THE SPEECH CRITICAL BAND (S-CB) IN COCHLEAR IMPLANT USERS: FREQUENCY RESOLUTION EMPLOYED DURING THE RECEPTION OF EVERYDAY SPEECH

Capstone Project

Presented in Partial Fulfillment of the Requirements for the

Doctor of Audiology

In the Graduate School of The Ohio State University

By

Jason Wigand, B.A.

*****

The Ohio State University
2013

Capstone Committee:

Eric W. Healy, Ph.D. - Advisor

Eric C. Bielefeld, Ph.D.

Christy M. Goodman, Au.D.

Approved by:

____
Advisor

Abstract

It is widely recognized that cochlear implant (CI) users have limited spectral resolution and that this represents a primary limitation. In contrast to traditional measures, Healy and Bacon [(2006) 119, J. Acoust. Soc. Am.] established a procedure for directly measuring the spectral resolution employed during processing of running speech. This Speech-Critical Band (*S*-CB) reflects the listeners' ability to extract spectral detail from an acoustic speech signal. The goal of the current study was to better determine the resolution that CI users are able to employ when processing speech. Ten CI users between the ages of 32 and 72 years using Cochlear Ltd. devices participated. The original standard recordings from the Hearing In Noise Test (HINT) were filtered to a 1.5-octave band, which was then partitioned into sub-bands. Spectral information was removed from each partition and replaced with an amplitude-modulated noise carrier band; the modulated carriers were then summed for presentation. CI subject performance increased with increasing spectral resolution (increasing number of partitions), never reaching asymptote. This result stands in stark contrast to expectation, as it indicates that increases in spectral resolution up to that of normal hearing produced increases in performance. Accordingly, it is concluded that CI users can access spectral information as high as that of normal hearing (NH) when presented with narrowband speech stimuli. These results have implications for the design of future devices that allow better representation of tonal languages, music, and speech in noise.

Acknowledgements

This research was supported by a grant from the National Institute on Deafness and Other Communication Disorders (NIDCD Grant No. DC008594).

I would like to express my sincere gratitude to my advisor, Eric Healy, Ph.D., for his inspiration and guidance. I would also like to thank my committee members, Eric C. Bielefeld, Ph.D. and Christy Goodman, Au.D. as well as Frederic Apoux, Ph.D., Sarah Yoho, Carla Youngdahl and Brittany Carter for their support, assistance and counsel. Finally, none of this could have been possible without the help of my family and the overwhelming love, encouragement and confidence provided me by my beautiful wife, Lori.

Vita

November 20, 1974…………………………………...…...……...…Born – Chicago, IL

May 1992…………………………………………………….………………Diploma
Madison Central High School, Richmond, KY

May 1996…………………………………………….……B.A. Economics, Spanish
University of Kentucky

December 2009 – September 2011…………..………………Graduate Research Assistant
Wexner Medical Center, Eye and Ear Institute
Department of Otolaryngology
The Ohio State University

September 2011 – May 2013…………….………..………………Graduate Teaching Assistant
Department of Speech and Hearing Sciences
The Ohio State University

Fields of Study

Major Field: Audiology

List of Tables

List of Figures

List of Abbreviations

CI	Cochlear implant
NH	Normal hearing
HI	Hearing impaired
cf	Center frequency
f0/F0	Fundamental frequency
CB	Critical band
S-CB	Speech critical band
Hz	Frequency
dB	Decibel
LNN	Low-noise noise
AM	Amplitude modulated
AIED	Autoimmune inner ear disorder
CIS	Continuous interleaved sampling
SPEAK	Spectral peak extraction
ACE	Advanced Combination Encoder

Table of Contents

Page

Chapter 1: Introduction

1.1. Cochlear Implants.

Cochlear implants are an increasingly common auditory intervention for individuals with significant hearing loss and for whom hearing aids are inadequate. Cochlear implant (CI) design has evolved from a single stimulating electrode to arrays that include multiple electrodes. Although CIs have provided audition to many severe and profoundly hearing impaired (HI) individuals, CI users show a great deal of variability in speech perception skills. Speech consists of several spectral components spanning a wide frequency range and containing many redundant cues for recognition. Previous research has shown that the speech signal can be degraded in various ways without causing significant reductions in recognition (Thibodeau and Van Tasell, 1987; ter Keurs et al., 1992, 1993; Shannon et al., 1995). The good speech recognition outcomes of many current CI users in quiet provide an excellent example of how resilient speech recognition is under these conditions of reduced cues.

In contrast to this often excellent speech recognition in quiet, studies in CI users show that diminished frequency resolution, combined with broad activation patterns and the limited ability to fully utilize spectral information, affects speech perception in noise (Friesen et al., 2001; Fu and Nogaki, 2005; Hong and Turner, 2006; Litvak et al., 2007; Nelson et al., 2003, 2008; Qin and Oxenham, 2005;

Stickney et al., 2004, 2007). Cochlear implant users' performance drops sharply compared to that of normal hearing (NH) listeners if there is noise or a competing background masker (Skinner et al., 1994; Battmer et al., 1997; Dorman et al., 1998a; Fetterman and Domico, 2002). Compared to NH listeners, CI users show a reduced ability to segregate the target voice from a background voice based on differences in fundamental frequencies (F0) such as voice gender difference (Stickney et al., 2004, 2007).

Further, CI users may not experience masking release, in which recognition improves when a background is modulated. In contrast, masking interference can sometimes occur, yielding poorer, rather than better, speech recognition in a modulated background (Nelson et al., 2003; Stickney et al., 2004). Fu and Galvin (2001) showed similar recognition of spectrally desynchronized sentences by CI users and NH listeners, implying that CI and NH listeners might utilize common temporal grouping mechanisms. In addition, recent studies have suggested that CI users have limited use of stream segregation to process non-speech stimuli (Chatterjee et al., 2006; Hong and Turner, 2006; Carlyon et al., 2007).

1.2. Frequency resolution and limitations in CI users.

A primary purpose of the CI is to provide spectral information electrically that has been lost due to reductions in normal cochlear processing. Therefore, the ability to discriminate between electrodes is critical for optimizing spectral resolution and allowing a CI user to take full advantage of the device. Poor frequency selectivity, frequency discrimination, and electrode discrimination contribute to poor speech perception in CI users. Furthermore, understanding CI users' susceptibility to noise remains a major challenge (Dowell et al., 1987; Hochberg et al., 1992; Kiefer et al., 1996; Müller-Deiler et al., 1995; Skinner et al., 1994).

Indeed, CI users are more susceptible to background noise than NH subjects listening to comparable CI simulations (Friesen et al., 2001), especially when the noise is dynamic such as competing speech (Fu and Nogaki, 2005; Müller-Deiler et al., 1995; Nelson et al., 2003). Cochlear implant users' increased susceptibility to noise is most likely due to the limited frequency resolution and the high degree of spectral smearing associated with channel interaction (Fu and Nogaki, 2005). Indeed, speech perception in CI recipients is strongly correlated with their ability to resolve spectral peaks (Henry and Turner, 2003; Litvak et al., 2007; Won et al., 2007). Consequently, much recent research and development has been aimed at increasing spectral resolution and reducing channel interactions.

Both reduced temporal resolution and reduced spectral, or frequency, resolution are thought to adversely affect speech perception in noise (Boothroyd et al., 1997; Dubno et al., 2003; Festen and Plomp, 1990; Glasberg and Moore, 1990; Glasberg and Moore, 1992). Baer et al. (1993) showed that poor spectral resolution is related to reduced masking release. Measurements simulating CIs also confirm the importance of spectral resolution to masking release (Nelson and Jin, 2004; Xu et al., 2005). However, ter Keurs et al. (1993) has shown that reduced frequency resolution in HI listeners is only loosely associated with speech intelligibility in noise, although a significant influence of spectral smearing on masking release was found. ter Keurs et al. concluded that even listeners with significantly broadened filters still have sufficient frequency resolution to resolve spectral cues important for speech intelligibility.

Although speech and music have somewhat different perceptual demands, previous studies have shown that the performance of speech correlates with music perception in CI users (Gfeller et al., 2002a, b). This suggests that the ability to resolve spectral changes contributes to both music and speech perception abilities.

Current CIs have a limited number of effective frequency channels, leading to poor pitch perception and speech recognition in noise (Fu et al., 1998; Qin and Oxenham, 2005; Stickney et al., 2004). Increasing the spectral resolution of the CI processor itself, by increasing the number of electrodes and decreasing the bandwidths of the analysis filters, yields some benefits, but only up to a certain point. For example, speech perception for NH listeners in noise continues to improve as the number of bands in a noise-excited envelope vocoder (see Shannon et al., 1995) increases from 4 to 20; however, speech perception for CI users tends to only improve with increasing number of electrodes up to about 8. Further, performance levels in CI users are well below those obtained by NH listeners (Friesen et al., 2001). This limitation is likely attributable to the spread of current from one electrode to the next and the resulting interactions that take place.

Many CI speech processing strategies exist, and many have a goal of increased spectral resolution. These strategies include continuous interleaved sampling (CIS, Wilson et al., 1991), and related spectral peak-picking schemes such as SPEAK (Seligman and McDermott, 1995) and ACE (Vandali et al., 2000). These widely used strategies, which present spectro-temporal information to the implanted electrode array in the form of pulse carriers modulated by amplitude envelopes extracted from a limited number of frequency bands, have achieved impressive speech intelligibility results. However, there are potentially important spectrally based aspects of speech perception that have received little attention. One such aspect is intonation, as conveyed by voice pitch variation. Although there is evidence that the availability of voice pitch information has little effect on performance in simple measures of speech perception such as vowel and consonant recognition (Faulkner et al., 2000), there is little doubt that intonation makes important contributions to running speech. In addition to being a major component of prosody, intonation is widely held to play an important role in early language development (Jusczyk, 1997). This is of particular significance given the increasing amount of implantation in very young children.

Pitch information remains extremely important in the perception of tonal languages, where it conveys semantic content. Cochlear implant processing strategies do not preserve important cues to pitch that are typically available to NH listeners. The limited spectral resolution means that the lower harmonics of speech that provide the principal pitch cues for NH are not resolved, as the coding schemes currently employed in CIs provide little or no representation of individual harmonics (Oxenham, 2008). Therefore, pitch perception in processed speech will largely depend upon deriving temporal pitch cues from modulations of the amplitude envelope at the voice F0. The extent to which this is possible depends upon several factors. Fundamental frequency must be passed by the envelope smoothing filter and the pulse rate must be high enough to represent modulations at F0. Psychophysical evidence suggest that such “rate pitch” is available up to roughly 300 Hz (see Rogers et al., 2006) and that accurate representation of the modulating envelope requires a carrier pulse rate at least 4 to 5 times the frequency of the modulation (McKay et al., 1994).

Results from studies of CI users with significant residual low frequency hearing show allowances for explicit preservation of F0, and potentially some F1 or F2 information (Kiefer et al., 1996). Both simulated and actual hybrid hearing studies (Kong et al., 2004; Turner et al., 2004) showed that the addition of low frequency acoustic information significantly improved speech performance in noise, particularly when the noise was a competing voice.

Studies involving a noise-excited envelope-vocoder technique, which simulates certain aspects of CI processing in NH listeners, have been used to investigate the link between F0 perception and the segregation of simultaneous vowels. Qin and Oxenham (2005) found that F0 discrimination was considerably poorer with harmonic complexes that were processed through an envelope vocoder. They tested this question by measuring the ability of NH listeners to identify pairs of simultaneous vowels with or without envelope-vocoder processing and found that no benefits were found for increasing the F0 difference between pairs of vowels, even with 24 channels of spectral information, far greater resolution than that found in current CIs. This suggests that even in the best situations, temporal envelope cues may not be sufficient to allow the perceptual segregation of simultaneous sounds based on F0 cues.

1.3. Channel interaction

As mentioned previously, the limit to the number of usable channels in CIs is thought to be determined by the extent to which electrodes stimulate non-overlapping populations of functional auditory neurons. This limitation to frequency resolution in CI users may be due to a number of factors, such as the spread of current produced by each electrode (channel interactions) and uneven neural survival patterns along the sensory epithelia, which are likely to vary between individual subjects (Kawano et al., 1998; Xu and Pfingst, 2003). Channel interaction can be reduced by decreasing current levels delivered to each electrode, through improved electrode positioning and design, or by using speech processing strategies that stimulate electrodes sequentially.

One major function of the processing of sound in the inner ear is spectral decomposition; in this respect, the peripheral auditory system can be modeled by a filter bank (Patterson and Nimmo-Smith, 1980). Each filter represents the response of a local population of nerve fibers (Glasberg and Moore, 1990; Moore, 2003). Sensorineural hearing loss results in a degradation of thresholds and a broadening of the auditory filters (Oxenham et al., 2004; Rose and Moore, 1997). Cochlear implant systems also perform a spectral decomposition. The sound processing algorithm extracts the waveform envelope in up to twenty-two frequency bands and maps the output tonotopically onto the electrode contacts. However, the electrode–neural interface imposes two major limitations on the number of analysis channels used to encode the frequency spectrum. First, the selectivity with which the individual nerve fibers can be targeted is limited by the large width of the electrical fields. The wider the electrical fields generated by individual contacts, the more diffuse the place pitch percept can be expected to be and the higher the probability that different electrode contacts will stimulate the same neural population, an effect known as nonsimultaneous channel interaction. Channel interactions have been shown in objective measures of the neural response, psychoacoustic performance, and speech perception (Townsend et al., 1987). It has been estimated that, although current electrode arrays have 12 to 22 contacts, the number of perceptually independent channels is limited to 10 (Fu et al., 1998). Secondly, even if each contact would stimulate a distinct population of fibers, twenty-two channels falls short of the number of auditory filters in normal hearing. For the adequate processing of speech in noise and of music, NH listeners make use of close to 40 (or approximately 30 across the speech-frequency range) auditory filters (Friesen, et al., 2001; Moore, 2003).

Current steering, or the creation of virtual channels between two adjacent electrodes, has been used to increase the number of spectral channels across the electrode array. Additionally, different types of stimulation are associated with different amounts of current spread and increased channel interaction. Alternate methods of current delivery and stimulation across electrode arrays have been used in an attempt to reduce electrode current spread and improve channel selectivity and spectral resolution.

1.4. Methods of electrode stimulation

Multiple electrodes can be configured to deliver stimulating currents to the auditory neurons in different ways. The three main configurations available with existing devices are known as monopolar, bipolar, and common ground. In clinical routine, modern CI devices all use a monopolar electrode configuration. In monopolar stimulation, the electrical current flows between one of the intracochlear electrode contacts and an extracochlear reference electrode, either in the casing of the implanted stimulator or in a ball electrode located under the temporalis muscle. The current pathway in monopolar stimulation, by definition, depends on the impedance of all structures between the active and the extracochlear reference electrodes, and is therefore largely uncontrolled. In CIs employing monopolar stimulation, it is important that the active electrodes be located close to the neural population so that, ideally, stimulation on each electrode excites a spatially distinct set of neurons and consequently elicits a perceptually discriminable auditory sensation.

In principle, the spatial separation of the stimulating current paths in CIs can be improved by using bipolar stimulation. In this configuration, current is passed between two electrodes, both of which are located relatively close to the auditory neurons. Several variations of the bipolar configuration may provide practical benefits in some conditions. In one variation, the separation between the two active electrodes can be increased, for example by activating pairs of electrodes that are separated by one or more inactive electrodes on the array. This usually results in a reduction of the current required to produce an audible sensation (i.e., threshold). Another variation involves arranging the electrodes spatially to direct the current flow in the cochlea more closely around a radial, rather than longitudinal, path. This is intended to increase the electrodes' spatial selectivity and reduce thresholds.

In the third type of electrode configuration, the common ground mode, one active electrode is selected, and many or all of the remaining intracochlear electrodes are used together as the return path for the stimulating current. In several respects, the common ground arrangement is intermediate between the bipolar and monopolar configuration.

Typically, CIs deliver stimulating currents to the active electrodes in a sequence of temporally nonoverlapping pulses. However, in all current CIs, it is possible for analog waveforms to be delivered simultaneously to several electrodes. Simultaneous stimulation by multiple electrodes may, in theory, have beneficial perceptual effects, particularly because it should enable the normal patterns of the auditory nerve responses to acoustic signals to be emulated more closely. Unfortunately, in past experiments with CIs, simultaneous stimulation has frequently been found to produce poor results; the complex summation of currents within the cochlea from multiple simultaneous active electrodes can result in reduced spatial selectivity of the neural excitation and poorer control of perceived loudness.

1.5. Measures of frequency resolution in CIs

A number of measures of frequency resolution have been applied to CIs. Perhaps the most direct are spatial tuning curves, where the current level needed to mask a brief low level signal is measured as a function of the spatial separation between the masker and the signal electrode. Cochlear implant users have roughly the same broad spatial frequency resolution at low to moderate stimulus levels that NH and HI listeners have at high stimulus levels. Accordingly, CI users can perform spectral resolution tasks under direct electrical stimulation as well as NH and HI listeners at high sound levels. Parameters of the speech processor and the electrode array, as opposed to spatial tuning characteristics, may be important factors limiting speech recognition for many CI users (Nelson et al., 2008).

Another measure involves spectral ripple discrimination, where a spectrally rippled stimulus is discriminated from another spectrally rippled stimulus, with the spectral (or spatial) positions of the peaks and valleys reversed (Henry and Turner, 2003; Litvak et al., 2007). The reasoning behind the test, which was originally developed to test normal hearing (Supin et al., 1994), is that the maximum ripple rate at which the original and phase-reversed stimuli are discriminable provides information regarding the limits of frequency resolution. Discriminating between different spectrally rippled broadband stimuli may be particularly relevant to speech tasks because the ripples are typically distributed over a wide spectral region and because the task of discriminating the positions of spectral peaks has some commonalities with identifying spectral features such as vowel formant frequencies in speech. Results have suggested that variability in spread of neural activation largely accounts for the variability in speech perception of CI listeners.

Previous studies suggest that reduced spectral resolution contributes to reduced speech perception in CI users. In acoustic hearing, when noise is added to speech, the auditory periphery is presented with the mixture of the two signals, and it is the task of the auditory system to segregate one from the other. In the electrical stimulation, however, noise presents as an addition to or alteration in the pattern of speech pulses. Such distortion is usually nonlinear and irrecoverable, yet another reason noise negatively impacts speech intelligibility more for a CI user than individuals with acoustic hearing (Nelson et al., 1996).

1.6. Critical Band

The concept of the critical band was developed during the 1930s at Bell Laboratories by Henry Fletcher (Fletcher, 1940). Fletcher used a band-narrowing procedure and showed that only a "critical band" (CB) of noise frequencies centered on a tone affected the tone's noise-masked threshold. A masker bandwidth at which the tone's thresholds began to decrease was used as an estimate of the CB. It was further shown that the CB increased in width (in Hz) as the frequency of the signal increased, implying a widening of the CB with increasing frequency. This and other observations were consistent with frequency resolution measured at the auditory periphery, and it was assumed that the CB reflected the biomechanical and neural processes occurring in the cochlea and auditory nerve. This concept has been extended to be modeled as a bank of bandpass filters (Fastl and Zwicker, 2007). This model is well matched with the essential function of the electrode array in the CI.

1.6.1. The Speech Critical Band

Healy and Bacon (2006) conducted measurements to determine the spectral resolution employed during the reception of running speech. This measure was called the *Speech* Critical Band (*S*-CB). It was known that CI simulations (noise-vocoded speech) having only four frequency channels can be well recognized despite their severely degraded spectral resolution (e.g., Apoux and Bacon, 2004; Shannon et al., 1995; Shannon et al., 1998). However, a single band of noise-vocoded speech is essentially unrecognizable (Shannon et al., 1995; Healy and Warren, 2003). This suggests that a minimum amount of spectral information is necessary to recognize speech (see Healy and Warren, 2003).

Healy and Bacon (2006) examined spectral resolution by partitioning various narrow bandwidths centered at 1500 Hz and replacing the information in each partition with carrier bands of noise that were amplitude-modulated by each speech partition's envelope. As the number of bands increased, so did spectral resolution. It was found that performance increased as spectral resolution increased, eventually reaching asymptote. The spectral resolution within the acoustic speech signal at performance asymptote reflects the resolution employed during processing of that signal, as no further increases in signal resolution could be used to improve performance. The *S*-CB is then defined as the resolution of the acoustic signal at intelligibility asymptote, and reflects the auditory system's ability to extract spectral detail from an acoustic speech signal. Results revealed that NH listeners benefited from increasing spectral information up to the limits of the peripheral auditory system and that the *S*-CB approximates the size of the psychophysical CB (Healy and Bacon, 2006). This result indicates that spectral resolution is determined not by limitations of an impoverished acoustic speech signal, but instead by limitations of the auditory periphery.

In the current study, the resolution that CI users are able to employ when processing everyday sentences was examined. The *S*-CB was measured in these users using a band of frequencies in the center of the speech spectrum, where resolution should be greatest, and by CI subjects using their everyday settings. The current results can be compared to those of Healy and Bacon (2006) involving NH listeners, for a comparison of speech frequency resolution across NH listeners and CI users.

Chapter 2: Methods

2.1. Experiment 1a

A preliminary experiment was performed to determine the relationship between bandwidth and intelligibility in CI users. The restriction of speech to a narrow band was required to reveal the maximum resolution of contrasting temporal patterns within the band. This information was then used to guide the selection of overall bandwidth used for measuring resolution in Experiment 1b.

2.1.1. Subjects

A group of ten CI users between the ages of 32 and 72 years of age (mean age = 54.3) participated in the experiment. Each participant received financial compensation for participation. All were postlingually hearing impaired. All were native speakers of American English and all had at least one year experience with their CI device and mapping strategy. All subjects used the Cochlear LTD. devices; two subjects utilized Nucleus Freedom speech processors and eight utilized Nucleus 5 (CP810) speech processors. The subjects were all implanted with the N24 Contour Advance electrode array with 22 electrodes, used the ACE strategy, monopolar (MP1+2) stimulation, a Q value of 20 and stimulation rates of 900 Hz or 1200 Hz/channel. All were considered average to above average users. All implant participants had previous experience in speech recognition experiments. Table 1 contains relevant information regarding gender, age, CI experience, age at diagnosis of hearing loss, etiology, bilateral or unilateral CI use, implanted ear used for testing, type of device and speech processor parameters. All CI listeners were tested unilaterally, utilizing the first ear implanted.

Table 1. Characteristics of CI participants.

Subject ID	Gender	Age (years)	CI experience (years)	Age at Dx of hearing loss	Etiology	Bilateral or unilateral	Implanted ear used for testing	Device
CISCB1	M	50	4	21	Sudden / unk	Bilateral	Right	Cochlear N5
CISCB2	M	65	4	12	Progressive	Bilateral	Right	Cochlear N5
CISCB3	F	72	4	45	Progressive/AIED	Unilateral	Left	Cochlear Freedom
CISCB4	F	61	7	40	Progressive	Bilateral	Right	Cochlear N5
CISCB5	F	36	2	7	Progressive	Unilateral	Left	Cochlear N5
CISCB6	M	62	3	50	Meniere's	Unilateral	Left	Cochlear N5
CISCB7	M	32	9	13	Progressive	Bilateral	Left	Cochlear N5
CISCB8	M	46	7	16	Hereditary	Unilateral	Left	Cochlear Freedom
CISCB9	F	65	7	40	Otosclerosis	Bilateral	Right	Cochlear Freedom
CISCB10	F	63	2	36	Hereditary	Unilateral	Right	Cochlear N5
Avg		55.20	4.90	28.00				

2.1.2. Stimuli

The stimuli were based upon the recordings of the everyday American speech sentences from the Central Institute for the Deaf (CID; Davis and Silverman, 1978; Silverman and Hirsh, 1955). Male-talker 22050-Hz, 16-bit digital files were utilized. Initially, all sentences were scaled to equate total RMS energy. The sentences were filtered to a single narrow band having a width of 1, 3/2, 2 or 3 octaves centered at 1500 Hz. Filtering was performed using a single pass through a 2000-order digital FIR filter implemented in MATLAB. These parameters produced extremely steep filter slopes, measuring over 1000 dB/octave.

Sentence recognition was conducted in free field in a double-walled sound treated booth (Industrial Acoustics Company, IAC; Winchester, Hampshire, UK). The processed digital signals were converted to analog form, amplified (Edirol UA-5, USB Digital Audio Capture), then presented at 65 dBA via a single powered loudspeaker (Mackie HR824). Presentation levels were set using a Larson Davis 824 sound level meter. Subjects were seated 1m from the loudspeaker and instructed to face the loudspeaker directly. Subjects were instructed to use their everyday volume and microphone sensitivity settings in their speech processors.

2.1.3. Procedure

Subjects were tested individually, seated with the experimenter in the audiometric booth. Prior to the first condition, each listener heard 5 practice sentences from recordings (male-talker 22050-Hz, 16-bit digital files) of the Speech Perception In Noise test (SPIN, Bilger, et al., 1984) in the broadband condition, followed by the same 5 sentences in the 3-octave condition. The listener then heard 5 CID sentences in each of the four bandwidths (sentences 81-100). Following this period of familiarization, each listener heard 10 sentences in each bandwidth condition in a randomized order. This set was then repeated using a new randomized order, for a total of 20 sentences per condition. Subjects were instructed to repeat as much of each sentence as possible aloud after hearing it. The listeners heard each sentence only once, received no feedback, and were encouraged to guess if unsure of the content. The experimenter was seated with the subject and controlled the presentation of sentences and scored the proportion of scoring keywords reported correctly.

2.1.4. Results

For each condition, scores from both 10-sentence groupings were averaged for a total intelligibility score. Figure 1 shows the intelligibility scores for the four speech bandwidths presented. Intelligibility increased with each bandwidth in most cases.

For Experiment 1b and in order to best evaluate the amount of information necessary for intelligibility, it was important to choose a bandwidth at which the subjects performed sufficiently well. Performance of at least 50% was obtained by 8 of the 10 subjects in the 3/2-octave condition. Therefore, 3/2 octaves was selected for further measurement of frequency resolution in assessing intelligibility.

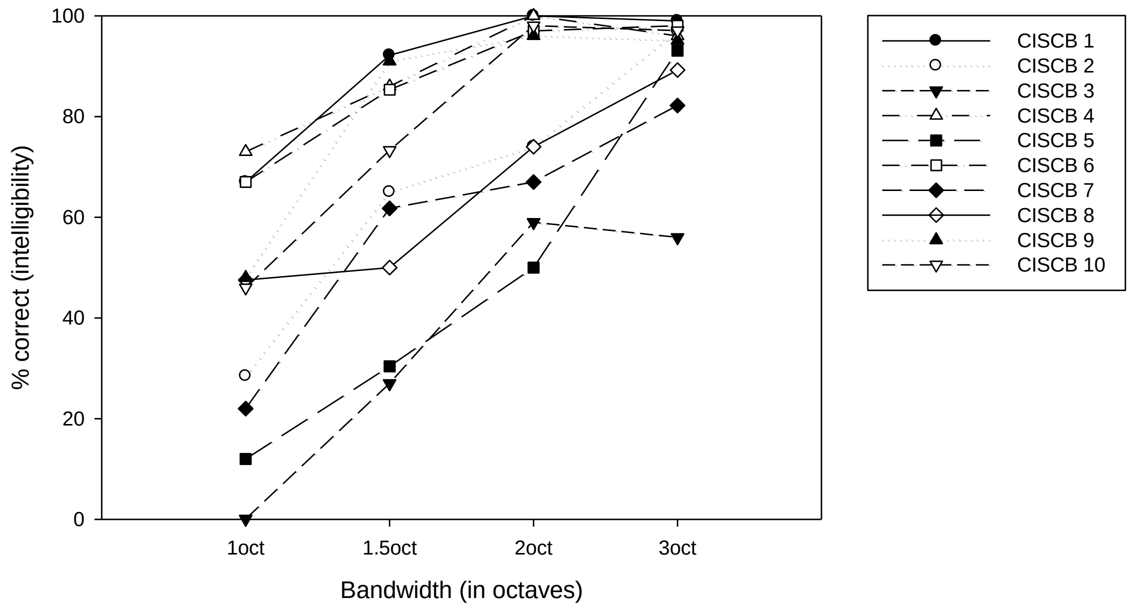

Fig. 1. Mean intelligibility scores on the four bandwidth conditions (1, 3/2, 2 or 3 octaves) centered at 1500 Hz.

2.2. Experiment 1b

In this experiment, spectral information within the narrow speech band was quantized by partitioning the band, and removing spectral information from each partition by replacing it with a carrier band that was amplitude modulated by the envelope of the corresponding speech partition.

2.2.1. Subjects and Stimuli

The same group of 10 CI users participated. The stimuli were the original standard recordings (original male-talker, 20161-Hz sampling, 16-bit resolution) of the Hearing In Noise Test (HINT, Nilsson et al., 1994). All sentences were scaled to equate total RMS energy. Based on the results of Experiment 1a, the 3/2-octave speech band was selected for the further measurement of frequency resolution. Five additional conditions were created by partitioning the 3/2-octave band into 2, 4, 6, 8 and 10 equal log-width sub-bands. The lowest and highest partitions were created using a low- or high-pass respectively; the inner bands were created using a bandpass filter. A 6000 FIR filter order was used for this processing to more sharply define the bands. Because the FIR filter is linear in phase, all component bands were exactly aligned in time.

As employed previously by Healy and Bacon (2006), low-noise noise (LNN) was selected for the carrier band. LNN is noise engineered to have extremely small fluctuations in amplitude (Pumplin, 1985; Hartmann and Pumplin, 1988). LNN can be created by restricting the relationship between noise components such that they will have related phase. The method proposed by Kohlrausch et al. (1997) was employed to generate the LNN. The method involves the division of the waveform by its envelope in a series of repetitions. Low-noise noise carriers were selected over Gaussian noise carriers because the random amplitude fluctuations of the narrow-band noises could potentially reduce the temporal details of the speech. These were selected over tonal carriers to allow spectral density to remain constant as the number of partition bands changed. Carrier bands having the same frequency composition as the speech partitions were created by summing sinusoidal components having appropriate amplitude and phase, and 0.5 Hz spacing. This component spacing produced a repeated noise having a duration that was sufficiently long to not substantially interfere with the perception of the sentences. The LNN carrier bands were separated by 0.5 Hz so that the array would have equal spacing of components.

The amplitude envelope was extracted from each speech partition by full-wave rectification and low-pass filtering (2000-order FIR, 100-Hz cutoff) and applied to a corresponding LNN carrier band. The AM LNN carriers were then post-filtered to ensure the restriction of the frequency region of the origin using the same filters employed to create the speech partitions.

The AM carriers comprising each condition (1, 2, 4, 6, 8 and 10 bands) were assembled for presentation to listeners. Because this method of manipulation preserved the relative overall level of each component band, this resulting array maintained the spectral profile of the original speech band. Each sentence in each condition was presented at a level of 65 dBA in the same manner and using the apparatus employed in the previous experiment.

Of these, only the 8-band condition was authored for the present study. All other conditions were produced previously (see Healy and Bacon, 2006).

2.2.2. Procedure

Again, subjects were tested individually and unilaterally as previously outlined in Exp. 1a, seated with the experimenter in the double-walled audiometric booth. Testing began with a period of familiarization where each listener heard 5 practice HINT sentences (from lists not utilized during testing) in the broadband condition, followed by the same 5 sentences in the 3/2 octave and 8-band conditions. The listener then heard 10 new practice sentences in each of the 3/2 octave and 8-

band conditions. For testing, listeners heard 5 rounds, each consisting of 5 sentences in each of the 6 conditions, blocked and presented in randomized order for a total of 150 sentences. Finally, and to allow for comparisons to results from normal hearing (NH) listeners by Healy and Bacon (2006), each CI listener heard a list of 10 sentences in the 1-band condition.

Subjects were instructed to repeat each sentence aloud after hearing it. They heard each sentence only once, received no feedback, and were encouraged to guess if unsure of the content. The experimenter controlled the presentation of sentences and scored the proportion of component words reported correctly.

2.2.3. Results

Intelligibility scores for the 3/2-octave bandwidth condition are shown in Figure 2. Within each group of randomized conditions, each 5-sentence bandwidth set was scored individually. Scores from the five blocks of six bandwidth conditions were averaged for a total intelligibility score. Finally, the 1-band condition was presented at the end of each session and scores were averaged.

Figure 3 shows mean intelligibility performance in the CI subjects grouped according to performance and plotted with the mean intelligibility scores from NH subjects from Healy and Bacon, 2006. CI subjects' performance increased with increasing number of partitions. However, performance does not reach asymptote as information is increased. Results from NH listeners tested by Healy and Bacon (2006) are overlaid. Performance in NH listeners was found to asymptote beyond presentations with 6 carrier bands.

A simple linear regression analysis was used to evaluate the effect of increasing spectral detail on performance in the highest- and lowest-performing CI subjects separately. The effect of spectral detail was linearly related to performance by both the highest-performing group, [r^2=.959, adjusted r^2=.948, ($p<.01$); $F(1, 5)$ = 60.833, $p < 0.01$] and by the lowest-performing group [r^2=.921, adjusted r^2=.901, ($p< 0.01$), $F(1, 5)$ = 60.833, p=0.002.]. The regression coefficient was statistically significant ($p < 0.01$).

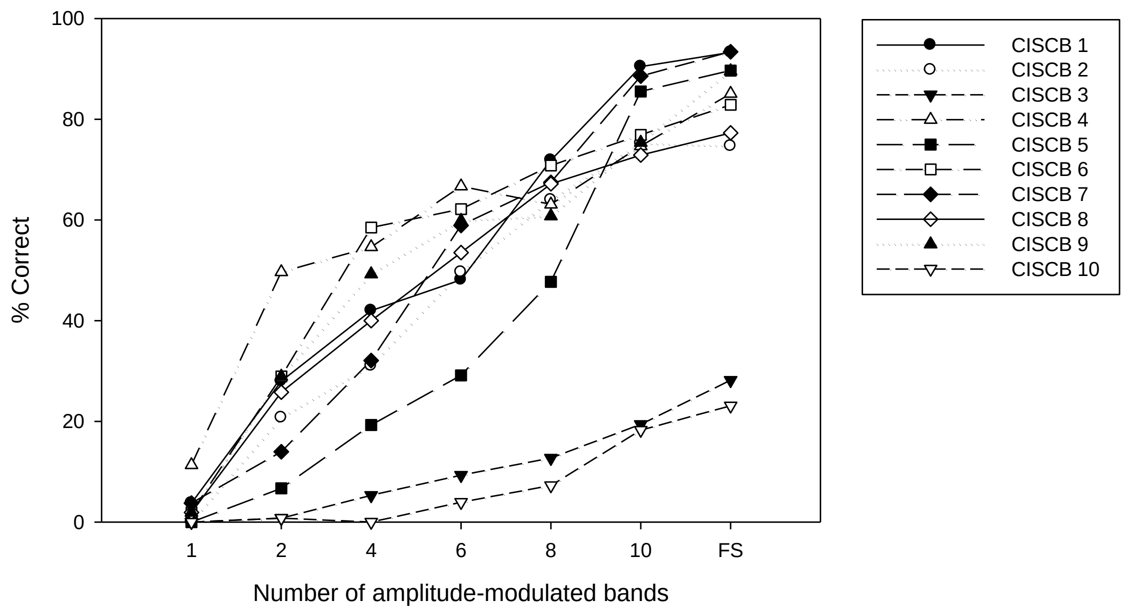

Fig. 2. Mean intelligibility scores for the 3/2-octave bandwidth condition having increasing frequency resolution.

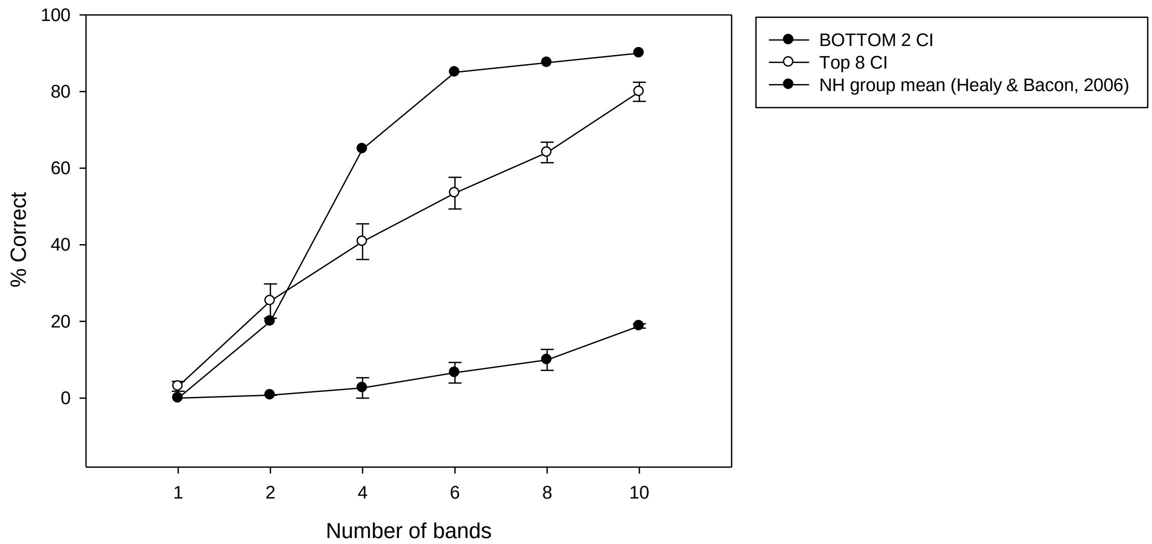

Fig. 3. Mean intelligibility scores of CI subjects grouped according to performance and plotted with the mean intelligibility scores from NH subjects from Healy and Bacon, 2006. NH subjects performance asymptotes at the 6-band condition. CI performance increases with increased resolution; neither group reaches asymptote with increased frequency resolution.

Chapter 3: General Discussion

Speech can be considered as an assembly of temporal patterns at different spectral locations. Work over the past 15 years has dramatically increased our appreciation of the spectral and temporal analysis that is normally associated with the perception of speech. Although this work has generally involved basic research and NH listeners, it has been motivated in part by CIs, which present a limited number of temporal patterns to corresponding fixed locations along the cochlea. This increased appreciation of spectral and temporal analysis is related to a greater understanding of the temporal information provided by the fluctuating amplitude patterns of speech (Thibodeau and Van Tasell, 1987; Rosen, 1992) and the ability of listeners to understand speech represented by a small number of temporal patterns at different spectral frequencies (Shannon et al., 1995).

It has been found that when NH subjects are presented with acoustic simulations by cochlear prosthesis, four to six independent channels of information are sufficient to achieve high levels of speech recognition under ideal situations (Shannon et al., 1995; Dorman and Loizou, 1997; Loizou et al., 1999). Under more difficult listening conditions, such as in the presence of background noise, the number of channels needed for the same levels of speech recognition was found to be much larger, depending on the SNR (Dorman et al., 1998b; Fu et al., 1998a; Friesen et al., 2001).

There is evidence that CI users can utilize as few as four to six channels effectively (Fishman et al., 1997; Dorman et al., 1998; Fu et al., 1998a; Friesen et al., 2001). Resolution may be limited due to a number of factors, such as the spread of current produced by each electrode and uneven neural survival patterns along the sensory epithelia, which are likely to vary between individual subjects (e.g., Hinojosa and Marion, 1983; Kawano et al., 1998; Xu and Pfingst, 2003). Fishman et al. (1997) evaluated the subjective benefit in implant patients of experimental processors that varied in the number of channels. They found that the subjective ratings of benefit increase up to 10 channels. Little further benefit was observed with 20 channels relative to 10.

One critical limitation to spectral resolution in the CI is the limited number of available physical channels. Additionally, spectral resolution is further limited when the channels are not independent. Cochlear implant users are often able to discriminate most electrodes. However, two adjacent electrodes that are discriminable may not provide independent channels of information. McDermott and McKay (1994) showed that different modulation rates delivered to two adjacent electrodes were perceived as having a pitch between the modulation rates, suggesting that although the electrodes were discriminable, they were not independent. Cochlear implant users' functional spectral resolution may be compromised by channel interactions resulting from current spread between electrodes; CI users having high channel interactions receive inputs to the auditory nerve with a high degree of spectral smearing (Fu and Nogaki, 2005). Fu and Nogaki (2005) further suggested that channel interaction was the limiting factor in CI performance by comparing performance to that of NH subjects listening to four broad noise bands to simulate channel interaction.

The influence of the number of active channels has been studied extensively. Though many studies have looked at speech understanding (Wilson et al., 1995; Friesen et al., 2005), others evaluated more focused measures such as CI users' ability to detect current level changes (Drennan and Pfingst, 2006), loudness perception (Macherey et al., 2006), effects of pulse duration and pulse rate (Shannon, 1989, Skinner et al., 2000; Kreft et al., 2004; Middlebrooks, 2004) and stimulation rates (Fu and Galvin, 2001; Holden et al., 2002). However, more accurate quantification of the amount of information implanted users need and can reliably utilize is needed.

Although the presence of current spread and channel interactions is well documented, it is not known to what extent CI users can extract spectral information from speech. This information is presumably of fundamental importance. In a preliminary experiment (Exp. 1a), the relationship between bandwidth and intelligibility in CI users was determined by restricting broadband speech to a narrow band centered at 1500 Hz. Based on this experiment, a 3/2-octave band was used for quantification of the frequency resolution employed by CI users. Spectral information within a 3/2-octave speech band was quantized into 1, 2, 4, 6, 8, and 10 sub-bands. With the spectral information restricted around 1500 Hz in this manner, thereby limiting the needed activation of the electrode array to very few channels, we observed increased performance with an increased amount of spectral information. When grouped according to overall level of performance, no change in resolution was observed between CI groups (see Fig.3). Further, no asymptote in performance was observed. Instead, performance increased as spectral resolution increased up the highest resolution employed. It is important to note that the resolution in the 10-band condition approximates current estimates of psychophysical tuning in NH at moderate levels. This observed spectral resolution was surprisingly high, given what is known about CI spectral resolution. These results indicate that CI users can access high levels of spectral resolution when the overall signal is restricted in bandwidth.

The current results are similar to those obtained by Galvin and Fu (2011). Melodic contour identification was measured in CI and NH subjects listening to piano samples that were bandpass-filtered into low, middle, high frequency ranges to preserve different amounts of F0 and harmonic information. It was found that NH listeners reached ceiling values for all filter ranges and performed much better than CI users. The best CI user performance was observed with the middle frequency range. This middle frequency band provided as good, and in some cases, better melodic contour identification than did the broadband signal. Galvin and Fu (2011) concluded that acoustic filtering may reduce potential mismatches between fundamental frequencies and harmonic components thereby improving CI users' melody perception.

One limitation of current CI devices involves the mismatch in frequency coding that exists between the source of information within the speech spectrum and the site of delivery of electrical stimulation inside the cochlea. Spectral information is shifted upward by CI subjects' clinical frequency allocations, with the greatest mismatch in the lower frequencies and reduced amounts in the middle to higher frequencies. This mismatch is a direct result of the inability to insert an electrode array deep into the cochlea and the resulting inability to stimulate sites that normally correspond to lower frequencies. In clinical CI speech processors, frequency allocation is generally optimized for speech recognition, with higher frequency resolution around 1500 Hz, where speech possesses the maximum density of information (ANSI, 1997). While broadband speech has the widest band of information and utilizes the most electrodes in the array to stimulate, our results suggest that greater resolution of speech might be possible using much narrower bandwidths and fewer stimulating electrodes, perhaps due to reduced degree of absolute mismatch within in the stimulated region.

Fundamental frequency (F0) information is also not adequately coded in current CI devices, due to the reduced spectral resolution of the speech information transmitted by the implant. The poor encoding of the harmonic structure affects CI users' performance in voice gender recognition (Stickney et al., 2004; 2007), music perception (Gfeller et al., 2005, 2002a; Kong et al., 2004; Laneau et al., 2006), recognition of prosodic aspects of speech, including intonation (Green et al., 2004, 2004; Peng et al., 2008) and lexical tone recognition (Ciocca et al., 2002; Luo et al., 2008; Peng et al., 2004; Wei et al., 2007). It is suggested that CI users may benefit from focused narrowband stimulations in order to gain additional access to this tonal information.

One way to increase spectral resolution in CIs is through current steering, which allows the creation of "virtual" channels between electrode contacts. Stimulation can be steered to sites between the electrodes by varying the proportion of current delivered simultaneously to adjacent electrodes, eliciting pitch percepts that are between the two electrodes. In theory, the number of distinct pitches that can be heard defines the number of spectral channels that can be perceived by the user. Reducing current spread would most likely also reduce channel interaction, which would, in turn, increase the number of independent spectral channels. However, virtual channel discrimination by way of current steering has not been shown to consistently benefit speech perception in CI users (Berenstein et al., 2008; Brendel et al., 2008; Busby et al., 2008; Saoji et al., 2009). Results from Landsberger et al. (2012) show that current steering and the creation of virtual channels do not provide greater benefit than single electrode activation. The current use of narrowband filtering has the potential to not only decrease channel interaction, but also more successfully generate virtual channels.

While improving implant and speech processor functionality, such as size, water resistance, battery life and accessories remains a priority to implant manufacturers, speech intelligibility in both simple and complex listening environments, through increased spectral resolution, should remain an ultimate goal in cochlear implant design. Results from the current study suggest that CI users do benefit from increasing amounts of spectral resolution when presented within narrow bands. The resolution employed by CI users in the current study matched that employed during NH and is therefore surprisingly high. This technique of providing increased spectral resolution holds promise for improving music perception, tonal language perception, intonation and prosodic cue use, as well as overall speech recognition in CI users.

Chapter 4: Summary and Conclusions

A preliminary experiment revealed the relationship between overall bandwidth and sentence intelligibility in CI users. Speech centered at 1500 Hz at four different bandwidths (1-, 3/2-, 2- and 3-octaves) was utilized to force listeners to employ the maximum resolution of contrasting temporal patterns within speech stimuli. For the main experiment, a 3/2-octave band speech was presented. Spectral information within the speech band was controlled by partitioning the band into six (1, 2, 4, 6, 8, 10) sub-bands, and removing spectral information from each partition by replacing it with a carrier band that was amplitude modulated by the envelope of the corresponding speech partition. Cochlear implant subjects' performance increased with increasing number of partitions, never reaching asymptote as spectral information increased. This is in stark contrast to expectation, as it indicates that increases in spectral resolution up to that of NH produced increases in performance. Accordingly, it is concluded that CI users can access spectral information as high as that of NH when presented with narrowband speech stimuli.

References

American National Standards Institute. (1997). *American National Standard Methods for the Calculation of the Speech Intelligibility Index.* (ANSI-S3.5). New York: American National Standards Institute.

Apoux, F., and Bacon, S. P. (2004). Relative importance of temporal information in various frequency regions for consonant identification in quiet and in noise. *The Journal of the Acoustical Society of America*, *116*, 1671.

Baer, T., Moore, B.C., and Gatehouse, S. (1993). Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times, *Journal of Rehabilitation Research and Development, 30,* 49–72.

Battmer, R. D., Reid, J. M., and Lenar, T. (1997). Performance in Quiet and in Noise with the Nucleus® Spectra 22 and the Clarion CIS/CA Cochlear Implant Devices. *Scandinavian audiology, 26(4),* 240-246.

Berenstein, C.K., Mens, L.H.M., Mulder, J.J.S., and Vanpoucke, F.J. (2008). Current steering and current focusing in cochlear implants: comparison of monopolar, tripolar and virtual channel electrode configurations. *Ear and Hearing, 29,* 250–260.

Bilger, R.C., Nuetzel, J.M., Rabinowitz, W.M., and Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research, 27,* 32-48.

Boothroyd, A., Mulhearn, B., Gong, J., and Ostroff, J. (1997). Simulation of sensorineural hearing loss: Reducing frequency resolution by uniform spectral smearing. In *Modeling Sensorineural Hearing Loss* (pp. 313-329). Mahwah, NJ: Erlbaum.

Brendel, M., Buechner, A., Krueger, B., Frohne-Buechner, C., and Lenarz, T. (2008). Evaluation of the Harmony sound processor in combination with the speech coding strategy HiRes120. *Otology and Neurotology, 29,* 199–202.

Busby, P.A., Battmer, R.D., and Pesch, J. (2008). Electrophysiological spread of excitation and pitch perception for dual and single electrodes using the Nucleus Freedom cochlear implant. *Ear and Hearing, 29(6),* 853-864.

Carlyon, R., Long, C., Deeks, J., and McKay, C. (2007). Concurrent Sound Segregation in Electric and Acoustic Hearing. *Journal of the Association for Research in Otolaryngology, 8(1),* 119-133.

Chatterjee, M., Sarampalis, A., and Oba, S,I. (2006). Auditory stream segregation with cochlear implants: A preliminary report. *Hearing Research, 222,* 100–107.

Ciocca, V., Francis, A. L., Aisha, R., and Wong, L. (2002). The perception of Cantonese lexical tones by early-deafened cochlear implantees. *The Journal of the Acoustical Society of America*, *111*, 2250.

Davis, H., and Silverman, S.R. (1978). *Hearing and deafness, 4th ed.* New York: Holt, Rinehart, and Winston.

Dorman, M.F., and Loizou, P.C. (1997). Speech intelligibility as a function of the number of channels of stimulation for normal-hearing listeners and patients with cochlear implants. *American Journal of Otolaryngology - Head and Neck Medicine and Surgery,* 18, 113–114.

Dorman, M.F., and Loizou, P.C. (1998a). The identification of consonants and vowels by cochlear implant patients using a 6-channel Continuous Interleaved Sampling processor and by normal-hearing subjects using simulations of processors with two to nine channels. *Ear and Hearing, 19,* 162–166.

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

Dowell, R. C., Seligman, P. M., Blamey, P. J., and Clark, G. M. (1987). Speech perception using a two-formant 22-electrode cochlear prosthesis in quiet and in noise. *Acta oto-laryngologica, 104(5, 6),* 439-446.

Drennan, W.R., and Pfingst, B.E. (2006). Current-level discrimination in the context of interleaved, multichannel stimulation in cochlear implants: effects of number of stimulated electrodes, pulse rate, and electrode separation. *Journal of the Association for Research in Otolaryngology, 7(3),* 308-316.

Fastl, H., and Zwicker, E. (2007). *Psychoacoustics: Facts and models (vol. 22).* Verlag: Springer.

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

Festen, J.M., and Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America, 88,* 1725–1736.

Fetterman, B. L., and Domico, E. H. (2002). Speech recognition in background noise of cochlear implant patients. *Otolaryngology--Head and Neck Surgery, 126(3),* 257-263.

Fishman, K.E., Shannon, R.V., and Slattery, W.H. (1997) Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *Journal of Speech, Language and Hearing Research, 40(5),* 1201-1215.

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics, 12,* 47–65.

Friesen, L.M., Shannon, R.V., Baskent, D., and 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*, 1150.

Friesen, L.M., Shannon, R.V., and Cruz, R.J. (2005). Effects of stimulation rate on speech recognition with cochlear implants. *Audiology and Neurotology*,*10*(3), 169-184.

Fu, Q. J., and Galvin III, J. J. (2001). Recognition of spectrally asynchronous speech by normal-hearing listeners and Nucleus-22 cochlear implant users.*The Journal of the Acoustical Society of America*, *109*, 1166.

Fu, Q.J., and Nogaki, G. (2005). Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology, 6,* 19–27.

Fu, Q. J., Shannon, R. V., & Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing.*The Journal of the Acoustical Society of America*, *104,* 3586.

Galvin, J. J., and Fu, Q. J. (2011). Effect of bandpass filtering on melodic contour identification by cochlear implant users. *The Journal of the Acoustical Society of America*, *129(2),* EL39.

Gfeller, K., Turner, C., and Mehr, M., Woodworth, G., Fearn, R., Knutson, J.F., Witt, S., and Stordahl, J. (2002a). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. Cochlear Implants International, 3, 31–55.

Gfeller, K., Witt, S., Woodworth, G., Mehr, M.A., and Knutson, J. (2002b). Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *Annals of Otology, Rhinology, and Laryngology, 111,* 349–356.

Gfeller, K., Olszewski, C., Rychener, M., Sena, K., Knutson, J. F., Witt, S., and Macpherson, B. (2005). Recognition of" real-world" musical excerpts by cochlear implant recipients and normal-hearing adults. *Ear and hearing, 26(3),* 237-250.

Glasberg, B.R., and Moore, B.C.J. (1990). Derivation of auditory filter shapes from notched-noise data, *Hearing Research, 47,* 103–138.

Glasberg, B.R., and Moore, B.C.J. (1992). Effects of envelope fluctuations on gap detection. *Hearing research, 64(1),* 81-92.

Green, T., Faulkner, A., and Rosen, S. (2004). Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. *The Journal of the Acoustical Society of America, 116,* 2298–2310.

Hartmann, W.M., and Pumplin, J. (1988). Noise power fluctuations and the masking of sine signals. *The Journal of the Acoustical Society of America, 83,* 2277–2289.

Healy, E.W., and Warren, R.M. (2003). The role of contrasting temporal amplitude patterns in the perception of speech. *The Journal of the Acoustical Society of America, 113,* 1676-1688.

Healy, E.W., and Bacon, S.P. (2006). Measuring the critical band for speech. *The Journal of the Acoustical Society of America, 119,* 1083-1091.

Henry, B.A., and Turner, C.W. (2003). The resolution of complex spectral patterns in cochlear implant and normal hearing listeners. *Journal of the Acoustical Society of America, 113,* 2861–2873.

Hinojosa, R., and Mitchell, M. (1983). Histopathology of profound sensorineural deafness. *Annals of the New York Academy of Sciences.* 405.1, 459-484.

Hochberg, I., Boothroyd, A., Weiss, M., and Hellman, S. (1992). Effects of noise and noise suppression on speech perception by cochlear implant users. *Ear and hearing, 13(4),* 263-271.

Holden, L.K., Skinner, M.W., Holden, T.A., and Demorest, M.E. (2002). Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy. *Ear and hearing, 23(5),* 463-476.

Hong, R.S., and Turner, C.W. (2006). Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients. *Journal of the Acoustical Society of America, 120,* 360–374.

Jusczyk, P.W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Kawano, A., Seldon, H. L., Clark, G. M., Ramsden, R. T., and Raine, C. H. (1998). Intracochlear factors contributing to psychophysical percepts following cochlear implantation. *Acta oto-laryngologica*, *118(3),* 313-326.

Kiefer, J., Müller, J., Pfennigdorff, T., Schön, F., Helms, J., Von Ilberg, C., Baumgartner, W., Gstöttner, W., Ehrenberger, K., Arnold, W., Stephan, K., Thumfart, W., and Baur, S. (1996). Speech understanding in quiet and in noise with the CIS speech coding strategy (MED-EL Combi-40) compared to the multipeak and spectral peak strategies (Nucleus). *ORL: Journal for Otorhinolaryngology and its Related Specialties, 58(3),* 127-135.

Kohlrausch, A., Fassel, R., van der Heijden, M., Kortekaas, R., van de Par, S., Oxenham, A. J., and Püschel, D. (1997). Detection of tones in lownoise noise: Further evidence for the role of envelope fluctuations, *Acta Acoustica, 83,* 659–669.

Kong Y., Cruz R., Jones J., and Zeng F. (2004). Music perception with temporal cues in acoustic and electric hearing. *Ear and Hearing, 25,* 173–185

Kreft, H.A., Donaldson, G.S., and Nelson, D.A. (2004). Effects of pulse rate and electrode array design on intensity discrimination in cochlear implant users. *The Journal of the Acoustical Society of America, 116,* 2258.

Landsberger, D. M., Padilla, M., & Srinivasan, A. G. (2012). Reducing current spread using current focusing in cochlear implant users. *Hearing Research, 284,*16-24.

Laneau, J., Wouters, J., and Moonen, M. (2006). Improved music perception with explicit pitch coding in cochlear implants. *Audiology and Neurotology, 11(1),* 38-52.

Litvak, L.M., Spahr, A.J., Saoji, A.A., and Fridman, G.Y. (2007). Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *Journal of the Acoustical Society of America, 122,* 982–991.

Loizou, P. C., Dorman, M., and Tu, Z. (1999). On the number of channels needed to understand speech. *The Journal of the Acoustical Society of America*, *106*, 2097.

Luo, X., Fu, Q.J., Wei, C.G., and Cao, K.L. (2008). Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users. *Ear and Hearing, 29,* 957–970.

Macherey, O., Van Wieringen, A., Carlyon, R.P., Deeks, J.M., and Wouters, J. (2006). Asymmetric pulses in cochlear implants: effects of pulse shape, polarity, and rate. *Journal of the Association for Research in Otolaryngology,7(3)*, 253-266.

McDermott, H.J., and McKay, C.M. (1994). Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea. *The Journal of the Acoustical Society of America*, *96,* 155.

McKay, C.M., McDermott, H.J., and Clark, G.M. (1994). Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. *The Journal of the Acoustical Society of America*, *96*, 2664.

Middlebrooks, J.C. (2004). Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds. *The Journal of the Acoustical Society of America*, *116,* 452.

Moore, B.C.J. (2003). An Introduction to the Psychology of Hearing, 5th ed. Academic, London.

Müller-Deiler, J., Schmidt, B.J., and Rudert, H. (1995). Effects of noise on speech discrimination in cochlear implant patients. *Annals of Otology, Rhinology, and Laryngology, 166,* 303–306.

Nelson, D. A., Schmitz, J. L., Donaldson, G. S., Viemeister, N. F., and Javel, E. (1996). Intensity discrimination as a function of stimulus level with electric stimulation. *The Journal of the Acoustical Society of America*, *100,* 2393.

Nelson, D.A., Donaldson, G.S., and Kreft, H. (2008). Forward-masked spatial tuning curves in cochlear implant users. *The Journal of the Acoustical Society of America*, *123*, 1522.

Nelson, P.B., Jin, S.H., Carney, A.E., and Nelson, D.A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *113,* 961–968.

Nelson, P.B., and Jin, S.H. (2004). Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America, 115,* 2286–2294.

Nilsson, M., Soli, S.D., and Sullivan, J.A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America*, *95,* 1085–1099.

Oxenham, A.J., Bernstein, J.G., and Penagos, H. (2004). Correct tonotopic representation is necessary for complex pitch perception. *Proceedings of the National Academy of Sciences (USA), 101,* 1421–1425.

Oxenham, A.J. (2008). Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends in amplification,12(4),* 316-331.

Patterson, R. D., and Nimmo-Smith, I. (1980). Off-frequency listening and auditory-filter asymmetry. *The Journal of the Acoustical Society of America,67*, 229.

Peng, S., Tomblin, J.B., Cheung, H., Lin, Y.S., and Wang, L.S. (2004). Perception and production of Mandarin tones in prelingually deaf children with cochlear implants. *Ear and Hearing*, *25,* 251–264.

Peng, S.C., Tomblin, J.B., and Turner, C.W. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear and Hearing*, *29(3),* 336-351.

Pumplin, J. (1985). Low-noise noise. *The Journal of the Acoustical Society of America*, *78,* 100–104.

Qin, M.K., and Oxenham, A.J. (2005). Effects of envelope-vocoder processing on F0 discrimination and concurrentvowel identification. *Ear and Hearing, 26,* 451–460.

Rogers, C.F., Healy, E.W., and Montgomery, A.A. (2006). Sensitivity to isolated and concurrent intensity and fundamental frequency increments by cochlear implant users under natural listening conditions. *Journal of the Acoustical Society of America*, *119*, 2276-2287.

Rose, M.M., and Moore, B.C.J. (1997). Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *102,* 1768–1778.

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects, *Philosophical Transactions of the Royal Society B: Biological Sciences, 336*, 367–373.

Saoji, A.A., Litvak, L., Spahr, A.J., and Eddins, D.A. (2009). Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners. , *Journal of the Acoustical Society of America, 126,* 955–958.

Seligman, P., and McDermott, H. (1995). Architecture of the Spectra 22 speech processor. *The Annals of Otology, Rhinology andLlaryngology. Supplement*, *166*, 139.

Shannon, R.V. (1989). A model of threshold for pulsatile electrical stimulation of cochlear implants. *Hearing Research*, *40(3),* 197-204.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270,* 303–304.

Silverman, S.R., and Hirsh, I.J. (1955). Problems related to the use of speech in clinical audiometry. *The Annals of Otology, Rhinology, andLlaryngology*, *64(4),* 1234.

Skinner, M.W., Clark, G.M., Whitford, L.A., Seligman, P.M., Staller, S.J., Shipp, D.B., Shallop, J.K., Everingham, C., Menapace, C.M., and Arndt, P.L. (1994). Evaluation of a new spectral peak coding strategy for the Nucleus 22 Channel Cochlear Implant System. *The American Journal of Otology*, *15*, 15.

Skinner, M.W., Holden, L.K., Holden, T.A., and Demorest, M.E. (2000). Effect of stimulation rate on cochlear implant recipients' thresholds and maximum acceptable loudness levels. *Journal of the American Academy of Audiology*,*11(4)*, 203.

Stickney, G.S., Zeng, F.G., Litovsky, R., and Assmann, P.F. (2004). Cochlear implant speech recognition with speech maskers. *The Journal of the Acoustical Society of America*, *116*, 1081–1091.

Stickney, G.S., Assmann, P.F., Chang, J., and Zeng, F.G. (2007). Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences *The Journal of the Acoustical Society of America*, *122,* 1069–1078.

Supin, A.Y., Popov, V.V., Milekhina, O.N., and Tarakanov, M.B. (1994). Frequency resolving power measured by rippled noise. *Hearing Research, 78,* 31–40,

ter Keurs, M., Festen, J.M., and Plomp, R. (1992). Effect of spectral envelope smearing on speech reception. I *The Journal of the Acoustical Society of America, 91*, 2872–2880.

ter Keurs, M., Festen, J.M., and Plomp, R. (1993). Effect of spectral envelope smearing on speech reception. II. *The Journal of the Acoustical Society of America, 93,* 1547–1552.

Thibodeau, L.M., and Van Tasell, D.J. (1987). Tone detection and synthetic speech discrimination in band-reject noise by hearing-impaired listeners, *The Journal of the Acoustical Society of America, 82*, 864–873.

Townsend, B., Cotter, N., van Compernolle, D., and White, R.L. (1987). Pitch perception by cochlear implant subjects. *The Journal of the Acoustical Society of America, 82*, 106–115.

Turner, C.W., Gantz, B.J., Vidal, C., Behrens, A., and Henry, B.A. (2004). Speech recognition in noise for cochlear implant listeners: benefits of residual acoustic hearing. *Journal of the Acoustical Society of America, 115*, 1729–1735.

Vandali, A.E., Whitford, L.A., Plant, K.L., and Clark, G.M. (2000). Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear and hearing, 21(6),* 608-624.

Wei, C., Cao, K., Jin, X., Chen, X., and Zeng, F.G. (2007). Psychophysical performance and mandarin tone recognition in noise by cochlear implant users. *Ear and Hearing, 28,* 62S–65S.

Wilson, B.S., Finley, C.C., Lawson, D.T., Wolford, R.D., Eddington, D.K., and Rabinowitz, W.M. (1991). Better speech recognition with cochlear implants. *Nature, 352(6332),* 236-238.

Wilson, B.S., Lawson, D.T., Zerbi, M., Finley, C.C., and Wolford, R.D. (1995). New processing strategies in cochlear implantation. *Otology and Neurotology, 16(5),* 669-675.

Won, J.H., Drennan, W.R., and Rubinstein, J.T. (2007) . Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *Journal of the Association for Research in Otolaryngology, 8,* 384–392.

Xu, L., and Pfingst, B. E. (2003). Relative importance of temporal envelope and fine structure in lexical-tone perception (L). *The Journal of the Acoustical Society of America*, *114*, 3024.

Xu, L., Thompson, C.S., and Pfingst, B.E. (2005). Relative contributions of spectral and temporal cues for phoneme recognition, *Journal of the Acoustical Society of America*, *117,* 3255–3267.