University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Computer Science and Engineering: Theses, Dissertations, and Student Research

Computer Science and Engineering, Department of

12-2016

The Effect of Frequency Resolution on Intelligibility Sentence and its Relevance to Cochlear Implant Design

Seth H. Roy
University of Nebraska-Lincoln, sethhroy@gmail.com

Follow this and additional works at: http://digitalcommons.unl.edu/computerscidiss

Part of the Computer Engineering Commons, Speech and Hearing Science Commons, and the Speech Pathology and Audiology Commons

Roy, Seth H., "The Effect of Frequency Resolution on Intelligibility Sentence and its Relevance to Cochlear Implant Design" (2016). *Computer Science and Engineering: Theses, Dissertations, and Student Research*. 122.
http://digitalcommons.unl.edu/computerscidiss/122

This Article is brought to you for free and open access by the Computer Science and Engineering, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Computer Science and Engineering: Theses, Dissertations, and Student Research by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

المنارة للاستشارات

www.manaraa.com

THE EFFECT OF FREQUENCY RESOLUTION ON INTELLIGIBILITY SENTENCE AND ITS RELEVANCE TO COCHLEAR IMPLANT DESIGN

By

Seth H. Roy

A THESIS

Present to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Computer Science

Under the Supervision of Professors Thomas Carrell and Ashok Samal

Lincoln, Nebraska

December, 2016

www.manaraa.com

THE EFFECT OF FREQUENCY RESOLUTION ON INTELLIGIBILITY SENTENCE AND ITS RELEVANCE TO COCHLEAR IMPLANT DESIGN

Seth Roy, M.S.

University of Nebraska, 2016

Advisor: Thomas Carrell and Ashok Samal

The purpose of this study is to understand how electrical stimulation (as opposed to acoustical stimulation) of the auditory nerve is used in cochlear implants. Speech is a complex signal that changes rapidly in time and frequency domains. Since phonemes (the smallest unit of speech that distinguishes words) depend on nuanced differences in frequency patterns, it would be expected that a signal with drastically reduced frequency information would be of limited value for conveying speech. Such a frequency-poor signal is the object to be investigated in the present work. It is also the basis of the way speech is represented in cochlear implants. How could sound in which most frequency information has been discarded by successfully used by so many thousands of individuals? There must be more information in the signal such as timing and amplitude that are important for the speech signal. In addition, semantic context and visual information play a significant role in speech intelligibility. It is the goal of this thesis is to examine how this information aggregates into the perception of speech signals limited by poor frequency resolution, such as cochlear implants.

To accomplish this goal, sentence lists were created with systematically varying levels of frequency resolution. Normally, hearing listeners were asked to identify the last word of each sentence presented to them at the different levels of frequency resolution. To examine the effect of context, half of the sentences ended with predictable words and half ended with unpredictable

المنارة للاستشارات

www.manaraa.com

words. The intelligibility of predictable and unpredictable words was compared at six different frequency resolutions. For this study, we used the standard R-SPIN sentences because each list was constructed to be equally intelligible with each of the other lists. The overall pattern of results showed that there were large effects of predictability and frequency resolution. There was an interaction between these two main effects that will be discussed below.

المنارة للاستشارات

www.manaraa.com

Acknowledgements

I would like to express my special gratitude to my advisers, Dr. Thomas Carrell and Dr. Ashok Samal, for their guidance, wisdom and support throughout my graduate school. I also wish to thank Bahar Shahsavarani for her help throughout this research.

I would like to thank specially Dr. Massimiliano Pierobon for his time on Masters' committee and for his comments and critique to improve this research work.

Lastly, I would like to thank my parents and my sister for being totally supportive of me throughout this journey.

www.manaraa.com

Contents

المنارة للاستشارات

www.manaraa.com

المنارة للاستشارات

www.manaraa.com

List of Figures

www.manaraa.com

List of Tables

المنارة للاستشارات

www.manaraa.com

Chapter 1

Introduction

1.1 Background

Graeme Clark developed a design for a cochlear implant in 1968. The first prototype was developed during 1970s. Clark's father was deaf and he was very inspired to come up with a device that could help his father to better understand words spoken to him. Clark is currently in Australia working for Cochlear, one of the major cochlear implant manufacturers in the world. The FDA approved Cochlear implant in 1980s [2].

According to some estimates, there are about 368 million people around the world who are deaf and hard of hearing which makes up about 5% of the world's population (World Health Organization). In the United States, it was reported that there are 34.25 million people with hearing loss. Out of that total, there are about 1,165,000 people in the US have severe-to-profound hearing loss [3]. Currently, there are about 324,200 individuals worldwide who are fitted with cochlear implants. The reason why the percentage is so low is that its cost is prohibitive. However, as of 2012, there are only about 96,000 people have cochlear implants in US [8].

Cochlear implants provide electrical stimulation (as opposed to acoustic) of the auditory nerve. Unlike a hearing aid, a cochlear implant circumvents damage to the cochlea and does not amplify the sound. Simple amplification is not effective for many individuals with hearing loss

www.manaraa.com

due to the different types of hearing loss. The three major categories of hearing loss are: conductive hearing loss, sensorineural and mixed hearing loss. Conductive hearing loss is caused by problems with the ear canal or the ear bones (malleus, incus and stapes). Sensorineural hearing loss is caused by nerve damage in the inner ear. Mixed hearing loss is caused by damage in the middle and the inner ear (cochlea) or the auditory nerve.

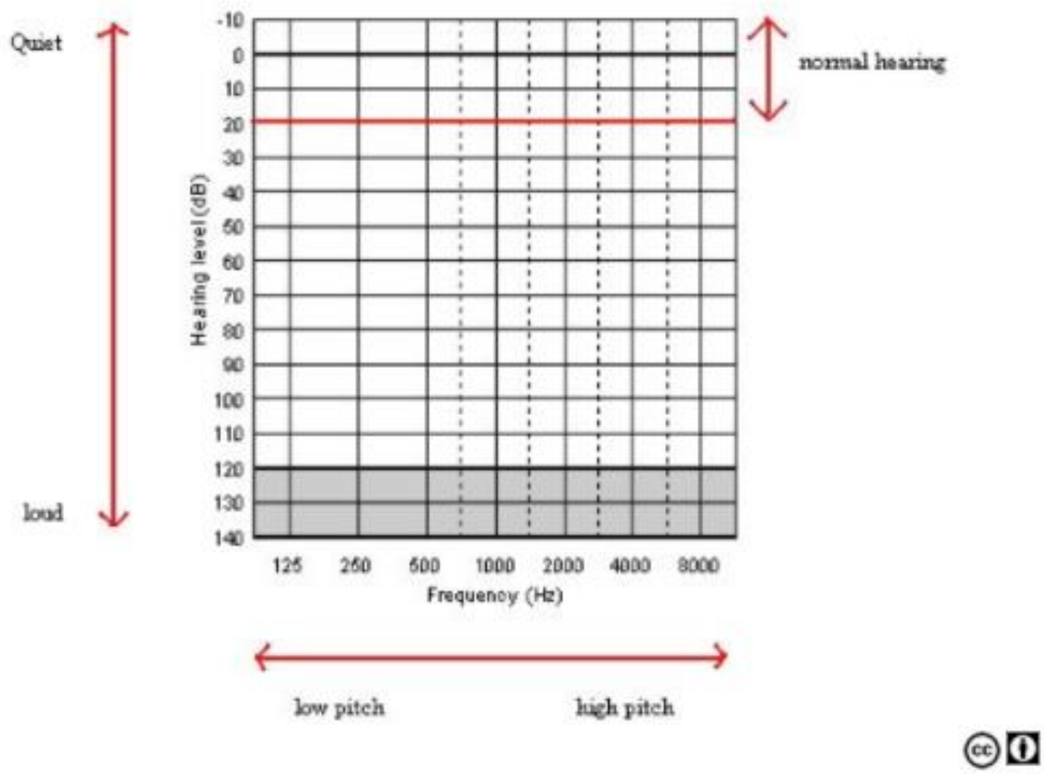

Figure 1: An audiogram is a graphical representation of an individual's hearing threshold. Note that smaller numbers represent better hearing.

Figure 1 shows a blank audiogram that is used to evaluate a patient's hearing threshold. The vertical axis represents the volume or loudness and is measured in decibels. Zero dB at 1000 hertz was set to be approximately the quietest sound which a healthy listener can hear. The thresholds at the remaining frequencies were derived from extensive human perceptual experimentation and have been codified in the ANSI-1969 standard. Note that the level required to reach the hearing

www.manaraa.com

threshold becomes louder from top to bottom as shown in Figure 1. The horizontal axis represents the frequency and is measured in hertz. The lowest pitch tested is at 125Hz while the highest is at 8000Hz. The average normal hearing ranges from -10 to 20 dB as shown in Table 1. A person with mild hearing loss (21 dB to 40 dB) would often have trouble understanding speech in noisy situations. If a person who has moderate hearing loss (41 dB to 70 dB), the person would be eligible to have amplification (or hearing aids). The person with this hearing loss would rely on speech-reading and facial gesture. A person who has a severe hearing loss (71 dB to 90 dB) will have very poor speech quality and often have a hard time hearing general forms of noise such as concerts or traffic. Hearing aids are suitable for a person with a moderate hearing loss. Finally, a person who has a profound hearing loss (> 91 dB) will be unable to hear almost any sound. Amplification may be useful to a limited degree but the person must rely on good communication tactics like lip-reading, signing and using subtitles/closed captioning.

Table 1: Frequency in Hertz (Hz) Table showing the sensitivity range for different listeners

	Frequency in Hertz (Hz) 125 Hz – 8000 Hz
-10 dB - 20 dB	Normal
21 dB - 40 dB	Mild
41 dB – 70 dB	Moderate
71 dB– 90 dB	Severe
91+ dB	Profound

To be eligible to receive a cochlear implant, the person must have a severe (71-90 dB) or profound hearing loss (+91 dB). The cause of deafness is often unknown; hearing parents can have a deaf

المنارة للاستشارات

www.manaraa.com

child as in my own case or deaf parents can either have a normal hearing child or a deaf child. In addition, hearing loss can be due to ear infections during infancy.

The three major cochlear implant companies are Cochlear, Med-El, and Advanced Bionics. Each of the cochlear implant companies has different configurations for electrodes. During the surgical procedure, the medical team makes a small incision behind the patient's ear. Once the incision is complete, the surgeon creates an air pocket through the skull that will expose the inner structure of the ear to allow an electrode array to be inserted. The electrode array can be in the range from 1mm to 1.5mm in length depending on the size or the type of electrode. It is composed of conductive metal alloy electrodes separated by flexible insulation. It is important that the electrode array fits inside the cochlea, a snail shaped structure (See Figure 2). Electrodes are placed to "tune" electrode firing to appropriate frequency representation on the basilar membrane, a structure in the inner cochlea. The inner cochlea contains between 17,000 to 24,000 hair cells that are responsible for hearing. Cochlear implant requires a very delicate surgery where the ear canal and the ear drum cannot be disturbed during the procedure.

The receiver/stimulator is implanted underneath the skin. Normally the length of stimulator ranges from 20 – 30 mm. The receiver/stimulator is the part of the implant that sits on the side of the skull. This is called the bedding preparation. It takes from two to four weeks for the cochlear implant user to recover from the surgery. Then the user typically undergoes a hearing rehabilitation process with an audiologist, often called mapping. Once the stimulator and the transmitter are activated, the stimulator receives signals from the processor and converts them into electrical

المنارة للاستشارات
www.manaraa.com

impulses. Electrode firing locations are based on sounds picked up by an external microphone. Figure 2 shows the detailed layout of the cochlear implant sitting inside the patient's ear.

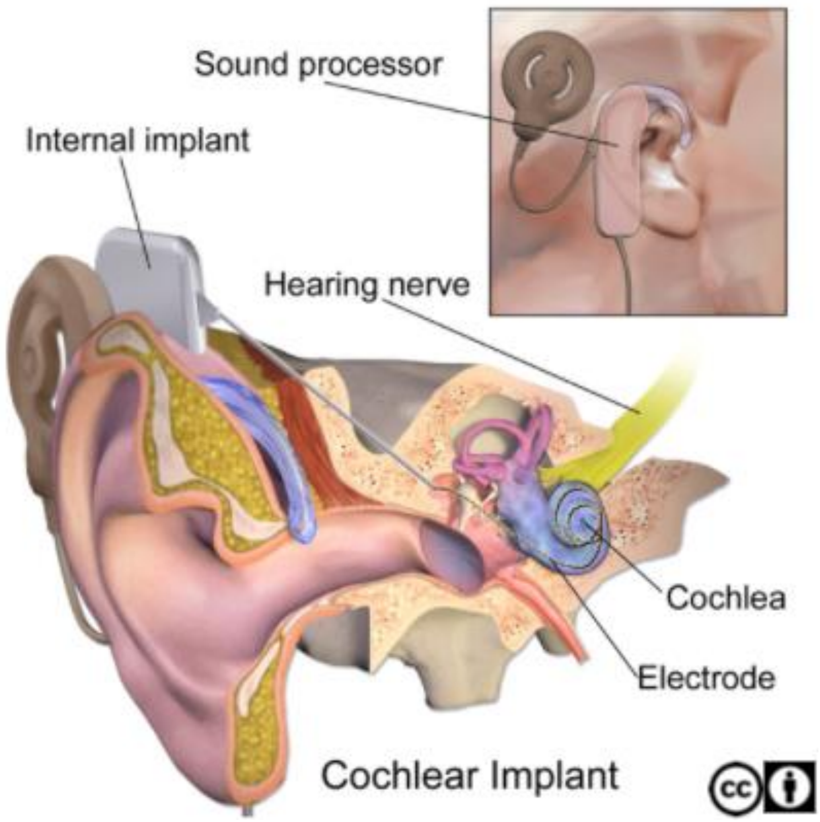

Figure 2: Structure of the inner ear. (Blausen, 2014)

المنارة للاستشارات
www.manaraa.com

The design of cochlear implants is still evolving. Their use is also on the rise and making significant impact in the deaf and hard of hearing world.

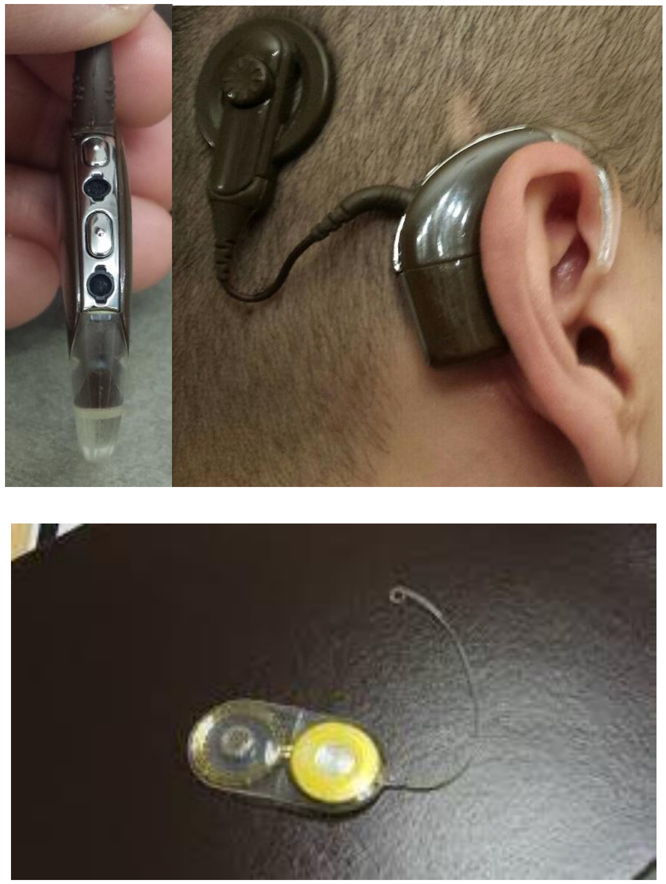

Figure 3: A Cochlear Implant on a User

Cochlear implants divide the sounds into channels and eventually drive electrodes. Each cochlear implant device has a different number of channels. Figure 3 shows a close-up of Nucleus 5 cochlear implant (manufactured by Cochlear). This cochlear implant has 22 channels. While Cochlear has 22 electrodes, the implant from MED-EL and Advanced Bionics, the other two leading manufacturers, have 12 and 16 channels, respectively. In comparison, the early cochlear implants had only one or two channels.

المنارة للاستشارات

www.manaraa.com

1.2 MOTIVATION

The primary goal of this research is to study the relationship between the precision of frequency information in cochlear implants combined with the usefulness of contextual predictability and mimic the sound aspect of cochlear implants by using low and high predictability sentences for the participants.

The goal of this research was to answer the following questions:

1. What is the contribution of the number of frequency channels for sentence intelligibility?
2. How does context interact with the number of channels? (Earlier research has shown that some CI listeners are not able to use context).
3. Is there a point of diminishing returns as the number of electrodes (frequency bands) is increased? Does context influence this number?

1.3 Related Work

In early 1970s, cochlear implants typically had a single channel with a frequency band of 340 and 2700 Hz. Scientists and medical staff were dubious of the benefits of a single channel device and many believed that it was only generating noise. Over time, it became evident that a single channel was not sufficient to improve the intelligibility of listeners. Amplitude envelope was more important than the original designers had suspected.

In the 1980s, scientists identified the key theoretical questions on reduced channel speech with multiple channels. “How many electrodes should be used? If one channel of stimulation is

www.manaraa.com

not sufficient for speech perception, then how many channels are needed to obtain high levels of speech understanding?" [7].

While Clark was working on his first cochlear implant in Australia, House 3M cochlear implant, designed by William House, was the first cochlear implant approved in the US by the FDA in 1980s (See Figure 4 below). Many cochlear implant companies use band pass filters to divide the incoming signals into various frequency-specific components and deliver these to specific regions of the cochlea. Cochlear implant users have a maximum of 7 or 8 independent spectral channels while any hearing person can maintain between 20 to 30 functional spectral channels [12].

Figure 4: House 3M Cochlear Implant

1.3.1 Signal Processing

Signal processing plays a crucial role in the design and functioning of cochlear implants. Signal processing methods are used to extract critical information from incoming acoustic signals.

المنارة للاستشارات
www.manaraa.com

Figure 5 shows typical processing steps in converting acoustic signals generated by a speaker to sound production.

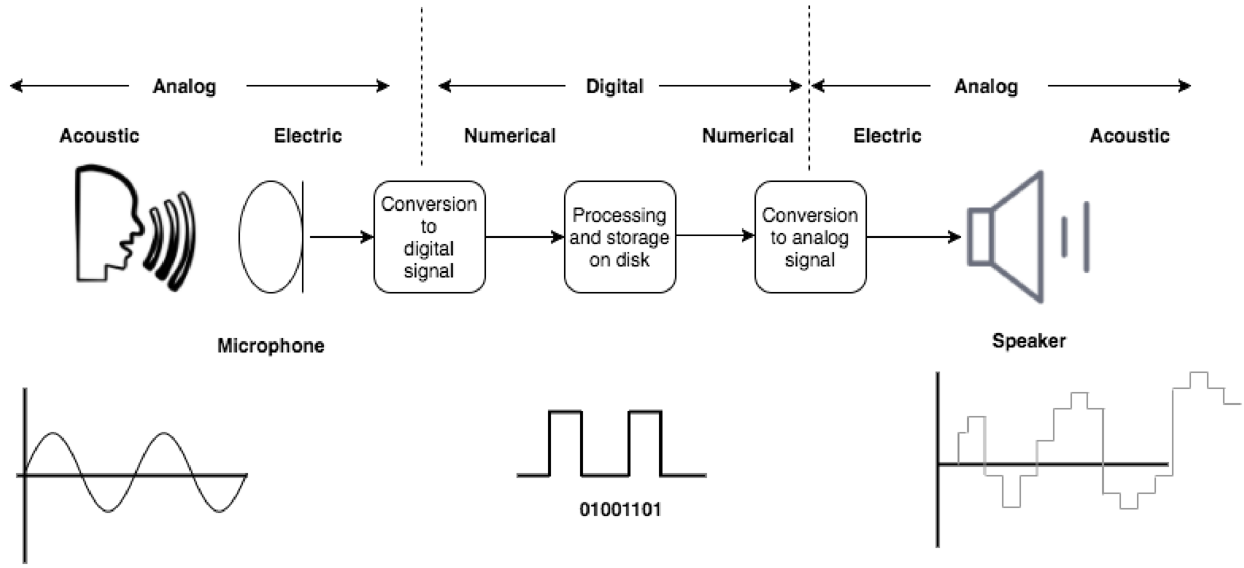

Figure 5: Analog to Digital Signal

When a person speaks through the microphone, the sound waves are converted to electrical waves and subsequently converted to a digital representation. Then the computer can act on the numerical representation for acoustic processing and making decisions. The signal processing algorithms are applied to the digital form of the speech. The digital signal is then converted back to analog format and fed to the speaker for sound production. In a cochlear implant, the analog to digital stages are employed. After processing the numerical representation is converted to the electrical pulses which are sent to an electrode array that has been placed in the cochlea.

1.3.2 Role of Context

Context refers to the surrounding words, phrases and paragraphs that convey the meaning of a word. In later chapters, we will talk about the low and high predictability sentences and how context can be an asset to comprehend the meaning of the word. Cochlear implants users struggle

www.manaraa.com

when it comes with spectral information than normal-hearing listeners [6]. Vowel recognition plays a huge role in speech perception in low and medium frequency noises. Sentence is perceived as a string of related words [12]. Studies have shown that younger and older adults use sentence context to compensate for decreased levels of hearing even though participants with hearing impairment showed better performance when the words were presented in a meaningful context [5].

1.3.3 Number of Channels

Several studies have shown that as the number of channels increases, the intelligibility also improves. “For consonant recognition, the improvement was the largest between one and six or eight channels with smaller improvements for greater number of channels” [6]. Cullington and Zeng controlled the spectral information by varying the number of channels between 1 and 16 while temporal information was controlled by varying the low pass cutoff frequencies of the envelope extractors from 1 to 514 Hz [12].

1.4 Thesis Outline

In Chapter 2, we describe the methodology that is used in this research. We discuss about the significance behind the filtering and the set-up procedure. Chapter 3 presents the results from our experiments. We also explain the statistical significance of the results. Finally, in Chapter 4, we conclude with a summary and directions for future research.

المنارة للاستشارات
www.manaraa.com

Chapter 2

Methodology

2.1 Participants

Twenty-eight listeners participated in this study (27 female and 1 male). Their ages ranged from 19 to 26. The average age was 21.3. Participants were all students at the University of Nebraska. Each participant met the following requirements: they had no hearing disability, they were currently taking no medications that could affect their hearing, and they were speakers of General American English. The participants were divided into four groups. Each group's participants were presented with the same set of sentences, but in different orders to reduce systematic learning effects.

2.2 Stimuli

Seven sentence lists formed the basis of the stimuli used in this experiment. These were taken from the Revised Speech Perception in Noise (SPIN) test [4]. Each list is 50 sentences long and has been normed to be equally intelligible with the other lists. Of the 50 sentences 25 have a high-predictability (HP) final word and 25 have a low-predictability (LP) final word. For example, one HP sentences was, "Kill the bugs with this spray" and one LP sentence was "Betty knew about the nap."

Each of the sentence lists was processed through a set of bandpass filters. For example, SPIN-list 1 was passed through a 2-channel filter, and SPIN-list 2 was passed through a 3-channel

المنارة للاستشارات
www.manaraa.com

filter. Table 2 shows how many frequency bands, or channels, each list was divided into. It also shows the corner frequencies demarking the channel boundaries. Six of the lists were used, those with 2,3,4,5,6, and 8 channels. A seventh list, with 12 bands was also created to familiarize listeners with the sounds and the task before the experiment itself began.

Table 2: Frequency Table

SPIN Sentence List Number	Number of Channels	Corner Frequencies
1	2	150; 1171; 5500
2	3	150; 660; 1997; 5500
3	4	150; 484; 1171; 2586; 5500
4	5	150; 397; 835; 1619; 3014; 5500
5	6	150; 345; 660; 1171; 1997; 3335; 5500
7	8	150; 287; 484; 766; 1171; 1751; 2586; 3783; 5500
6	12	150; 236; 345; 484; 660; 885; 1171; 1535; 1997; 2586; 3335; 4288; 5500

The corner frequencies were selected to correspond to equally spaced intervals along the basilar membrane using the Greenwood function.

$$F = 165.4(10^{.06x} - 1) \qquad \textit{(Equation 1).}$$

Where F is the character's frequency of the sound measured in Hertz and *x* is the distance (measured in mm) from the apex, assuming a basilar-membrane length of 35 mm. Note that no frequencies below 150Hz or above 5500 Hz are represented in this model. A pseudo-code example of the stimuli generation method used to filter sentences is shown in Figure 6. All signals had a sampling rate of 11,025 Hz and had the acoustic range of 150 Hz to 5500 Hz.

The function "generateStimuli" generates a reduced channel signal from a given sentence. Given the number of channels, it first determines the specification of the band pass (lower and upper cutoff frequency) filters for each channel. This is computed by using the Greenwood

www.manaraa.com

Function. Then the method computes the enveloped shaped noise to match the dimension of the sentence. The signal is then passed through each filter, multiplied by the noise and filtered again. Finally, the filtered bands are added together to generate the reduce channel signal.

```
samprate=11025;
%Read and input the sentences from the original folder
for i = 2:2 folder number = 1 to 8
    for j = 1:50 %50 sentences per folder
         if j < 10
             filename = sprintf('SPIN_L%d_S0%d.wav', i,j);
             %  = sprintf('SPIN_L%dS0%d.wav',i,j);
          else
             filename = sprintf('SPIN_L%d_S%d.wav',i ,j);
         end

        [x, samprate] = audioread(filename);
        %x = [sig, samprate];
        % x is the original sentence variable

        randarray = rand(mrows, ncols);

        for jj = 1:mrows
            if randarray(jj) > 0.5
                randarray(jj) = 1;
            else
                randarray(jj) = -1;
            end
        end

        %to normalize the cutoff frequencies
        d1 = fdesign.bandpass('N,F3dB1,F3dB2,Ast', 6, 150/(fs/2), 1171/(fs/2), 60);
        d2 = fdesign.bandpass('N,F3dB1,F3dB2,Ast', 6, 1171/(fs/2), 5500/(fs/2), 60);
        hd1 = design(d1,'ellip');
        hd2 = design(d2,'ellip');
         y1 = filter(hd1, x);   % x is the original signal and y is the filtered signal
        y2 = filter(hd2, x);
        ysum = y1 + y2;

        %Create ESN (Envelope Shape Noise)
        esn_sentence1 = y1 .* randarray;
        esn_sentence2 = y2 .* randarray;
        esn_sigf1 = filter(hd1, esn_sentence1);
        esn_sigf2 = filter(hd2, esn_sentence2);
        %sum
        2-channel-sentence = esn_sigf1 + esn_sigf2;
 end
```

Figure 6:: Stimulus Generation Procedure

المنارة للاستشارات

www.manaraa.com

2.3 Envelope Shaped Noise

After frequency-band filtering, each channel was processed by multiplying the time-changing RMS level of that channel by white noise. The resulting signal is called "Envelope-Shaped Noise" (ESN). When a signal has gone through this process, no frequency information remains, only the RMS level, as it changes over time, remains. If the high- and low-band of a signal are "ESNed" separately and then recombined, the resulting signal is spectrally very poor but the amplitude and timing information remain.

Figure 7 is an example of a sentence that has been bandpass filtered, each filter was ESNed, and the resulting ESNed files were then re-filtered. The original sentence (top panel) shows thousands of frequencies; the resulting sentence shows the original timing and amplitude information but only four frequency ranges are represented. 1). 0 – 800 Hz, 2). 800-1600 Hz, 3). 1600-2800Hz, and 4). 2800-6000 Hz. Thousands of frequencies are represented in natural sentences. At a 4 channel, only 4 frequencies bands are being represented. Envelope shape noise multiplies the amplitude by white noise. How does it work? It creates a set of random numbers. Every single sample is get multiplied by 1 or -1.

www.manaraa.com

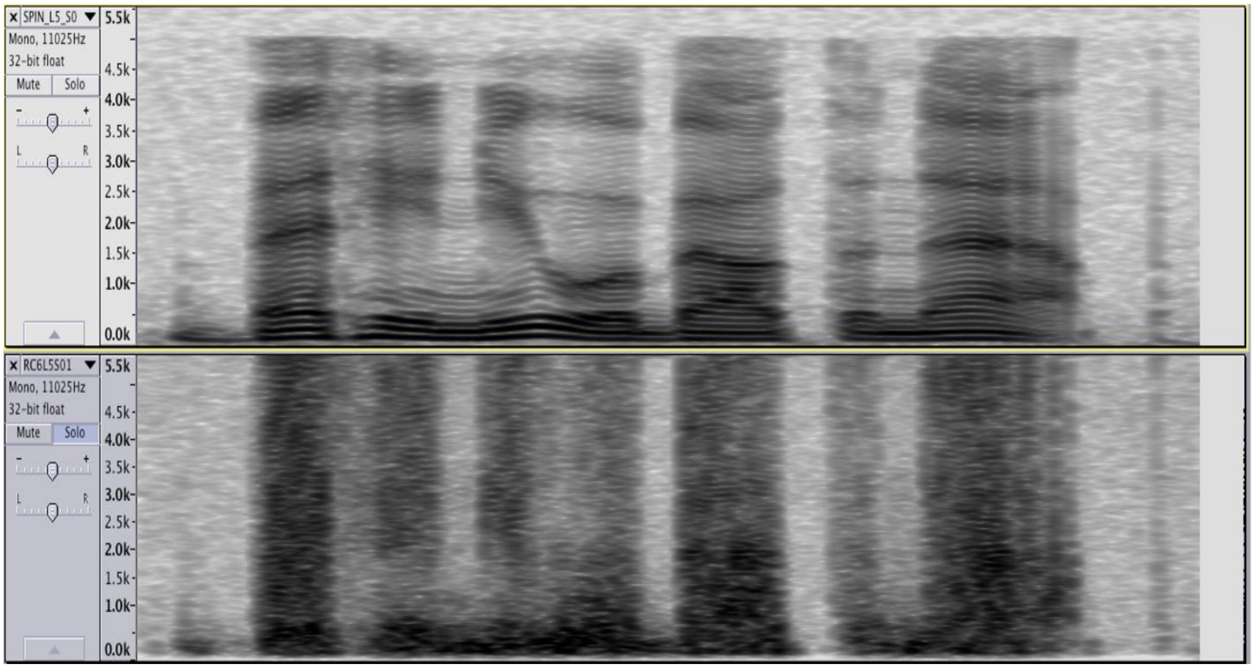

Figure 7: Spectrogram of original and 4-band reduced-channel sentence. Top Spectrogram is the original sentence while the bottom spectrogram represents the filtered and ESNed speech.

The top spectrogram in Figure 7 shows the original sentence while the bottom spectrogram shows the filtered or RC sentence. Spectrogram is a photographic or visual way of representing loudness of a signal over a period. Figure 8 is a graphical summary of the entire process of generating a vocoded sentence from a natural sentence.

www.manaraa.com

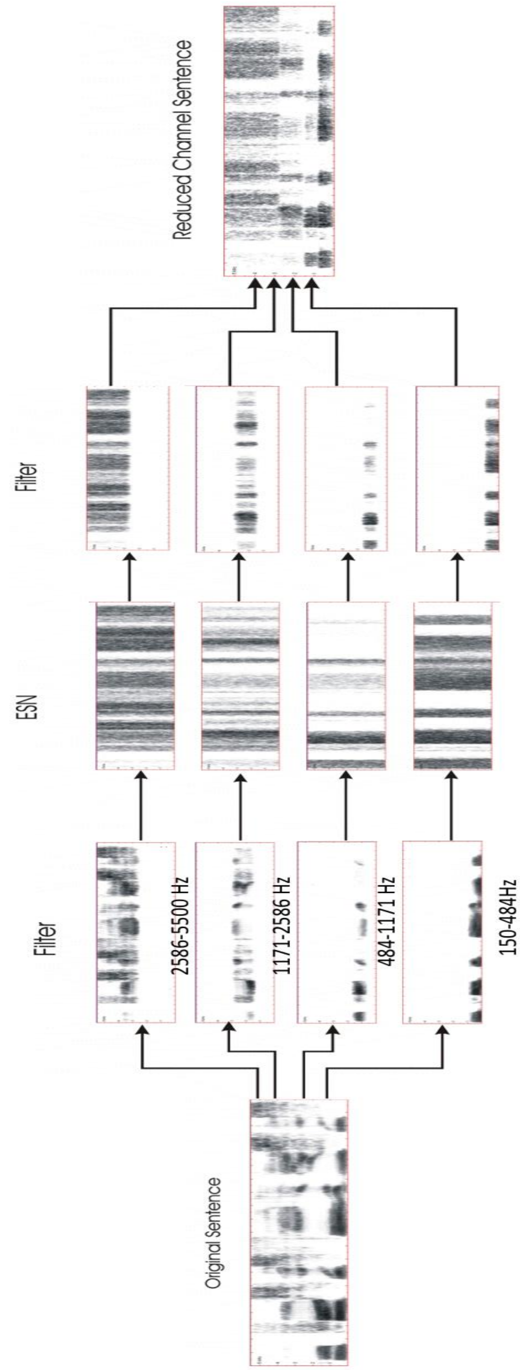

Figure 8: Example of ESN Diagram

المنارة للاستشارات
www.manaraa.com

2.4 Procedure

The listeners were tested in groups of one to four individuals in a test suite at the Speech Perception Laboratory at University of Nebraska-Lincoln. All the stimuli were equilibrated internally to have the same RMS level as one another with a sampling rate of 11,025 Hz. The sentences were presented binaurally via headphones with a maximum loudness of approximately 68 dB. Custom software created at the University of Nebraska controlled the timing and sequencing of the sentences.

Table 3: Stimulus order for each subject group

List Number:	F	1	2	3	4	5	6
	Number of Frequency Bands						
Group 1	12	2	5	3	6	4	8
Group 2	12	8	4	6	3	5	2
Group 3	12	2	3	4	5	6	8
Group 4	12	8	6	5	4	3	2

Four groups of listeners were presented with the word lists in four different orders. The groups were counterbalanced in a modified Latin Square design to avoid a list-by-frequency confound. Participants were familiarized with the task with the first list (labeled “F” above). Then they were presented with the sentence lists in the order shown in Table 3. There were 50 familiarization sentences and 300 experimental sentences. The listeners wrote the final word (or their best guess) during an inter-trial interval of 3 seconds. The entire experiment, from greeting to departure took participants under 40 minutes. The participants were seated at individual listening stations in a sound treated testing room. The ambient room noise level was 28 dB SPL.

المنارة للاستشارات

www.manaraa.com

The stimuli were presented via Sennheiser HD 280 PRO closed-ear circumaural headphones. The experiment was controlled and its progress monitored in a second room.

www.manaraa.com

Chapter 3

Results

3.1 Analysis

Analyses were conducted to determine how predictability and frequency resolution influenced word intelligibility. Descriptive statistics and 2 x 6 repeated-measures ANOVAs were used to evaluate the relationships.

3.1.1 Words Correct

The initial analysis showed that increasing the frequency content in the signal improved the intelligibility of the final (key) word in the sentences. Specifically, as the number of channels increased, the proportion correct word responses increased. In addition, predictable words were identified more accurately than unpredictable words. These results are shown graphically in Figure 8 and in numerically in Table 4. Visual inspection indicates that any interaction between number of channels and predictability was likely due to the two-channel condition in which both predictable and unpredictable words were correctly identified less than one percent of the time.

المنارة للاستشارات

www.manaraa.com

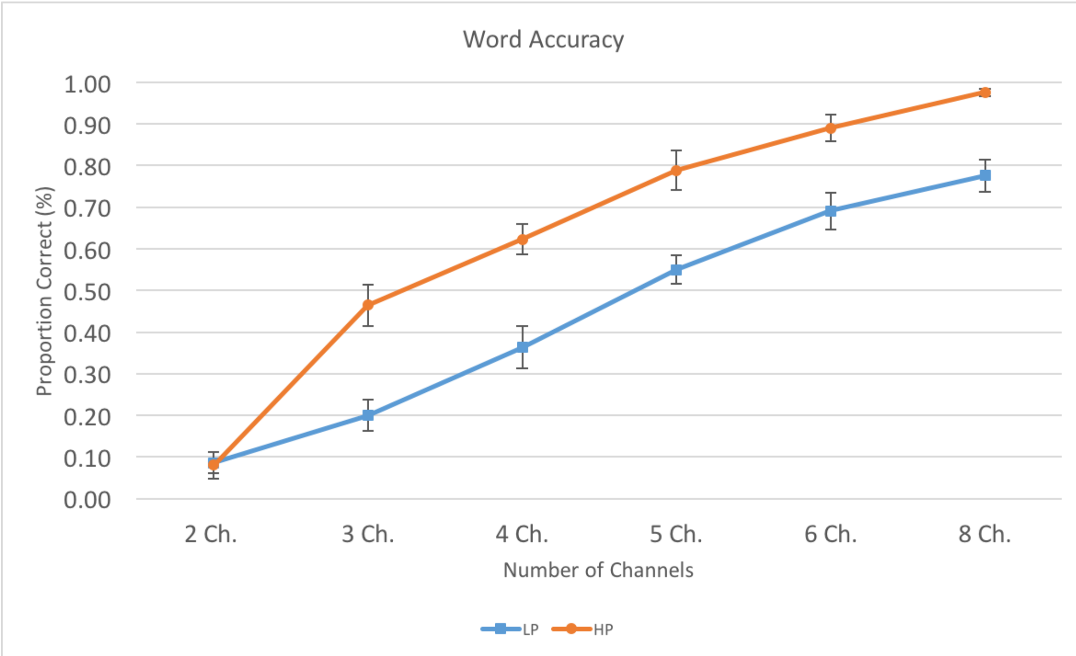

Figure 9: Word Accuracy

In Figure 8, Proportion of sentence-final words perceived correctly based on number of frequency channels and word predictability. The error bars indicate 95% confidence intervals.

Table 4: Mean Proportion Correct (N=28)

	Number of Frequency Channels					
Predictability	2	3	4	5	6	8
Low	0.09	0.20	0.36	0.55	0.69	0.78
High	0.08	0.46	0.62	0.79	0.89	0.98

3.1.2 ANOVA Results

A 2 x 6 (predictability-by-frequency) repeated-measures ANOVA was performed on the accuracy data to determine whether the effects that were visible in the graph would also be statistically meaningful. The results shown in Table 5 show that both main effects and the interaction are highly significant and greatly exceeded the $p < 0.01$ alpha level set for this experiment. Main effect 1: Words with higher predictability are more intelligible than those with lower. Main effect 2: Words with more channels of frequency information (i.e., greater frequency

www.manaraa.com

resolution) are more intelligible than those with fewer channels. Interaction: The predictability-based intelligibility improvement is not present at every number of frequency channels. Inspection of Figure 9 indicates that the interaction is due to the intelligibility of the lowest-quality signals, the two-channel stimuli.

Table 5: Two Way ANOVA

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Predictability	37316.42	1	37316.42	270.72	1.22E-44	3.87
NBands	300298.69	5	60059.74	435.72	1.88E-141	2.24
PredX NBands	8956.80	5	1791.36	13.00	1.55E-11	2.24
Error	44659.99	324	137.84			
Total	391231.91	335				

The purpose of the ANOVA is to assess whether observed differences among sample means are statistically significant. Based on the p-values, they clearly were in the present investigation. The smaller the p-value, the stronger the evidence against the null hypothesis and in favor of the alternative hypothesis. ANOVA uses F statistics to calculate p-value to evaluate the null hypothesis.

3.1.3 Phonemes Correct

The phonetic analysis was more fine-grained that the word analysis. For example in the sentence, “The scarf was made of shiny silk.”. When a listener wrote “shuck” instead of “silk” they would receive a score of 0 out of 1 (0%) based on words correct but 1 out of 4 (25%) based on phonemes correct. Table 6 shows the proportion of correct phonemes for the listeners. Note that the same general pattern of results was found for word and phoneme intelligibility.

المنارة للاستشارات
www.manaraa.com

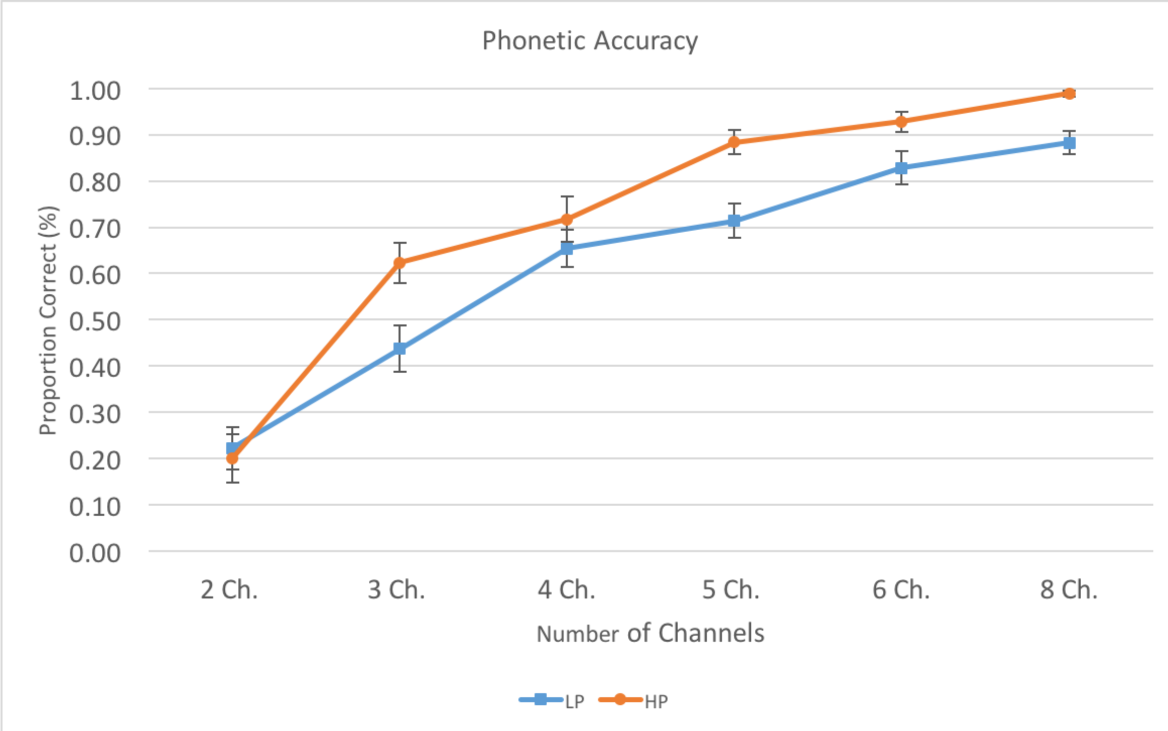

Figure 10: Phonetic Accuracy

Table 6: Proportions Correct for Phonetics

	Number of Frequency Channels					
Predictability	2	3	4	5	6	8
Low	0.22	0.44	0.65	0.71	0.83	0.88
High	0.20	0.62	0.72	0.88	0.93	0.99

3.1.4 Rau Transformation

It is possible that the interaction effect is due to a “floor effect” from the poorest quality stimuli. This is a common problem with proportion correct data and the “rau” arcsine transformation is often used to “rationalize” the nonlinearities of proportion correct data near 0 and 100 percent. There are three types of effects in transformation 1). Normalizing the distribution

المنارة للاستشارات
www.manaraa.com

of the data, 2). Producing a scale in which the real effects are linear and additive and 3). Providing a mean value that is a true estimate of the mean level of the measurements [10].

$$T = \arcsin\sqrt{X/(N+1)} + \arcsin\sqrt{(X+1)/(N+1)} \qquad \textit{(Equation 2)}$$

$$R = 1.46\,(31.83098861T - 50) + 50 \qquad \textit{(Equation 3)}$$

R is the score in raus, where T is the transform of arcsine from the above Equation 3.

All of the scores for the second ANOVA have been transformed according to Equation 3.

3.1.5 One-Way ANOVA Results

The observations for one way ANOVA have been computed by taking the difference of the transformed High Predictability and Low Predictability data (Table 7).

Table 7: One-Way ANOVA

SUMMARY				
Groups	*Count*	*Sum*	*Average*	*Variance*
2 Ch.	28	-45.52	-1.63	194.32
3 Ch.	28	720.61	25.74	111.51
4 Ch.	28	701.84	25.07	345.94
5 Ch.	28	671.43	23.98	260.23
6 Ch.	28	666.79	23.81	130.09
8 Ch.	28	825.8	29.49	116.7

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
NBands	17913.6	5	3582.72	18.55074525	1.51E-14	2.27
Error	31287.18	162	193.13			
Total	49200.78	167				

The results of the ANOVA indicated that even when the data had been rau-transformed the effect of context varied across different number of frequency channels in the stimuli.

المنارة للاستشارات
www.manaraa.com

In other results, this difference of HP and LP Graph shows that the correctness data that there is context of High Predictability and Low Predictability sentences among the listeners (Figure 10).

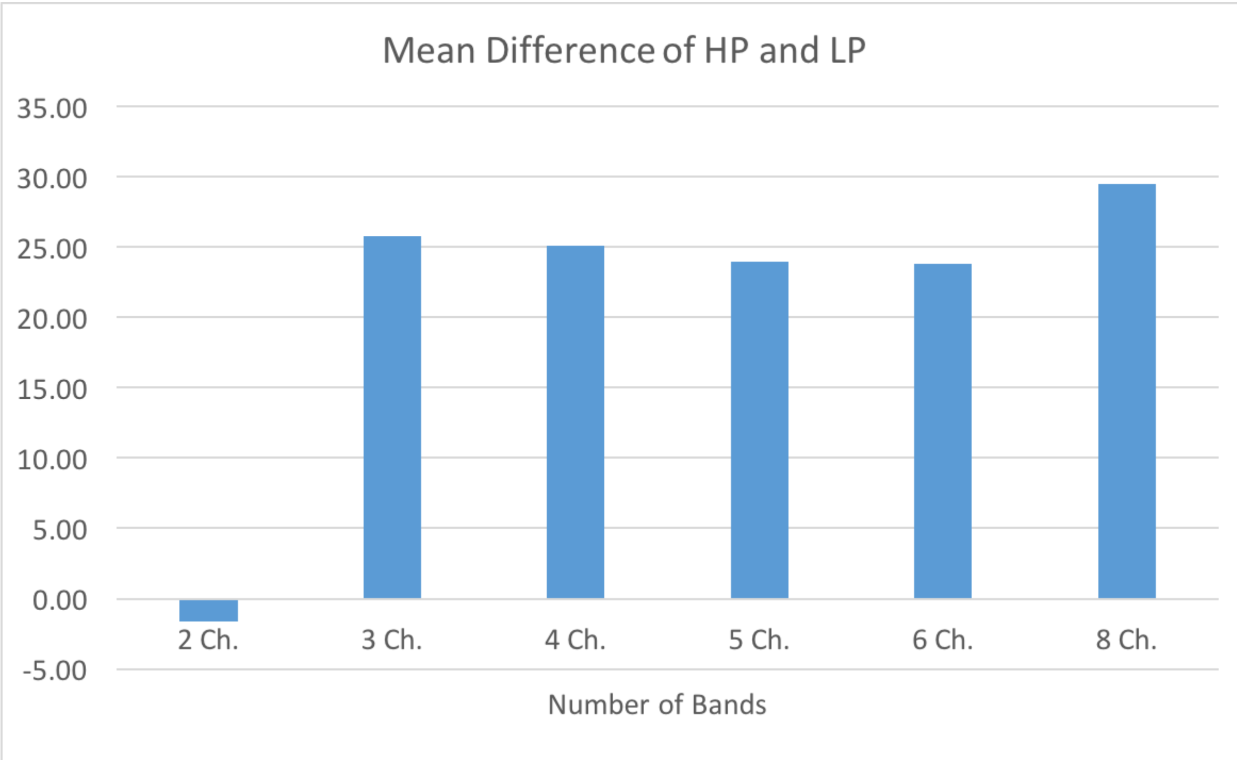

Figure 11: The Mean Difference of High and Low Predictability.

www.manaraa.com

Chapter 4

Summary and Future Work

4.1 Summary

In this thesis, we examined effect of frequency resolution on speech intelligibility. A set of standard R-SPIN sentences were used to generate the stimuli presented to a set of listeners. Using a set of band pass filters distorted versions of acoustic signals were produced. After initial experimentation, we presented the listeners sentences with 6 different channels ranging from 2-8. To examine the effect of context, the subjects were presented with high and low predictability words at the end of each sentence. The results of the study show that sentence intelligibility is significant different for different number of channels. The intelligibility increases as the number of channels increases. Furthermore, context plays a significant role in most levels of distortion. The results of the study can be used to inform the design of cochlear implant devices.

4.2 Future Work

Our work can be extended in a number of different ways. Instead of using last word of each sentence, the participants can record whole sentence. This will be more challenging and more work to the participants, but it can provide additional insight into intelligibility. The experiment has generated a large volume of data about the phonemes. We could mine the data to obtain interesting patterns that may prove useful in understanding the effect of reduced channel speech. Clustering and association analysis will be particularly useful in this regard.

المنارة للاستشارات
www.manaraa.com

Appendix A

المنارة للاستشارات
www.manaraa.com

A.1 Demographic Questionnaire

Demographic Questionnaire v2

Subject Number: ______________________

Date: ______________________________

Age: __________________________

I am 19 years of age or older (circle one): YES NO

Gender (circle one)

Female Male Other (or no response)

Handedness (circle one)

Right-handed Left-handed Both

Major (if known): ____

Please describe your language background:

First language spoken: ____

Birth place (e.g., state or country): _______________________________

____

Years of musical training:_______________________________

Other languages spoken: ____

____

Have you ever had a hearing problem? If so, please describe briefly: ____________________

____

____

Have you taken NSAIDs (e.g., aspirin) or antibiotics (e.g., erythromycin) within the past 24 hours?

YES NO

www.manaraa.com

A.2 Instructions

Nebraska UNIVERSITY OF Lincoln

WRITTEN INSTRUCTIONS

DEPARTMENT OF SPECIAL EDUCATION AND COMMUNICATION DISORDER

Experiment Identification: The Effect of Frequency Resolution on Intelligibility Sentence and its Relevance to Cochlear Implant Design.

Welcome to the speech perception lab! Today you will hear six conditions. We have made these sentences very difficult to hear on purpose so that we can better understand how the brain processes speech. There will be two phases of this experiment. The first will be the familiarization phase, where we will play the different types of sentences and give you the answers. The second phase will be the experimental phase, during which we'll play the sentences for you and you will write down the last word of each sentence without our giving you the answer.

Directions for the Familiarization Phase

This part of the experiment will give you a chance to become familiar with the types of sentences you'll be listening to. You do not need to write anything down or press any buttons. You will hear each sentence three times. As you listen, try to decide what you're hearing. After this, the sentence will be shown to you on your computer screen and you will hear it three more times. This process will be repeated for two sentences for each of the four types of sentences, for a total of eight sentences.

Do you have any questions?

Directions for Experimental Task

For this experiment, you will hear a total of 300 sentences. These sentences will be in groups of 50, with short breaks between each group. Before each new sentence is played, you will hear a warning tone directing your attention to carefully listen to what is to follow. You will hear each sentence one time. On the answer form, you will write the last word of the sentence (or your best guess). Please write down anything you heard, even if it sounds silly or the word doesn't necessarily make sense. For example, if you think you hear something that sounds like "ubee", go ahead and write it down just how it sounds to you. After you have written your response, wait for the tone to signal the next sentences.

After 50 sentences, you will have a brief break during which you will receive new response forms for the next group of sentences. We will repeat this experimental task for four different types of sentences.

Some General Tips

- We're not worried about spelling—just do your best. Spell it well enough that we know the word you meant.
- If the word sounds like a word that doesn't exist, you can write down what you heard (for example, if you think you hear something that sounds like "pemurs", just write it down). You should know that all the sentences are comprised of familiar English words.
- Listen to each sentence carefully before you decide what to write.

Do you have any questions?

المنارة للاستشارات

www.manaraa.com

Bibliography

[1] Audacity Team (2016). Audacity (Version 2.1.1) [Computer Program]. Retrieved December 3, 2016, from http://audacityteam.org/

[2] Australian Academy of Science (2016). Professor Graeme Clark, otolaryngologis [Online]. Available: https://www.science.org.au/learning/general-audience/history/interviews-australian-scientists/professor-graeme-clark. Accessed: Dec. 4, 2016.

[3] Blanchfield, B.B., Feldman, J.J., Dunbar, J.L., Gardner E.N. (2001). "The Severely to Profoundly Hearing-Impaired Population in the United States: Prevalence Estimates and Demographics," *Journal of American Academy of Audiology*, 12:183-189.

[4] 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 Researc*h, 27:32-48.

[5] Conway, C.M., Deocampo, J. A., Walk, A. M., Anaya, E. M., and Pisoni, D.B. (2014). "Deaf Children with Cochlear Implants Do Not Appear To Use Sentence Context to Help Recognize Spoken Words," *Journal of Speech, Language, and Hearing Research*, 57(6): 2174-2190.

[6] Cullington, H.E., and Zeng, F.-G. (2008). "Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects." *The Journal of the Acoustical Society of America*, 123:450-461.

[7] Laizou, P.C. (1999). "Signal-processing Techniques for Cochlear Implants." *IEEE Engineering in Medicine and Biology Magazine* 18:34-46.

[8] National Institutes of Health (2016). Health information [Online]. Available: https://www.nidcd.nih.gov/health/cochlear-implant. Accessed: Dec. 4, 2016.

[9] Rosen, S., and Howell, P. (1991). Signals and Systems for Speech and Hearing. Academic Press, London, UK.

[10] Studebaker, G. A. (1985). "A Rationalized Arcsine Transform," *Journal of Speech, Language, and Hearing Research*, 28(3): 455-462.

[11] World Heath Organization (2016). Deafness and hearing loss [Online]. Available: http://www.who.int/mediacentre/factsheets/fs300/en/. Accessed: Dec. 4, 2016.

المنارة للاستشارات

www.manaraa.com

[12] Xu, L., Thompson, C.S., and Pfingst, B.E. (2005). "Relative Contributions of Spectral and temporal cues for phoneme recognition. *The Journal of the Acoustical Society of America*, 117(5): 3255-3267.

[13] Miller, G.A., Heise, G.A., Lichten, W, (1950). "The Intelligibility of speech as a function of the context of the test materials."

للاستشارات

www.manaraa.com

www.manaraa.com