

**EFFECT OF CONTEXT ON BRAINSTEM ENCODING OF
SPEECH**

Registration No: 09AUD014

A Dissertation Submitted in Part Fulfillment for the Degree of
Master of Science (Audiology)
University of Mysore, Mysore

ALL INDIA INSTITUTE OF SPEECH AND HEARING

MANASAGANGOTHRI

MYSORE – 570006

June 2011

CERTIFICATE

This is to certify that this dissertation entitled '**Effect of Context on Brainstem Encoding of Speech**' is the bonafide work submitted in part fulfillment for the Degree of Master of Science (Audiology) of the student with Registration No. : 09AUD014. This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore

June, 2011

Dr. S.R. Savithri

Director

All India Institute of Speech and Hearing,

Manasagangothri, Mysore-570006

CERTIFICATE

This is to certify that the dissertation entitled '**Effect of Context on Brainstem Encoding of Speech**' has been prepared under my supervision and guidance. It is also certified that this has not been submitted earlier in any other university for the award of any Diploma or Degree.

Mysore
June, 2011

Dr. Sandeep M.
Guide
Lecturer in Audiology,
Department of Audiology,
All India Institute of Speech and Hearing,
Manasagangothri, Mysore - 570006


DECLARATION

This is to certify that this Master's dissertation entitled '**Effect of Context on Brainstem Encoding of Speech**' is the result of my own study under the guidance of Dr. Sandeep M, Lecturer in Audiology, Department of Audiology, All India Institute of Speech and Hearing, Mysore, and has not been submitted in any other University for the award of any Diploma or Degree.

Mysore

Registration No. **09AUD014**

June, 2011



DEDICATED
TO MY DEAR
MAMMA,
DADDA,
DEENA AND
RIZA

*Words can never thank you enough
For all your love and everything you have done for me.
I want you to know that I love you both
More than words can say,
And I want to thank you for turning me into
The person I am today.
I couldn't have asked Allah
For better parents than you,
so **thank you, Mamma and Dadda**
For everything you do!!!*

*"Being a sister is not a big thing,
It's a million little things."
Thank you My dearest **Deena**
for the million little things you've done for me
And for those you continue to do each day.
Thank you so so much!
I love you!!!*

*"Once in a long while,
someone special walks into your life
and really makes a difference.
They take the time
to show you in so many little ways
that you matter.
They see and hear the worst in you,
but don't walk away;
in fact they may care more about you.
Their heart breaks with yours,
their tears fall with yours,
their laughter is shared with yours"
Thank you Riza, for walking into my life and staying!
I am amazed by how much of yourself you give;
not just to me, but to all.
I am inspired by your strength, courage and grace
And touched by your love!!!*

ACKNOWLEDGEMENT

I thank and praise Allah for all the kindness and mercy showered upon me. This dissertation could not have been composed and would not now be completed but for the efforts of a very large number of people to whom I offer my heartfelt thanks.

*At the onset, I express my pronounced indebtedness to my Guide **Dr. Sandeep M**, for his invaluable assistance, endless inspiration and expert guidance during the course of this study. It gives me great pleasure to express my gratitude for his continuous encouragement, inspiration and support throughout the course of my study. Thank you Sir!!! 😊*

I am grateful to Late Dr. Vijayalaxmi Basavaraj and Dr. S. R. Savithri for permitting me to carry out this dissertation.

*I am thankful to **Dr. P. Manjula**, HOD, Department of Audiology, for permitting me to carry out this study and for providing the facilities to carry out this study.*

*I express my deep sense of gratitude, respect, and heartfelt thanks to **Sujith Sir**, for his timely help in installing all the necessary softwares into my laptop and teaching me how it is to be used. I am also thankful to, **Ms Erika Skoe**, Northwestern University, Chicago, for permitting to use the Brainstem Tool Box software.*

*I am grateful to **Jaykumar Sir**, for helping me generate the stimulus for this study. Thank you Sir, for all your valuable time and advice. I am also thankful to **Prof. Y. V. Geetha**, for permitting me to use the speech science lab.*

*I am thankful to **Hemanth sir and Ganapathy Sir**, for always being ready to open the department on weekends so that we could finish the hurdle of our data collection 😊 and also for their help in analyzing my data.*

*I sincerely thank **Prof. Asha Yathiraj, Dr. Manjula. P, Dr. Rajalaxmi, Dr. Animesh, Dr. Vijay Kumar Narne, Ms. Mamatha, Ms. Devi, Mr. Niraj Singh, Mr. Prawin Kumar, Ms. Geetha, Dr. vasanthalaxmi, Dr. Ajeesh Abraham and Dr. Venkatesh**, for not only enthusiastically teaching us the subject but also inculcating discipline and shaping our professional skills.*

*I am also thankful to **Revathi Maam and Dhanlaxmi Maam, Baba Sir and Ramadevi Maam** for making my postings at LTU and HAT a learning and enjoyable experience 😊*

*This study could never have reached its completion if it wasn't for **my classmates and juniors** who took their time off and became my subjects for this study.*

*I would like to sincerely thank all my teachers at Manipal, for introducing me to the field and molding me into what I am today. Its my immense pleasure to thank **Dr. Rajashekhar**, Dean, Manipal College Of Allied Health Sciences. He is a brilliant, friendly and inspiring teacher. Sir, I am lucky and proud to have been your student.*

I am also thankful to Gopee sir, Shivani Maam, Rama Maam, Kishan Sir, Kanaka Maam, Veena Maam, Aithal Sir, Usha Maam, Santosh Sir, Geehta Maam and Dennis Sir. I sincerely appreciate the effort and hard work you have put into teaching me.

*"The task of an excellent teacher is to stimulate "apparently ordinary" people to unusual effort". I am very happy to thank two of my most favorite teachers, **Peter Sir and Hari Sir**. You both have been inspirational to me. Peter Sir, your class notes which are now a good 6 years old ☺ are still the only books that I use to clear my Diagnostic Audio doubts and the only books I need to open before the viva. Thank you for being such an incredible teacher.*

Hari Sir, I can't thank you enough for all the help and support you have provided me throughout this course. I consider myself lucky to have had you for guiding my CC and the paper for ISHACON. It would have been a disaster if it wasn't for you ☺ I truly appreciate your help for preparing us for the masters entrance at AIISH and also for your valuable inputs while selecting the topic for my dissertation. Thank you Sir!!!

*I would also like to thank all my teachers at **Jabriya Indian School, Kuwait, and MGM PU College, Udupi.***

Now, this is the part I have been most looking forward to. This dissertation is probably the only bound book that I would ever be an author of, considering how ultra of hard working or dedicated that I am☺ I bet, all my friends can vouch by my laziness☺

Here again I would like to thank my guide Sandeep Sir, if it weren't for him I would probably postpone the submission of this dissertation to the next year or the year after. Thank you Sir, for never getting angry and for making the dissertation an enriching and enjoyable experience!

I would also like to **thank all my seniors**, both at AIISH and Manipal, who assured me that no matter what I would finish this dissertation. Thank you **Shreela chichi**, for being so sweet and kind. Thank you **Dhanya maam** for helping me with my references. Thank you **Vijay sir**, for helping me figure out that converting waveforms to ASCII and then FFT is not all that rocket science ☺ Thank you **Gayathri, Disha, Medha, Sandhya** and **sister Reeja**, the first seniors I was ever posted with and how much I've learnt from you all. One of many things that you taught me was that the hair cell isn't actually a cell which produces hair. ☺

Thank you **Sheetal di, Mansi di, Niyathi di, Venessa, Marissa, Deepthi, Sahana, Remya, Sabir, Neish and Ayas** for all the beautiful memories of masthi at the postings ☺

I am thankful to **all my juniors**, again both at Aiish and Manipal, thank you **Mandira, Pooja, Garvitha, Parul, Anushka, and Kutti**... You guys are superb ☺

I was saving the best for the end. It is said by someone real smart, that "life is partly what we make it and partly what is made by the friends we choose". I am fortunate to have the best people on the planet for friends.

I don't know what I'd be without you guys, thank you **Wally, Leann and Velroy**. They say that school friends are forever, and it's so true with you. Thank you my darlings, **Chryselles, Samira and Priya**, I can't find fancy words to tell you how much you guys mean to me, I know I can count on you guys no matter what. Love you so much. Mwah!

Another smart person said that, "a blessed thing for a human is to have a friend, one soul you can trust utterly, who knows the best and worst of us, and who loves us in spite of our all faults". Thank you **Priky** for being there with me always. Thank you **Namita** for being who you are and making me believe I was smart, lol.

Sangu & Lydia, I can always trust you to be just a phone call away. **Bhargu, Ashu, Angel** (My true angels, I love you :*) **Sowmya, Megha, Hanaa, Bala, Mritunjay, Snithin, Divya, Geeta, Amana, Anjali and Kusum**, the time we've spent in college together were some of my best memories. I am so glad I was put in a class with all wonderful people like you!

Swapnaja and Sonu, I've known you for just a past two years, but somehow you guys have turned out to be so very special to me. Swapna, I don't know how I am going to react when I wake up and not be able to drop in your room. Lol. I'll miss you! and Sonu, I don't think I'd have managed to pass all these exams if I didn't have you studying

with me. I just discovered that all drafts look just like Chinese if I don't study them with you. You both are adorable. Remain this way!

***Midulla** (my Eda), you are awesome. I love you!*

***Kavita**, I ve never met someone who loves as much shopping as u do. And never met someone with the same taste for clothes, accessories or food. I ll miss you girl.*

*Life at hostel would not have been this fun if it wasn't for you guys, **Deepthi, Gargi, Maduban, Pallavi, Jyoti, Priyanjli & Hema.***

*Each day at the postings for two continuous years, I ve been lucky to have **Sharanya, Akash and Anisha**, as my posting partners. I am going to miss the breaks ,the sharing of disappointment while staring at so many files in HE and HAT, and studing together just before the test, and running for hours in LTU. I had an awesome time with you guys. Thank you! I would also like to thank **Praveen, Pragathi, Achaiah, Pavan, Vivek, OT, Shruthi, Bhamini, Mukesh, Jyoti, Dhathri, Nike, Mukesh, Sindu and Anju** for making postings and classes fun and cool!*

*Special thanks to **Suma** for being my subject for god knows how many times, if it wasn't for your cooperation I would never have started my data collection. Thank you **Ranjeet and Prabhash** for being there in need. I d like to thank each and every person in my class, for being so amazing and supportive throughout these two years. You guys all **ROCK!!!***

*Last but not the least, I would like to thank my Big and super awesome **FAMILY**. It would take another book on its own if I have to mention each of your names☺ Thank you all for your dua s and advices and love!!!*

TABLE OF CONTENTS

<i>S. No.</i>	<i>Contents</i>	<i>Page No.</i>
1	Chapter 1: Introduction	1-4
2	Chapter 2: Review of Literature	2-25
3	Chapter 3:Method	26-39
4	Chapter 4: Results	40-51
5	Chapter 5: Discussion	52-60
6	Chapter 6 : Summary and Conclusion	61-63
7	References	64-73

LIST OF TABLES

<i>Table No.</i>	<i>Title</i>	<i>Page No.</i>
3.1	Spectral and temporal characteristics of the syllable /ba/	29
3.2	Spectral and temporal characteristics of the syllable /ta/	30
3.3	Mean behavioral thresholds in (dBSPL) of the synthetically generated syllables /da/, /ba/ and /ta/ and, the white noise	32
3.4	Stimulus and acquisition parameters used to elicit the Brainstem Responses	36
4.1	The percentage of occurrence of each wave across four conditions	41
4.2	The mean and standard deviation (SD) of peak latency of wave V and A, and V to A peak to peak amplitude across four test conditions	43
4.3	Pair-wise comparison of mean latency of wave V	44
4.4	Pair-wise comparison of mean latency of wave A	44
4.5	The mean and standard deviation (SD) of peak latency and peak amplitude of waves D, E and F across four test conditions	46
4.6	Pair wise comparison of mean latency of wave E	48
4.7	Pair wise comparison of mean latency of wave F	48
4.8	The mean and standard deviation (SD) of amplitude of fundamental frequency (F0), first formant (F1) and higher harmonics (HF)	49
4.9	Pair wise comparison of mean amplitude of higher harmonics (HF)	50
4.10	The mean and standard deviation (SD) of peak latency and amplitude of wave O across four test conditions	50

LIST OF FIGURES

<i>Figure No.</i>	<i>Legend</i>	<i>Page No.</i>
2.1	Time-amplitude waveform and broadband spectrogram of a 40 ms synthesized speech stimulus /da/	8
2.2	Time amplitude waveform of a 40 ms stimulus /da/ and its corresponding response in the time domain (courtesy Nina Kraus, NWU, Chicago).	9
3.1	Time-amplitude waveform and broadband spectrogram of a 40 ms synthesized speech stimulus /da/.	27
3.2	Time amplitude waveform and spectrogram of (A) synthetically generated syllable /ba/, (B) synthetically generated syllable /ta/ and (C) white noise.	30-31
3.3	Representation of different stimulus conditions used.	35
3.4	Brainstem response elicited by syllable /da/ in a representative subject	38
3.5	Fast Fourier transform analysis of the brainstem response to the stimulus /da/ in a representative subject.	39
4.1	Onset responses elicited by syllable /da/ across 4 conditions in a representative subject.	43
4.2	Sustained response elicited by syllable /da/ across 4 conditions in a representative subject.	46

Chapter 1

INTRODUCTION

Historically, Cortex is agreed upon as the location at which speech specific processing occurs and the brainstem is considered as a part of brain unworthy of interest when it comes to language processing (Krishnan & Gandour, 2009). Conventionally it is believed that the processing operations that occur in the relay nuclei of the brainstem and thalamus is general to all sounds, and speech-specific operations do not begin until the signal reaches the cerebral cortex (Scott & Johnsrude, 2003).

However, in the past decade, there has been an increasing interest in the role of auditory brainstem in speech processing. Experimenters have measured the neural transcription of complex auditory stimuli such as speech; non-invasively from the auditory brainstem (Johnson, Nicol, Zecker & Kraus, 2008; Hornickel, Skoe, Nicol, Zecker & Kraus, 2009; Tzounopoulos & Kraus, 2009). Results of the studies comparing the brainstem responses to speech and non speech stimuli have suggested a possibility of differential processing of speech and non speech stimuli at the level of brainstem (King, Warrier, Hayes & Kraus, 2001; Wible, Nicol & Kraus, 2004). Two models are being debated to account for these results. One is the corticofugal model which states that the brainstem function is modified by the top-down feedback via the corticofugal efferent network (Suga, 2008: Suga, Xiao, Ma & Ji, 2002; Zhang & Suga, 1997). The other model is the local reorganization model, which states that the brainstem function is modified over a long timescale, that is, the brainstem is reorganized to promote the encoding of frequently encountered sounds (Krishnan & Gandour, 2009; Krishnan, Swaminathan & Gandour, 2009).

The local reorganization model has gained support by recent studies which revealed that the auditory brainstem function is malleable as a result of lifelong language experience (Krishnan, Xu, Gandour & Cariani, 2005; Xu, Krishnan & Gandour, 2009) and music experience (Musacchia, Sams, Skoe & Kraus, 2007; Wong, Skoe, Russo, Dees & Kraus, 2007), as well as short-term training (Russo, Nicol, Zecker, Hayes & Kraus, 2005). However, very little is understood regarding the corticofugal model.

1.1 Justification for the Study

Auditory processing at the brainstem is reported to involve an interplay between the sensory and cognitive systems, which is mediated by the feed-forward and feed-back pathways (Tzounopoulos & Kraus, 2009). The massive efferent connections from the cortex to the subcortical structures is reported to be the basis for such feedback-related top-down control (Winer, 2005). Although the functional role of these efferent connections is currently not very well understood, a recent study has hypothesized that corticofugal feedback may provide significant benefits in noisy environments by selectively amplifying relevant information in the signal, and inhibiting irrelevant information at the earliest stages of auditory processing (Luo, Wang, Kashini & Yan, 2008).

The functional significance of these feedback pathways is known to increase with auditory experience, but the exact mechanism by which such plastic changes take place is currently unclear. There is a need to understand, if plasticity occurs as a result of long term experience or whether it is a continuous process. Hence to obtain information

on the extent to which plasticity is operational online, there is a need to study the brainstems sensitivity to ongoing contextual demands.

Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009), studied the effect of context on auditory brainstem responses. In their study, they used a procedure of eliciting and comparing the brainstem responses to /da/ in two conditions. That is, a repetitive condition and a highly variable context condition. In the repetitive condition, 6300 sweeps of /da/ were presented with the probability of 100% and in the variable context condition, 2100 sweeps of /da/ were presented randomly with a probability of 12.5% in context of seven other speech sounds, which differed from /da/ with respect to formant structure, duration, voice onset time and fundamental frequency. The total time taken for one recording of the brainstem responses in both the contexts was approximately 30 minutes. The result of the study showed that there was a significant difference between the brainstem responses elicited in the two contexts. Brainstem representation of features related to voice pitch were enhanced in the repetitive context condition relative to the variable context condition. This result was attributed to the possible corticofugal modulation, where in the cortex continuously modifies the brainstem function. It cannot be attributed to the effect of long term experience, because the test duration was only 30 minutes. The different behavior of the brainstem in the two contexts was inferred to be helpful in perception of speech in noise.

The present study was taken to systematically further understand the influence of stimulus context on brainstem processing. In this study, the difference in the brainstem response elicited to a repeated stimulus was compared to that elicited when the repeated sequence of stimulus presentation was disturbed by interference of another stimulus

(contextual stimuli). Here, both speech as well as noise stimuli served as contexts. This was based on the assumption that the resultant responses may show differential corticofugal modulation as proposed by Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009), if any to speech and noise contexts. Considering noise as irrelevant stimulus, the brainstem may inhibit its effect on the core speech stimulus. In addition, when speech is used as the contextual stimulus, the brainstem may perhaps change its role based on the extent of spectral and temporal similarity of the contextual stimulus to the core stimulus. Studying the brainstem's response in these varied contexts may help in better understanding the complex mechanisms in the brainstem. Hence the present study was taken up.

1.2 Objective of the Study

The sole objective of the study was to study the effect of context on the brainstem responses elicited by stimulus /da/.

Chapter 2

REVIEW OF LITERATURE

Electrical activity evoked by auditory stimuli has been extensively recorded in both humans and animals, in order to assess the functioning of the auditory system (Picton, Hillyard, Krausz & Galambos, 1974). Auditory evoked potentials are classified as early, middle, or late latency responses, based on the time at which the response occurs after the auditory stimulus. Auditory Brainstem Response (ABR), is one of the early responses and it collectively describes the electrical activity originating in the auditory nerve and brainstem (Moller & Jannetta, 1985; Chandrasekaran & Kraus, 2010). They arise within the first 10 milliseconds after the onset of auditory stimulus (Mason et al., 1988) and are the far field recordings of the synchronous electrical activity at the generator site. Although Sohmer and Feinmesser in 1967 first recorded these responses from humans, description of the ABR was given by Jewett and Williston in 1971. Jewett reported that a response elicited using a high intensity click consists of a series of seven waves when recorded from a vertex and ipsilateral mastoid electrode configuration.

The origin of these waves is understood based on the studies done on animal (Achor & Starr, 1980), individuals with brainstem disorders (Stockard & Rossiter, 1977) and comparison of far field and near field recordings (Hashimoto, 1982). Wave I and II are thought to arise from the proximal and distal portions of the auditory nerve respectively; wave III from the cochlear nucleus; wave IV from the superior olivary complex and wave V from the lateral lemniscus. The inferior colliculus is considered as the source of wave VI and VII (Mason, McCormick & Wood, 1988).

2.1 Stimuli used to Elicit ABR

ABR is an onset response. The stimulus to evoke ABR should be fairly abrupt so as to produce a high degree of synchronous firing of the nerve fibers. Thus, click is the most widely used stimuli to elicit ABR. Click is generated by passing a 100 microsecond electrical pulse through a headphone. Spectrally, most of the energy is contained below the first spectral zero. The first spectral zero occurs at $1/\text{duration}$. Hence, the 100 microsecond click has energy below 10 kHz. Since the headphones used clinically begin to roll off at 5 kHz and since hearing evaluation is limited to 8 kHz or below, click evoked ABR has evolved as a popular means to access auditory function in clinical settings.

Click evoked ABR also shows clear pattern of development in early years of life, which include maturation of morphology of the response, decrease in peak latency and increase in the amplitude of the response. Hence, it has emerged as a versatile, objective and reliable tool for the evaluation of hearing and integrity of auditory pathway in newborns, infants and other vulnerable populations (Starr & Don, 1988).

However, click is not an ideal stimulus to assess the neural integrity of the auditory pathway when compared to a complex stimulus such as speech because, Clicks are not good approximates of behaviorally relevant sounds that are encountered outside the laboratory such as speech, music, environmental sounds and non speech vocal sounds (Skoe & Kraus, 2010a). Unlike speech, clicks are short, non-periodic sound containing a broad range of frequencies. Hence, they do not lend themselves to the extraction of information about encoding of stimulus onset, offset, and periodicity. But most

importantly clicks are unable to detect the subtle abnormalities of the auditory pathway, as seen in individuals with auditory processing disorder (Song, Banai, Russo & Kraus, 2006).

Hence, there is a gradual transition in the stimuli being used to evoke ABR, from being simple to the more complex stimuli. Among the complex stimuli, are temporally and spectrally modified syllables (Cunningham, Nicol, King, Zecker & Kraus, 2002), syllables that are synthesized with Mandarin tones (Song, Skoe, Wong & Kraus, 2008) or presented in background noise (Anderson, Chandrasekaran, Skoe & Kraus, 2010), words (Galbraith et al., 2004) and musical notes (Musacchia, Sams, Skoe & Kraus, 2007).

2.2 Speech Evoked ABR

Among these complex stimuli, the most extensively used stimulus is the consonant-vowel (CV) syllable /da/. A Klatt cascade/parallel formant synthesizer (Klatt, 1980) was used to synthesize syllable /da/ of 40 ms at a sampling rate of 10 kHz. The stimulus consists of an onset burst frication at F_3 , F_4 , and F_5 during the first 10 ms, followed by 30 ms F_1 and F_2 transitions ceasing immediately before the steady-state portion of the vowel. The stimulus does not contain a steady-state portion, but it is psychophysically perceived as a consonant-vowel speech syllable. The stimulus was chosen such that it is short enough in duration to minimize test time while still containing key acoustic phonetic information. Time-amplitude waveform and broadband spectrogram of 40 ms synthesized speech stimulus /da/ and is shown in Figure 2.1.

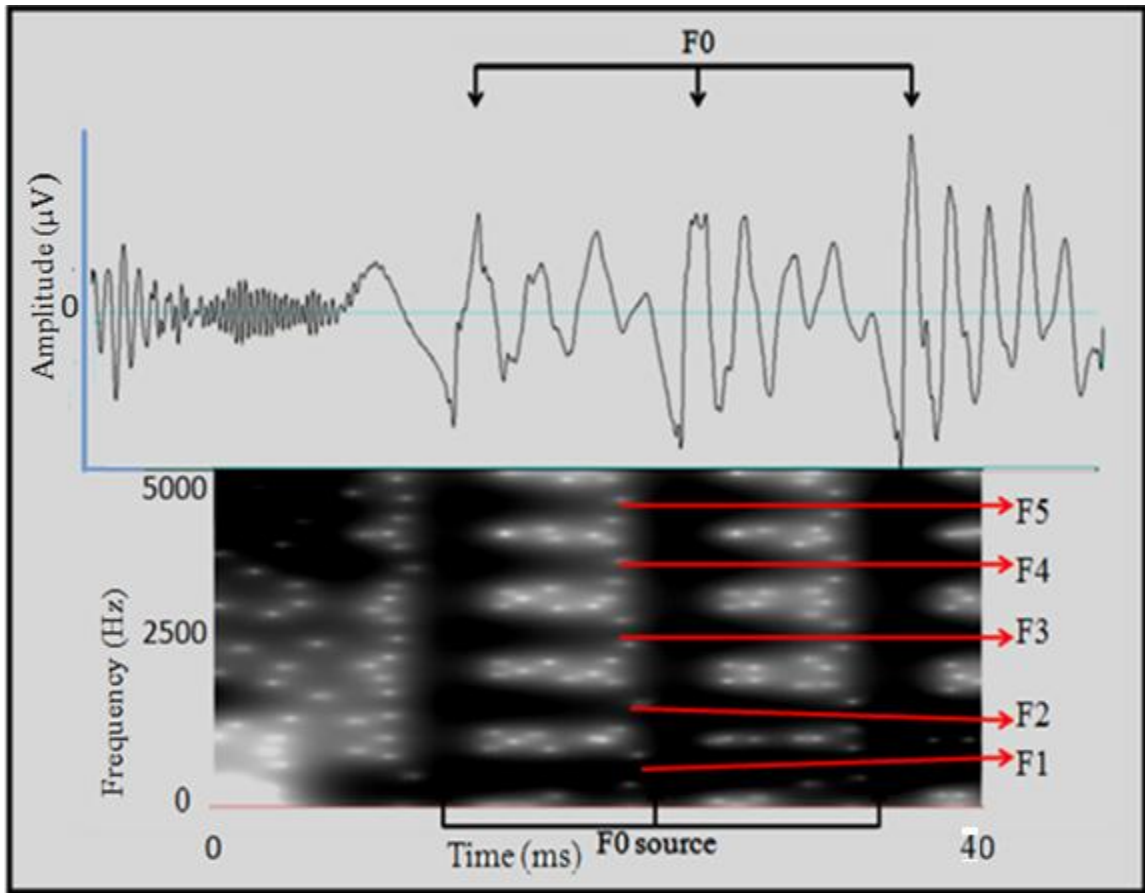


Figure 2.1: Time-amplitude waveform and broadband spectrogram of a 40 ms synthesized speech stimulus /da/.

In order to establish a valid and reliable means to differentiate the neural activity of speech evoked ABR, it is essential to understand the functional relationship between the acoustic structure of speech and the brain stem response to speech (Cunningham, Nicol, Zecker, Bradlow & Kraus, 2001; Russo, Nicol, Musacchia & Kraus, 2004). Because speech is inherent with rapid temporal fluctuations and complex spectral distributions, both transient and sustained measures are used to describe the response.

Transient responses give information about the accuracy with which brainstem neurons synchronously respond to a sound and the reliability with which the response mimics the stimulus or the degree to which it is degraded by background noise. While,

the sustained responses provide information about the fidelity with which the brainstem responds to sound and also represents the range of spectral components within the brainstem responding to the sound. Time amplitude waveform of a 40 ms stimulus /da/ and its corresponding response in the time domain is shown in Figure 2.2.

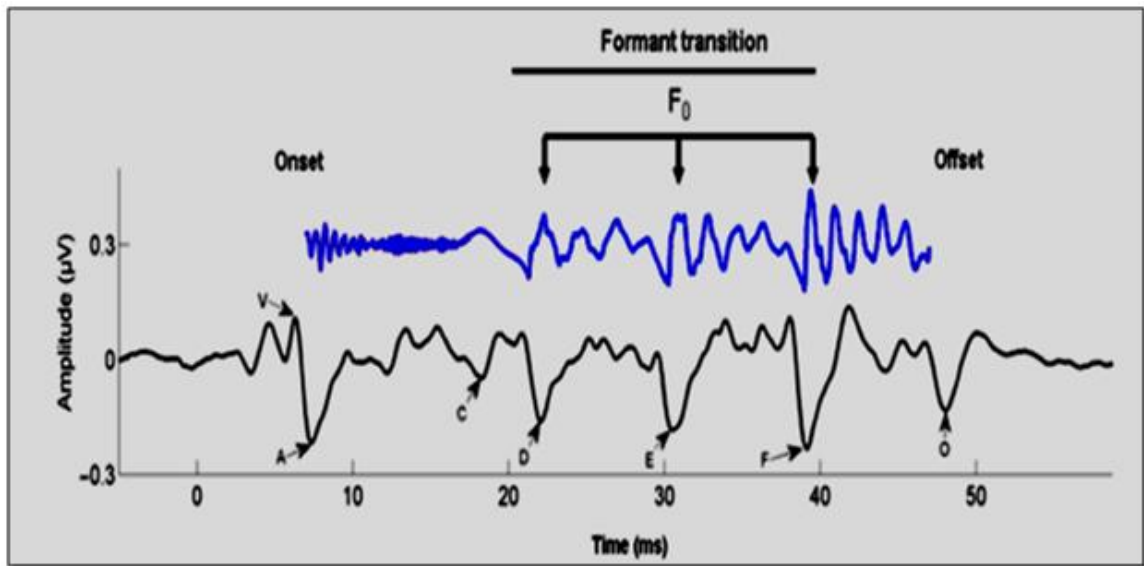


Figure 2.2: Time amplitude waveform of a 40 ms stimulus /da/ and its corresponding response in the time domain (courtesy Nina Kraus, NWU, Chicago).

In the response, there is a positive wave V of ABR that occurs at the latency of 6-10 ms after the onset of the stimulus. A negative wave A immediately follows the wave V. Wave C marks the transition. Waves D, E, and F represent the periodic portion of the syllable from which the fundamental frequency of the stimulus can be extracted. Lastly, Wave O marks the offset of the stimulus (Russo et al., 2004).

The onset response of the speech-evoked ABR is comparable to that elicited by clicks (Song, Banai, Russo & Kraus, 2006). However, the FFR does not appear in the click-evoked response. The encoding of the fundamental frequency and harmonic composition of speech stimulus is reflected faithfully by the FFR (Moushegian, Rupert &

Stillman, 1973). Hence an advantage of the speech evoked ABR is that it represents both transient and sustained portions of the stimulus, which allows for objective assessment at the level of the brainstem. Response replicability and test-retest reliability is established in both quiet and background noise conditions (Russo et al., 2004). Among the aforementioned responses, the transient response of peak latencies and VA slope are reported to be highly detectable with little variability.

2.3 Sensitivity and Specificity of Speech Evoked Brainstem Response

2.3.1 Sensitivity and Specificity in individuals with Learning Impairment

Banai, Nicol, Zecker and Kraus (2005) reported 40% sensitivity using speech stimulus /da/ in detecting individuals with a subgroup of learning deficit (LD). These individuals exhibited abnormal brainstem timing. The authors suggested that speech evoked ABR may serve as a reliable tool for detection of a subgroup of individuals with LD. Song, Banai, Russo and Kraus (2006), compared ABR evoked by clicks and synthesized speech syllable /da/ of 40 ms, in normal-learning children and in children with language-based learning problems. Results showed normal click evoked ABRs in both the groups, however there was a significant delay in speech evoked ABR of children with language based learning problems compared to no delay in the normal learning group. The specificity of speech evoked ABR was 84.38% and sensitivity was 27.73% in detecting children with language based learning disorder.

King, Warrier, Hayes and Kraus (2002), compared the speech evoked ABR in normal learning children and children with language learning deficit (LLD) and reported that children with LLD present abnormalities of both the onset response and the

magnitude of FFR. Twenty of the 54 listeners with LLD showed a delay of 1 SD in the Wave A latency, along with delayed waves C and E. Wible, Nicol and Kraus (2004), reported that in children with LLD, the transition between wave V and A was shallower as compared to normal learning children, suggesting a sluggish response in the group with LLD. The amplitude of FFR in frequency region corresponding to F1 of the /da/ stimuli was found to be lesser in children with LLD. Similar reduction in F1 magnitude in children with LLD was reported by Cunningham, Nicol, Zecker, Bradlow & Kraus (2001) when the stimulus was presented in background noise. However, the magnitude of response in the F0 range was found to be normal in both quiet (Wibel et al., 2004) and noise (Cunningham et al., 2001). Speech evoked ABR represents both source (F0) and filter (onset, offset, and formant transition) characteristics of speech signals. The aforementioned studies have shown that the brainstem coding of the filter may be more impaired in individuals with learning deficits than the source coding (Cunningham et al., 2001).

2.3.2 Sensitivity in individuals with Autism

Intriguingly, a reverse trend using speech stimulus in individuals with autism spectrum disorders (Russo et al., 2008). He reported impaired coding of F0 at the brainstem in 20% of children with autism, demonstrated by pitch-encoding deficits greater than 1.65 standard deviations. This finding was consistent with the clinical symptom of impaired prosody in this population.

2.3.3 Sensitivity in individuals with Phonological Disorders

Sensitivity of 56% and specificity of 94.4% for detecting phonological disorders using speech evoked ABR was reported by Goncalves, Wertzner, Samelli and Matas (2011).

2.3.4 Sensitivity in individuals with Sensorineural Hearing Loss

Comparison of click evoked ABR and speech burst evoked ABR in individuals with mild to moderate sensorineural hearing loss (SNHL) showed deviant wave V latency in ears with SNHL while the click evoked ABRs were within normal limits (Khaladkar, Karthik & Vanaja, 2006).

Conventional click-evoked ABRs provide an indication of the integrity of the cochlea and the ascending auditory pathway; they do not provide information about encoding of more temporally complex signals. However, speech evoked ABR provides objective information about how the sound structure of speech syllables is encoded by the auditory system, it can be used to diagnose auditory processing deficits despite normal processing of click stimuli. Hence, brainstem responses to both sounds provide objective and complementary information about sound encoding in the auditory system.

2.4 Generators of the Speech Evoked ABR

In the past few years, there has been an increase in interest for identifying the sources of the scalp-recorded brainstem response to speech stimulus. Even though there are limited studies on the onset response to speech, given that the response latency occurs within the first 10 ms, it is clear that these responses are of brainstem origin. However,

the neural source of the sustained FFR that mimics the periodicity of the input stimulus is unclear.

Multiple origins have been proposed for scalp-recorded FFR (Galbraith, 1994). But there is less agreement on the precise source of the FFR (Gardi, Merzenich, & McKean, 1979; Moushegian, Rupert & Stillman, 1973; Sohmer, Pratt, & Kinarti, 1977).

2.4.1 Evidence to Support Neural Origin of FFR

Worden and Marsh (1968) conducted experiments to delineate FFR from cochlear microphonics (CM) and provided evidence for a neural basis for the FFRs. They reported that the latency of FFR for even simple pure tones occurs beyond 5 ms, suggesting a site of generation beyond the cochlea and also proving that the FFR is not a reflection of stimulus related artifacts. Adding on, the authors reported that the FFR is not an exact reproduction of the input stimulus, since the response consists of considerable amplitude and phase fluctuation which is unlike the CM which perfectly replicates the input stimulus. The CM can be recorded even under anoxia. However, the FFR shows reduction in amplitude which is consistent with other neural potentials. Also, the CM is not susceptible to change in stimulation rate where as the FFR shows shifts in latency with increasing rates. Marsh, Brown, and Smith (1974) obtained FFRs in cats and reported of precise phase correspondence between the FFRs and electrical activity at the cochlear nucleus (CN), trapezoid body, and superior olivary complex (SOC), thus, indicating that the FFR is an ensemble response reflecting phase-locked activity from multiple generator sites within the auditory brainstem. Collectively, the aforementioned

studies support for a neural origin of FFR and evidently delineate the phase-locked activity reflected by the FFR from CM or stimulus-artifact-related activities.

2.4.2 Evidence to Support Brainstem Origin of FFR

In most studies, FFRs are recorded using vertical ipsilateral montage compared to the horizontal montage. Though the FFRs are recorded from the vertex, there are abundant reasons that suggest that the FFRs reflect activity of the brainstem rather than cortex. Most studies using speech syllables to evoke FFR, report of the response amplitude within 1microvolt. Hoorman, Falkenstein, Hohnsbein and Blanke (1992), obtained FFRs across different stimulus frequencies and reported that the FFR amplitudes were largest between 320 and 380 Hz. 400 nanovolts was found to be the mean FFR amplitude in these frequency ranges. This is in contrast to the usual cortical responses, which are much larger in amplitude ranging over several microvolts. This difference in size between brainstem and cortical responses, leads to difference in the number of averages required to obtain the optimum morphology for each of these responses. FFRs need at least 1000 averages to attain its distinctive morphology, whereas, only about 75–100 averages are sufficient for cortical responses.

The FFR amplitude is stable with increase in stimulus repetitions (Johnson, Nicol & Kraus, 2008), whereas reduction in amplitude with increase in repetition rate, a phenomenon known as neural adaptation occurs in cortical potentials (Grill-Spector, Henson, & Martin, 2006). These differences support the evidence which suggests that at the single neuron level, stimulus specific adaptation is seen more in cortical neurons than

subcortical neurons (Ulanovsky, Las & Nelken, 2003). Taken together these studies support brainstem origin of FFR as compared to cortical origin.

It is evident that the FFR is generated by the brainstem. However, the acquisition of FFR is dependent on various factors.

2.5 Factors Affecting Speech Evoked ABR

Speech evoked ABR is affected by the factors related to the subject, stimulus parameters and recording parameters. Each of these factors would be discussed in the following subsections.

2.5.1 Subject Related Factors

Age is proven to effect the encoding of speech at the brainstem (Vander Werff & Burns, 2010; Johnson, Nicol & Kraus, 2008). Unlike clicks, a developmental pattern was observed in the brainstem response to speech across age groups between 3 to 12 years. The onset response and FFR was found to be significantly delayed in 3-4 years group relative to 5-12 years group (Johnson et al., 2008). This data suggests an effect of age in both temporal and frequency domains of speech evoked ABR and also suggests a possibility of experience-dependent plasticity in the human auditory brainstem. Similarly, when the mean latencies and amplitudes of speech evoked ABR was compared between young normal hearing adults group (age range of 20- 26 years) and an older normal hearing group (age range of 61-78 years). A significant delay was reported in the offset response along with reduction in the amplitude of onset and offset response in the older adult group relative to the younger adult group (Vander Werff & Burns, 2010). The authors reported that these effects were different from those of simply decreasing the

overall stimulus level, which causes significant shifts in latencies of all waves evoked by speech stimulus.

Another biologically inherent factor reported to have an effect on brainstem encoding of speech is the native language of the subject. Experience with one's native language is reported to not only shape speech perception but auditory processing in general. Therefore, native speakers of Mandarin (in which pitch provides important information) were found to have better processing of pitch contours even in a nonlinguistic context, compared to native speakers of English (Bent, Bradlow & Wright, 2006). At the physiological level, Mandarin speakers show more robust encoding of the pitch of Mandarin sounds at the brainstem suggesting that language experience fundamentally changes the neural circuitry of the auditory pathway (Krishnan, Xu, Gandour & Cariani, 2005).

2.5.2 Stimulus Related Factors

The stimulus factors which are proven to have an effect on speech evoked ABR are ear of stimulation, the type of transducer used for presentation of the stimulus, stimulus intensity, stimulus polarity, repetition rate and the number of stimulus.

Hornickel, Skoe and Kraus (2009), recorded brainstem responses to /da/ syllable, which was presented monaurally to the right and left ears in adults with symmetrical interaural click-evoked responses. Right ear responses were reported to have earlier latencies for peaks D and F, than the left ear. Further, robust encoding of F1 was observed when the stimulus was presented to the right ear than the left ear. The authors suggested a possibility of right ear advantage for speech stimulus. Thereby showing that

right-ear advantage for speech is evident at the brainstem. Additionally, Majority of the studies recommend electromagnetically shielded insert earphones for presentation of stimulus relative to circumaural headphones. This is because there is an increased chance for stimulus artifact contamination while using circumaural headphone. Furthermore, the intensity of the speech stimulus was also found to have an effect on the onset and sustained response (Akhoun et al., 2008). When the syllable /ba/ was varied as a function of intensity from 0 to 60 dB SL, in 10 dB increments. As seen in clicks and tones, both response components showed orderly latency shifts with increasing intensity. The onset response and FFR latencies decreased with increasing stimulus intensity, with a greater rate for FFR (-1.4 ms/10 dB) than for onset response (-0.6 ms/10 dB).

Two different methods have been recommended while recording speech evoked ABR. The first method recommends recording of response to any one stimulus polarity (Krishnan, 2007). The second method suggests recording responses to both polarities and either adding (Russo et al., 2004 & Akhoun et al., 2008) or subtracting the responses (Krishnan, 2002) to the two stimulus polarities. The process of adding will emphasize the low frequency components of the response which includes phase locking to the amplitude envelope and minimizing stimulus artifact and the CM. Subtracting will increase the high frequency components by maximizing the spectral response, and also maximize artifact contamination. Hence, the Alternating polarity is most preferable. Adding on, the number of sweeps required for speech stimulus, to obtain robust and reliable responses are comparatively greater than that required for clicks and tones. A general principle of EP signal averaging is that the SNR is proportional to the square root of the number of sweeps (Hood, 1998; Hall 2006). Thus, the overall SNR increases quickly at first and

then begins to plateau with more sweeps. However, each component of the speech evoked ABR may show its own SNR progression with different response components requiring greater or fewer sweeps. Most studies use an approach of collecting responses to more than one stimulus trials, typically 2000 to 3000 per polarity and adding the responses. This strategy helps to determine response repeatability and track how the responses evolved over time. Also, subtle response characteristics and small group differences may not be apparent until additional sweeps are collected and repeatability is confirmed.

Lastly, the length of the stimulus and the inter stimulus interval (ISI) are also important to decide the repetition rate. It is an established fact that, changing the ISI can modify the perception of sound. Also, if the ISI is short, the response to one stimulus may not fully conclude before the next stimulus is presented. Hence, the ISI and the analysis time should be sufficiently long enough to allow for the response to return to baseline.

Krizman, Skoe and Kraus (2010), conducted a study to determine the effects of stimulation rate on ABR, they recorded evoked responses to both clicks and speech syllable /da/ presented at three rates (15.4, 10.9 and 6.9 Hz). The results showed that the latency of click evoked response was constant over the three repetition rates. But, latency of the onset response to /da/ varied systematically, increasing in peak latency as presentation rate increased. The FFR was also found to be rate dependent. It was found that the magnitude of the high frequency components of the response reduced with increasing rate.

2.5.3 Response Acquisition Related Factors

The response acquisition factors that are proved to have an effect on speech evoked ABR are the analysis time, sampling rate, electrode montage, filter setting, and amplification. Among these, the analysis time window is recommended to be long enough to include a pre stimulus baseline period, a response period, and a post stimulus period. The post stimulus period needs to account for the stimulus transmission delay and neural conduction time. Hence, a post stimulus period between 10 and 50 msec is recommended to ensure that the response returns to baseline (Skoe & Kraus, 2010a). Another important factor is the sampling rate. Sampling rates determines the frequency of digitization of the neural signal by the recording system. According to the Nyquist theory, the sampling frequency should be twice that of the highest frequency in the stimulus. Hence studies using speech as stimulus to evoke brainstem responses, have made use of sampling rates ranging from 7 to 50 kHz (Musacchia, Sams, Skoe & Kraus, 2007; Akhoun et al, 2008; Banai, Hornickel, Skoe, Nicol, Zecker & Kraus, 2009). A higher sampling rate not only reduces sampling error but also increases the temporal precision of the recording and allows for finer differentiation of response peaks.

Additionally to record speech evoked ABR, majority of the studies have used the vertical, one channel montage. This configuration requires three electrodes corresponding to the active (noninverting), reference (inverting), and ground electrode. The preferred electrode placements are Cz for active electrode, ipsilateral earlobe or nape of the neck for reference electrode, and forehead or contralateral earlobe as ground. A non cephalic site is preferred over the mastoid as reference because it leads to fewer artifacts from bone vibration (Hall, 2006). Furthermore, filtering is an essential technique to isolate

activity evoked by the subcortical structures from cortical structures and to increase the SNR of the response. The band-pass filter for speech evoked ABR and FFR falls in the range of 100 to 3000 Hz (Skoe & Kraus, 2010a). This frequency range has been found to increase the detection of the high frequency transient peaks, such as wave V, which has a sharp slope. Lastly, Since the response of interest originates from the brainstem. The response amplitude is in the order of several nanovolts. Hence, amplification of the response is essential. A gain of 100000 is found to be sufficient.

Thus by combining the knowledge of these subject related, stimulus related and, acquisition related factors affecting speech evoked ABR, one can obtain a reliable response to speech stimulus. This reliability of the brainstems response of speech has helped in investigation of the various complex mechanisms used by the brainstem to encode speech.

2.6 Efferent Modulation of the Brainstem

Recent studies have given evidences suggesting malleability in the brainstem representation of speech (Kraus & Nicol, 2005; Banai Nicol, Zecker & Kraus, 2005). Long-term and short-term auditory experiences have been shown to enhance the brainstem responses to complex, behaviorally relevant sounds.

2.6.1 Modulation of Brainstem Physiology Following Long-term Experience to Language and Music

Krishnan, Xu, Gandour and Cariani (2005) carried out a cross-language study, and showed that long-term experience with linguistic pitch contours enhanced pitch representation at the brainstem as reflected by the FFR. The native speakers of Mandarin had significantly better brainstem representation of linguistic pitch contours compared to

Native American English speakers. Such plasticity appears to be particular to the nature of the long-term experience, as only naturally occurring Mandarin tones elicited experience dependent effects in native speakers, and not their linear approximates.

Krishnan, Swaminathan and Gandour (2008) reported that plasticity is not specific to speech stimuli as long as linguistic relevance is maintained. They conducted a cross-language study using iterative ripple noise (IRN) to simulate Mandarin tones. The IRN stimuli is non speech in nature, but preserved the complex pitch information. Mandarin participants represented pitch better at the level of the brainstem compared to English speakers, suggesting that brainstem plasticity is not specific to speech. Rather, it is specific to dimensions that occurred in natural speech.

Long-term experience with music has also been shown to provide an advantage in the brainstem representation of speech (Musacchia, Sams, Skoe & Kraus, 2007; Strait, Skoe, Kraus & Ashley, 2009; Wong, Skoe, Russo, Dees & Kraus, 2007).

FFRs obtained from musicians and non musicians showed an advantage for musicians in processing native speech sounds (Musacchia et al., 2007). Musicians had earlier and larger brainstem responses than non musician controls to both speech and music stimuli. Perception of pitch, reflected by Phase locking to stimulus periodicity was enhanced in musicians and also, strongly correlated with length of musical practice. Musicians showed faithful and robust encoding of nonnative linguistic pitch compared to non musicians (Wong et al., 2007). Strait, Skoe, Kraus & Ashley (2009) reported similar enhancement in musicians for emotionally salient vocal sounds. These studies clearly indicate that plasticity at the level of the brainstem is not specific to the context of the long term experience.

2.6.2 Modulation of Brainstem Following Short-term Experience

Short-term auditory training has shown to improve the timing of the FFR to the syllable /da/. Brainstem response to /da/ was obtained from children with learning problem in both quiet and in presence of noise. Results showed that children who underwent an auditory training program exhibited brainstem responses that were more resistant to the deleterious effects of background noise (Russo, Nicol, Zecker, Hayes, & Kraus, 2005).

Song, Skoe, Wong & Kraus (2008) examined whether short-term training improves brainstem representation of lexical pitch contours. Participants in the study who were non-Mandarin -speakers underwent a short-term word learning training program in which they were taught to lexically incorporate Mandarin pitch contours embedded in non words. FFRs were recorded before and after training. The eight-session training program showed significant improvement in the brainstem representation of the Mandarin dipping tone.

Madhok and Sandeep (2010), obtained speech evoked ABR from ten normal hearing adults before and after training for frequency and intensity discrimination and temporal modulation identification. Results showed enhancement in the onset and sustained responses to speech stimulus following training. Taken together, these studies suggest that the adult brainstem is indeed malleable to short-term training.

2.6.3 Mechanisms Underlying Experience Dependent Plasticity

All the above mentioned studies have shown the dynamic nature of encoding of the auditory brainstem, and also reflect the short-term and long-term auditory experience,

but the neurobiological mechanism that contributes to this plasticity is unknown. Presently, two hypotheses on the nature of experience dependent brainstem plasticity are being debated (Krishnan & Gandour, 2009). One is the corticofugal model (Suga, Xiao, Ma, & Ji, 2002, Suga, 2008) and the other is the local reorganization model (Krishnan & Gandour, 2009).

The corticofugal model states that top-down feedback via the corticofugal efferent network modifies brainstem function (Suga, 2008; Suga et al., 2002). The corticofugal model predicts moment-to-moment changes in brain function as a result of top-down feedback. On the other hand, the local reorganization model states that the brainstem function is modulated over a longer timescale, that is, the brainstem is reorganized to promote the encoding of frequently encountered sounds (Krishnan and Gandour, 2009). Both models require top-down modulation and are not mutually exclusive. The local reorganization model predicts top-down modulation of brainstem circuitry during learning, after which top-down feedback is no longer required. Thus, both models predict plasticity in relevant feature representation, but the timescales are vastly different. There are good reasons to implicate a corticofugal tuning mechanism. Since there are massive efferent connections from the cortex to subcortical structures, these connections could form the basis of feedback-related top-down projections (Kral & Eggermont, 2007). Also, efferent connections are present between layers of the auditory cortex which provide excitatory and inhibitory control over the inferior colliculus (Keuroghlian & Knudsen, 2007). Repeated stimulation by stimuli that are behaviorally relevant (Chowdhury & Suga, 2000), electrical stimulation of structures of the forebrain (Ma & Suga, 2008; Zhang & Suga, 2005), and auditory fear conditioning (Gao & Suga, 2000) have all been

shown to induce plastic changes to the neuronal response properties in the IC in animals (Suga, 2008; Suga et al., 2002). Importantly, these changes in the IC are restricted when the forebrain structures are inactivated, suggesting that some kind of cortico-collicular tuning shapes response properties of the IC. Taken together, these animal studies strongly support the view that corticofugal modulation changes the neuronal properties of sub-cortical structures in a behaviorally relevant manner.

Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009) elicited brainstem response to speech syllable /da/ in two conditions; variable and repetitive conditions. The results showed that there was a significant difference between the brainstem responses elicited in the two conditions. The response elicited in the repeated condition was enhanced in the lower harmonics and first formant range relative to the variable context condition. This was attributed to corticofugal modulation. The results cannot be attributed to the effect of long term experience, since the test duration was only 30 minutes. Similarly Skoe and Kraus (2010b) monitored the response elicited to a repeating melody and repeating note within a melody over a time course of 1.5 hours. The response to the note which is repeated was enhanced relative to the response to the note which does not repeat itself. The authors attributed their results to online corticofugal modulation of the brainstem.

It is suggested that as an individual becomes a good listener through long-term or short-term auditory experience , he becomes more efficient in utilizing the corticofugal feedback mechanism (Banai, Hornickel, Skoe, Nicol, Zecker & Kraus, 2009; Chandrasekaran et al., 2009; Song et al., 2008; Wong et al., 2007). On the other hand, in individuals with deficits in reading and speech in noise, faulty corticofugal mechanism

results in deficient encoding (Chandrasekaran et al., 2009). Hence, there is a critical need to understand the complex, bidirectional interactions between higher level cognitive processing and lower level sensory encoding in good listeners as well as those with auditory processing disorders. Cognitive and sensory processes are thus inextricably linked, and scalp-recorded brainstem responses may provide a comprehensive view of the consequences of these processes.

Summary of Literature

Brainstem response to speech has offered a unique window into understanding how the brainstem represents the component of speech signals. The brainstem response to speech has two unassociated components, the onset and a sustained frequency following response (FFR). Collectively, these components loyally represent the source and filter characteristics of the speech signal. The generators of the FFR can be distinguished from cochlear and cortical activity. Multiple evidences including ablation studies and developmental information, in addition to the phase-locking capabilities of the auditory brainstem, strongly imply a brainstem origin for the scalp-recorded FFR. Even though the scalp recorded onset response and the FFR reflect the activity at numerous sources (LL, CN, IC), they offer a noninvasive technique to study the sub-cortical encoding of speech, as well as the effect of experience on the representation of speech at the brainstem. Furthermore, the dynamic nature of the brainstem response to speech provides a means to examine corticofugal modulation in the humans, which is yet to be completely explored.

Chapter 3

METHOD

In the present study, true experimental design was used to assess the effect of context on brainstem encoding of speech. The study was based on the hypothesis that context has no effect on brainstem encoding of speech. To verify this hypothesis, the following method was adopted.

3.1 Subjects

Fifteen human adults participated as subjects in the study. All of them were in the age range of 18 to 25 years, with the mean age of 21.2 years. All the subjects had pure tone thresholds within normal limits (<15dBHL) at octave frequencies between 250Hz and 8kHz. They had type 'A' tympanogram with the presence of ipsilateral and contralateral reflexes that ensured normal middle ear functioning (Jerger, 1970). They did not complain of any difficulty listening in adverse listening conditions and had more than 60% score in the speech in noise test performed at 0dB SNR. All the subjects were meritorious students from different parts of the country, pursuing their bachelors and masters degree in Speech and Hearing. A written consent was taken from each subject prior to their inclusion in the study.

3.2 Stimulus Generation

Four different stimuli were used to record brainstem responses, of which one was a core stimulus and the other three were used as contexts. A synthetically generated syllable /da/ was the 'core stimulus'. It was called so, as only the response, recorded for syllable /da/ were of importance in the present study. The three contextual stimuli were a

synthetically generated syllable /ta/ (which differs from /da/ with respect to voicing), a synthetically generated syllable /ba/ (which differs from /da/ with respect to place of articulation) and White noise.

The Five-formant synthesized /da/ (core stimulus) was obtained from Professor Nina Kraus, Principal investigator at the Auditory Neuroscience lab, Northwestern University, Chicago. Time-amplitude waveform and broadband spectrogram of a 40-ms synthesized speech stimulus /da/ are shown in Figure 3.1.

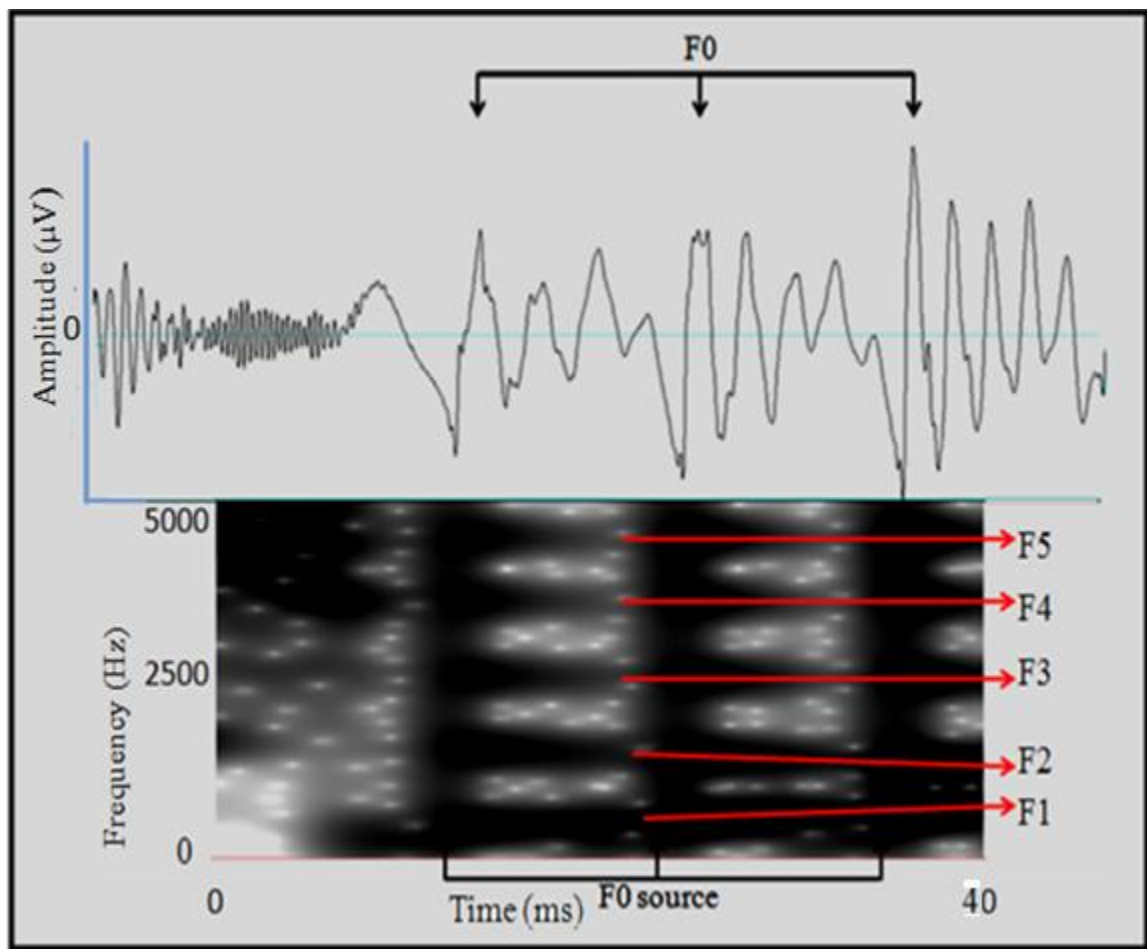


Figure 3.1: Time-amplitude waveform and broadband spectrogram of a 40 ms synthesized speech stimulus /da/.

Fundamental frequency, a source cue, is visible in the spectrograms vertical striations corresponding in time to the major F0 peaks in the time–amplitude plot. The five formant (F1–F5) - high-energy bands resulting from the filter characteristics of the vocal cavity are the darker horizontal stripes seen from 10 ms to 40 ms (the largest and lowest in frequency, F1, is also indicated in the time–amplitude plot). The fundamental frequency was ramped up from 103 to 121 Hz, while F1 was ramped up from 220 to 720 Hz. The F2 and F3 were ramped down from 1700 to 1240 Hz, and from 2580 to 2500 Hz, respectively. The F4 and F5 remained constant at 3600 and 4500 Hz. Rise and fall times were 5 ms as quoted in King, Warrier, Hayes & Kraus, (2002). In the time–frequency plot, the consonant burst is the diffuse dark area in the high-frequency range of the first 10 ms.

The syllables /ba/ and /ta/ were synthesized in the Speech Science Lab of the All India Institute of Speech and Hearing, Mysore. The author made use of a parametric synthesis method, namely the ‘Formant synthesis’ to synthesize the stimuli. Syllables /ba/ and /ta/ were uttered by an adult male speaker and recorded using a directional microphone into a computer using PRAAT software (Version 4.5.18). Sampling frequency of 16000 Hz and 16 bit resolution was used during the recording. The sampling frequency of 16000Hz was necessary to facilitate its analysis using the Speech Science Lab which does not allow analysis and synthesis of stimuli beyond the sampling frequency of 16000Hz. The recorded syllables were then analyzed using the Formant based analysis by synthesis of the ACOFON I module of the Speech Science Lab (SSL): Professional Edition (Version 4.1). For each syllable, voicing, fundamental frequency, first four formant frequencies and first four formant bandwidths were noted at

every 10 ms for a duration of 200 ms. The target syllables were then synthesized by feeding the analyzed values of the naturally uttered syllables into Hybrid (modified Klatt's model) of the ACOPHON II module of SSL. The syllables /ba/ and /ta/ were synthesized such that their duration is approximately 40 ms, so as to match their duration to afore mentioned syllable /da/. The spectral and temporal characteristics of the syllables /ba/ and /ta/ are given in Table 3.1 and 3.2 respectively. The Time amplitude waveforms and spectrogram of the synthetically generated syllables /ba/, /ta/, and the white noise are shown in the Figure 3.2

Table 3.1: *Spectral and temporal characteristics of the syllable /ba/*

Parameter	Target Values					
Discrete duration of signal (ms)	0	10	26	28	38	48
Voicing	S	V	B	V	V	V
Duration (ms)	10	16	2	10	10	10
Frequency (Hz)	0	124	0	124	124	124
Intensity (dB)	0	70	70	70	70	65
F1 (Hz)	0	244	0	314	544	581
B1 (Hz)	0	80	0	80	80	80
F2 (Hz)	0	1568	0	1120	1159	1147
B2 (Hz)	0	100	0	100	100	100
F3 (Hz)	0	2376	0	2121	2391	2407
B3 (Hz)	0	120	0	120	120	120
F4 (Hz)	0	3463	0	3277	3456	3433
B4 (Hz)	0	120	0	120	120	120
G1 (Hz)	0	10	0	10	10	10
G2 (Hz)	0	7	0	7	7	7
G3 (Hz)	0	5	0	5	5	5
G4 (Hz)	0	2	0	2	2	2

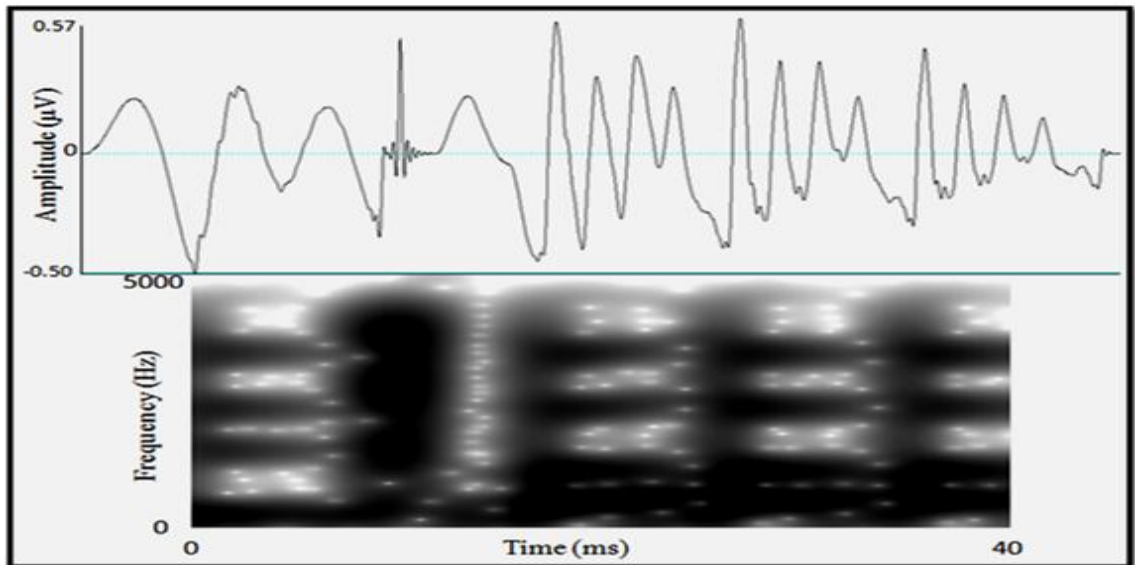
Note: S= Silence; B= Burst; V= Voicing

Table 3.2: Spectral and temporal characteristics of syllable /ta/

Parameter	Target Values					
Discrete duration of signal (ms)	0	10	16	31	41	51
Voicing	S	B	V	V	V	S
Duration (ms)	10	6	15	10	9	10
Frequency (Hz)	0	0	124	124	124	0
Intensity (dB)	0	70	70	70	70	0
F1 (Hz)	0	0	511	566	572	0
B1 (Hz)	0	0	80	80	80	0
F2 (Hz)	0	0	1655	1334	1260	0
B2 (Hz)	0	0	100	100	100	0
F3 (Hz)	0	0	2206	2186	2161	0
B3 (Hz)	0	0	120	120	120	0
F4 (Hz)	0	0	3584	3509	3469	0
B4 (Hz)	0	0	120	120	120	0
G1 (dB)	0	0	69	69	69	0
G2 (dB)	0	0	65	65	65	0
G3 (dB)	0	0	42	42	42	0
G4 (dB)	0	0	30	30	30	0

Note: S= Silence; B= Burst; V= Voicing

(A)



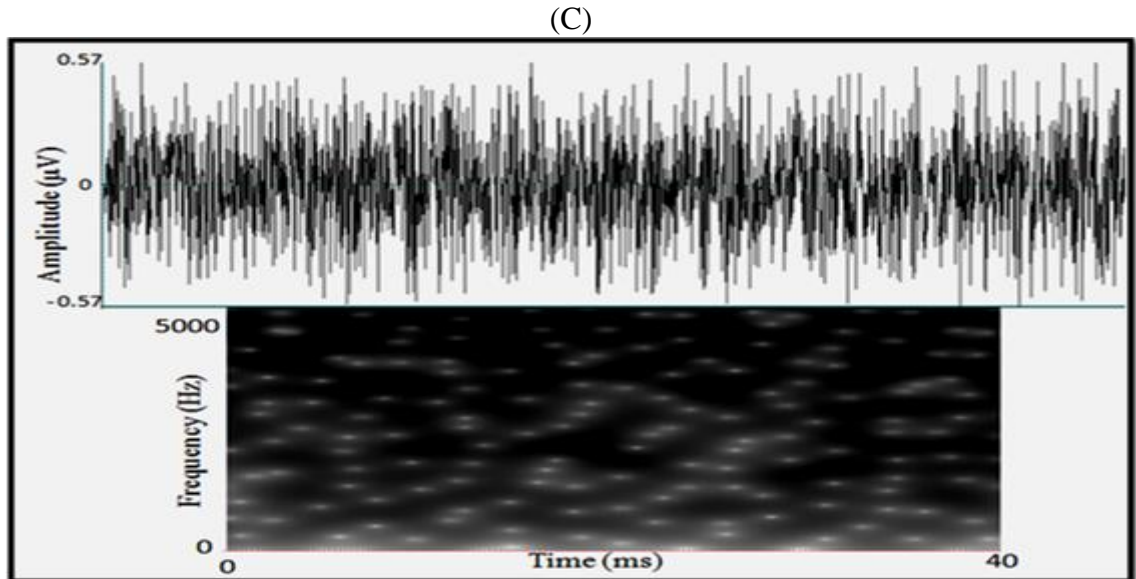
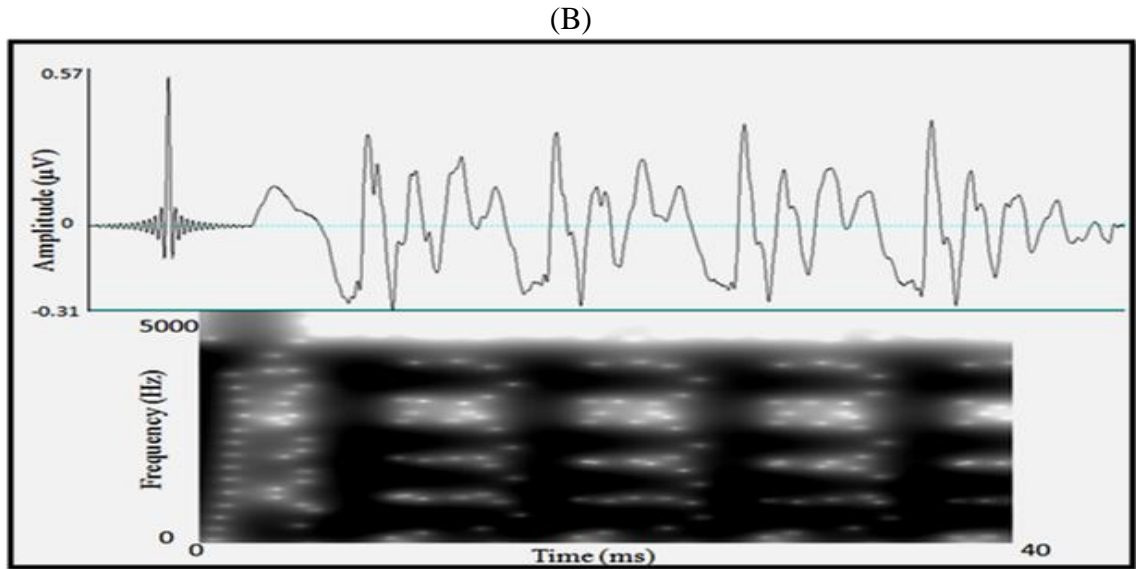


Figure 3.2: Time amplitude waveform and spectrogram of (A) synthetically generated syllable /ba/, (B) synthetically generated syllable /ta/ and (C) white noise.

The sampling frequency of both the synthetically generated syllables (/ba/ and /ta/) was converted from 16000Hz to 48000Hz, using the Cool Edit software (Version 2). This was done so as to facilitate the loading of the stimuli into the Biologic Navigator Pro software (Version 7.0). White noise, the third contextual stimulus was generated using Adobe audition (Version 1.0) at a sampling rate of 48000Hz and 16 bit resolution. The

syllables were then edited using Adobe Audition (Version 1.0) with respect to duration, so as to maintain a constant duration of 40ms across all four stimuli.

All the four stimuli were individually normalized and then group normalized to obtain equal average RMS power of 93.4 dB SPL. They were then loaded into the personal computer with the Bio-Logic Navigator Pro AEP Software (Version 7.0). The synthetic speech syllables /ba/, /ta/ and /da/ were subjected to a subjective rating for naturalness and quality from 10 sophisticated listeners with normal hearing. All the three stimuli were rated as natural and reported to be of good quality.

Before using these stimuli for recording ABR, intensity was calibrated into dBnHL. To do this, all the four stimuli were presented at a repetition rate of 10.9/s through the broad band insert receivers of the Bio-Logic Navigator Pro AEP system. Twenty normal hearing subjects listened to the four stimuli and their mean behavioral thresholds were obtained in dB SPL. Since the mean behavioral threshold obtained was not a whole number. It was approximated to the nearest whole number which was then considered as 0 dBnHL. The mean behavioral thresholds in dB SPL are given in Table 3.3.

Table 3.3: *Mean behavioral thresholds in (dB SPL) of the synthetically generated syllables /da/, /ba/ and /ta/ and, the white noise*

Stimulus	Mean Behavioral Thresholds	Approximated Mean Behavioral Thresholds
Synthetically generated syllable /da/	26.2 dB SPL	25 dB SPL
Synthetically generated syllable /ba/	30.83 dB SPL	30 dB SPL
Synthetically generated syllable /ta/	31.6 dB SPL	30 dB SPL
White noise	30.5 dB SPL	30 dB SPL

3.3 Test Environment

The stimulus recording and all the audiological evaluation were carried out in a sound treated room. The ambient noise in the room was within the permissible limits as recommended by ANSI (S3.1; 1991).

3.4 Test Procedure

3.4.1 Pure Tone Audiometry

Behavioral air conduction and bone conduction thresholds were tracked with modified Hughson and Westlake procedure (Carhart & Jerger, 1959) using a calibrated GSI- 61 clinical audiometer coupled to impedance matched TDH 39 earphones with MX-41/ AR ear cushions and a bone vibrator (Radio ear B-71). Air-conduction thresholds were found from 250Hz to 8 kHz, while the bone conduction thresholds were tracked for frequencies from 250Hz to 4 kHz.

3.4.2 Speech Perception in Noise

The above mentioned audiometer was also used to obtain the Speech identification scores (SIS). Standardized monosyllabic-words in English developed by Rout & Yathiraj, (1996) were presented at 40dBSL (with reference to the Puretone Average of 500 Hz, 1 kHz and 2 kHz) and 0dBSNR. The corresponding SIS was obtained from each subject monaurally for both ears.

3.4.3 Tympanometry

A calibrated microprocessor based automatic Immittance meter (Grason - Stadler GSI-TS) with a visual display was used for tympanometry. Middle ear evaluation was

done using 226Hz probe tone. Immittance test was carried out by sweeping the pressure from +200 to -400dapa. In reflexometry, both ipsilateral and contralateral acoustic reflex thresholds were measured for 500Hz, 1000Hz, 2000Hz, and 4000Hz pure tone at the peak pressure. The threshold was defined as the minimum intensity of the signal that leads to 0.03ml change in compliance.

3.4.4 Auditory Brainstem Response

An evoked potential system (Bio-Logic Navigator Pro AEP Software (Version 7.0) was used to record the brainstem responses. The subjects were made to sit on a reclining chair. The skin surface at the vertex (Cz), nape of the neck and the upper forehead were cleaned using the skin preparing gel. Gold plated disc electrodes along with the conduction paste were placed over the cleaned skin surface and secured at its place using tape to obtain impedance of less than 5kOhms at each electrode site. Single channel vertical ipsilateral montage was used for recording the response. Subjects were instructed to relax and avoid any body movements.

Brainstem response to synthetically generated syllable /da/ was recorded in four different conditions; which included one repetitive condition and in three different contexts. First, the response to the repetitive condition was obtained for 1500 sweeps of /da/. Only the stimulus /da/ was used in this paradigm. Then the brainstem responses were recorded in three different stimulus contexts using the MMN/P300 protocol, wherein the infrequent stimulus was /da/ presented with a probability of 33% and the frequent stimulus was either /ba/ or /ta/ or white noise presented with the probability of

66%. The order of recording was randomized to avoid the order effect. The Figure 3.3 represents the different stimulus conditions.

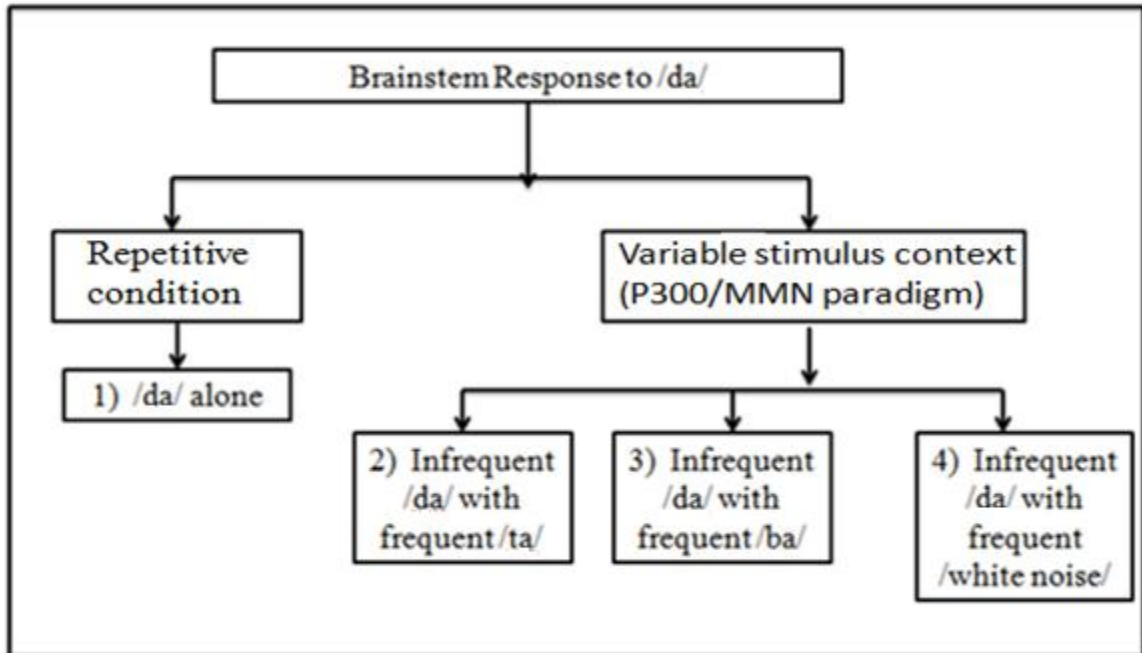


Figure 3.3: Representation of different stimulus conditions used.

For each stimulus condition, the response was recorded twice to ensure the replicability and reliability. Brainstem response for each of the four stimulus conditions was collected using the stimulus and acquisition parameters given in Table 3.4.

Table 3.4: Stimulus and acquisition parameters used to elicit the Brainstem Responses

Stimulus Parameters		
1	Stimuli	<i>Repetitive context: /da/</i>
		<i>Variable context:</i> Frequent stimuli (66%): /ba/ or /ta/ or noise Infrequent stimuli (33%): /da/
2	Frequent to infrequent ratio	2:1
3	Ear	Right ear/ left ear
4	Duration of stimuli	40 ms
5	Intensity	70 dBnHL(100dBSPL for /ba/, /ta/ and white noise and, 95dBSPL for /da/)
6	Repetition rate	10.9/s
7	Polarity	Alternating
8	Number of sweeps	1500
Acquisition Parameters		
1	Analysis time	64 ms
2	Electrode montage	Vertical
3	Number of points	1024
4	Amplification	100000
5	Artifact rejection	+ or – 23.8 μ V
6	Filter setting	100-2000Hz

3.4.5 Response Analysis

The averaged response obtained for the /da/ stimulus in the repetitive condition was compared with the averaged response obtained for /da/ stimuli in each of the three stimulus contexts. Responses elicited for /ba/, /ta/, and white noise were not analyzed in this study. Analysis of the response for /da/ was carried out both subjectively and objectively. Both the transient and sustained portions of the responses were analyzed. The peak latency and peak amplitude of wave V, A, B, C, D, E, F and, O and the V to A amplitude were the measures considered for comparison. The subjective analysis was carried out by two experienced Audiologists. The right end of the wave with the largest amplitude around 6 ms following the stimulus onset was marked as wave V. The immediate negative trough following the wave V was marked as wave A. V to A amplitude was obtained from the voltage difference between the wave V and wave A. The replicable negative waves occurring at the expected peak latencies with large amplitude were marked as wave B, C, D, E, F and O.

The response amplitude between 0 and 1 ms was approximately 0 μ V across all the recordings. Hence there was no baseline amplitude correction used and the peak amplitude of waves B, C, D, E, F and O were obtained directly from the instrument by placing the cursor on the respective waves. Brainstem response elicited by syllable /da/ in a representative subject is shown in Figure 3.4.

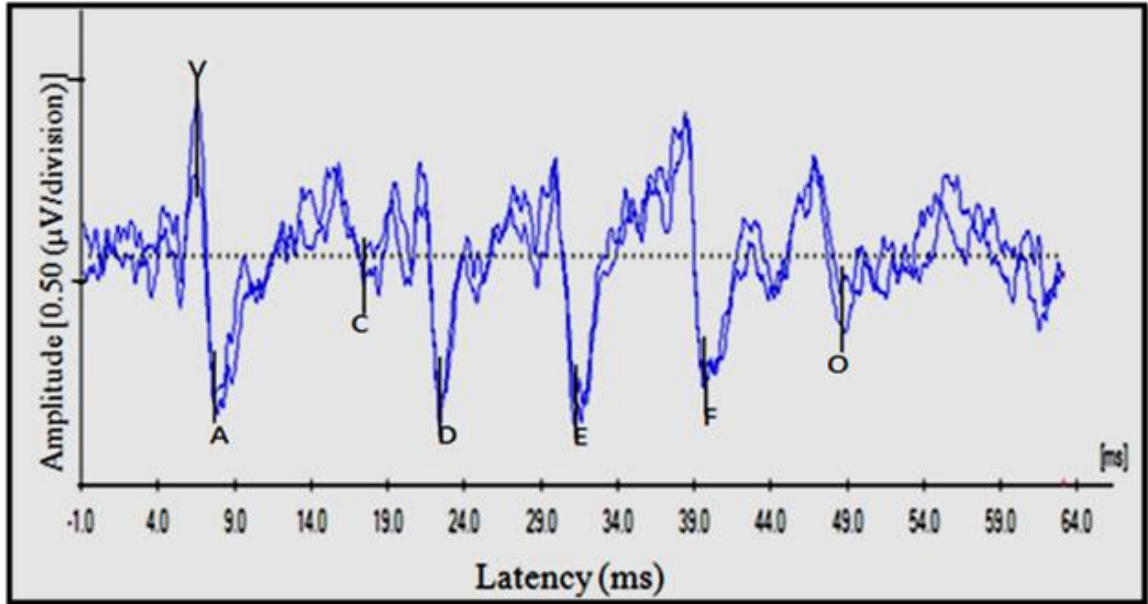


Figure 3.4: Brainstem response elicited by syllable /da/ in a representative subject.

Additionally, to evaluate the spectral composition of the response, Fast Fourier transform (FFT) analysis of the sustained response of the speech evoked ABR was done. This was executed using the MATLAB R 2009a platform and software (Brainstem toolbox) developed by Nina Kraus (2004) at Northwestern university. Fourier analysis was performed on the 11.4–40.6 ms epoch of the FFR in order to assess the amount of activity occurring over three frequency ranges; (103– 121Hz), (454-719Hz) and (721-1155Hz). These frequency ranges were chosen because the neural responses at these frequencies would correspond to the Fundamental frequency, first formant and higher harmonics of the stimulus /da/ respectively (Johnson, Nicol, Zecker, Bradlow, Skoe, E & Kraus, 2008). A 2 ms on- 2 ms off Hanning ramp was applied to the waveform (to avoid the spectral splatter). Zero-padding was employed to increase the number of frequency points where spectral estimates were obtained. The raw amplitude value of the F0, F1 and higher frequency component of the response FFR was then measured. Fast Fourier

transform analysis of the brainstem response to the stimulus /da/ in a representative subject is shown in Figure 3.5.

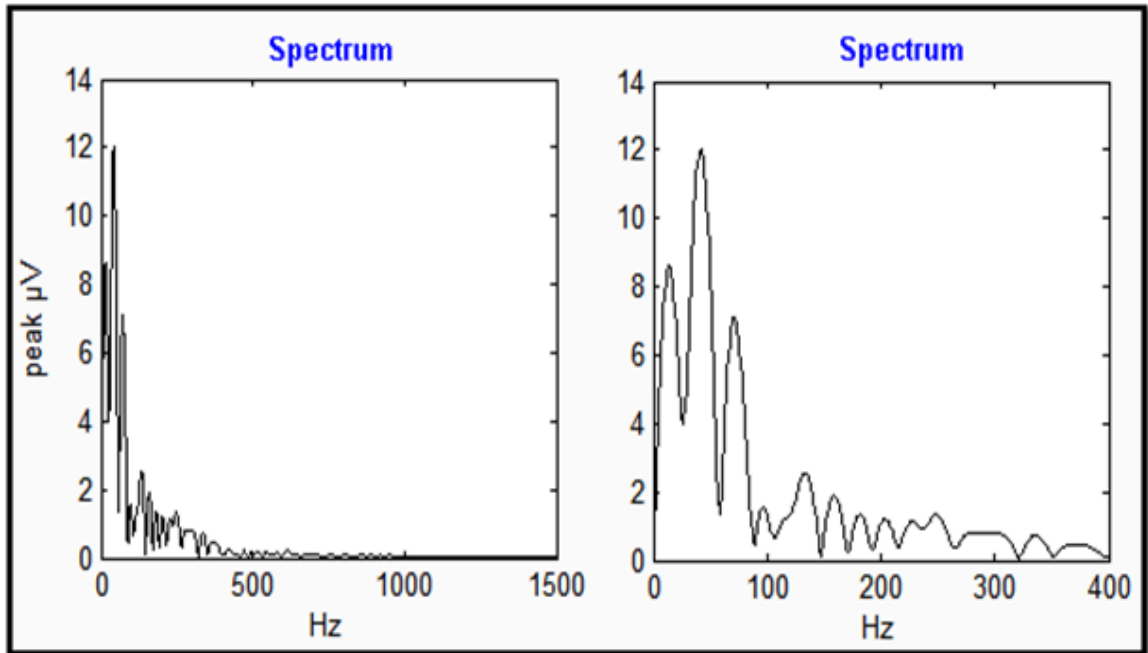


Figure 3.5: Fast Fourier transform analysis of the brainstem response to the stimulus /da/ in a representative subject.

3.4.6 Data Analysis

The data thus obtained was used for the following comparisons to test the hypothesis of the study:

- 1) Comparison between the responses obtained in repetitive stimulus condition and in different stimulus contexts.
- 2) Comparison across the responses obtained in different contexts.

Chapter 4

RESULTS

The study was aimed to assess the effect of context on brainstem encoding of speech. To do so, brainstem responses to a synthetically generated syllable /da/ was recorded in four different conditions, which included one repetitive condition and in three different context conditions. In condition 1, a repetitive stimulus paradigm was used, wherein brainstem responses were elicited only for the stimulus /da/. Variable stimulus context paradigm was used in conditions 2, 3, and 4, wherein responses were elicited for the stimulus /da/ presented infrequently in the context of another frequent stimulus. In condition 2, the frequent stimulus used was a synthetically generated syllable /ta/. Similarly, synthetically generated syllable /ba/ was used as the frequent stimulus in condition 3, while white noise was used as the frequent stimulus in condition 4.

The averaged onset and sustained response obtained for the stimulus /da/ in the repetitive condition was compared with the averaged onset and sustained response obtained for the stimulus /da/ in each of the three stimulus context conditions. The data was tabulated and statistical analysis was carried out using Statistical Package for Social Science software (version 17.0). Descriptive statistics, Kolmogorov–Smirnov test, Paired t- test, Repeated measures ANOVA and Sidak Post-Hoc test were the statistical tests used.

To ensure that the data obtained is normality distributed, each target measure (dependent variable) tested on Kolmogorov–Smirnov test of normality. Results of the test indicated that all the data were normally distributed ($p>0.05$).

Because the test was done monaurally in the two ears, initially the ear difference was tested. Paired t- test was done to compare the mean difference between the right and left ear. The mean peak latency and peak amplitude of the waves V, A, B, C, D, E, F, and O, and the peak amplitude of fundamental frequency, first formant and higher harmonics were compared between the ears. The results showed no significant effect ($p>0.05$) of ear in any of the measures. Hence for all further statistical procedures, the data of the two ears were combined.

4.1 Percentage of occurrence of Onset and Sustained Responses

The responses were subjectively analyzed to identify the waves V, A, B, C, D, E, F, and O. It was observed in the analysis that not all the recordings had all the aforementioned waves. The percentage of occurrence of each wave across the four conditions is given in Table 4.1.

Table 4.1: *The percentage of occurrence of each wave across four conditions*

Condition	Percentage of occurrence of waves							
	V	A	B	C	D	E	F	O
1	100%	100%	40%	70%	86.6%	100%	100%	96.6%
2	100%	100%	50%	50%	83.3%	100%	93.3%	83.3%
3	100%	100%	43.3%	63.3%	93.3%	93.3%	96.6%	86.6%
4	100%	100%	40%	56.6%	83.3%	86.6%	93.3%	76.6%

It is evident from Table 4.1, that the wave V and A are present in all the ears in all four conditions. But, the percentage of occurrence of other target waves did vary across the four conditions. Condition 1 had higher percentage of occurrence of the target waves C, D, E, F, and O compared to condition 4. Additionally when compared among the waves, waves V, A, D, E, F, and O had higher percentage (> 83%) compared to waves B and C (< 70%). This was true in all four conditions. This finding is suggestive of influence of condition as a variable in determining the morphology of responses, although elicited by the same stimulus. Because waves B and C were obtained in less than 75% of the ears tested and were present < 50% in some conditions, they were not analyzed on repeated measures ANOVA.

4.2 Results of Onset Response

In the onset response, the peak latency of waves V and A, and the V to A peak to peak amplitude were analyzed. The onset response elicited by the stimulus /da/ across the four conditions in a representative subject is shown in Figure 4.1. The mean and standard deviation (SD) of latency and amplitude obtained from 30 ears across four conditions is given in Table 4.2.

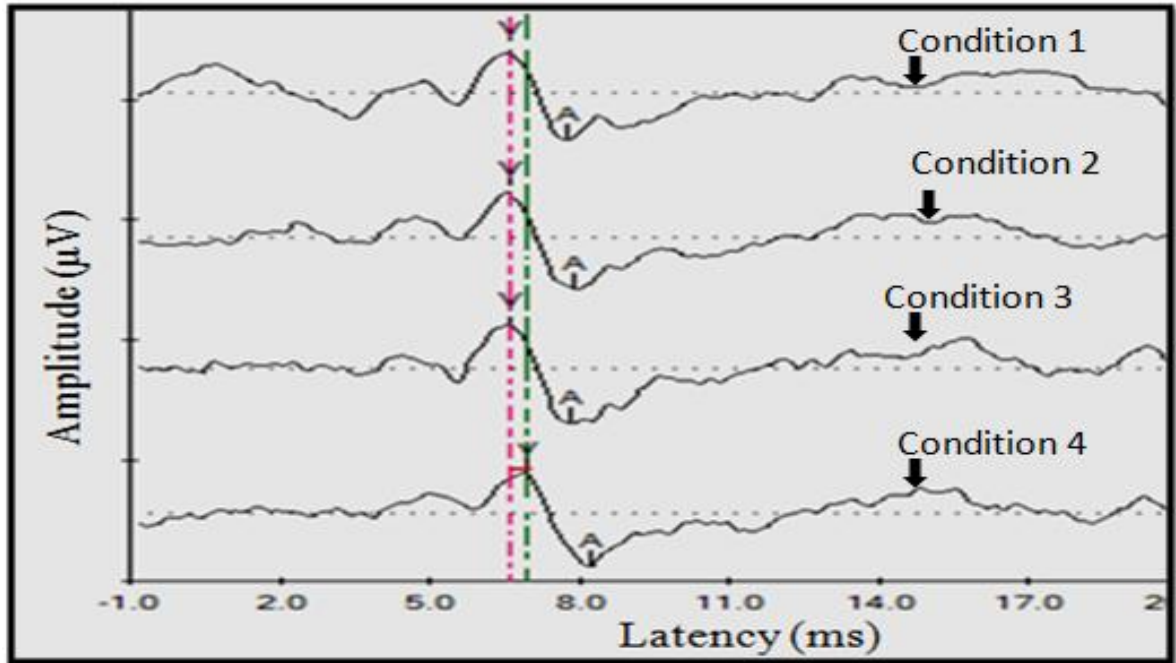


Figure 4.1: Onset responses elicited by syllable /da/ across 4 conditions in a representative subject.

Table 4.2: The mean and standard deviation (SD) of peak latency of wave V and A, and V to A peak to peak amplitude across four test conditions

Condition	Measure					
	Latency of wave V (ms)		Latency of wave A (ms)		Amplitude (V to A) (µV)	
	Mean	SD	Mean	SD	Mean	SD
1	6.39	0.27	7.42	0.34	0.62	0.16
2	6.39	0.29	7.42	0.32	0.59	0.17
3	6.41	0.31	7.50	0.32	0.59	0.19
4	6.75	0.31	7.81	0.44	0.57	0.17

It is evident from Table 4.2 that the mean peak latency was same between condition 1 and 2, while it differed in condition 3 and 4. The mean peak latency of wave

V and A was prolonged in condition 4 compared to conditions 1, 2 and 3. A trend of reduction in V to A amplitude can also be seen in the conditions 2, 3 and 4 relative to that in condition 1.

The data was further subjected to Repeated measures ANOVA, to test if the mean difference observed were statistically significant. Results revealed a significant main effect of condition on the mean latency of wave V [$F(3, 87) = 104.75, p < 0.01$] and wave A [$F(3, 87) = 52.28, p < 0.01$] across the four conditions. But there was no significant difference [$F(3, 87) = 0.97, p > 0.05$] in the mean V to A amplitude across the four conditions.

Because there was a significant main effect of stimulus condition on the latencies of the onset response, pair-wise comparison of the data was done using the Sidak Post-Hoc test. The results of the Post-Hoc test for latency of waves V and A are represented in Table 4.3 and 4.4 respectively.

Table 4.3: *Pair-wise comparison of mean latency of wave V*

Conditions	2	3	4
1	NS	NS	S
2	NS	NS	S
3	NS	NS	S

Note: $S = p < 0.05$; $NS = p > 0.05$

Table 4.4: *Pair-wise comparison of mean latency of wave A*

Conditions	2	3	4
1	NS	S	S
2	NS	NS	S
3	NS	NS	S

Note: $S = p < 0.05$; $NS = p > 0.05$

Results of the Sidak Post-Hoc test can be summarized as follows:

1. There was a significant increase in the mean latency of wave V in the condition 4 compared to that in the conditions 1, 2 and 3. There was no significant difference in the mean latency of wave V obtained in conditions 1, 2 and 3.
2. There was a significant increase in the mean latency of wave A in the condition 4 compared to that in conditions 1, 2 and 3. Also, there was a significant difference in the mean latency of wave A between condition 1 and 3. But there was no significant difference in the mean latency of wave A across other conditions.

4.3 Results of the Sustained Frequency Following Responses (FFR)

The sustained frequency following response was analyzed both subjectively and objectively. The results of the two are reported separately.

4.3.1 Results of the Subjective Analysis of Sustained FFR

The sustained responses were subjectively analyzed to note down the peak latency and amplitude of waves D, E and F were analyzed. The sustained response elicited by the stimulus /da/ across the four conditions in a representative subject is shown in Figure 4.2. The mean and standard deviation of peak latency and amplitude obtained from the subjects across the four test conditions are given in Table 4.5.

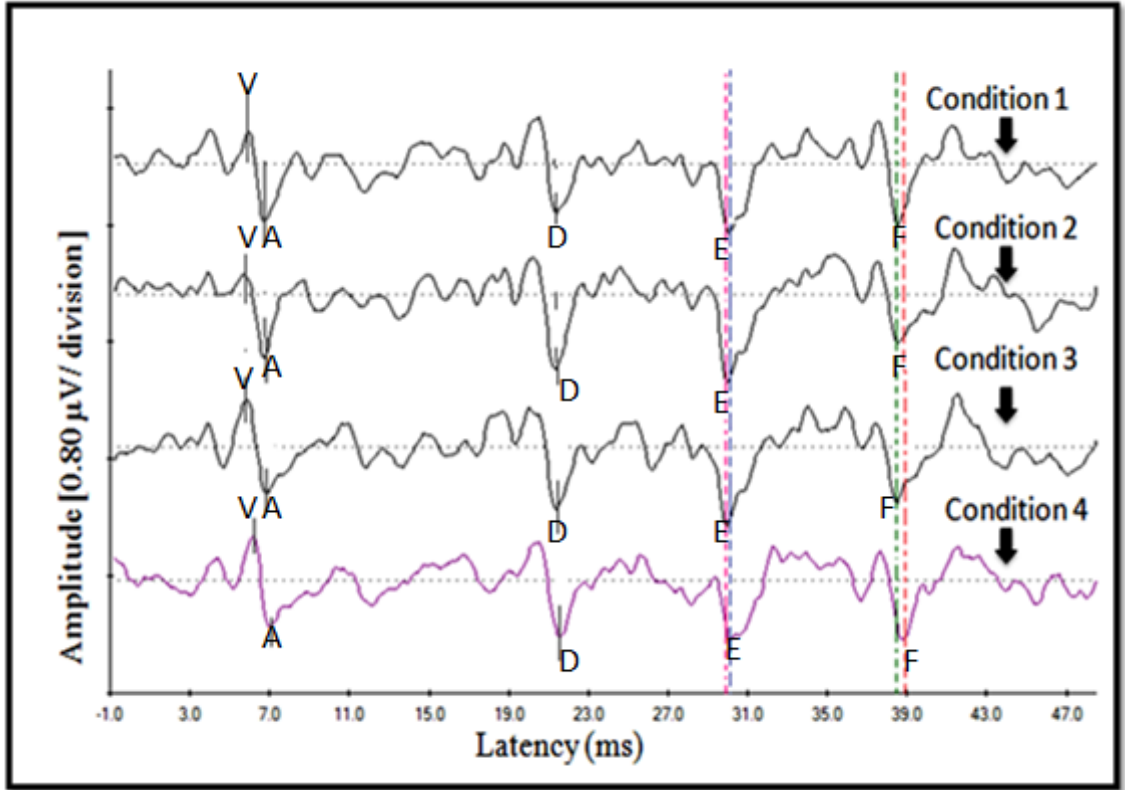


Figure 4.2: Sustained response elicited by syllable /da/ across 4 conditions in a representative subject.

Table 4.5: The mean and standard deviation (SD) of peak latency and peak amplitude of waves D, E and F across four test conditions

Condition	Peak Latency of Wave (ms)			Peak Amplitude of Wave (μ V)		
	D	E	F	D	E	F
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
1	22.51 (0.75)	30.77 (0.45)	39.27 (0.38)	0.23 (0.08)	0.33 (0.09)	0.28 (0.08)
2	22.51 (0.71)	30.86 (0.51)	39.28 (0.43)	0.25 (0.12)	0.32 (0.12)	0.26 (0.12)
3	22.53 (0.66)	30.95 (0.55)	30.30 (0.43)	0.23 (0.10)	0.32 (0.09)	0.26 (0.10)
4	22.74 (0.77)	31.29 (0.74)	39.59 (0.49)	0.22 (0.10)	0.31 (0.10)	0.25 (0.12)

It is evident from the data in Table 4.5 that the mean peak latency of waves D, E and F in condition 1 differed from that in conditions 2, 3 and 4. The mean peak latency of wave D, E and F were prolonged in the conditions 3 and 4 compared to that in condition 1. Additionally, it is evident that the mean peak amplitude of waves D, E and F is reduced in condition 4 compared to that in condition 1.

The data was further subjected to Repeated measures ANOVA, to check if the mean differences observed in latency and amplitude of sustained response were statistically significant. Results revealed the presence of significant main effect of condition on the mean latency of wave E [$F(3, 72) = 9.89, p < 0.01$] and wave F [$F(3, 75) = 9.53, p < 0.01$], while the mean difference in wave D was not statistically significant [$F(3, 60) = 1.75, p > 0.05$]. Additionally, there was no significant difference in the mean amplitude of either wave D [$F(3, 60) = 0.48, p > 0.05$]; E [$F(3, 78) = 0.67, p > 0.05$] or F [$F(3, 69) = 0.57, p > 0.05$].

Because there was a significant main effect of stimulus context on the latency of the waves E and F of the sustained response, pair-wise comparison was done using the Sidak Post-Hoc test. The results of the Post-Hoc test for the mean peak latency of waves E and F are represented in Table 4.6 and 4.7 respectively. The post-Hoc test results are as follows:

- 1) There was a significant increase in the mean peak latency of wave E in condition 4 compared to that in conditions 1 and 2. There was no significant increase in the mean peak latency wave E in condition 4 compared to condition 3. There was no significant difference in the mean peak latency of wave E across conditions 1, 2 and 3.

- 2) There was a significant increase in the mean peak latency of wave F in condition 4 compared to that in the conditions 1, 2 and 3. There was no significant difference in the mean peak latency of wave F across the conditions 1, 2 and 3.

Table 4.6: *Pair wise comparison of mean latency of wave E*

Conditions	2	3	4
1	NS	NS	S
2	NS	NS	S
3	NS	NS	NS

Note: S= $p < 0.05$; NS= $p > 0.05$

Table 4.7: *Pair wise comparison of mean latency of wave F*

Conditions	2	3	4
1	NS	NS	S
2	NS	NS	S
3	NS	NS	S

Note: S= $p < 0.05$; NS= $p > 0.05$

4.3.2 Results of the Objective Analysis of Sustained FFR

Brainstem Toolbox was used to carry out the Fast Fourier Transformation (FFT) of the sustained responses. The peak amplitude at the frequencies corresponding to the fundamental frequency (F0), first formant (F1) and the higher harmonics (HF) of the stimulus was derived from the FFT analysis. The mean and standard deviation (SD) of amplitude of F0, F1 and HF obtained from 30 ears across the four test conditions are given in Table 4.8.

Table 4.8: *The mean and standard deviation (SD) of amplitude of fundamental frequency (F0), first formant (F1) and higher harmonics (HF)*

Conditions	Amplitude of F0 (arbitrary dB)		Amplitude of F1 (arbitrary dB)		Amplitude of HF (arbitrary dB)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
1	5.12	2.29	0.35	0.09	0.14	0.02
2	4.92	2.26	0.40	0.09	0.16	0.02
3	5.52	2.92	0.38	0.08	0.17	0.03
4	5.15	2.67	0.37	0.08	0.18	0.07

It is evident from Table 4.8 that the mean amplitudes of F0, F1 and HF was consistently more in condition 4 than that in condition 1. To test if the mean differences were statistically significant, Repeated measures ANOVA was done. Results showed a significant main effect of condition on the mean amplitude of HF [$F(3, 87) = 5.20, p < 0.05$]. Significant main effect was absent for mean amplitude of F0 [$F(3, 87) = 0.20, p > 0.05$] and F1 [$F(3, 87) = 2.53, p > 0.05$].

Because there was a significant main effect of stimulus context on the amplitude of HF of the sustained response, pair-wise comparison of the responses was done using the Sidak Post-Hoc test. The results of the Post-Hoc test for mean amplitude of HF is shown in Table 4.9.

Table 4.9: *Pair wise comparison of mean amplitude of higher harmonics (HF)*

Conditions	2	3	4
1	S	S	S
2	NS	NS	NS
3	NS	NS	NS

Note: S= $p < 0.05$; NS= $p > 0.05$

Results showed a significant reduction in the mean amplitude of higher harmonics (HF) in condition 1 compared to the mean amplitude of HF in conditions 2, 3 and 4. There was no significant difference in the mean amplitude of HF across conditions 2, 3 and 4.

4.4 Results of Offset Response

In the offset response, the peak latency and amplitude of wave O was analyzed. The mean and standard deviation (SD) of latency and amplitude obtained from 30 ears across four conditions is given in Table 4.10.

Table 4.10: *The mean and standard deviation (SD) of peak latency and amplitude of wave O across four test conditions*

Conditions	Measures			
	Latency of wave O		Amplitude of wave O	
	Mean	SD	Mean	SD
1	48.14	0.44	0.25	0.08
2	48.22	0.44	0.20	0.07
3	48.11	0.51	0.20	0.08
4	48.28	0.45	0.20	0.08

From Table 4.10, it is evident that the mean peak latency of wave O is prolonged in the condition 4 compared to condition 1. Also, the mean peak amplitude is reduced in

condition 4 compared to condition 1. To check if the mean difference observed were statistically significant, the data was further subjected to statistical analysis using Repeated measures ANOVA. Results revealed no significant main effect of condition on either the mean peak latency [$F(3, 60) = 2.35, p > 0.05$] or peak amplitude [$F(3, 63) = 2.67, p > 0.05$] of wave O.

To summarize, results showed that the latencies of onset and sustained responses were prolonged in the stimulus context conditions when compared to repetitive condition. This was particularly and consistently true with condition 4. There were no differences found in the amplitude of the waves across all four conditions. Results of the FFT (HF) on the other hand showed significantly poorer responses in the repetitive condition than the contextual conditions.

Chapter 5

DISCUSSION

In the present study it was hypothesized that context has no effect on the brainstem responses. However, results obtained did not support this hypothesis. Both onset and sustained responses varied depending on the context. In the subsequent subsections, the reasons that could be attributed to the results obtained (details in chapter 4) will be discussed.

5.1 Percentage of occurrence of Onset and Sustained responses

Results of the percentage of occurrence of the onset and sustained responses elicited for the syllable /da/ across the four conditions showed that the waves V and A were invariable across all four condition. However the target waves B, C, D, E, F, and O showed a trend, with higher percentage of occurrence in condition 1 compared to condition 4. In condition 1, the syllable /da/ was presented repeatedly, while in condition 4, between any 2 presentations of /da/, there was more frequently occurring short duration whitenoise (contextual stimulus). The stimulus used to elicit the responses was the same syllable /da/ and number of sweeps of /da/ averaged, was constant in all four conditions. Hence, it is expected that the percentage of occurrence of response should be same across the four conditions. The observed variation between conditions cannot be attributed to trial-trial variability because within each condition, no such variation was observed on multiple trials. That is, if the wave D was present in the first trial, it was present in the successive trials too and vice versa in each condition. The observed decrement in the percentage of occurrence of responses in condition 4 compared to

condition 1 is suggestive of a possible influence of condition as a variable in determination of response morphology. The reasons for these differences for individual waves, across all four conditions, would be discussed in the following sections.

5.2 Effect of Stimulus Condition on the Onset responses

The mean peak latency of the waves V and A were found to be significantly prolonged in condition 4 compared to condition 1, 2 & 3. Also, the mean peak latency of wave A was found to be significantly prolonged in condition 3 relative to condition 1.

One possible explanation for delayed onset response in the context of noise could be attributed to the effect of forward masking of the stimulus /da/ by the white noise. Noise induced disruptions in the transient responses elicited for the stimulus /da/ was reported by Russo Nicol, Musacchia and Kraus 2004. The latency delay reported in the study by Russo and colleagues (2004) in presence of background noise was 0.53 ms for wave V and 0.87 ms for wave A, while the delay in latency in context of noise was 0.36 ms for wave V and 0.39 ms for wave A in the present study. However, the findings of the present study cannot be compared to the study by Russo and colleagues (2004), as the noise was presented as a background stimulus in their study and as a contextual stimulus in the present study. Moreover, the effect of forward masking is dependent on the masker duration and the gap between the masker and signal (Elliott, 1962 & Elliott, 1967). The effect of forward masking is reported to last for 30 ms following the end of the masker relative to the beginning of the maskee (Howard & Angus, 2009). In the present the masker duration is very brief of 40 ms and the gap between the masker and

signal is approximately around 64 ms, hence forward masking cannot account for the results in the present study.

Another possible reason for delay of the onset response in condition 4 relative to the other conditions could be due to the effect of noise *as the contextual stimulus*. The disturbance in the repeated representation of stimulus /da/ may be weakening the online plasticity of the brainstem physiology. If the same stimulus is repeated, then the response is enhanced due to online plasticity, regulated by the corticofugal pathway. This inference is based on the findings of Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009), in which they reported enhanced harmonics when stimulus was presented repeatedly.

The corticofugal pathway originates at the auditory cortex and forms multiple feedback loops to modulate auditory signal processing at the brainstem nuclei such as the inferior colliculus (IC) and medial geniculate body (MGB). The auditory cortex and the corticofugal pathway evoke small short term changes in the subcortical nuclei, in response to a sound that is repeatedly delivered (Suga, Xiao, Ma & Ji, 2002). These changes are reported to be specific to the parameters characterizing the sound. That is, when the sound becomes relevant by associative learning, its responses are enhanced by the activation of the corticofugal pathway.

In this case, because the contextual stimulus is noise, it is probably deemed as a dissimilar and insignificant stimulus on associative learning, and since it is also causing interruption in the repetitive presentation of the stimulus, the corticofugal pathways are not facilitating the online plasticity. As a result, there may be reduction in the

synchronous firing which is evident in the form of increase in the latency and reduction in the amplitude of onset responses in condition 4 compared to condition 1.

However, if the interruption in repetitive sequence was the cause for delayed responses, one would expect similar delays in condition 2 and 3, as that found in condition 4 compared to condition 1. But the findings of the present study show that the latency of the onset response V and A is significantly delayed in condition 4 and not significantly delayed in condition 2 and 3 compared to condition 1. This is probably because, corticofugal pathway is able to distinguish between speech and non-speech stimulus. The online plasticity is facilitated if the target speech stimulus occurs in the context of speech but not in the context of noise.

In the present study, no significant delay was observed in the latency of either wave V or A in condition 2 relative to the condition 1, while the latency of wave A was prolonged in condition 3 relative to condition 1. The possible explanation for these variations of the brainstems behavior to contextual speech stimuli could be the spectral proximity of the contextual speech stimulus to the core stimulus /da/. The stimulus /ta/ used as contextual stimulus in condition 2 is spectrally equivalent to the stimulus /da/, relative to /ba/. Only the temporal feature of pre-voicing is different between /ta/ and /ba/. There is no pre-voicing in /ta/ whereas it is present in /ba/. But, the syllable /ba/ differs spectrally from /da/, that is with respect to the burst spectrum and F2 transition. The burst spectra of /ba/ has primary concentration of energy at the low frequencies (500-1500Hz) while /da/ has a relatively flat or high frequency (>4000Hz) concentration (Halle, Hughes & Radley, 1957). The F2 transition in /ba/ is falling in nature while it is reported to be rising in /da/ (Delattre, Liberman, & Gerstman, 1954).

Therefore it is possible that the corticofugal modulation is capable of detecting and discriminating the spectral variations between the core stimulus and the contextual stimulus, and thus influencing the response of the brainstem. On the other hand, the corticofugal pathway may not be sensitive to temporal variation between two speech stimulus such as /ta/ and /da/, and hence does not influence the brainstem's response. Hence, the spectral similarity may be playing a role while higher centers interpret the context. However, the reasons for observing such differences only in wave A are not clear.

No significant effect of context on the latency of wave V and A was obtained in the study by Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009). This result is partly in line with the results obtained in the present study, wherein significant effect of context on onset response was not found between condition 1 and condition 2. However in the present study, context was found to influence the onset response in condition 3. This result in the present study has been attributed to the hypothesis that the discrimination of the spectral deviances between the core stimulus and contextual stimulus by the corticofugal pathway determines the brainstem response.

The possible reason for the absence of this contextual effect on the latencies of the onset response could be due to the stimulus paradigm adopted in their study. In their study the stimulus /da/ was presented with the probability of 12.5% in context of seven speech sounds, while in the present study /da/ was presented with a probability of 33% in context of only one speech or noise stimulus. It is possible that the number of contextual stimulus used would have an effect on the corticofugal network. The present study

utilized just one contextual stimulus, while multiple (seven) contextual stimuli were used in the study by Chandrasekaran et al.,(2009).

Another reason could be the differences in the length of the stimulus used; it was 170 ms in the study by Chandrasekaran et al.,(2009) while it was 40ms in the present study. This is suggestive that the brainstem's capacity to differentiate two signals may be limited, such that it can discriminate two closely occurring signals and not signals which are separated by longer time duration.

Hence it can be concluded that the latencies of the onset response, that is the waves V and A are influenced by the corticofugal system, as the generators (CN, IC and LL) of these waves are within the feedback loop of the corticofugal pathway (Chandrasekaran & Kraus, 2010). Moreover, the brainstem response is modulated by the ability of the corticofugal pathway to identify spectral differences between the core stimulus and contextual stimulus. The spectrally similar contextual stimulus does not influence the latency of the onset response while the spectrally dissimilar contextual stimulus influences the latency of the onset response. The influence of spectrally dissimilar contextual stimulus is found to prolong the latency of the onset response.

5.3 Effect of Stimulus Condition on the Sustained Frequency Following Response (FFR)

The sustained response was analyzed both subjectively and objectively.

5.3.1 Effect of stimulus condition on Subjective measures of Sustained (FFR)

The results of the subjective analysis of the FFR showed a significant prolongation in the latency of wave E in condition 4 compared to condition 1 and 2.

Similarly, significantly prolonged latency of wave F was obtained in the condition 4 relative to condition 1, 2 and 3.

The possible reason for delay of wave F in condition 4 relative to the other conditions and prolongation of wave E in condition 4 compared to conditions 1 and 2, could be due to the effect of noise as the contextual stimuli. The generators of waves E and F (IC and LL) fall within the network of the corticofugal pathway (Marsh, Brown & Smith, 1974). The disturbance in the repeated representation of stimulus /da/ would possibly cause reduction in the online plasticity of the brainstem.

Moreover, the spectrum of noise is quite different from that of speech. As hypothesized if the corticofugal pathway is capable of detecting the spectral variations between the speech and contextual noise stimulus, while modulating the brainstem response, then the delay in latency of waves E and F in condition 4 can be attributed to the influence of corticofugal pathway on the brainstem. Thus, it is suggested that the corticofugal pathway is inherently assesses the spectral variation between two incoming signals.

In the present study no significant difference was obtained in the wave E latency in condition 3 and condition 4. The contextual stimulus used in condition 3 is the syllable /ba/ and; in condition 4 it is noise. Like the noise, the syllable /ba/ is spectrally different from the stimulus /da/. Hence, the corticofugal network may identify this difference and thus influence the brainstems response.

Hence it can be concluded that the spectral difference between the core stimulus and contextual stimulus influences the latency of frequency following response (FFR) via

the corticofugal pathway. The spectrally similar contextual stimulus does not degrade the latency of FFR evoked by speech stimulus while the spectrally dissimilar contextual stimulus influences the latency of FFR. The influence observed in the presence of a spectrally dissimilar stimulus is the prolongation of the latencies of the FFR, especially the waves E and F.

5.3.2 Effect of Stimulus Condition on the Objective Measures of Sustained FFR

The peak amplitude at the frequencies corresponding to the fundamental frequency (F0), first formant (F1) and the higher harmonics (HF) of the stimulus was derived from the FFT analysis. In the present study, the peak amplitude of the higher harmonics (HF) corresponding to (721-1155Hz) was found to be significantly lower in the condition 1 compared to all other conditions.

However, these findings are not in accordance to that found in the study by Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009), wherein the peak amplitudes corresponding to lower harmonics H2 and H4 (200Hz and 400Hz) and F1 range (400-720 Hz) differed between the repetitive and variable context conditions. The spectral amplitude of H2 and H4 and over the F1 range was enhanced in the repetitive condition compared to the variable context condition. The authors attributed their findings to the corticofugal modulation of the brainstem. The exact reason for the reduction in amplitude of HF in the present study is not clear.

5.4 Effect Stimulus Condition on the Offset Response

In the present study, both the peak latency and peak amplitude of the offset response, wave O, elicited to speech stimulus /da/ did not show a significant difference

between the conditions. This finding is suggestive that the stimulus context does not affect the coding of the offset response at the brainstem. This finding is in congruence with the earlier study by Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009), who demonstrated no significant effect of stimulus context on the latency and amplitude of the offset response. This may be probably because the generators of wave O may not be well connected by the corticofugal pathway.

Chapter 6

SUMMARY AND CONCLUSIONS

In the present study, true experimental design was used to assess the effect of context on brainstem encoding of speech. To do so, the difference in the brainstem response elicited to a repeated stimulus was compared to that elicited by the same stimulus presented in the context of other stimuli. The present method used both speech as well as noise stimuli as contextual stimulus.

Brainstem responses were obtained from both ears of 15 normal hearing subjects (mean age of 21.2 years), to a synthetically generated syllable /da/ in four different conditions, which included one repetitive condition and in three different context conditions. In condition 1, a repetitive stimulus paradigm was used, wherein brainstem responses were elicited only for the stimulus /da/. Variable stimulus context paradigm was used in conditions 2, 3, and 4, wherein responses were elicited for the stimulus /da/ presented infrequently in the context of another frequent stimulus. In condition 2, the frequent stimulus used was a synthetically generated syllable /ta/. Similarly, synthetically generated syllable /ba/ was used as the frequent stimulus in condition 3, while white noise was used as the frequent stimulus in condition 4.

Results of the study showed that the latencies of onset and sustained responses were prolonged in the stimulus context conditions when compared to repetitive condition. This was particularly and consistently true with condition 4. There were no differences found in the amplitude of the waves across all four conditions. Results of the FFT (HF)

on the other hand showed significantly poorer responses in the repetitive condition than the contextual conditions.

The ability of the corticofugal pathway to identify spectral differences between the core stimulus and contextual stimulus is hypothesized to influence the latencies of the brainstem responses. This is because the generators of the onset and the sustained responses (CN, LL and IC) falls within the feedback loop of the corticofugal pathway. Hence, the latencies on the brainstem response to speech are influenced by the online plasticity regulated by corticofugal network.

It was observed that the presence of a spectrally similar contextual stimulus does not influence the latency of brainstem response while the spectrally dissimilar contextual stimulus influences the latency of the brainstem response. In the current study the influence of spectrally dissimilar contextual stimulus is found to prolong the latency of the brainstem responses. However, the exact reason for the reduction in amplitude of HF in the present study is not clear.

The results of this study are not in accordance to the previous study assessing the effect of context on speech ABR by Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009). The differences in the results between the two studies could be attributed to the differences in stimulus paradigm. The number of contextual stimulus used would have an effect on the corticofugal pathway. The present study utilized just one contextual stimulus, while, in the study by Chandrasekaran et al., (2009) multiple (seven) contextual stimuli were used.

Implications of the Findings

- 1) The results of the present study threw more light on the mechanisms involved in corticofugal modulation of brainstem physiology.
- 2) It also helps in understanding the role of the brainstem in speech processing and perception, particularly in the presence of noise.

Future Topics for Investigation

- 1) Effect of contextual stimulus either presented dichotically or contralesionally could be studied. This would help in understanding the role olivocochlear bundle, and laterality of corticofugal network in modulation of brainstem.
- 2) Similar stimulus paradigm could be used to study the influence of the corticofugal pathway on the plasticity of brainstem in individuals known to have abnormal auditory processing such as in individuals with maturation delay, learning disability, and autism. Further, this paradigm could be used to compare the effect of training on the plasticity of the brainstem structure.

REFERENCES

- Achor, L. J., & Starr, A. (1980). Auditory brainstem response in the cat: Intracranial and extracranial recordings. *Electroencephalography and Clinical Neurophysiology*, 48, 154-173.
- Akhoun, I., Gallego, S., Moulin, A., Menard, M., Veuillet, E., Berger-Vachon, C., Collet, L., & Thai-Van, H. (2008). The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. *Clinical Neurophysiology*, 119, 922–933.
- Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010). Neural Timing is Linked to Speech Perception in Noise. *Journal of Neuroscience*, 30, (14), 4922-4926.
- Banai, K., Nicol, T., Zecker, S., & Kraus, N. (2005). Brainstem timing: Implications for cortical processing and literacy. *Journal of Neuroscience*, 25, (43), 9850-9857.
- Banai, K., Hornickel, J. M., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral Cortex*, 19, (11), 2699-2707.
- Bent, T., Bradlow, A.R., & Wright, B.A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and non-speech sounds. *Journal of Experimental Psychology: Human Perception & Performance*, 32, 97–103.

- Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., & Kraus N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: Implications for developmental dyslexia. *Neuron*, *64*, 311-319.
- Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*, *47*, 236–246.
- Chowdhury, S. A., & Suga, N. (2000). Reorganization of the frequency map of the auditory cortex evoked by cortical electrical stimulation in the big brown bat. *Journal of Neurophysiology*. *83*, 1856–1863.
- Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. *Clinical Neurophysiology*, *112*, 758-767.
- Cunningham, J., Nicol, T., King, C. D., Zecker, S. G., & Kraus, N. (2002). Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. *Hearing Research*, *169*, 97-111.
- Delattre, P. C., Liberman, A. M., & Gerstman, L. J. (1954). The role of selected stimulus-variables in the perception of stop and nasal consonants. *Psychological Monographs*, *68*, (8), 379.
- Elliott, L. L. (1962). Backward masking: Monotic and dichotic conditions. *Journal of the Acoustical Society of America*, *34*, (8), 1108-1115.
- Elliott, L. L. (1967). Development of auditory narrow band frequency controls. *Journal of the Acoustical Society of America*, *42*, (1), 143-153.

- Galbraith, G. C. (1994). Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. *Electroencephalography and Clinical Neurophysiology*, *92*, 321–330.
- Galbraith, G. C., Amaya, E. M., Diaz de Rivera, J. M., Donan, N. M, Duong, M. T., Hsu J. N., Tran, K., & Tsang, L. P. (2004). Brain stem evoked response to forward and reversed speech in humans. *NeuroReport*, *15*, (13), 2057–2060.
- Gao, E., & Suga, N. (2000). Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: Role of the corticofugal system. *Proceedings of the National Academy of Sciences*, *97*, 8081–8086.
- Gardi, J., Merzenich, M., & McKean, C. (1979). Origins of the scalp recorded frequency-following response in the cat. *Audiology*, *18*, 358–381.
- Gonçalves, I. C., Wertzner, H. F., Samelli, A. G., & Matas, C. G. (2011). Speech and non-speech processing in children with phonological disorders: an electrophysiological study. *Clinics*, *66*, (2), 293-298.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*, 14–23.
- Hall, J. W. (2006). *New Handbook of Auditory Evoked Responses*. Boston, Allyn and Bacon.
- Halle, M., Huges, G. W., & Radley, J. P. (1957). Acoustic properties of consonants. *Journal of Acoustical Society of America*, *29*, (1), 107-116.

- Hashimoto. (1982). Auditory evoked potentials from the human midbrain: Slow brainstem responses. *Electroencephalography and Clinical Neurophysiology*, 53, 652-657.
- Hood, L. J. (1998). *Clinical Applications of the Auditory Brainstem Response*. San Diego, CA: Singular Publication Group.
- Hoormann, J., Falkenstein, M., Hohnsbein, J., & Blanke, L. (1992). The human frequency-following response (FFR): Normal variability and relation to the click-evoked brainstem response. *Hearing Research*, 59, 179–188.
- Hornickel, J. M., Skoe, E., & Kraus N. (2009). Subcortical lateralization of speech encoding. *Audiology Neurotology*, 14, 198-207.
- Howard, D. M., & Angus, J. A. S (.2009). *Acoustics and Psychoacoustics*. (4th edition), UK; Linacre House, Jordan Hill.
- Jerger, J. (1970). Clinical experience with impedance audiometry. *Archives of Otolaryngology*, 92, 311- 324.
- Jewett, D. L., Romano, N. M., & Williston, J. S. (1970). Human auditory evoked potentials: possible brain stem components detected on the scalp. *Science*, 167, 1517–1518.
- Jewett, D.L., & Williston, J.S. (1971). Auditory-evoked far fields averaged from the scalp of humans. *Brain*, 94, 681-696.

- Johnson, K. L., Nicol, T., & Kraus, N. (2005). The brainstem response to speech: a biological marker. *Ear and Hearing, 26*, (5), 424-433.
- Johnson, K.L., Nicol, T., & Kraus, N. (2008). Developmental plasticity in the human auditory brainstem. *Journal of Neuroscience, 28*, (15), 4000-4007.
- Keuroghlian, A. S., & Knudsen, E. I. (2007). Adaptive auditory plasticity in developing and adult animals. *Progress in Neurobiology, 82*, 109– 121.
- Khaladkar, A. A., Karthik, N., & Vanaja, C. S. (2006). Speech burst and click evoked ABR. http://www.audiologyonline.com/articles/article_id=1373.
- King, C., Warrier, C. M., Hayes, E., & Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neuroscience Letters, 319*, 111–115.
- Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America, 67*, 971-995.
- Kral, A., & Eggermont, J. (2007). ‘What’s to lose and what’s to learn: Development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Research Reviews, 56*, 259– 269.
- Kraus, N., & Nicol, T. (2005). Brainstem origins for cortical "what" and "where" pathways in the auditory system. *Trends in Neurosciences, 28*, 176-181.
- Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Research, 25*, 161–168.

- Krishnan, A. (2007). Frequency-Following Response. In R. F. Burkard, J. J. Eggermont, M. Don (Eds.). *Auditory Evoked Potentials: Basic Principles and Clinical Application* (pp. 313–335). Philadelphia, PA: Lippincott Williams & Wilkins.
- Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experiencedependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience*, *21*, 1092–1105.
- Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*, *110*, 135–148.
- Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experience dependent Enhancement of Linguistic Pitch Representation in the Brainstem is Not Specific to a Speech Context. *Journal of Cognitive Neurosciences*, *21*, 1092-1105.
- Krizman, J., Skoe, E., & Kraus, N. (2010). Stimulus rate and subcortical auditory processing of speech. *Audiology Neurotology*, *15*, 332-342.
- Lou, F., Wang, Q., Kashini, A., & Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *Journal of Neuroscience*, *28*, 11615-11621.
- Ma, X., & Suga, N. (2008). Corticofugal modulation of the paradoxical latency shifts of inferior collicular neurons. *Journal of Neurophysiology*, *100*, 1127–1134.
- Madhok, P., & Sandeep, M. (2010). Neurophysiological consequence of auditory training: Subcortical and cortical structures. *Unpublished dissertation*, University of Mysore, India.

- Marsh, J. T., Brown, W. S., & Smith, J. C. (1974). Differential brainstem pathways for the conduction of auditory frequency following responses. *Electroencephalography and Clinical Neurophysiology*, 36, 415–424.
- Mason, S., McCormick, B., & Wood, S. (1988). Auditory brainstem response in paediatric audiology. *Archives of Disease in Childhood*, 63, 465-467.
- Moller, A. R., & Jannetta, P. (1985). Neural generators of the auditory brainstem response. In Jacobsen J, The auditory brainstem response, (pp. 13-32), San Diego, CA: College Hill Press.
- Moushegian, G., Rupert, A. L., & Stillman, R. D. (1973). Scalp-recorded early responses in man to frequencies in speech range. *Electroencephalography and Clinical Neurophysiology*, 35, (6), 665–667.
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Science*, 104, (40), 15894-15898.
- Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials. I: Evaluation of components. *Electroencephalography and Clinical Neurophysiology*, 36, 179-190.
- Rout , A., & Yathiraj, A. (1996). Perception of monosyllabic words (English) in Indian children. *Unpublished dissertation*, University of Mysore, India.
- Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology*, 115, 2021-2030.

- Russo, N., Nicol, T., Zecker, S., Hayes, E., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioral Brain Research, 156*, 95-103.
- Russo, N. M., Bradlow, A. R., Skoe, E., Trommer, B. L., Nicol, T., Zecker, S., & Kraus, N. (2008). Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clinical Neurophysiology, 119*, (8), 1720-1731.
- Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences, 26*,(2), 100–107.
- Skoe, E., & Kraus, N. (2010a). Auditory Brain Stem Response to Complex Sounds: A Tutorial. *Ear and Hearing, 31*, (3), 0-0.
- Skoe, E., & Kraus, N. (2010b). Hearing it again and again: on-line subcortical plasticity in humans. *PlosONE, 5*(10), e13645.
- Sohmer, H., & Feinmesser, M. (1967). Cochlear action potentials recorded from the external ear in man. *Annals of Otology, Rhinology and Laryngology, 76*, 427-435.
- Sohmer, H., Pratt, H., & Kinarti, R. (1977). Sources of frequency following responses (FFR) in man. *Electroencephalography and Clinical Neurophysiology, 42*, 656–664.
- Song, J. H., Banai, K., Russo, N. M., & Kraus, N. (2006). On the relationship between speech and nonspeech evoked auditory brainstem responses. *Audiology and Neuro-Otology, 11*, 233-241.

- Song, J. H., Skoe, E., Wong, P. C. M., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, *20*, (10), 1892-1902.
- Starr A., & Don, M. (1988). *Brainstem potentials evoked by acoustic stimuli*; in Picton TW: *Handbook of Electroencephalography and Clinical Neurophysiology* (pp. 97–150), Amsterdam, Elsevier.
- Stockard, J. J., & Rossiter, M.A. (1977). Clinical and pathologic correlates of brainstem auditory response abnormalities. *Neurology*, *27*, 316-325.
- Strait, D. L., Skoe, E., & Kraus, N., & Ashley, R. (2009). Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience*, *29*, 661-668.
- Suga, N., Xiao, Z., Ma, X., and Ji, W. (2002). Plasticity and corticofugal modulation of hearing in adult animals. *Neuron*, *36*, 9–18.
- Suga, N. (2008). Role of corticofugal feedback in hearing. *Journal of Comparative Physiology. A, Sensory, Neural, and Behavioral Physiology*, *194*, 169–183.
- Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron*, *62*, 463-469.
- Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of lowprobability sounds by cortical neurons. *Nature Neuroscience*, *6*, 391–398.
- Vander Werff, K. R., & Burns, K. S. (2010). Brain Stem Responses to Speech in Younger and Older Adults. *Ear and Hearing*, *31*, (6), 761-768.

- Wible, B., Nicol, T., & Kraus, N. (2004). Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. *Biological Psychology*, *67*, 299-317.
- Winer, J. A. (2005). Decoding the auditory corticofugal systems. *Hearing research*, *207*, 1-9.
- Worden, F., & Marsh, J. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography and Clinical Neurophysiology*, *25*, 42-52.
- Wong, P., Skoe, E., Russo, N., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420-422.
- Xu, Y., Krishnan, A., & Gandour, J. T. (2006). Specificity of experiencedependent pitch representation in the brainstem. *NeuroReport*, *17*, 1601-1605.
- Zhang, Y., Suga, N., & Yan, J. (1997). Corticofugal modulation of frequency processing in bat auditory system. *Nature*, *387*, 990-903.
- Zhang, Y., & Suga, N. (2005). Corticofugal feedback for collicular plasticity evoked by electric stimulation of the inferior colliculus. *Journal of Neurophysiology*, *94*, 2676-2682.