the 6-year-olds in **Supplementary Table S6** shows some randomly distributed errors, in addition to 6 (35%) of the /y:/ stimuli repeated as either /i:/ or /ʉ:/. There were fewer vowel misperceptions for the 13-year-olds than for the 6-year-olds, but even so, 3 (25%) of the /y:/ stimuli were repeated as /i:/, as displayed in **Supplementary Table S9**.

Perception of Consonant Features Compared by Information Transmission and Chi Square Statistics Between the Pre- and Postlingually Deaf

Figure 3 shows that nasality versus non-nasality had the highest information transmission, and voicing versus non-voicing had the lowest. The information transmission of speech features did not display large differences between pre- and postlingually deaf participants.

Chi square testing showed no statistically significant differences between the transmission of voicing and non-voicing ($\chi^2 = 1.16$; $p = 0.28$), nor between the transmission of nasality and non-nasality ($\chi^2 = 0.41$; $p = 0.52$), nor between the transmission of stops and fricatives ($\chi^2 = 1.12$; $p = 0.29$). **Supplementary Table S10** displays the three 2 × 2 matrices that these analyses are based on.

DISCUSSION

The objective of this study was to assess the effectiveness of CIs by obtaining a measure of speech sound confusions in well-performing children and adolescents with CIs, using an NSRT, and to investigate whether the perception of speech features differs between the pre- and postlingually deaf. The study was cross-sectional, and it included 36 participants with CIs and 2 reference groups (17 NH 6-year-olds and 12 NH 13-year-olds).

An important finding was that unvoiced consonants were significantly less confused than voiced consonants for the participants with CIs. Moreover, there was a devoicing bias for the stops; unvoiced stops were confused with other unvoiced stops and not with voiced stops, and voiced stops were confused with both unvoiced stops and other voiced stops. Another major finding was that there was no significant difference between the perception of speech sound features for pre- and postlingually deaf CI users.

A central issue when assessing consonant confusions in participants with CIs is to investigate the underlying reasons. Are the confusions caused by limitations in the implants, are they due to immature cognitive development, or can they be explained by other factors? The difference between the NSRS and the HIST real-word monosyllable score suggests that the participants with CIs rely substantially on their language proficiency and the top-down processing introduced by lexical content present in real-word stimulus material. The finding is in line with a study on NH individuals by Findlen and Roup (2011), who investigated dichotic speech recognition performance for nonsense and real-word CVC syllables, and found that performance with nonsense CVC syllables was significantly poorer. Findlen and Roup's study is to the authors' knowledge the only previous investigation of recognition differences between real-word and nonsense CVC syllable stimuli that have similar phonetic content but differ in lexical content.

The moderate correlation between NSRS-C_{voi} and HIST monosyllables suggests that problems with perceiving the real-word monosyllables could partly be explained by difficulties in perceiving the voiced consonants.

TABLE 8 | Confusion matrix of vowel repetitions in the bVb context for participants with CIs.

	Response										
Stimulus	**/bɑ:b/**	**/be:b/**	**/bi:b/**	**/bu:b/**	**/bʉ:b/**	**/by:b/**	**/bæ:b/**	**/bø:b/**	**/bɔ:b/**	**U**	**N**
/bɑ:b/	35									1	36
/be:b/		35	1								36
/bi:b/			36								36
/bu:b/				36							36
/bʉ:b/			2		30	4					36
/by:b/			24		1	11					36
/bæ:b/	1						35				36
/bø:b/		2	1		5	1		26		1	36
/bɔ:b/				1					35		36
									Total sum		324

U = unclassified.

The Results of the Participants With CIs Related to Those of the NH Reference Groups

As expected, the scores on the NSRT were higher for the NH 13-year-olds than for the NH 6-year-olds for all variables. However, the differences were not significant for NSRS-C_{uCu}, NSRS-C_{unvoi}, and NSRS-V, probably because NH 13-year-olds usually have a more developed phonemic lexicon and higher phonemic awareness, or because of age-related differences in attentiveness during the task. We compared the scores of the participants with CIs only to those of the NH 6-year-olds, as these two groups are closest in hearing age. Significant differences were found between the groups of NH 6-year-olds and CI users for all variables except for the NSRS-V, which was just as high for both groups. This may be due to the long duration and high energy of the vowels in the NSRT.

For the NH groups, there were no statistically significant differences in any of the comparisons, except for unvoiced versus voiced consonant score for the NH 6-year-olds. Since this difference was not found for the NH 13-year-olds, this can probably be explained by language immaturity and fatigue.

For the participants with CIs, the difference between voiced and unvoiced consonant scores seems to be mostly due to the fact that unvoiced stops in Norwegian, /p, t, k/, are strongly aspirated and hence have a substantially longer voice onset time (VOT)[2] than the voiced stops, /b, d, ɡ/ (Halvorsen, 1998). For both CI users and the NH 6-year-olds, the low, voiced consonant score is likely due to the nasals, /m, n, ŋ/, being confused with one another, and by /l/ having a low recognition score.

The Most Common Confusions of Consonants and Vowels for Participants With CIs

Most consonant confusions observed in the present study can be explained by acoustic similarity in manner and voicing, a conclusion that has also been reached in many previous studies (e.g., Fant, 1973; Dorman et al., 1997; Dinino et al., 2016).

A bias toward unvoiced stops was found, a phenomenon that only occurred for the CI group and hence probably is implant related. This may be related to two main issues: (1) implants convey the F0 in voiced sounds rather poorly due to missing temporal information in the electrical signal for most implant models and to the electrode's insertion depth possibly being too shallow to cover the whole cochlea (Hamzavi and Arnoldner, 2006; Svirsky et al., 2015; Caldwell et al., 2017) and (2) the VOT makes the unvoiced stops much easier to perceive than the voiced stops due to the aspirated pause between the stop and the following vowel in the VCV syllables.

The subgroups of voiced and unvoiced stops can be distinguished by the presence of a silent gap in the unvoiced stops (Lisker, 1981). For Norwegian unvoiced stops, as for unvoiced stops in most Germanic languages, aspiration is a salient feature: a distinct final auditory breathy pause that is created by closing the vocal cords from a maximally spread position, lasting longer than the occluded phase of the stop articulation (Kristoffersen, 2000). Stops can be difficult to identify, since they are very short and unvoiced stops have little acoustic energy. In identifying stops, CI users usually rely considerably on the spectral properties of the surrounding vowels, such as locus and length of the formant transitions, spectral height and steepness, and VOT (Välimaa et al., 2002).

Moreno-Torres and Madrid-Cánovas (2018) found a voicing bias for the stops for children with CIs, which is the opposite of the results of the present study. Their study design is, however, considerably different from the present study, as the children were Spanish-speaking and were tested with added, speech-modulated noise, which may create a perception of voicing. Also, Spanish does not have aspiration as a salient feature of unvoiced stops, as Norwegian has. Studies with English and Flemish participants have found a devoicing bias similar to our study (e.g., van Wieringen and Wouters, 1999; Munson et al., 2003).

The least correctly repeated consonant was the lateral [l], which elicited many confusions in the unclassified category of the CMs and had the largest difference in correct scores between the participants with CIs and the NH 6-year-olds. Since all the NH participants were recruited from the same dialect area, Standard East Norwegian, many of them confused [l] with [ɭ], which is also part of their speech sound inventory. Remarkably, [l] was almost never confused with the nasals for any of the participant groups.

The nasals, [m, n, ŋ], were often confused with one another by the participants with CIs, and this – together with the [l]-confusions – comprise most of the difference between the NSRS-C_{voi} and NSRS-C_{unvoi}. It seems that nasality adds a new obstacle to consonant recognition. This may be due to the prominence of low frequencies around 250 Hz in the nasals' spectrum; the nasal murmur, also called the nasal formant (F1). The CIs render low frequencies rather poorly compared to high frequencies (Caldwell et al., 2017; D'Alessandro et al., 2018). Perceptual experiments with NH listeners have shown that nasal murmur and the formant transitions are both important for providing information on place of articulation (e.g., Kurowski and Blumstein, 1984). The transitions of F2 are particularly important; [m] is preceded or succeeded by an F2 transition toward a lower frequency, [n] provides little transition change, and [ŋ] is preceded or succeeded by an F2 transition toward a higher frequency.

Although the NH 6- and 13-year-olds perceived almost all consonants and vowels correctly, they confused /ŋ/ with /m/ in 19.6 and 19.4% of the cases, respectively. This confusion was almost exclusively found in the uCu-context. The reason for this tendency might be twofold. First, the tongue body is very retracted for the Norwegian [uː], with a narrow opening of the mouth and in a position close to the tongue position of [ŋ], making the formant transition audibly indistinct. Second, the listeners might primarily be focused on recognizing letters when performing this type of task. There is no unique letter in Norwegian rendering the speech sound [ŋ], and participants may not on the spur of the moment consider this speech sound an alternative, and instead decide on the one that they find

---

[2] VOT is the time between air release and vocal-cord vibration.

acoustically more similar to the other nasals, [m] and [n], which both correspond to single letters of the alphabet.

The most prevalent vowel confusion for the participants with CIs was [y:] perceived as [i:]. The main reason for this confusion is probably that the F1s of these vowels are low (∼250 Hz) and almost coinciding, and the F2 of [i:] is only slightly higher than of [y:]. These vowels are thus closely located in the vowel chart in **Figure 1**. However, [i:] was never perceived as [y:], probably because [i:] in Norwegian is about 10 times more prevalent than [y:] (Øygarden, 2009, p. 108), and when in doubt, the participants would be likely to choose the most common of the two speech sounds.

Vowels are known to be more easily perceived than consonants, due to their combination of high intensity and long duration. Norwegian vowels are distinguishable by F1 and F2 alone, as opposed to vowels in other languages, which may also be distinguished by higher formants. Vowels are never distinguished by F0.

Comparison Between the Pre- and Postlingually Deaf Participants

Between the pre- and postlingually deaf participants, we found no significant differences for the consonant and vowel scores, and no significant differences for the speech feature contrasts voicing versus non-voicing, nasality versus non-nasality, and stopping versus frication. All but three participants were provided with CIs in their optimal ($N = 28$) or late ($N = 5$) sensitive period. Four of the prelingually deaf participants who received CIs in their late sensitive period had used bilateral HAs and developed language in the period between onset of deafness and implantation, and their auditory pathways had presumably been effectively stimulated in this period.

For postlingually deaf CI users, the vowel formants conveyed by the implant tend to be misplaced in the cochlea compared to its natural tonotopy. This may be a reason why acoustically similar vowels are more easily confused for the CI users than for the NH listeners.

The mechanisms of brain plasticity and the consequences of age at onset of deafness (pre-, peri-, and postlingual) are important factors for both auditory and linguistic development. Buckley and Tobey (2011) found that the influence of cross-modal plasticity on speech perception ability is greatly influenced by age at acquisition of severe to profound (pre- or postlingual) deafness rather than by the duration of auditory deprivation before cochlear implantation. In our study, brain plasticity at implantation may be a more relevant prognostic factor for the development of speech perception skills than age at onset of deafness, because of the large individual variations in age at implantation and HA use before implantation.

The Impact of Vowel and Consonant Context on Recognition

The results of the perception of consonants in different vowel contexts indicated that formant transitions played a larger role for the participants with CIs than for the NH participants, since the influence of vowel context on the consonant score was statistically significant for the CI group but not for the NH groups. This is in accordance with Donaldson and Kreft (2006), who found that the average consonant recognition scores of adult CI users were slightly but significantly higher (6.5%) for consonants presented in an aCa or uCu context than for consonants presented in an iCi context. The vocal tract is more open for [ɑ:] than for [i:] and [u:], making the formant transition more pronounced and the consonants therefore more easily perceptible. The Norwegian [u:] is much more retracted than the English [u:], and thus closer to the velar speech sounds, making their formant transitions more challenging to perceive.

The nine long vowels were presented in only one consonant context, with /b/, as vowel perception is based on steady-state formants rather than on formant transitions.

Inclusion Criteria and Test Design

By only including well-performing participants with CIs (score above 50% on the HIST monosyllable test and 100% correct spontaneous pronunciation score of all the Norwegian speech sounds), we were able to reveal systematic details in speech sound confusions. If poorer-performing participants with CIs had been included, a great deal of noise would have been added to the CMs, as the unclassified category would have become much larger.

In the present study, other higher language skills are of minor importance, as the NSRT is limited to speech sounds and syllables. We therefore had no inclusion criterion regarding language skills. Since the participants with CIs and the NH 6-year-olds had a similar mean hearing age, some perception problems may be related to their developmental stage in speech perception ability, in addition to being implant related.

As our study required that the participants respond verbally, a closed-set test was not a practical option. Moreover, we consider an open-set test design to be more ecologically valid than a closed-set test design, as repetition of unknown syllables is a common activity for children and one with which they are familiar when acquiring new vocabulary in their everyday life.

Limitations and Strengths

As expected, we obtained ceiling effects on both the vowel and consonant scores for the NH reference groups. For the participants with CIs, there were ceiling effects only on the vowel scores. This explains lack of statistical significance in many of the comparisons, and is in line with previous studies. For instance, Rødvik et al. (2018) have shown that NSRTs rarely result in ceiling effects when measuring consonant perception for CI users but may do so for vowel perception. It is well known that vowels are easier to perceive than consonants, due to longer duration and higher intensity. All nine Norwegian vowels exist in a long and a short version, and in the NSRT, only long vowels were included, making them audibly very distinct.

An important reason for the ceiling effect on the vowel and HIST scores for the participants with CIs is probably our criterion of only including well-performing CI users who had scores above 50% on HIST. The ceiling effect on the HIST score has probably also weakened the correlations with consonant and vowels scores in the CI users.

Since the test lists of the NSRT counted as many as 90 CVC and VCV words, fatigue and lack of concentration may have influenced the results, especially for the younger children. We randomized the word order to prevent the same words from always appearing at the end of the test list and thus avoiding systematic errors.

This study used a convenience sample due to a limited time window for recruiting participants, who were assessed in conjunction with their regular CI checkup. This design has limitations as far as internal matching regarding, for instance age, gender, age at onset of deafness, duration of implant use, age at implantation, or implant model is concerned. Using a convenience sample may, however, also be considered a strength, as the participants represent a completely random sample of Norwegian-speaking children with CIs, since all implanted children in Norway have received their CI at the same clinic, Oslo University Hospital.

The two groups of pre- and postlingually deaf participants are very different in size, and the participants are very different with regard to level of hearing loss after onset of deafness, HA use before implantation, and age at implantation. Ideally, these factors should have been controlled for, so the evidence present to compare these groups may therefore have been weak.

Recommendations for Future Research and Clinical Use

This study provides information regarding typical misperceptions of speech sounds in participants with CIs, which may be useful as a basis for further research, focusing on its consequences for CI programming. The information will also be very useful when planning listening and speech therapy for the implantees.

The study might also be used as a basis for the development, validation, and norming of a simplified version of the NSRT to be included in the standard test battery in audiology clinics. Children with CIs tested regularly with the NSRT would be provided with individual feedback on what needs to be targeted in the programming of their CIs and in their listening therapy sessions. Pre- and post-testing with the NSRT can be used as a quality control tool of the programming. A clinical NSRT would also meet the increasing challenge of assessing speech perception in patients with different language backgrounds, as it can be adjusted for different languages by modifying it to only include speech sounds existing in a particular language.

A close examination of the CMs of each individual CI user may possibly be employed when deciding whether to reprogram the CIs or simply adjust the approach in listening therapy, since speech sounds within the same manner-group in the CMs are in general more acoustically similar than speech sounds in different manner groups. Hence, a rule-of-thumb may be that in case of confusions within the same manner-group, start with listening therapy, and in case of confusions between two manner-groups, reprogramming of the implant may be useful.

CONCLUSION

For the participants with CIs, consonants were mostly confused with consonants with the same voicing and manner. In general, voiced consonants were more difficult to perceive than unvoiced consonants, and there was a devoicing bias for the stops. The vowel repetition score was higher than the consonant repetition score. Additionally there was a [iː]-[yː] confusion bias, as [yː] was perceived as [iː] twice as often as [yː] was repeated correctly.

The subgroup analyses showed no statistically significant differences between consonant repetition scores for the pre- and postlingually deaf participants.

Although the children with CIs obtained scores close to 100% on vowels and real-word monosyllables, none of them obtained scores for voiced consonants above 78%. This is likely to be related to limitations in CI technology for the transmission of low-frequency sounds, such as insertion depth of the electrode and ability to convey temporal information.

AUTHOR'S NOTE

Preliminary results from this study were presented at the CI conference: CI 2017 Pediatric 15th Symposium on Cochlear Implants in Children, in San Francisco, United States, July 26–29, 2017.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The study was approved by the Regional Ethical Committees for Medical and Health Research Ethics – REC South East, Oslo, Norway. This study was carried out in accordance with the recommendations of "helseforskningsloven" (Health Research Law) §9, §10, §11, and §33, and cf. "forskningsetikkloven" (Research Ethics Law) §4, approved by the REC South East, with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the REC South East. Additional considerations regarding vulnerable populations such as minors: the speech perception testing of the children included in the project implied no risk for them, and no additional measures were necessary.

AUTHOR CONTRIBUTIONS

AR designed the study, analyzed the data, and wrote the manuscript. OT was responsible for the analyses and for technical matters regarding the CI, JT was responsible for methodological, structural, and linguistic matters, OW was responsible for audiological and educational matters, IS was responsible for phonetic and speech therapeutic matters, and JS was responsible for study design and medical matters. All authors discussed the results and suggested revisions of the manuscript at all stages.

ACKNOWLEDGMENTS

The authors gratefully acknowledge all the children and their parents participating in this project and statistician Marissa LeBlanc at Oslo University Hospital.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01813/full#supplementary-material

REFERENCES

Arisi, E., Forti, S., Pagani, D., Todini, L., Torretta, S., Ambrosetti, U., et al. (2010). Cochlear implantation in adolescents with prelinguistic deafness. *Otolaryngol. Head Neck Surg.* 142, 804–808. doi: 10.1016/j.otohns.2010.02.016

Boersma, P., and Weenink, D. (2018). *Praat: Doing phonetics by computer [Computer program]. Version 6.0.37*. Available at: http://www.praat.org (accessed March 14, 2018).

Bruijnzeel, H., Ziylan, F., Stegeman, I., Topsakal, V., and Grolman, W. (2016). A systematic review to define the speech and language benefit of early (<12 months) pediatric cochlear implantation. *Audiol. Neurotol.* 21, 113–126. doi: 10.1159/000443363

Buckley, K. A., and Tobey, E. A. (2011). Cross-modal plasticity and speech perception in pre- and postlingually deaf cochlear implant users. *Ear Hear.* 32, 2–15. doi: 10.1097/AUD.0b013e3181e8534c

Caldwell, M. T., Jiam, N. T., and Limb, C. J. (2017). Assessment and improvement of sound quality in cochlear implant users. *Laryngoscope Investig. Otolaryngol.* 2, 119–124. doi: 10.1002/lio2.71

Ching, T. Y., Zhang, V. W., Flynn, C., Burns, L., Button, L., Hou, S., et al. (2018). Factors influencing speech perception in noise for 5-year-old children using hearing aids or cochlear implants. *Int. J. Audiol.* 57, 70–80.

Clopper, C. G., Pisoni, D. B., and Tierney, A. T. (2006). Effects of open-set and closed-set task demands on spoken word recognition. *J. Am. Acad. Audiol.* 17, 331–349. doi: 10.3766/jaaa.17.5.4

Coady, J. A., and Aslin, R. N. (2004). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *J. Exp. Child Psychol.* 89, 183–213. doi: 10.1016/j.jecp.2004.07.004

D'Alessandro, H. D., Ballantyne, D., Boyle, P. J., De Seta, E., DeVincentiis, M., and Mancini, P. (2018). Temporal fine structure processing, pitch, and speech perception in adult cochlear implant recipients. *Ear Hear.* 39, 679–686. doi: 10.1097/AUD.0000000000000525

Darley, F. L., and Winitz, H. (1961). Age of first word: review of research. *J. Speech Hear. Disord.* 26, 272–290. doi: 10.1044/jshd.2603.272

Dettman, S. J., Dowell, R. C., Choo, D., Arnott, W., Abrahams, Y., Davis, A., et al. (2016). Long-term communication outcomes for children receiving cochlear implants younger than 12 months: a multicenter study. *Otol. Neurotol.* 37, e82–e95. doi: 10.1097/MAO.0000000000000915

Dinino, M., Wright, R. A., Winn, M. B., and Bierer, J. A. (2016). Vowel and consonant confusions from spectrally manipulated stimuli designed to simulate poor cochlear implant electrode-neuron interfaces. *J. Acoust. Soc. Am.* 140, 4404–4418. doi: 10.1121/1.4971420

Donaldson, G. S., and Kreft, H. A. (2006). Effects of vowel context on the recognition of initial and medial consonants by cochlear implant users. *Ear Hear.* 27, 658–677. doi: 10.1097/01.aud.0000240543.31567.54

Dorman, M. F., Loizou, P. C., and Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *J. Acoust. Soc. Am.* 102, 2403–2411. doi: 10.1121/1.419603

Doyle, K. J., Mills, D., Larky, J., Kessler, D., Luxford, W. M., and Schindler, R. A. (1995). Consonant perception by users of nucleus and clarion multichannel cochlear implants. *Am. J. Otol.* 16, 676–681.

Drullman, R. (2005). *D-7-1: Speech Recognition Tests for Different Languages Report FP6–004171 HEARCOM Hearing in the Communication Society*. Sankt Augustin: Fraunhofer Institute for Applied Information Technology.

Elley, W. B. (1992). *How in the World Do Students Read? IAE Study of Reading Literacy*. The Hague: International Association for the Evaluation of Educational Achievement.

Engdahl, B., Tambs, K., Borchgrevink, H. M., and Hoffman, H. J. (2005). Screened and unscreened hearing threshold levels for the adult population: results from the Nord-Trøndelag hearing loss study. *Int. J. Audiol.* 44, 213–230. doi: 10.1080/14992020500057731

Fant, G. (1973). *Speech Sounds and Features*. Cambridge, MA: MIT Press.

Feng, G., Ingvalson, E. M., Grieco-Calub, T. M., Roberts, M. Y., Ryan, M. E., Birmingham, P., et al. (2018). Neural preservation underlies speech improvement from auditory deprivation in young cochlear implant recipients. *Proc. Natl. Acad. Sci. U.S.A.* 115, E1022–E1031. doi: 10.1073/pnas.1717603115

Findlen, U. M., and Roup, C. M. (2011). Dichotic speech recognition using CVC word and nonsense CVC syllable stimuli. *J. Am. Acad. Audiol.* 22, 13–22. doi: 10.3766/jaaa.22.1.3

Fischer, B., and Hartnegg, K. (2004). On the development of low-level auditory discrimination and deficits in dyslexia. *Dyslexia* 10, 105–118. doi: 10.1002/dys.268

Freed, D. (2001). *Speech Utility [Computer Program]*. San Francisco, CA: House Ear institute.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., and Emslie, H. (1994). The children's test of nonword repetition: a test of phonological working memory. *Memory* 2, 103–127. doi: 10.1080/09658219408258940

Geers, A. E., Brenner, C., and Davidson, L. (2003). Factors associated with development of speech perception skills in children implanted by age five. *Ear Hear.* 24, 24S–35S. doi: 10.1097/01.aud.0000051687.99218.0f

Halvorsen, B. (1998). *Timing Relations in Norwegian Stops*. Ph. D dissertation. University of Bergen: Bergen.

Hamzavi, J., and Arnoldner, C. (2006). Effect of deep insertion of the cochlear implant electrode array on pitch estimation and speech perception. *Acta Oto Laryngol.* 126, 1182–1187. doi: 10.1080/00016480600672683

Harrison, R. V., Gordon, K. A., Mount, R. J., and Maurer, D. (2005). Is there a critical period for cochlear implantation in congenitally deaf children? Analyses of hearing and speech perception performance after implantation. *Dev. Psychobiol.* 46, 252–261. doi: 10.1002/dev.20052

Henry, K. S., and Heinz, M. G. (2012). Diminished temporal coding with sensorineural hearing loss emerges in background noise. *Nat. Neurosci.* 15, 1362–1365. doi: 10.1038/nn.3216

Kral, A. (2013). Auditory critical periods: a review from system's perspective. *Neuroscience* 247, 117–133. doi: 10.1016/j.neuroscience.2013.05.021

Kral, A., and Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends Neurosci.* 35, 111–122. doi: 10.1016/j.tins.2011.09.004

Kristoffersen, G. (2000). *The Phonology of Norwegian*. Oxford: Oxford University Press.

Kurowski, K., and Blumstein, S. E. (1984). Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants. *J. Acoust. Soc. Am.* 76, 383–390. doi: 10.1121/1.391139

Leake, P. A., and Hradek, G. T. (1988). Cochlear pathology of long term neomycin induced deafness in cats. *Hear. Res.* 33, 11–33. doi: 10.1016/0378-5955(88)90018-4

Leigh, J., Dowell, R., Briggs, R., and Dettman, S. (2013). Communication development in children who receive a cochlear implant by 12 months of age. *Otol. Neurotol.* 34, 443–450. doi: 10.1097/MAO.0b013e3182814d2c

Lisker, L. (1981). On generalizing the rapid-rabid distinction based on silent gap duration. *Haskins Lab. Status Rep. Speech Res.* 65, 251–259.

Litovsky, R. Y. (2015). "The human auditory system: fundamental organization and clinical disorders," in *Handbook of Clinical Neurology*, ed. G. G. Celesia (Amsterdam: Elsevier Science).

Liu, H., Liu, S., Kirk, K. I., Zhang, J., Ge, W., Zheng, J., et al. (2015). Longitudinal performance of spoken word perception in Mandarin pediatric cochlear implant users. *Int. J. Pediatr. Otorhinolaryngol.* 79, 1677–1682. doi: 10.1016/j.ijporl.2015.07.023

Locke, J. L. (1983). *Phonological Acquisition and Change*. London: Academic Press.

Maxon, B. A., and Hochberg, B. I. (1982). Development of psychoacoustic behavior: sensitivity and discrimination. *Ear Hear.* 3, 301–308. doi: 10.1097/00003446-198211000-00003

Miller, G. A., and Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *J. Acoust. Soc. Am.* 27, 338–352. doi: 10.1121/1.1907526

Moreno-Torres, I., and Madrid-Cánovas, S. (2018). Recognition of Spanish consonants in 8-talker babble by children with cochlear implants, and by children and adults with normal hearing. *J. Acoust. Soc. Am.* 144:69. doi: 10.1121/1.5044416

Munson, B., Donaldson, G. S., Allen, S. L., Collison, E. A., and Nelson, D. A. (2003). Patterns of phoneme perception errors by listeners with cochlear implants as a function of overall speech perception ability. *J. Acoust. Soc. Am.* 113, 925–935. doi: 10.1121/1.1536630

Myhrum, M., Strøm-Roum, H., Heldahl, M. G., Rødvik, A. K., Eksveen, B., Landsvik, B., et al. (2017). Sequential bilateral cochlear implantation in children: outcome of the second implant and long-term use. *Ear Hear.* 38, 301–313. doi: 10.1097/AUD.0000000000000383

National Institute for Occupational Safety, and Health [NIOSH] (1996). *DRAFT: Criteria for a Recommended Standard: Occupational Noise Exposure Revised Criteria*. Cincinnati, OH: National Institute for Occupational Safety, and Health.

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., et al. (2010). Spoken language development in children following cochlear implantation. *J. Am. Med. Assoc.* 303, 1498–1506.

Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., and Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *J. Acoust. Soc. Am.* 84, 1316–1324. doi: 10.1121/1.396630

Øygarden, J. (2009). *Norwegian Speech Audiometry*. Ph. D Dissertation. Norwegian University of Science and Technology: Trondheim.

Rødvik, A. K., Torkildsen, J. V. K., Wie, O. B., Storaker, M. A., and Silvola, J. T. (2018). Consonant and vowel identification in cochlear implant users measured by nonsense words: a systematic review and meta-analysis. *J. Speech Lang. Hear. Res.* 61, 1023–1050. doi: 10.1044/2018_JSLHR-H-16-0463

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Techn. J.* 27, 379–423.

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002a). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear.* 23, 532–539. doi: 10.1097/00003446-200212000-00004

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002b). Rapid development of cortical auditory evoked potentials after early cochlear implantation. *Neuroreport* 13, 1365–1368. doi: 10.1097/00001756-200207190-00030

Sharma, A., Spahr, A. J., Dorman, M. F., and Todd, N. W. (2002c). Early cochlear implantation in children allows normal development of central auditory pathways. *Ann. Otol. Rhinol. Laryngol.* 111, 38–41. doi: 10.1177/00034894021110s508

Sheffield, B. M., and Zeng, F. G. (2012). The relative phonetic contributions of a cochlear implant and residual acoustic hearing to bimodal speech perception. *J. Acoust. Soc. Am.* 131, 518–530. doi: 10.1121/1.3662074

Sommers, S. M., Kirk, I. K., and Pisoni, B. D. (1997). Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners. I: the effects of response format. *Ear Hear.* 18, 89–99. doi: 10.1097/00003446-199704000-00001

Svirsky, M. A., Fitzgerald, M. B., Sagi, E., and Glassman, E. K. (2015). Bilateral cochlear implants with large asymmetries in electrode insertion depth: implications for the study of auditory plasticity. *Acta Oto Laryngol.* 135, 354–363. doi: 10.3109/00016489.2014.1002052

Tingleff, H. (2002). *Norsk Fonemtest [Norwegian Phoneme Test]*. Oslo: Damm.

Tobey, E. A., Thal, D., Niparko, J. K., Eisenberg, L. S., Quittner, A. L., and Wang, N.-Y. (2013). Influence of implantation age on school-age language performance in pediatric cochlear implant users. *Int. J. Audiol.* 52, 219–229. doi: 10.3109/14992027.2012.759666

Trehub, S. E., Schneider, B. A., Morrongiello, B. A., and Thorpe, L. A. (1988). Auditory sensitivity in school-age children. *J. Exp. Child Psychol.* 46, 273–285. doi: 10.1016/0022-0965(88)90060-4

Tye-Murray, N., Lowder, M., and Tyler, R. S. (1990). Comparison of the F0F2 and F0F1F2 processing strategies for the Cochlear Corporation cochlear implant. *Ear Hear.* 11, 195–200. doi: 10.1097/00003446-199006000-00005

Tye-Murray, N., Spencer, L., and Gilbert-Bedia, E. (1995). Relationships between speech production and speech perception skills in young cochlear-implant users. *J. Acoust. Soc. Am.* 98, 2454–2460. doi: 10.1121/1.413278

Tyler, R. S. (1990). Speech perception with the nucleus cochlear implant in children trained with the auditory/verbal approach. *Am. J. Otol.* 11, 99–107.

Tyler, R. S., and Moore, B. C. J. (1992). Consonant recognition by some of the better cochlear-implant patients. *J. Acoust. Soc. Am.* 92, 3068–3077. doi: 10.1121/1.404203

Välimaa, T. T., Määttä, T. K., Löppönen, H. J., and Sorri, M. J. (2002). Phoneme recognition and confusions with multichannel cochlear implants: consonants. *J. Speech Lang. Hear. Res.* 45, 1055–1069. doi: 10.1044/1092-4388(2002/085)

van Wieringen, A., and Wouters, J. (1999). Natural vowel and consonant recognition by Laura cochlear implantees. *Ear Hear.* 20, 89–103. doi: 10.1097/00003446-199904000-00001

Wie, O. B. (2010). Language development in children after receiving bilateral cochlear implants between 5 and 18 months. *Int. J. Pediatr. Otorhinolaryngol.* 74, 1258–1266. doi: 10.1016/j.ijporl.2010.07.026

Wilson, B. S., and Dorman, M. F. (2008). Cochlear implants: current designs and future possibilities. *J. Rehabilitat. Res. Dev.* 45, 695–730. doi: 10.1682/jrrd.2007.10.0173

Wolfe, C. J., Schafer, C. E., John, C. A., and Hudson, C. M. (2011). The effect of front-end processing on cochlear implant performance of children. *Otol. Neurotol.* 32, 533–538. doi: 10.1097/MAO.0b013e318210b6ec

World Health Organization [WHO] (2019). *Grades of Hearing Impairment*. Geneva: World Health Organization.

World Medical Association [WMA] (2017). *WMA Declaration of Helsinki-Ethical Principles for Medical Research Involving Human Subjects*. Ferney-Voltaire: World Medical Association.

Wouters, J., Francart, T., and Mcdermott, H. J. (2015). Sound coding in cochlear implants: from electric pulses to hearing. *IEEE Signal Process. Magaz.* 32, 67–80. doi: 10.1109/msp.2014.2371671

Yoon, Y. S., Li, Y., and Fu, Q. J. (2012). Speech recognition and acoustic features in combined electric and acoustic stimulation. *J. Speech Lang. Hear. Res.* 55, 105–124. doi: 10.1044/1092-4388(2011/10-0325)

Zeitler, D. M., Anwar, A., Green, J. E., Babb, J. S., Friedmann, D. R., Roland, T. Jr., et al. (2012). Cochlear implantation in prelingually deafened adolescents. *Arch. Pediatr. Adolesc. Med.* 166, 35–41.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary text 1. Recording and preparation of the nonsense syllable repetition test.

1. Sound editing

The test words were spoken by a Norwegian female speech therapist and recorded with a Zoom H4n hard disk recorder in an anechoic chamber. The recorded wav files were imported into the Praat computer program (Boersma and Weenink, 2018). The files were cut and split so they all were of equal length, ~1.0 s, and then imported into Adobe Audition CS6, Ver. 5.0, Build 708 (San José, CA, USA, 2012), for noise reduction and editing.

We decided not to normalize the sound files, as the energy level varies naturally between different speech sounds; for instance, [uː] contains less energy than [ɑː]. A normalizing procedure would have made some speech sounds unnaturally loud and some unnaturally soft. Instead, we used the “Match volume” command in Adobe Audition and adjusted the blocks of data of the recorded aCa, iCi, and uCu words to keep the volume at an even level between the blocks of speech sounds in the same vowel context.

2. Equipment and test setup

We used the Matlab computer program (Mathworks, Nantick MA, 2013) to randomize the nonsense syllables to be played in permuted sequential blocks consisting of 9 aCa syllables, 9 iCi syllables, 9 uCu syllables, and 9 bVb syllables. This randomization was to prevent the transcribers from learning the order of the nonsense syllables and thus being biased in their transcription. By dividing the nonsense syllables into blocks of 9 similar units, we aimed to make the test less repetitive and easier for the participants than a complete randomization would be.

We presented the nonsense syllables with the SpchUtil, ver. 5 computer program (Freed, 2004). This program enables each sound file to be called upon, and the sound level for each file can be individually adjusted. Because the sound level varies across each word, we measured the maximum value of the instantaneous sound pressure level, F_{max}, of each word in the listening position. F_{max} was registered manually on a Norsonic 110 sound level meter (Tranby, Norway). We decided to keep 65 dB(A) in the listening position as a desired average sound pressure level, and we measured the F_{max} of the nonsense syllable sound files in 2 series. We found that the average F_{max} of all the sound files in both series was 60 dB(A). We then added 5 dB to all the nonsense syllables in SpchUtil, as we maintained the assumption that the mutual loudness differences between the recorded nonsense syllables was equal to their natural loudness differences.

Table S1 | Demographics of participants with CIs

Participant no.	Age (Years)	Gender (Male/ Female)	Age at implantation (Years)	Duration of implant use (Years)	°Implant model	Stimulation strategy	Modality (1CI—1/ 2CI—2)	Pre- or postlingually deaf at implantation
1	7.5	Female	4.0	3.4	Nucleus CI512	ACE	2	Post-
2	11.5	Female	3.2	8.2	Nucleus CI24RE(CA)	ACE	2	Pre-
3	11.6	Male	1.0	10.5	Nucleus CI24RE(CS)	ACE	2	Pre-
4	11.1	Male	1.6	9.4	Pulsar ci100	FSP	2	Pre-
5	16.0	Female	2.6	13.2	Nucleus CI24R(ST)	ACE	2	Pre-
6	13.0	Male	1.6	11.2	Nucleus CI24R(CS)	ACE	2	Pre-
7	9.1	Male	0.8	8.2	Nucleus CI24RE(CA)	ACE	2	Pre-
8	10.3	Male	5.1	5.1	Nucleus CI24RE(CA)	ACE	2	Pre-
9	10.2	Female	4.2	5.1	Sonata i100	FSP	2	Pre-
10	10.6	Female	1.4	9.1	Pulsar ci100	FSP	2	Pre-
11	11.7	Female	1.0	10.6	Combi 40+	FSP	2	Pre-
12	15.8	Male	2.4	13.3	Nucleus CI24M	ACE	2	Pre-
13	14.0	Female	5.1	8.8	Nucleus CI24RE(CA)	ACE	2	Pre-
14	9.4	Female	0.5	8.8	Nucleus CI24RE(CA)	ACE	2	Pre-
15	15.3	Female	3.7	11.5	Combi 40+	CIS+	2	Post-
16	8.2	Female	2.6	5.5	Nucleus CI24RE(CA)	ACE	2	Post-
17	5.9	Female	1.7	4.1	Nucleus CI512	ACE	2	Pre-
18	15.0	Male	7.8	7.1	Nucleus CI24RE(CA)	ACE	2	Post-
19	12.4	Female	11.7	0.6	Nucleus CI24RE(CA)	ACE	1	Post-

° Combi 40+, Pulsar ci100, Sonata i100, and Concerto Mi1000 are manufactured by Med-El. Nucleus CIxx is manufactured by Cochlear. The participants are ordered by their test date.

(Continued on the following page)

Table S1 | Demographics of participants with CIs

Participant no.	Age (Years)	Gender (Male/ Female)	Age at implantation (Years)	Duration of implant use (Years)	°Implant model	Stimulation strategy	Modality (1CI—1/ 2CI—2)	Pre- or postlingually deaf at implantation
20	8.8	Male	0.6	8.1	Pulsar ci100	FS4	2	Pre-
21	13.1	Male	1.3	11.7	Combi 40+	CIS+	2	Pre-
22	6.4	Male	1.2	5.0	Nucleus CI512	ACE	2	Pre-
23	14.9	Male	6.7	8.1	Pulsar ci100	FSP	2	Pre-
24	8.1	Female	1.0	7.0	Sonata ti100	FS4	2	Pre-
25	13.0	Female	3.6	9.3	Pulsar ci100	FSP	2	Pre-
26	10.8	Male	1.7	9.0	Nucleus CI24RE(CA)	ACE	2	Pre-
27	14.7	Female	2.4	12.2	Nucleus CI24R(CS)	ACE	2	Pre-
28	9.5	Female	3.0	6.4	Nucleus CI24RE(CA)	ACE	2	Pre-
29	10.1	Male	2.6	7.4	Nucleus CI24RE(CA)	ACE	2	Pre-
30	15.7	Female	11.7	3.9	Nucleus CI512	ACE	2	Post-
31	6.1	Male	0.9	5.1	Nucleus CI512	ACE	2	Pre-
32	8.7	Female	0.7	7.9	Nucleus CI24RE(CA)	ACE	2	Pre-
33	14.6	Male	4.8	9.8	Pulsar ci100	FS4	2	Post-
34	15.2	Male	2.1	13.0	Combi 40+	FSP	2	Pre-
35	13.9	Male	3.2	10.5	Pulsar ci100	FS4	2	Pre-
36	14.8	Male	2.1	12.5	Nucleus CI24RE(Ca)	ACE	2	Pre-

Table S2 | Test results of participants with CIs

Patient no.	HIST Monosyllable score (%)	NSRS (%)	NSRS-C (%)	NSRS-V (%)
1	88	69.3	70.8	66.7
2	86	72.0	68.8	77.8
3	86	82.7	79.2	88.9
4	94	74.7	72.9	77.8
5	90	77.3	70.8	88.9
6	90	57.3	52.1	66.7
7	90	72.0	68.8	77.8
8	72	56.0	50.0	66.7
9	86	65.3	58.3	77.8
10	76	78.7	72.9	88.9
11	90	78.7	72.9	88.9
12	92	85.3	77.1	100.0
13	98	81.3	77.1	88.9
14	84	84.0	81.3	88.9
15	96	80.0	75.0	88.9
16	82	60.0	56.3	66.7
17	90	78.7	72.9	88.9
18	94	77.3	70.8	88.9
19	86	65.3	64.6	66.7
20	94	81.3	70.8	100.0
21	80	65.3	58.3	77.8
22	80	69.3	64.6	77.8
23	92	76.0	68.8	88.9
24	80	80.0	75.0	88.9
25	94	86.7	79.2	100.0
26	78	70.7	60.4	88.9
27	82	70.7	60.4	88.9
28	86	70.7	72.9	66.7
29	84	80.0	68.8	100.0
30	100	82.7	79.2	88.9
31	96	76.0	68.8	88.9
32	82	80.0	68.8	100.0
33	88	89.3	83.3	100.0
34	76	80.0	75.0	88.9
35	80	77.3	64.6	100.0
36	88	74.7	72.9	77.8

CI, Cochlear Implant; HIST, Høgskulen i Sør-Trøndelag (Sør-Trøndelag University College); NSRS, Nonsense Syllable Repetition Score; NSRS-C, Nonsense Syllable Repetition Score—Consonants; NSRS-V, Nonsense Syllable Repetition Score—Vowels.

Table S3 | The VCV and CVC nonsense syllables included in the test

Number°	aCa-syllables	iCi-syllables	uCu-syllables	bVb-syllables
1	['ɑːbɑ]	['iːbi]	['uːbu]	[bɑːb]
2	['ɑːdɑ]	['iːdi]	['uːdu]	[beːb]
3	['ɑːfɑ]	['iːfi]	['uːfu]	[biːb]
4	['ɑːgɑ]	['iːgi]	['uːgu]	[buːb]
5	['ɑːhɑ]	['iːhi]	['uːhu]	[bʉːb]
6	['ɑːʝɑ]	['iːʝi]	['uːʝu]	[byːb]
7	['ɑːkɑ]	['iːki]	['uːku]	[bæːb]
8	['ɑːlɑ]	['iːli]	['uːlu]	[bøːb]
9	['ɑːmɑ]	['iːmi]	['uːmu]	[bɔːb]
10	['ɑːnɑ]	['iːni]	['uːnu]	
11	['ɑːpɑ]	['iːpi]	['uːpu]	
12	['ɑːsɑ]	['iːsi]	['uːsu]	
13	['ɑːtɑ]	['iːti]	['uːtu]	
14	['ɑːvɑ]	['iːvi]	['uːvu]	
15	['ɑːʃɑ]	['iːʃi]	['uːʃu]	
16	['ɑːŋɑ]	['iːŋi]	['uːŋu]	
17	['ɑːɾɑ]	['iːɾi]	['uːɾu]	
18	['ɑːçɑ]	['iːçi]	['uːçu]	
19	['ɑːɖɑ]	['iːɖi]	['uːɖu]	
20	['ɑːɽɑ]	['iːɽi]	['uːɽu]	
21	['ɑːɭɑ]	['iːɭi]	['uːɭu]	
22	['ɑːʈɑ]	['iːʈi]	['uːʈu]	
23	['ɑːɳɑ]	['iːɳi]	['uːɳu]	
24	['ɑːɲɑ]	['iːɲi]	['uːɲu]	
25	['ɑːʎɑ]	['iːʎi]	['uːʎu]	
26	['ɑːʁɑ]	['iːʁi]	['uːʁu]	
27	['ɑːrɑ]	['iːri]	['uːru]	

° Numbers 1–16 contain speech sounds common in all Norwegian dialects, and were included in the analyses.

Table S4 | Confusion matrix for NH 6-year-olds (N = 17); consonant repetitions in the aCa, iCi, and uCu contexts added together.

			Response																			
			Unvoiced							**Voiced**												
			S			**F**				**S**			**F**		**Na**			**L**				
Stimulus			**/p/**	**/t/**	**/k/**	**/s/**	**/ʃ/**	**/f/**	**/h/**	**/b/**	**/d/**	**/g/**	**/j/**	**/v/**	**/n/**	**/m/**	**/ŋ/**	**/l/**	**U**	**Sum**
Unvoiced	**S**	**/p/**	44					5		2										51
		/t/		49															2	51
		/k/		2	44							2							3	51
	F	**/s/**				47		1											3	51
		/ʃ/				2	47												2	51
		/f/	1					49											1	51
		/h/						1	47				1						2	51
Voiced	**S**	**/b/**	3							42	1			5						51
		/d/								2	48								1	51
		/g/			1							48							2	51
	F	**/j/**	1						2				44						4	51
		/v/				1		1		1				47			1			51
	N	**/n/**									1				41	1		1	7	51
		/m/								2				1	3	42	2		1	51
		/ŋ/												1	2	10	32		6	51
	L	**/l/**																38	13	51

S = stops; F = fricatives; Na = nasals; L = the lateral [l]; U = unclassified.

Table S5 | Confusion matrix for the NH 6-year-olds (N = 17); consonant repetitions collapsed with regard to manner and place of articulation (percentage of stimulus feature in each cell).

			Response (%)								
			Unvoiced		Voiced						
			S	F	S	F	Na	L			
Stimulus			/p/ /t/ /k/	/s/ /ʃ/ /f/ /h/	/b/ /d/ /ɡ/	/j/ /v/	/n/ /m/ /ŋ/	/l/	U	Sum (%)	N
Unvoiced	S	/p/ /t/ /k/	90.8	3.3	2.6				3.3	100	153
	F	/s/ /ʃ/ /f/ /h/	0.5	95.1		0.5			3.9	100	204
Voiced	S	/b/ /d/ /ɡ/	2.6		92.2	3.3			2.0	100	153
	F	/j/ /v/	1.0	3.9	1.0	89.2	1.0		3.9	100	102
	N	/n/ /m/ /ŋ/			2.0	1.3	86.9	0.7	9.2	100	153
	L	/l/						74.5	25.5	100	51

S = stops; F = fricatives; Na = nasals; L = the lateral [l]; U = unclassified; N = sample size.

Table S6 | Confusion matrix for NH 6-year-olds (N = 17); vowel repetitions in the bVb context

	Response										
Stimulus	**/bɑːb/**	**/beːb/**	**/biːb/**	**/buːb/**	**/bʉːb/**	**/byːb/**	**/bæːb/**	**/bøːb/**	**/bɔːb/**	**U**	**Sum**
/bɑːb/	17										17
/beːb/		15	1				1				17
/biːb/			16							1	17
/buːb/				17							17
/bʉːb/					16					1	17
/byːb/			4		2	11					17
/bæːb/	2						14			1	17
/bøːb/					3			14			17
/bɔːb/									17		17

U = unclassified.

Table S7 | Confusion matrix for NH 13-year-olds (N = 12); consonant repetitions in the aCa, iCi, and uCu contexts added together

			Response																	
			Unvoiced							**Voiced**										
			S			**F**				**S**			**F**		**Na**			**L**		
Stimulus			**/p/**	**/t/**	**/k/**	**/s/**	**/ʃ/**	**/f/**	**/h/**	**/b/**	**/d/**	**/g/**	**/j/**	**/v/**	**/n/**	**/m/**	**/ŋ/**	**/l/**	**U**	**Sum**
Unvoiced	**S**	**/p/**	36																	36
		/t/		35		1														36
		/k/			36															36
	F	**/s/**				33													3	36
		/ʃ/					36													36
		/f/						36												36
		/h/							35										1	36
Voiced	**S**	**/b/**								35									1	36
		/d/									35								1	36
		/g/										36								36
	F	**/j/**											36							36
		/v/												36						36
	N	**/n/**													36					36
		/m/													2	33			1	36
		/ŋ/													1	7	25		3	36
	L	**/l/**																34	2	36

S = stops; F = fricatives; Na = nasals; L = the lateral [l]; U = unclassified.

Table S8 | Confusion matrix for NH 13-year-olds; consonant repetitions collapsed with regard to manner and place of articulation (percentage of stimulus feature in each cell)

			Response (%)																		
			Unvoiced							Voiced											
			S			F				S			F		Na			L			
Stimulus			/p/	/t/	/k/	/s/	/ʃ/	/f/	/h/	/b/	/d/	/g/	/j/	/v/	/n/	/m/	/ŋ/	/l/	U	Sum (%)	N
Unvoiced	S	/p/	99.1			0.9														100	108
		/t/																			
		/k/																			
	F	/s/				97.2													2.8	100	144
		/ʃ/																			
		/f/																			
		/h/																			
Voiced	S	/b/								98.1									1.9	100	108
		/d/																			
		/g/																			
	F	/j/											100.0							100	72
		/v/																			
	N	/n/													96.3				3.7	100	108
		/m/																			
		/ŋ/																			
	L	/l/																94.4	5.6	100	36

S = stops; F = fricatives; Na = nasals; L = the lateral [l]; U = unclassified.

Table S9 | Confusion matrix for NH 13-year-olds (N = 12); vowels in the bVb context

	Response									
Stimulus	**/bɑːb/**	**/beːb/**	**/biːb/**	**/buːb/**	**/bʉːb/**	**/byːb/**	**/bæːb/**	**/bøːb/**	**/bɔːb/**	**U**	**Sum**
/bɑːb/	12										12
/beːb/		12									12
/biːb/			11			1					12
/buːb/				12							12
/bʉːb/					12						12
/byːb/			3			9					12
/bæːb/							12				12
/bøːb/					1			11			12
/bɔːb/									12		12

U = unclassified.

TABLE S10 | Crosstab for the perception of speech features (*N* = 36).

Speech Feature Contrast	Age at onset of deafness	Correct Repetitions	Incorrect Repetitions
Voicing vs. nonvoicing	Prelingually deaf	1,171	39
	Postlingually deaf	289	6
Nasality vs. nonnasality	Prelingually deaf	1,202	8
	Postlingually deaf	292	1
Stops vs. fricatives	Prelingually deaf	921	18
	Postlingually deaf	114	4