

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

Shruthi G. N.

Register No.: 16AUD028



A Dissertation Submitted in Part Fulfillment of Degree of Master of Science
(Audiology)
University of Mysore
Mysuru

ALL INDIA INSTITUTE OF SPEECH AND HEARING
MANASAGANGOTHRI, MYSURU-570 006

April, 2018

CERTIFICATE

This is to certify that this dissertation entitled '**Effect of noise on context dependent brainstem encoding of speech**' is a bonafide work submitted in part fulfillment for degree of Master of Science (Audiology) of the student Registration Number: 16AUD028. 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.

Mysuru

April, 2018

Dr. S. R. Savithri,

Director

All India Institute of Speech and Hearing

Manasagangothri, Mysuru-570006

CERTIFICATE

This is to certify that this dissertation entitled '**Effect of noise on context dependent brainstem encoding of speech**' has been prepared under my supervision and guidance. It is also certified that this dissertation has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru
April, 2018

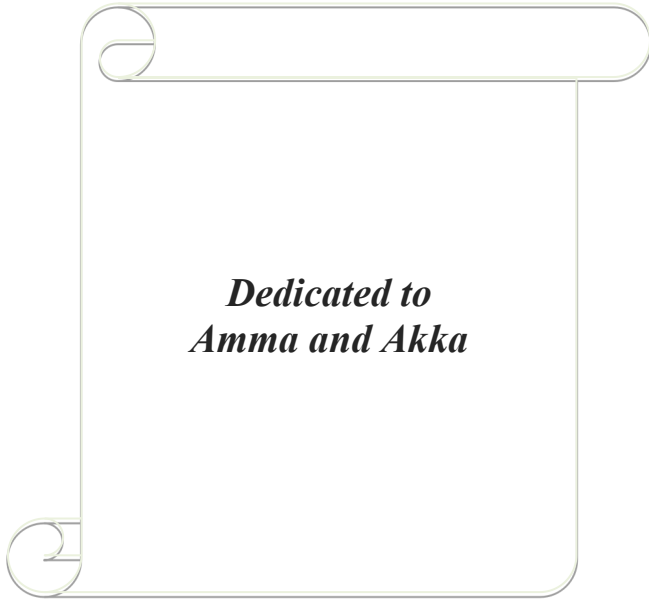
Guide
Dr. Sandeep M.
Reader in Audiology,
All India Institute of Speech and Hearing
Manasagangothri, Mysuru-570006

DECLARATION

This is to certify that this dissertation entitled '**Effect of noise on context dependent brainstem encoding of speech**' is the result of my own study under the guidance of Dr. Sandeep M, Reader in Audiology, Department of Audiology, All India Institute of Speech and Hearing, Mysuru, and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

*Mysuru,
April, 2018*

Registration No. 16AUD028



*Dedicated to
Amma and Akka*

ACKNOWLEDGEMENTS

First and Foremost I would like to express my sincere gratitude to my guide Dr. Sandeep M. Sir who is the main reason why I am still here and for my dissertation completion. Sir without you I wouldn't have done anything. Thank you sir for all your support and patience and also forgiving me for many of my mistakes and giving me another chance. You have always inspired me with your personality and I feel privileged to have you as my mentor and role-model.

I would like to thank Nike sir for his tremendous support. I can't thank you enough for all your efforts for helping me complete my dissertation. Your knowledge and foresight inspired me throughout.

I thank Dr. S.R Savithri, Director, All India Institute of Speech and Hearing, Mysuru for permitting me to carrying out my dissertation.

I thank Dr. Sujith Sinha, HOD, All India Institute of Speech and Hearing, Mysore for permitting me to use the department.

I would also like to thank all the lecturers and staff of Audiology, All India Institute of Speech and Hearing, for helping and guiding me throughout my Master's course.

Many thanks to Anoop Sir, Shreyank Sir and Subha Tak ma'm for opening the department in the evenings and even on Saturdays and Sundays. Without your help, I would never have finished my data collection.

I thank Rajalakshmi ma'm, Sharath sir, Vikas sir for their valuable inputs.

This is only for YOU MUMMY, even though you are not there with me physically, I can feel your presence and I know you are blessing me each and every minute. I will do everything you wanted me to do maa. I love you more than anything else in this world, I love you more than anybody else, I need you more and more mommmy, Miss you each and every minute.

I am never complete without you dear. Thank you soulmate for being there for me always. You have been like my mother throughout. Thank you for always listening to me, supporting me and encouraging me throughout the years. I am glad you are my sister. Thank you so much.

I thank to my world.... Aanya, Vihaan and Jeremy...when I did not have a reason to smile you guys came into my world...Aanya I see my mom in you...you made my life better...thank you..

I thank you Ramesh anna for encouraging me throughout and you have been inspiring me with your personality and I am so much thankful for having you as my brother.

Thank you Pappa and Puni for being there for me.

I thank kappe, the best thing i have ever got. Thanks for being there for these 22 years and thanks for helping me to do my dissertation work.

Thanks to these five, Chai, Anu, Suppu, Deka, Megha, and neha for making my life beautiful and for your constant support.

Special thanks to these three idiots, Anu, deka and chai for helping me in everything, love you the most.

My friends who I want to thank for their unconditional support and motivation. Nikhil, Shezeen, Vishali, Varsha, Rakshith, Aishu, Adi, Chandu, Vishwa, Harish, Prithivi, Sabee, Suju, Pinki, Rashmi, Jyo, and so on.

I thank you manoj, for tolerating me and for the constant motivation, you have been the best.

I would love to thank you pavana akka , you have been helping me in every difficult step of mine, love you the most. Thanks for everything pavu, thank you so much.

I like to thank you preethu for being there for me, for reminding about my dissertation each and every day and thanks for encouraging me to do all my works.

I thank to my sister, shanthu, Nobody else could ever understand the pain of what I was going through, you were there for me throughout and I am proud to say that you are my first and the best senior ever.

I thank Jay , who kept on motivating to do things and help me in each and every step. Thank you Jay, thanks alot.

I am thankful to the best juniors I have ever got. Thank you Bhoomi, Kasthuri, Rohini Liki, Suppi, Rashmi, Sushmitha, Manasa, Harshitha. I would definitely miss you guys, love you so much.

I thank Keerthi akka, Meenu di, Priya akka, Ankitha, Akshaya akka, Shubha ganga akka, Robina di for your support.

I am thankful to all my participants of this study, without whom my study was impossible to complete. I thank them for being patient enough and spending their valuable time for the study.

I would like to thank for all my classmates, my dear juniors and seniors.

There are numerous other people who may have been missed out, but your contributions directly or indirectly are never trivial.

TABLE OF CONTENT

SL. No	Chapter	Page No.
1	INTRODUCTION	1-4
2	REVIEW OF LITERATURE	5-13
3	METHOD	14-22
4	RESULTS	23-36
5	DISCUSSION	37-44
6	SUMMARY AND CONCLUSIONS	45-46
	REFERENCES	47-52

LIST OF TABLES

Table No.	Title of the Table	Page No.
3.1	Spectral characteristics of syllables, /bi/, /bu/, /gi/ & /da/	17
3.2	Stimulus and acquisition parameters used to record FFRs	21
4.1	Harmincs, F value, Degree of Freedom, Error factor, significance value, Partial Eta Square and Observed power ^a in condition , noise and the interaction between noise and condition	24
4.2	Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) of FFR recorded in repetitive paradigm in quiet and noise conditions	26
4.3	Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) of FFR in variable paradigm in quiet and noise conditions	27
4.4	Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) in quiet condition	29
4.5	Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) in noise condition	30
4.6	Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) of context dependent encoding in quiet and noise conditions	33

LIST OF FIGURES

Figure No.	Title of the Figure	Page No.
3.1	The waveforms and Spectrograms of syllables /bi/, /bu/, /gi/ and /da/, generated and used in the present study	16
3.2	Nineteen bisyllables displayed on the computer screen during SNR-50 estimation using Smriti-Shravan module	19
3.3	Representation of different stimulus conditions used in present study	21
4.1	The grand average spectra of FFR in repetitive paradigm in quiet and noise conditions (WON - Without noise condition, WN - With noise condition)	25
4.2	The grand average spectra of FFR in variable paradigm in quiet and noise conditions (WON - Without noise condition, WN - With noise condition)	27
4.3	The grand average spectra of FFR in repetitive and variable paradigms in quiet	28
4.4	The grand average spectra of FFR in repetitive and variable paradigms in noise condition	29
4.5	H1 amplitude of individual participants in repetitive and variable paradigms, in quiet and noise conditions	30
4.6	H2 amplitude of individual participants in repetitive and variable paradigms, in quiet and noise conditions	31
4.7	H3 amplitude of individual participants in repetitive and variable paradigms, in quiet and noise conditions	32
4.8	H4 amplitude of individual participants in repetitive and variable paradigms, in quiet and noise conditions	32

4.9	Scatter plots depicting SNR-50 in dB (Y-axis) and Harmonics in uV in quiet (X-axis). p-significant value, r-correlation coefficient, A-H1, B-H2, C-H3 and D-H4	35
4.10	Scatter plots depicting SNR-50 in dB (Y-axis) and Harmonics in uV in noise condition (X-axis). p-significant value, r-correlation coefficient, A-H1, B-H2, C-H3 and D-H4	36

Chapter 1

INTRODUCTION

Human communication rarely occurs in optimal listening environments, rather we are often surrounded by background noise. Despite the frequent presence of noise, humans are remarkably adept at disentangling target sounds from a complex soundscape. A key mechanism thought to underlie accurate perception in noise is the auditory system's ability to extract regularities from an ongoing acoustic signal (Chandrasekaran, Hornickel, Skoe, Nicol & Kraus, 2009; Winkler, Denham & Nelken, 2009).

Recordable 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. Though the scalp recorded onset response and the FFR reflect the activity at numerous sources (LL, CN, IC), they are noninvasive tools to study the subcortical 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 allows for a means to examine how plasticity occurs at the level of brainstem in humans.

FFR, a component of the auditory brainstem response, reflects neural phase-locking to F0 and its harmonics (Chandrasekaran & Kraus, 2010). The FFR closely mimics the incoming signal. When FFR waveform recorded in response to words is played back, subjects could identify the words with greater-than-chance accuracy (Galbraith et al., 1997). Studies have demonstrated that FFR can serve as an

index of long-term and training-related plasticity (Krishnan & Gandour, 2009; Krishnan et al., 2005). Native speakers of a tonal language in which changes to voice pitch alone can change word meaning, representation of the voice pitch in the brainstem responses were more robust than that in nonnative speakers.

Recently, brainstem responses have been shown to represent context-dependent encoding of the stimulus (Chandrasekaran et al., 2009). This was done by comparing the FFRs elicited in two different stimulus contexts: a predictable context versus a highly variable context. Chandrasekaran et al. (2009) showed that repetition induces improved neural representation of cues that are relevant for perceiving voice pitch, an important cue for segregating sound sources in noisy environments. Importantly, repetition-induced plasticity in representation of voice pitch was strongly associated with behavioral performance on speech-in-noise tests. This result suggests that the ability to fine-tune brainstem encoding of repeating elements in the auditory environment is important for speech-in-noise perception.

In a similar study done by Tonse (2011), 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 stimulus). 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, et al. (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 was 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. The result revealed that stimulus context influences the neural processing of speech at the brainstem and such influences are determined by the spectral differences between the target and the contextual stimulus. The findings of the study show evidence of online plasticity in the brainstem encoding which may be important for speech perception in noise.

1.1 Justification for the Study

The ability to tag the repeating elements in the auditory environment is important in determining success in accurately perceiving speech in noise (Ahissar et al., 2006). In a study done by Russo, Nicol, Musacchia and Kraus (2004) measures of transient and sustained components of the brainstem response to speech syllables were robust in quiet. The background noise disrupted the transient responses whereas the sustained response was more resistant to the deleterious effects of noise. The context dependent encoding of FFR has been shown to be related to speech perception in noise. The present study aimed to test the context dependent encoding in the presence of noise. It was of interest to study whether fine tuning to repetitive stimuli persists even in the presence of noise. One can speculate that FFR being immune to background noise, shall continue to show context dependent encoding. If the underlying mechanism of context dependent encoding has role in speech perception in noise, then the context dependent encoding should persist even in the presence of noise. However, this notion needs scientific support. Therefore the present study was taken up.

1.2 Aim of the Study

The present study aimed to test the context dependent encoding of speech in the presence of noise.

1.3 Objectives of the Study

There were two specific objectives of this study:

1. To compare between context dependent encoding of FFR in quiet, and in noise of 10dBSNR, in normal hearing individuals
2. To find the relation between context dependent encoding of FFRs in noise and speech perception in noise, in normal hearing individuals

Chapter 2

REVIEW OF LITERATURE

In the recent past, speech evoked auditory brainstem responses (ABRs) have drawn interest of researchers across the globe alike, due to the enormous information about neurophysiology that they provide. The focus of this review of literature is to orient the reader on the significance of the information provided by speech elicited ABR, the phenomena of context dependent encoding, along with its role in speech perception and the factors that influence context dependent encoding.

The most commonly used stimulus to record speech ABR is syllable /da/. The stimulus consists of an onset burst frication at F3, F4, and F5 during the first 10ms, followed by 30 ms of F1 and F2 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 syllable. The stimulus was chosen such that it is short enough in duration to minimize test time while still containing key acoustic phonetic information.

In order to establish a valid and reliable means to differentiate the neural activity within speech evoked ABR, it is essential to understand the functional relationship between the acoustic structure of the stimulus and the corresponding brainstem response (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. Whereas, the sustained responses (Frequency following response-FFR) are known to provide information about the fidelity with which the brainstem responds to sound and also represents the range of spectral components within the brainstem.

2.1 Application of Speech Evoked Brainstem Responses

Khaladkar, Karthik and Vanaja (2005) used two stimuli i.e., an acoustic clicks and /t/ syllable (burst portion). Brainstem responses were recorded for the two stimuli from 20 ears with Sensorineural hearing loss. Results showed poorer representation of speech burst ABR compared to click evoked ABR in individuals with Sensorineural hearing loss. Sumesh and Barman (2007) found similar results and also reported that sustained responses were affected more compared to transient responses, due to hearing loss.

Banai, Nicol, Zecker and Kraus (2005) recorded speech ABR in individuals with learning difficulties and found that 40% of them showed abnormalities in the responses. Deviations in the brainstem responses to speech have been reported in children with learning problems (King, Warrier, Hayes & Kraus, 2002; Cunningham et al., 2001; Wible et al., 2004), autism spectrum disorders (Russo, et al., 2008), and phonological disorders (Goncalves, Wertzner, Samelli, & Matas, 2011).

2.2 Efferent Modulation of the Brainstem Responses

The existing evidences suggest malleability in the brainstem representation of speech (Kraus & Nicol, 2005, Banai Nicol, Zecker & Kraus, 2005). Both long-term and short-term auditory experiences have been shown to enhance the brainstem responses to complex behaviorally relevant sounds.

2.2.1 Modulation of Brainstem Physiology Secondary to 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 influences the 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 determined by 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 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). Stait, Skoe, Kraus and Ashley (2009) reported similar enhancement in musicians for emotionally salient veal sounds. These studies clearly indicate that brainstem is malleable and the experience can modulate the brainstem physiology.

2.2.2 Modulation of Brainstem Following Short-term Experience

Short-term auditory training has been shown to improve the timing of the FFR (Russo, Nicol, Zecker, Hayes & Kraus, 2005). Brainstem response to /da/ was obtained from children with learning problems 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.

Song, Skoe, Wong and 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 the 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 response to speech stimulus, following training. Taken together, these studies suggest that the adult brainstem is indeed malleable to short-term training.

2.3 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, Suga, 2011) and the other is the local reorganization model (Krishnan & Gandour, 2009).

The corticofugal model state 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 & Gandour, 2009). Both models require 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. Considering there are massive efferent connections from the cortex to subcortical structures, these connections could form the basis of feedback-related top-bottom projections (Kral & Eggermont, 2007). Furthermore, 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 subcortical 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 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 (2010) 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, Harnickel, 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 procedures.

2.4 Context Dependent Brainstem Encoding of Speech

Tonse and Maruthy (2012) conducted a study to check if plasticity occurs as a result of long term experience or whether it is a continuous process. To obtain information on the extent to which plasticity is operational online, brainstem response to speech syllable /da/ was elicited in four conditions, which included one repetitive condition and three stimulus context conditions. Results showed that the latencies of onset and sustained responses were prolonged in the stimulus context conditions when compared to repetitive condition. Since the generators of the onset and the sustained responses (CN, LL & IC) fall within the feedback loop of the corticofugal pathway, the ability of the corticofugal pathway to identify spectral differences between the target stimulus and contextual stimulus is hypothesized to influence the brainstem responses. The results of the study suggested the possibility of online plasticity at the level of brainstem regulated by the corticofugal network.

Clark, Strait and Kraus (2011) showed enhanced speech perception in noise is seen in musicians. In this study they aimed to examine how the encoding of same speech syllable presented in predictable and variable conditions differs and speech perception in the presence of noise in 31 musicians and non-musicians. The results showed that musicians exhibited robust neural encoding of fundamental frequency of speech seen in predictable condition relative to the variable condition than that in non-musicians. The amount of neural encoding in predictable condition correlated with participant's musical practice histories and also with their speech perception in noise abilities. They concluded that subcortical sensitivity to speech regularities is shaped by musical training and may contribute to speech perception in noise in them.

Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009) examined context dependent brainstem encoding in children with developmental dyslexia using auditory brainstem responses to a speech syllable presented in a repetitive and variable conditions. Children without developmental dyslexia showed enhanced brainstem responses with respect to voice pitch in the repetitive condition compared to the variable context. In contrast, children with developmental dyslexia did not show the ability to modify representation in predictable contexts. They also found that the extent of context dependent brainstem encoding correlated positively with behavioral speech in noise perception. The ability to sharpen representation of repeating elements is crucial for speech perception in noise, since it allows superior "tagging" of voice pitch, which is an important cue for segregating sound streams in background noise. The dysfunction of this mechanism contributes to a deficit in noise exclusion, which is a hallmark symptom in developmental dyslexia.

2.5 Summary of Review of Literature

Overall, this focused review of literature reveals that brainstem response to speech offers a unique window into understanding how the brainstem represents the components of speech signals. The brainstem response to speech has two unassociated components, the transient and sustained FFRs. Collectively, these components 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 strongly imply predominant brainstem origin for the scalp recorded FFRs. Even though the scalp recorded onset response and the FFR reflect the activity at numerous sources (LL, CN, IC), they offer a noninvasive techniques to study the subcortical encoding of speech, as well as the effect of experience on the representation of speech at the brainstem. Recent studies have focused on context dependent brainstem encoding using predictable and variable stimulus conditions. Literature review suggests that context dependent brainstem encoding has the scope to serves as an index of behavioral speech perception in noise. However, there is a dearth of literature on the precise underlying mechanisms of context dependent encoding and the influence of noise on context dependent encoding.

Chapter 3

METHOD

The primary aim of the study was to test the context dependent encoding of speech in the presence of noise. The secondary aim was to study the role of context dependent encoding in noise on speech perception in noise. It was hypothesized that there is no significant difference between context dependent encoding of speech in quiet and in noise. A quasi experimental research design with purposive sampling was used in the present study. The following method was used to test the hypothesis.

3.1 Participants

Fifteen normal hearing adults in the age range of 18 to 25 years participated in the study. They were ensured for normal hearing sensitivity, normal middle ear functioning and outer hair cell functioning using puretone audiometry, immittance evaluation and otoacoustic emissions respectively. Individuals with a present/past history of middle ear pathologies and/or neurological disorders were excluded from the study. None of the participants had a complaint of difficulty in understanding speech in noisy situations. Musicians were excluded from the study in view of maintaining homogeneity among the participants.

All the participants were students, pursuing their bachelors and masters degree in Speech and Hearing. An informed consent was obtained from all the participants prior to their inclusion in the study and the procedures used in the study conformed to the ethical guidelines for bio behavioral research in human participants.

3.2 Test Stimuli

Four different stimuli were used to record context dependent brainstem encoding. Of these four, one was a target stimulus and the other three were used as contextual stimuli. A synthetically generated syllable /da/ was the target stimulus. Only the response recorded for syllable /da/ was of importance in the present study. The other three syllables that served as contextual stimuli were, /bu/, /bi/ and /gi/. The contextual stimuli differed from /da/ syllable in terms of burst of the stop, second formant transition and the vowel.

The syllable /da/ was of 100ms. Longer duration was preferred because the spectral information was better represented and better FFR were being recorded than that with shorter duration /da/ (40ms). The syllables /da/, /bu/, /bi/ and /gi/ were synthesized in the Electrophysiology Lab of the All India Institute of Speech and Hearing, Mysore. These syllables were uttered by an adult male who was a native speaker of Kannada. Utterances were recorded using a dynamic microphone placed six inches distance away from the speaker's mouth. The microphone output was routed to the Stim2 hardware (Compumedics-Neuroscan, Charlotte, NC, USA) and recorded with a resolution of 16 bits at a sampling frequency of 44,100 Hz in the Sound module of the Stim2 software suite. The duration of the syllables were edited to restrict the duration to 100ms. This was done by deleting the vowel cycles beyond 100 ms at the nearest zero-crossing.

The recorded syllables were initially analyzed using Speech Processing and Synthesis toolboxes incorporating a linear predictive coding algorithm. This was done in order to extract and modify the different acoustic parameters independently. The modified linear predictive coding parameters were then used to synthesize the CV

stimuli of 100 ms using the toolboxes. Later these synthetic speech syllables (/bi/, /bu/, /gi/ & /da/) were subjected to a perceptual rating. The syllables were rated for naturalness and quality by 10 sophisticated listeners. Based on the ratings of the listeners, linear predictive coding parameters were modified to resynthesize the stimuli with higher naturalness. The waveforms and spectrograms of the four stimuli used in the present study are shown in Figure 3.1. The spectral characteristics are shown in Table 3.1.

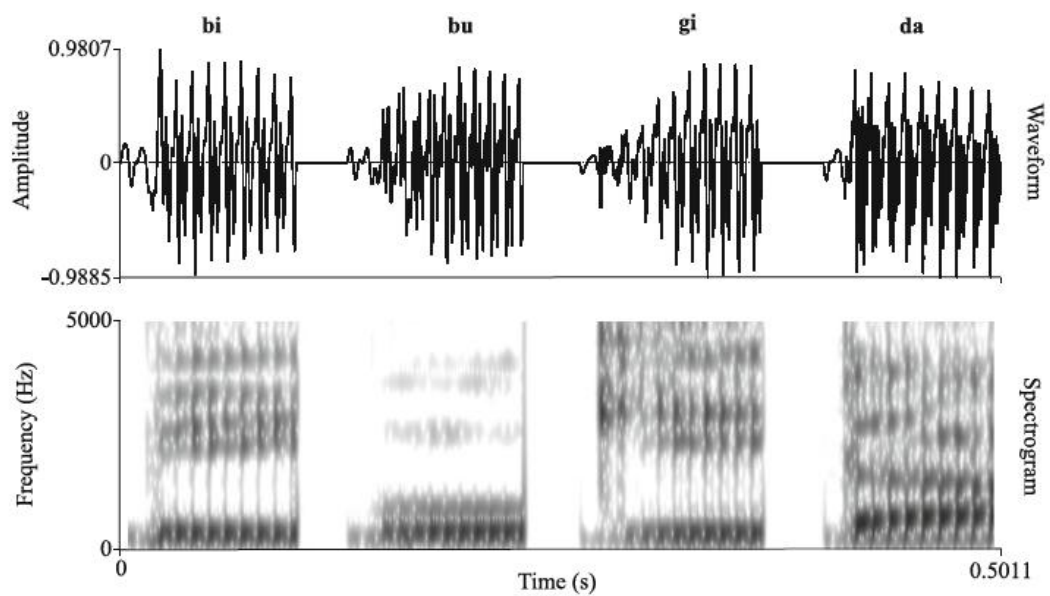


Figure 3.1: The waveforms and Spectrograms of syllables /bi/, /bu/, /gi/ and /da/, generated and used in the present study.

Table 3.1: Spectral characteristics of syllables, /bi/, /bu/, /gi/ & /da/ used in the present study

Stimulus	F0 (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	F5 (Hz)
/bi/	100.29	563 to 630	1168 to 1193	2488 to 2566	3690 to 3748	Steady 5091
/bu/	117.58	324 to 328	836 to 845	2533 to 2534	3667 to 3746	Steady 5331
/gi/	113.07	267 to 295	2213 to 2377	3042 to 3147	4049 to 4015	Steady 4846
/da/	100.24	563 to 692	1453 to 1281	2510 to 2475	3285 to 3287	Steady 3472

3.3 Test Environment

All the tests were carried out in an acoustically and electrically shielded room wherein the ambient noise levels were well within the permissible limits (ANSI S.3, 1991).

3.4 Test Procedure

3.4.1 Preliminary Evaluations

Detailed audiological assessment was performed for all the participants before recruiting them for the study. Modified Hughson and Westlake procedure was used to carryout pure tone audiometry. All selected participants had pure tone hearing thresholds of less than 15dBHL at octave frequencies between 250Hz and 8000Hz. A calibrated two channel Inventis *piano* diagnostic audiometer with TDH-39 headphone was used for puretone audiometry. Their Speech recognition thresholds (SRT) were within ± 12 dB of pure tone average (average threshold of 500 Hz, 1000 Hz, 2000 Hz

& 4000 Hz) and the speech identification scores were greater than 80% at 40 dB SL (ref SRT) for phonetically balanced word lists.

Immittance of the middle ear was tested for a 226Hz probe tone using Grason Stadler Inc. Tymptstar immittance meter. Normal functioning of middle ear was indicated by bilateral type 'A' tympanogram and, normal acoustic reflex thresholds (ipsilateral & contralateral) at 500, 1000 and 2000 Hz.

Normal OHC functioning was assessed by recording transient evoked otoacoustic emissions (TEOAEs). ILO-Echoport plus (version 6) was used for recording TEOAEs. TEOAEs were recorded monaurally for clicks presented around 75dBpkSPL. Nonlinear stimulus paradigm was used to elicit TEOAEs and the response was acquired using the standard protocol of the ILO V6 equipment. The SNR of TEOAEs was more than 6dB in all the octave and mid-octave frequencies between 1kHz and 6kHz, in all the participants, in their both the ears.

3.4.2 Experimental Test Procedure

The individuals who satisfied all the inclusion criteria served as participants of the study. Such participants were tested for their speech perception in noise (SNR-50) and context dependent brainstem encoding using frequency following responses (FFRs).

Estimation of Speech Perception in Noise (SNR-50): This test estimated minimum ratio required for 50% identification of monosyllables (SNR-50). SNR paradigm from Smriti-Shravan developed by Kumar and Maruthy (2016) was used for this purpose. A HDA-200 headset was used to deliver the test stimuli. The module consisted of 19 bisyllables mixed with broadband noise at varying SNRs. A one-down

one-up procedure was used for finding SNR-50. The test started with SNR of 2 dB. The SNR was subsequently decreased by 2 dB for every correct response and increased by 2 dB for every incorrect response. Stimuli were played from a laptop and were delivered through Sennheisser HAD-200 headphones. Participants were instructed to listen carefully, recognise the bisyllable heard and indicate the response by clicking on the respective bisyllable among the 19 bisyllables displayed on a computer screen. A total of 10 reversals were used and the average of last six reversals were taken as the SNR-50.

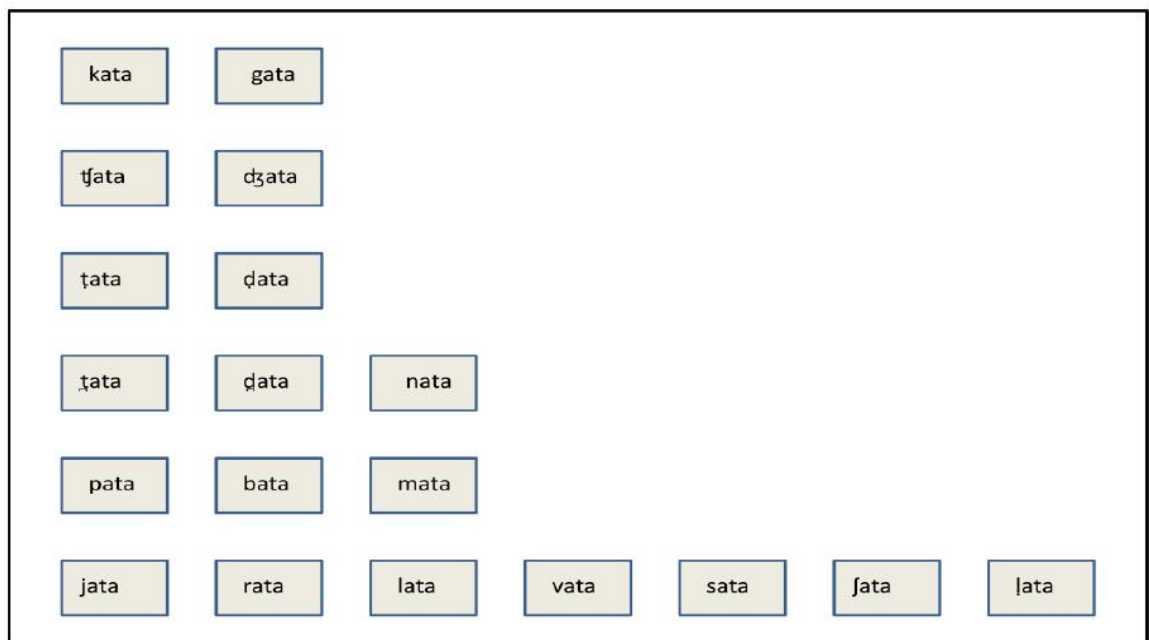


Figure 3.2: Nineteen bisyllables displayed on the computer screen during SNR-50 estimation using Smriti-Shravan module.

Recording FFRs and Context-dependent Brainstem Encoding: The participants were seated in a sound treated and electrically well shielded room. Intelligent Hearing system was used to record FFRs. The electrode sites were prepared with skin preparation gel. Gold plated electrodes were placed with conducting gel and adhesive tape to obtain impedance of less than 5 kOhms at each

electrode site. Single channel vertical ipsilateral montage was used for recording the response. The participants were asked not to move, and relax during the testing. The speech stimuli were delivered through an insert ear phone and ipsilateral noise was presented through the inventis piano audiometer. The parameters listed in Table 3.2 were used to record FFRs in repetitive and variable paradigm, with and without ipsilateral noise. For each stimulus condition, the responses were recorded twice to ensure the replicability.

Brainstem response to synthetically generated syllable /da/ was recorded in four different conditions. This included recording FFRs in two stimulus paradigms in quiet and at 10dBSNR. The two stimulus paradigms were, the repetitive paradigm and the variable paradigm.

In condition 1, the FFRs to the repetitive paradigm was obtained for 1000 sweeps of /da/. Only the syllable /da/ was presented in this paradigm, repeatedly. Then in condition 2, brainstem responses were recorded in variable stimulus paradigm by making use of MMN/P300 protocol of Smart EP, wherein /bu/ stimulus was considered as the frequent stimulus presented with a probability of 50% and other three stimuli /da/, /bi/ and /gi/ as the infrequent stimuli with the probability of 30%, 10% and 10% respectively. In this case again, FFRs were recorded for 1000 presentations of /da/. In condition 3 and 4, FFRs were recorded for repetitive and variable paradigms (as described above) but in the presence of ipsilateral speech noise. Figure 3.3 represents the different stimulus paradigms used in the present study. The stimulus and acquisition parameters used to record FFRs are given in Table 3.2. In each stimulus condition, FFRs were recorded twice to ensure the replicability of the response.

Table 3.2: Stimulus and acquisition parameters used to record FFRs

<i>Stimulus Parameters</i>	
Stimuli	Repetitive paradigm: /da/ only Variable paradigm: /da/, /bi/, /bu/, /gi/ The above syllables were presented in quiet and at 10dBSNR
Ear	Right/left
Duration of stimulus	100 ms
Intensity	70 dBnHL
Repetition rate	7.1/s
Polarity	Rarefaction Polarity
Number of sweeps	1000
<i>Acquisition Parameters</i>	
Analysis time	128 ms
Electrode montage	Vertical
Amplification	100000
Artifact rejection	25 μ V
Filter setting	30-3000 Hz

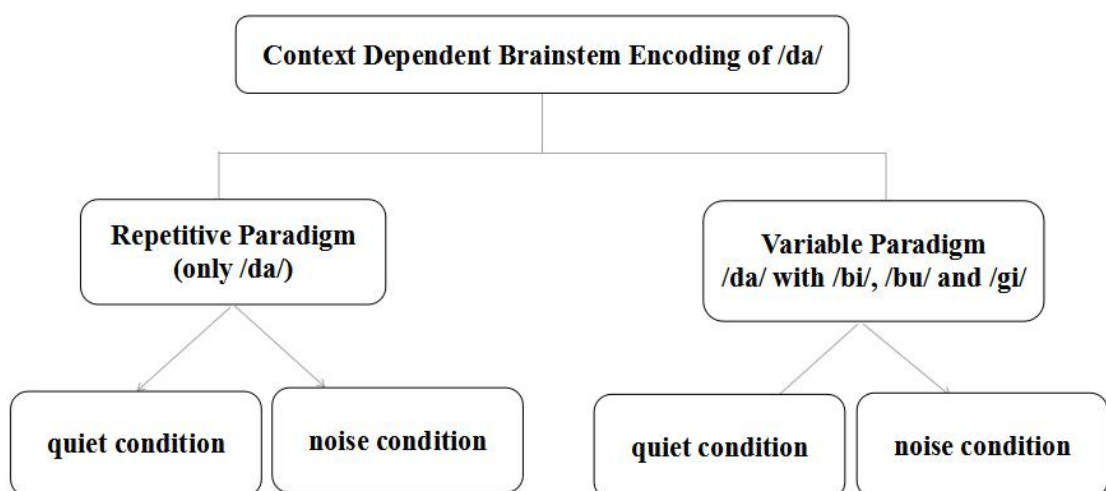


Figure 3.3: Representation of different stimulus conditions used in the present study.

3.4.3 Response Analysis

The averaged response obtained for syllable /da/ in the four stimulus conditions were objectively analysed using Fast Fourier Transform (FFT). This was to derive the spectral composition of the response in its sustained portion (FFRs). The averaged responses were subjected to spectral analysis to analyse the amplitudes at the spectral components corresponding to the fundamental frequency (H1 - 100 Hz), second harmonic (H2 - 200 Hz), third harmonic (H3 - 300 Hz) and fourth harmonic (H4=400) of the stimulus. This was done in a custom written program in Matlab 2014a platform developed at Northwestern University. The raw amplitude value of the F0, H2, H3 and H4 of the FFR was then measured. The waveforms were windowed from 15 to 100ms using a 10% tapered Tukey window and zero-padded up to a total duration of 1 s to increase the spectral resolution to 1Hz. The zero-padded waveforms were then subjected to FFT. The magnitudes at H1, H2, H3 and H4 were then analyzed by averaging the magnitudes of ten bins (1 Hz wide) around the H1, H2, H3 and H4 frequencies. These spectral magnitudes were used as the index of brainstem encoding.

The data thus obtained was used for the comparison between the responses obtained in repetitive and variable stimulus paradigms, in quiet and ipsilateral noise conditions.

Chapter 4

RESULTS

The present study was aimed to test the context dependent encoding of speech in the presence of noise. In the present study, the independent variable was noise and context dependent encoding in FFR was the dependent variable. The measures of context dependent encoding were differences in the H1, H2, H3 and H4 of FFR between repetitive and variable paradigms.

Paired t test was done to compare between repetitive and variable paradigms in quiet as well as noise conditions. Paired t test was chosen on the results of Shapiro Wilk's normality test which showed normal distribution of the data in all the variables. Later correlation of SNR-50 and context dependent encoding was done separately for quiet and noise conditions using Spearman's correlation. In the results, we found some interesting findings which are reported under the following headings;

1. Results of repeated measures ANOVA
2. Comparison of FFR recorded in repetitive paradigm between quiet and noise conditions
3. Comparison of FFR recorded in variable paradigm between quiet and noise conditions
4. Comparison of FFR recorded between repetitive and variable paradigm in quiet condition
5. Comparison of FFR recorded between repetitive and variable paradigm in noise condition
6. Comparison of context dependent encoding between quiet and noise conditions
7. Relation between SNR-50 and context dependent encoding in quiet and noise

4.1 Results of Repeated Measures ANOVA

To assess the effect of stimulus paradigm (repetitive versus variable) and condition (quiet versus noise) on FFRs, the data was subjected to repeated measures ANOVA taking stimulus paradigm and condition as repeating variable. Table 4.1 shows the results of ANOVA which revealed that there was a significant main effect of condition on H1, H2, H3 and H4. There was no significant main effect of noise on H1, H2, H3 and H4. There was a significant interaction between condition and noise seen, however only in H3.

Table 4.1: Haromincs, F value, Degree of Freedom, Error factor, significance value, Partial Eta Square and Observed power ^a in condition , noise and the interaction between noise and condition

Variable	Measure	F	df (error)	p	Effect size	Observed Power ^a
Condition	H1	0.247	(1,14)	0.627	0.017	0.75
	H2	0.070	(1,14)	0.729	0.005	0.57
	H3	0.005	(1,14)	0.946	0.000	0.50
	H4	0.539	(1,14)	0.475	0.037	0.105
Paradigm	H1	19.94	(1,14)	0.001	0.587	0.985
	H2	34.01	(1,14)	0.000	0.708	1.000
	H3	7.86	(1,14)	0.014	0.360	0.742
	H4	8.58	(1,14)	0.011	0.380	0.778
Condition *	H1	0.819	(1,14)	0.381	0.055	0.135
	H2	0.245	(1,14)	0.628	0.017	0.075
	H3	6.247	(1,14)	0.025	0.309	0.643
	H4	0.361	(1,14)	0.558	0.025	0.087

The comparisons between the conditions and the stimulus paradigms was further tested using paired t test in view of clear presentation of the results.

4.2 Comparison of FFR Recorded in Repetitive Paradigm in Quiet and Noise Conditions

Figure 4.1 shows the grand average spectra of the FFR obtained in repetitive paradigm in quiet and noise conditions. Table 4.2 shows the mean and standard deviation (SD) of amplitude of H1, H2, H3 and H4 of FFR obtained in repetitive paradigm in quiet and noise conditions. There was no common pattern in the way the mean amplitudes varied between the two conditions. The data did not show mean differences in H2 and H4, but in the quiet condition H1 was lower and H3 was higher compared to noise condition.

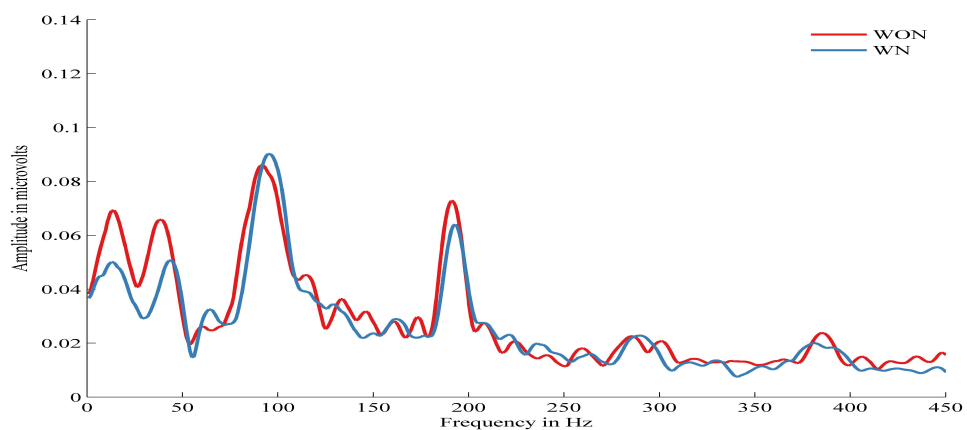


Figure 4.1: The grand average spectra of FFR in repetitive paradigm in quiet and noise conditions (WON - Without noise condition, WN - With noise condition).

Table 4.2: Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) of FFR recorded in repetitive paradigm in quiet and noise conditions

Harmonics	Condition	Mean (uV)	SD	t	df	p
H1	Quiet	0.07	0.02	-0.95	14	0.361
	Noise	0.08	0.05			
H2	Quiet	0.04	0.01	-0.23	14	0.82
	Noise	0.04	0.02			
H3	Quiet	0.02	0.008	2.349	14	0.03
	Noise	0.01	0.004			
H4	Quiet	0.01	0.004	0.213	14	0.834
	Noise	0.01	0.005			

Paired t test was carried out to test whether the mean differences between quiet and noise conditions of FFR recorded in repetitive paradigm are significantly different. The results revealed no significant difference between these conditions in H1, H2 and H4 amplitudes ($p > 0.05$) but there was a significantly higher amplitude in H3 ($p < 0.05$) in quiet condition compared to noise condition.

4.3 Comparison of Variable Paradigm in Quiet and Noise Conditions

Figure 4.2 shows the grand average spectra of the FFR obtained in variable paradigm in quiet and noise conditions. Table 4.3 depicts the mean and standard deviation (SD) of amplitude of H1, H2, H3 and H4 of FFR obtained in variable paradigm in quiet and noise conditions. The data showed no mean differences in H1, H3 and H4, but there was higher amplitude of H2 in quiet than in noise condition.

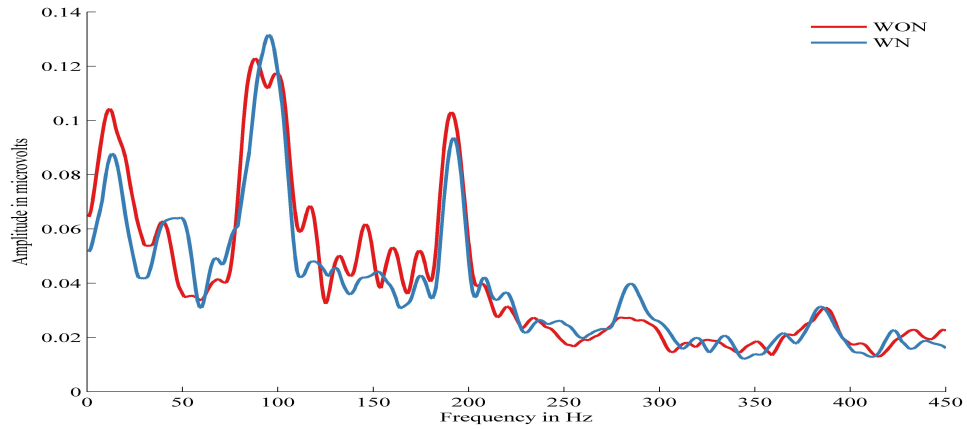


Figure 4.2: The grand average spectra of FFR in variable paradigm in quiet and noise conditions (WON - Without noise condition, WN - With noise condition).

Table 4.3: Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) of FFR in variable paradigm in quiet and noise conditions

Harmonics	Condition	Mean (uV)	SD	t	df	p
H1	Quiet	0.12	0.05	-0.04	14	0.96
	Noise	0.12	0.06			
H2	Quiet	0.06	0.02	0.438	14	0.66
	Noise	0.05	0.02			
H3	Quiet	0.02	0.008	-1.242	14	0.23
	Noise	0.02	0.01			
H4	Quiet	0.02	0.007	0.83	14	0.42
	Noise	0.02	0.004			

Paired t test revealed that there is no significant difference between quiet and noise conditions in any of the harmonics ($p > 0.05$) of FFR of variable paradigm.

4.4 Comparison between Repetitive and Variable Paradigms in Quiet Condition

The context dependent encoding was derived by comparing the repetitive and variable paradigm. This comparison was done separately in quiet and noise conditions. Figure 4.3 shows the grand average spectra of the FFR obtained in quiet condition in the repetitive paradigm and variable paradigms. Table 4.4 gives the mean and standard deviation (SD) of amplitude of H1, H2, H3 and H4 of FFR obtained in quiet condition in the repetitive and variable paradigms. The data showed no difference in the mean H3, but mean H1, H2 and H4 amplitudes were more in variable paradigm compared to repetitive paradigm.

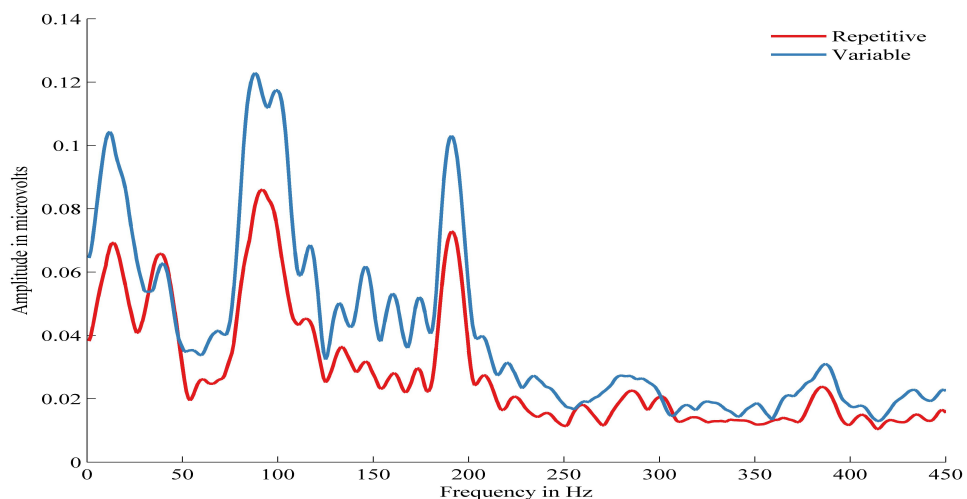


Figure 4.3: The grand average spectra of FFR in repetitive and variable paradigms in quiet.

The statistical significance of observed differences in the mean amplitudes was tested using paired sample t test. Results showed that there was significantly higher amplitudes of H1, H2 and H4 ($p < 0.05$) in the variable paradigm while there was no significant difference in H3 ($p > 0.05$) between repetitive and variable paradigms.

Table 4.4: Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) in quiet condition

Harmonics	Paradigm	Mean (uV)	SD	t	df	p
H1	Repetitive	0.07	0.02	-4.85	14	0.00
	Variable	0.12	0.05			
H2	Repetitive	0.04	0.01	-3.91	14	0.00
	Variable	0.06	0.02			
H3	Repetitive	0.02	0.008	-0.29	14	0.77
	Variable	0.02	0.008			
H4	Repetitive	0.01	0.004	-2.27	14	0.04
	Variable	0.02	0.007			

4.5 Comparison between Repetitive and Variable Paradigm in Noise Condition

Figure 4.4 shows comparison of the grand average spectra of the FFR between repetitive paradigm and variable paradigms, obtained in noise condition. Table 4.5 gives the mean and standard deviation (SD) of amplitude of H1, H2, H3 and H4 in noise condition between repetitive and variable paradigm. The data showed that all the harmonics had higher amplitude in variable paradigm compared to repetitive paradigm in noise condition.

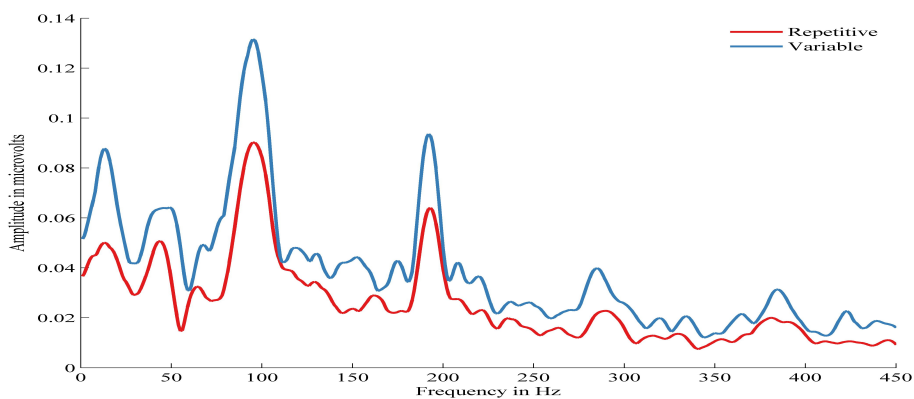


Figure 4.4: The grand average spectra of FFR in repetitive and variable paradigms in noise condition.

Table 4.5: Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) in noise condition

Harmonics	Paradigm	Mean(uV)	SD	t	df	p
H1	Repetitive	0.08	0.05	-3.01	14	0.00
	Variable	0.12	0.06			
H2	Repetitive	0.04	0.01	-3.01	14	0.00
	Variable	0.05	0.02			
H3	Repetitive	0.01	0.004	-3.12	14	0.00
	Variable	0.02	0.01			
H4	Repetitive	0.01	0.005	-2.199	14	0.04
	Variable	0.02	0.004			

The differences in the mean amplitudes between the two paradigms were tested for their statistical significance using paired sample t test. The results showed significant differences in H1, H2, H3 and H4 ($p < 0.05$) between variable paradigm and repetitive paradigm.

4.6 Comparison of the Amplitudes of Harmonics across the Four Stimulus Conditions in Each Participant

Figure 4.5, 4.6, 4.7 and 4.8 depicts the individual data of H1, H2, H3 and H4 respectively, across the four stimulus conditions.

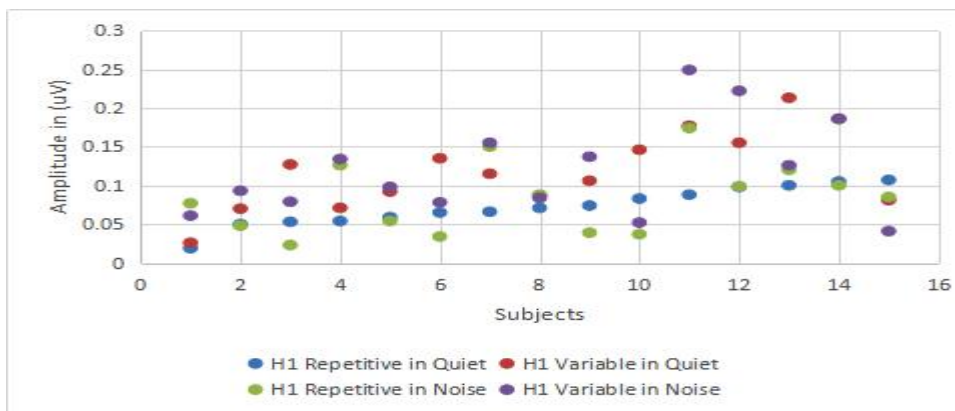


Figure 4.5: H1 amplitude of individual participants in repetitive and variable paradigms, in quiet and noise conditions.

In Figure 4.5, it is clearly shown that except the 15th participant, all other participants had higher amplitude of H1 in variable paradigm than repetitive paradigm in quiet condition. In noise condition except three participants (1, 8 and 15) all other participants had higher amplitude in variable paradigm than repetitive paradigm. No particular pattern was evident between quiet and noise conditions.

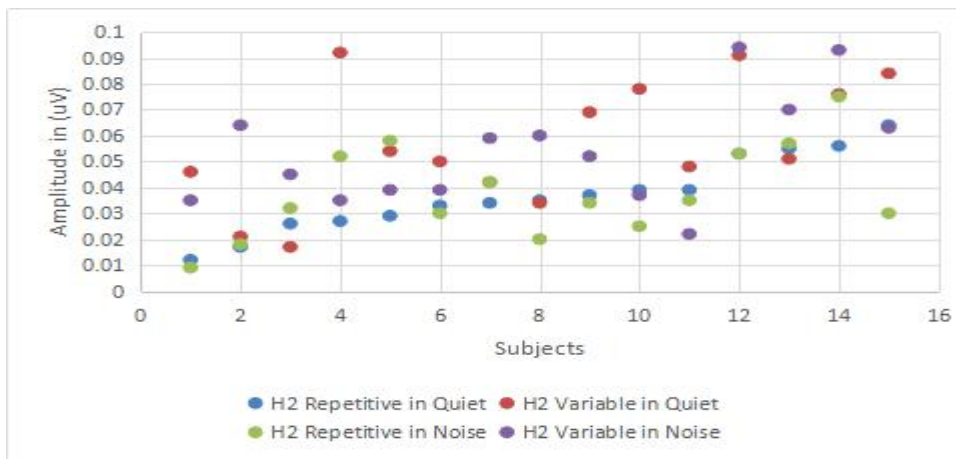


Figure 4.6: H2 amplitude of individual participants in repetitive and variable paradigms, in quiet and noise conditions.

Figure 4.6 shows that, except three participants (4, 7 & 11) participant, all other participants had higher H2 amplitude in variable paradigm than repetitive paradigm in quiet condition. Similarly, except three participants (8, 10 & 15) all other participants had higher amplitude in variable paradigm than repetitive paradigm in noise condition. No particular pattern was observable between quiet and noise conditions.

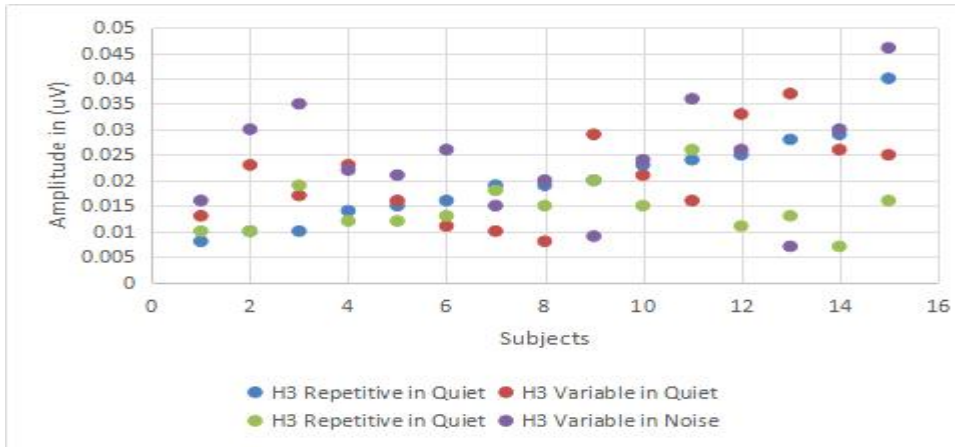


Figure 4.7: H3 amplitude of individual participants in repetitive and variable paradigms, in quiet and noise conditions.

Figure 4.7 shows that eight participants had higher amplitude of third harmonic frequency in variable paradigm than repetitive paradigm in quiet condition. In noise condition except four participants (2, 4, 14 & 15) all other subjects had higher amplitude in variable paradigm than repetitive paradigm.

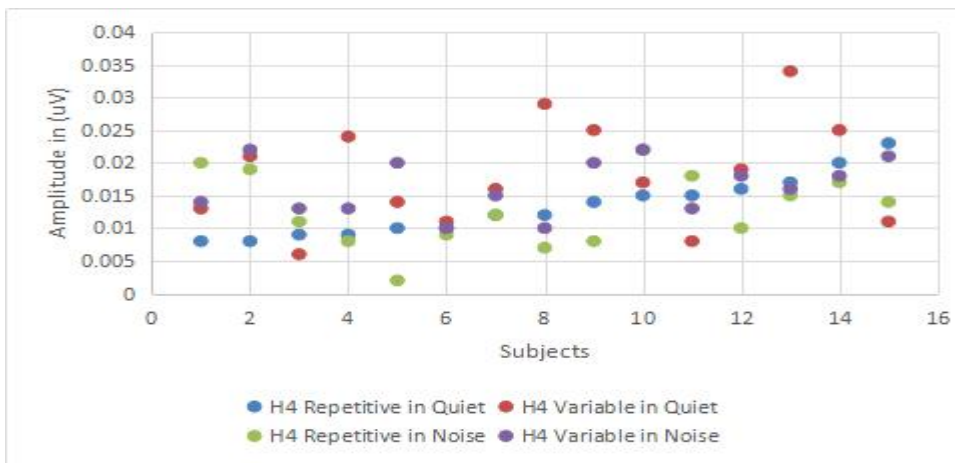


Figure 4.8: H4 amplitude of individual participants in repetitive and variable paradigms, in quiet and noise conditions.

Figure 4.8 shows that except three participants (5, 9 & 13), all other participants had higher amplitude of second harmonic frequency in variable paradigm than repetitive paradigm in quiet condition. In noise condition, except two participants (3

and 13) all other subjects had higher amplitude in variable paradigm than repetitive paradigm.

4.7 Comparison of Context Dependent Brainstem Encoding between Quiet and Noise Conditions

Context dependent brainstem encoding at H1, H2, H3 and H4 were derived by subtracting the amplitude of variable paradigm by repetitive paradigm in all the harmonics in both quiet and noise conditions. Table 4.6 gives the mean and standard deviation (SD) of amplitude of H1, H2, H3 and H4 of context dependent encoding in quiet and noise conditions. The data showed no mean differences in H2 but there was lower mean of H1, H4 and higher mean of H3 in quiet condition than that in noise conditions.

Table 4.6: Mean and standard deviation (SD) of amplitude of first (H1), second (H2), third (H3) and fourth harmonic (H4) of context dependent encoding in quiet and noise conditions

Harmonics	Paradigm	Mean (uV)	SD	t	df	p
H1	Quiet	-0.05	0.04	-0.90	14	0.38
	Noise	-0.04	0.05			
H2	Quiet	-0.02	0.02	-0.49	14	0.63
	Noise	-0.02	0.02			
H3	Quiet	-0.0007	0.008	2.49	14	0.02
	Noise	-0.01	0.01			
H4	Quiet	-0.005	0.008	-0.60	14	0.56
	Noise	-0.003	0.006			

Paired t test was carried out to test whether the mean difference between the two conditions is statistically significant in context dependent brainstem encoding. The results revealed no significant difference between quiet and noise conditions in

the context dependent encoding of H1, H2 and H4 amplitudes ($p>0.05$) of FFR but there was a significant difference in the amplitude of H3 ($p<0.05$) between quiet and noise conditions.

4.8 Results of Correlation

In the present study, SNR-50 was correlated with context dependent brainstem encoding. Context effects at H1, H2, H3 and H4 were derived by subtracting the amplitude of variable paradigm by repetitive paradigm in all the parameters. Correlation was assessed separately for each of the parameters in both quiet and noise conditions. Spearman's correlation coefficient was used. The Correlation between SNR-50 and harmonics in quiet and noise conditions are depicted in the Figure 4.9 and 4.10 respectively. Results indicated that none of the parameters showed significant correlation with SNR-50 ($p>0.05$) in quiet as well as noise conditions.

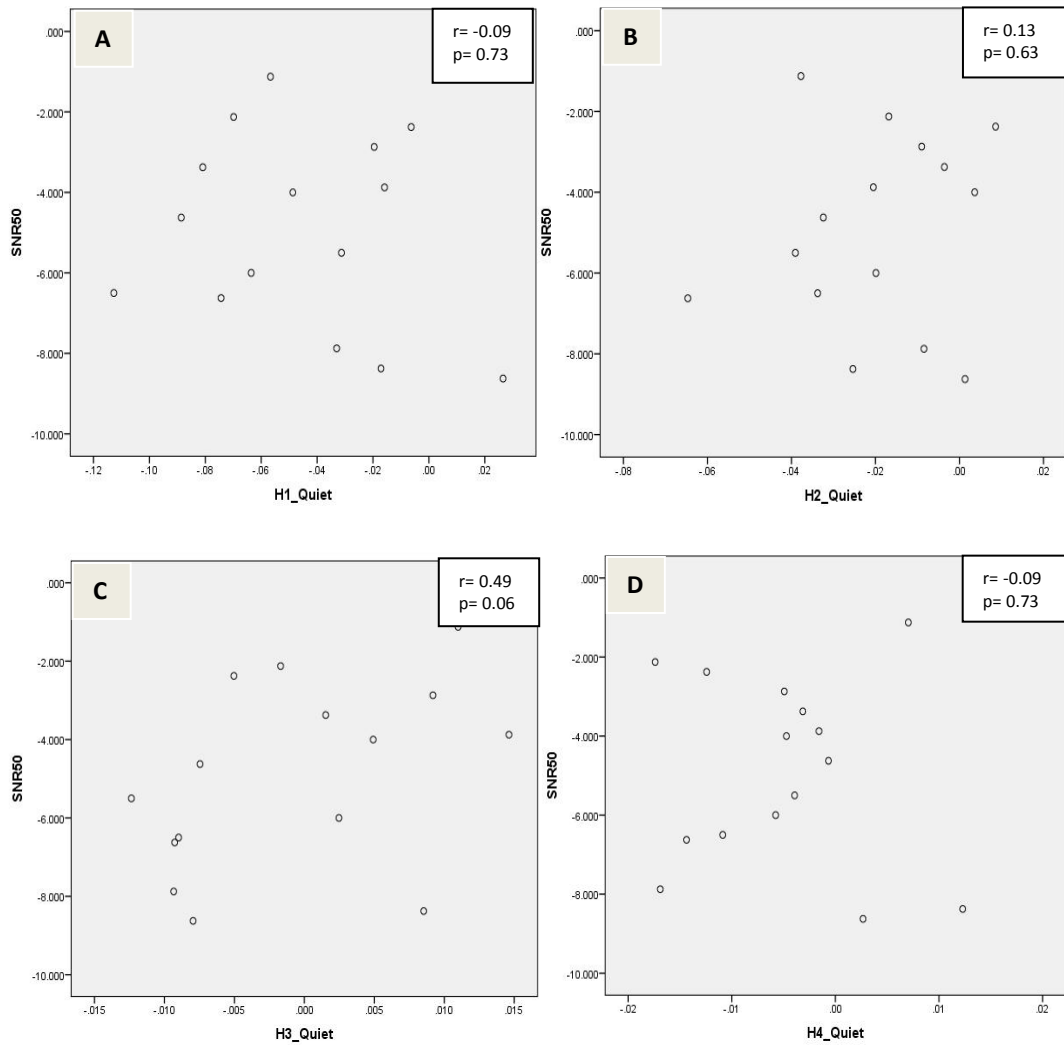


Figure 4.9: Scatter plots depicting SNR-50 in dB (Y-axis) and Harmonics in μV in quiet (X-axis). p -significant value, r -correlation coefficient, A-H1, B-H2, C-H3 and D-H4.

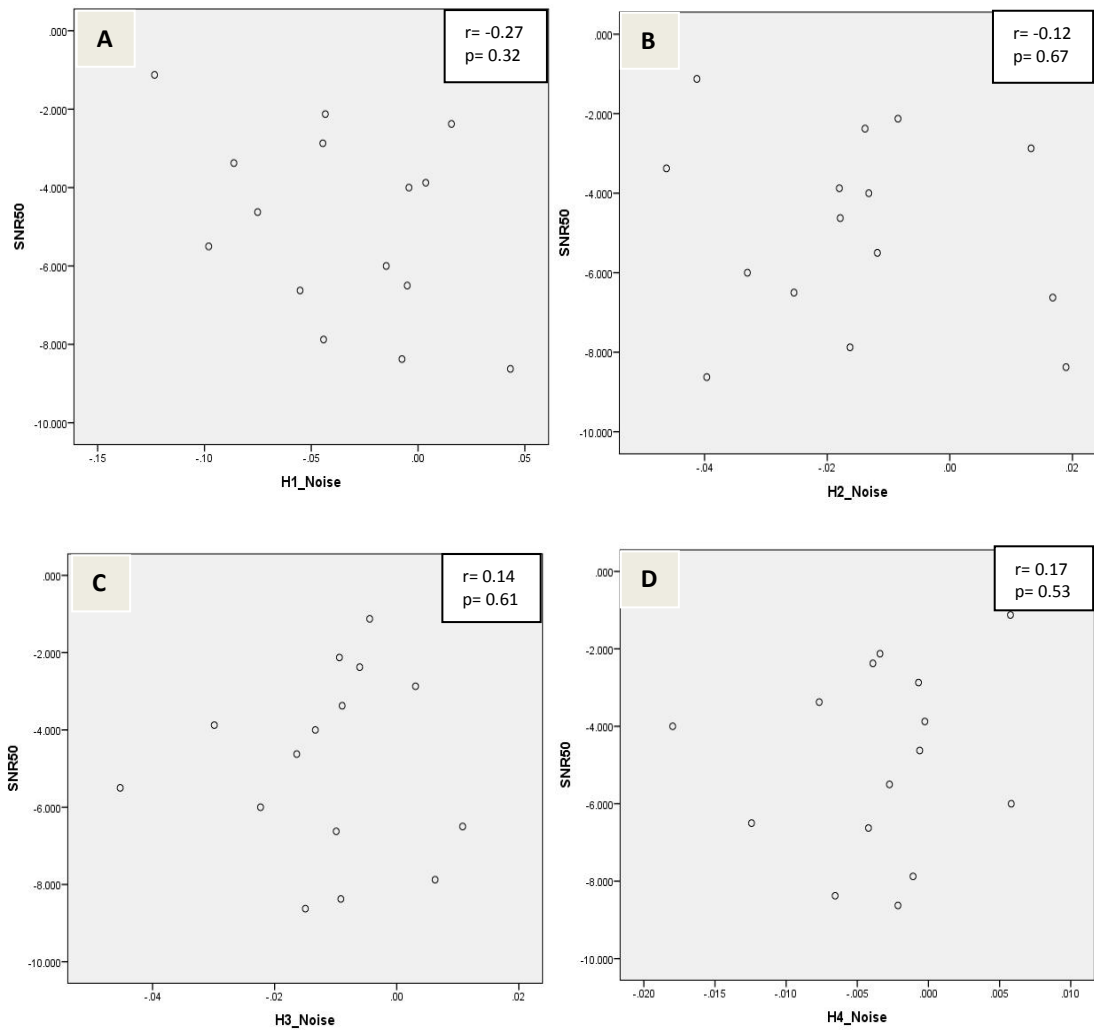


Figure 4.10. Scatter plots depicting SNR-50 in dB (Y-axis) and Harmonics in μV in noise condition (X-axis). p -significant value, r -correlation coefficient, A-H1, B-H2, C-H3 and D-H4.

Chapter 5

DISCUSSION

The primary aim of the study was to test the context-dependent brainstem encoding of speech in the presence of noise. The secondary aim was to assess the relationship between context-dependent brainstem encoding of speech in noise and the behavioral speech perception in noise. Spectral analysis was carried out to analyze brainstem encoding of speech at the frequencies corresponding to the fundamental frequency (H1) and the higher harmonics of (H2, H3 and H4) of an 100 ms /da/ syllable. This was performed in the repetitive and variable paradigm in quiet and in the presence of noise.

A thorough review of literature shows that FFR recorded in the repetitive paradigm is enhanced relative to that in variable paradigm. When listening to speech, we try to pay attention to certain features of sounds and we extract some regularly occurring pattern to understand speech. These repeating elements are helpful in extraction speech even in the presence of noise. Since there are no studies on context-dependent brainstem encoding in noise and how it is correlating with behavioral speech perception in noise the present study was taken. In the results of this study we found some interesting findings which are discussed under the following headings.

1. Context dependent encoding of FFRs recorded in Quiet
2. Context dependent encoding of FFRs recorded in the presence of noise
3. Relation between context dependent encoding and speech perception in noise

5.1 Context Dependent Encoding of FFR's Recorded in Quiet

In quiet condition, except H3, all the FFR amplitudes (at fundamental frequency and the three harmonics assessed) had significantly higher amplitudes in variable condition compared to repetitive condition. The finding was also true at the individual level, where the majority of the individuals had higher amplitudes in variable condition compared to repetitive condition. Thus, it can be inferred that the context significantly modulated the FFRs.

The FFRs are primarily known to be generated by the nuclei in the brainstem, especially the inferior colliculus, with contributions from the cochlear nucleus (Galbraith et al., 2000; Marsh, Brown, & Smith, 1974; Ping, Li, Galbraith, Wu, & Li, 2008). The contextual effect on the FFR suggests online modulation of the neural input at the brainstem based on the ongoing sound statistics. Such a modulation mechanism could be attributed to the local modulation of the neural activity by the brainstem nuclei, or the efferent input from the higher cortical regions. Studies suggest that such a local modulation is unlikely to occur in a single recording session, rather such a change would be due to long-term potentiation and experience (Chandrasekaran & Kraus, 2010; Strait, Hornickel & Kraus, 2011). It has thus been suggested that such a context dependent effect on the FFR might be a result of the modulatory effects of the cortico-fugal pathway (Chandrasekaran, Hornickel, Skoe, Nicol & Kraus, 2009; Maruthy, Kumar & Gnanateja, 2017; Skoe & Kraus, 2010).

Though there was a significant effect of context on the FFR in the current study, the direction of effect did not follow the same pattern as in the previous studies. In the previous studies (Chandrasekaran et al., 2009; Gnanateja, Ranjan, Firdose, Sinha & Maruthy, 2013; Maruthy et al., 2017; Parbery-Clark, Strait & Kraus, 2011; Skoe &

Kraus, 2010; Strait et al., 2011) the FFRs recorded with the repetitive stimuli showed higher spectral magnitudes than those recorded in the context of other stimuli. These studies suggested a pattern of repetition-related enhancement in the FFRs. In the current study however, an opposite pattern was found. The FFR spectral magnitudes were higher in the context of other syllables when compared to repetitive presentation of the same syllable.

The difference in the pattern of results in comparison to the earlier studies may be attributed to methodological differences. One major difference in method between the earlier studies and the current study is the stimulus polarity. Previous studies have used alternating polarity stimuli, while in the current study a single polarity was used. Averaging FFRs recorded for rarefaction and condensation polarities (as in alternating polarity) results in enhancement of envelope following responses and suppression of spectral following responses (Aiken & Picton, 2008). With FFRs recorded in alternate polarity it is hard to comment about the actual spectral features of the stimulus, as they are canceled out due to the alternating polarity of the stimulus. It could be speculated that the repetition related FFR enhancement evidenced in the previous studies might be a result of enhanced encoding of the envelope of the stimulus. While in the current study, the spectral following responses in the FFR were not eliminated due to the use of a single polarity. Considering the findings of the current study and those in previous studies, it can be speculated that context dependent effect may manifest as; a) enhancement in the envelope encoding of repetitive stimuli (repetitive condition) and, b) enhancement in the spectral encoding of the non-repetitive/novel stimuli (variable condition).

The enhancement of responses in the variable condition is well-known phenomenon in the cortical responses to sound. Earlier studies have shown increased N1 and P50 responses to target stimuli presented in the context of other stimuli (Boutros, Gjini, Urbach & Pflieger, 2011; Malmierca, Sanchez-Vives, Escera & Bendixen, 2014). It is suggested that the human auditory system is responsive to changes in the acoustic scene. The auditory pathways are laced with neurons which are sensitive to novelty detection and multiple levels (Slabu, Grimm & Escera, 2012). Thus, an increase in the FFR amplitudes in the variable condition might be attributed to novelty detection mechanism in the auditory pathway. Such a novelty detection mechanism is possibly modulated by the corticofugal pathways innervating the brainstem (Suga, Gao, Zhang, Ma, & Olsen, 2000; Terreros & Delano, 2015; Yan & Suga, 1998).

5.2 Context Dependent Encoding of FFRs Recorded in the Presence of Noise

The context dependent effect in the FFR has been suggested to be important for better temporal encoding of the frequently heard stimulus in the auditory scene. This increased encoding of the frequently heard stimulus has been suggested to help in tagging onto the target speech and separating out the background noise (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011). Such a separation of the target speech from the background noise has been suggested to be one of the mechanisms that aid in speech perception in the presence of background noise.

Though the previous studies make claims about the context effect as an important mechanism for speech perception in noise, none of the studies have actually investigated the context effect in the presence of noise. In the current study the effect of a background noise on the context-dependent encoding was investigated. As the

previous studies suggest that the context-dependent effect is important in separating the target speech from the contextual syllables, it was hypothesized that the context dependent effect to repetitive stimuli would be higher in the presence of noise.

The contextual effect seen in the presence of noise was large similar to that observed in quiet. In the presence of noise. The context dependent effect was seen at the H1, H2 H3 and H4, while in quiet, the context effect was present only at H1, H2 and H3. Additionally, comparison of the magnitude of contextual effect showed higher effects in noise compared to quiet. Examination of the individual data also showed that the in most of the individuals contextual effect was more.

In line with the hypothesis, higher contextual effect was observed in the presence of noise. However the direction of effect was not the same as was hypothesized. Based on previous studies it was expected that a repetition-related enhancement would be higher in the presence of noise. However, the FFRs in the variable condition showed higher amplitudes than the FFRs in the repetitive condition. The reason for such an effect difference in the direction of the effect has been discussed in the previous section.

Based on literature (Chandrasekaran et al., 2009; Malmierca et al., 2014; Skoe & Kraus, 2010; Slabu et al., 2012) and the current findings, it can be inferred that the novelty detection mechanism (mediated by the cortico-fugal pathway) may play an important role in the stimulus encoding in the presence of noise. The preferential encoding of novel stimuli might thus aid in better speech encoding. Such a preferential brainstem encoding of sound in the presence of noise suggests that the mechanisms to improve speech encoding in noise based on ongoing sound statistics are present at multiple levels in the auditory neural pathway.

In a study done by Russo, Nicol, Musacchia and Kraus (2004), there was reduction only in the amplitude of onset response but there was no change in the sustained responses. The background noise disturbed only the transient responses whereas the sustained response was more resistant to the presence of noise. The reason for reduction of amplitude in the transient response in the presence of background noise was attributed to the reduced neural synchrony, while the intact encoding reflected in the sustained portion enabled the vowel perception. In our study also the FFRs recorded in repetitive as well variable paradigms were not influenced by the presence of 10dBSNR noise.

5.3 Relation between Context Dependent Encoding and Speech Perception in Noise

Previous studies on context dependent brainstem encoding of speech have suggested that the contextual effect is related to speech perception in noise (Chandrasekaran et al., 2009; Maruthy et al., 2017; Parbery-Clark et al., 2011; Strait et al., 2011). In the current study it was the contextual effects on FFRs which were studied in with and without the background noise. It was hypothesized that the contextual effect on FFRs in the presence of noise might be closely related to the speech perception in noise. It was expected that the relationship between contextual effect in noise and speech perception in noise would be stronger than the relationship between contextual effect in quiet and speech perception in noise.

Contrary to the proposed hypothesis and the findings in the previous studies, the contextual effect did not show any correlation with speech perception in noise. This was true for the contextual effect in quiet and in noise. This was a surprising finding, considering that all the previous studies showed significant correlation

between context-dependent encoding and speech perception in noise. The difference in the results may be due difference in methodology between the current study and the previous studies.

The first major difference in methodology between the current study and the previous studies, as has been discussed earlier is the use of a single polarity. It could be speculated that the contextual effect on the spectral following responses might not directly related to speech perception, while the contextual effect on the envelope following responses (FFRs in alternate polarity) might be related to speech perception in noise. The exact reason as to why such a phenomenon would happen is not clear. Future studies should try to probe into the influence of stimulus polarity and contextual effects and their relationship with speech perception in noise.

Another major difference in methodology between the current study and the previous studies is the material used for assessing speech perception in noise. The previous studies used hearing in noise test (HINT) (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011) and Kannada sentence identification test (Maruthy et al., 2017) and correlated the speech perception in noise with the contextual brainstem encoding. In the current study however, nonsense monosyllables were used to assess speech perception in noise. The duration and the extrinsic redundancy of the speech materials were different between the current study and the previous studies. As the contextual effect requires the repetition of certain desired characteristics in the sound, it could be inferred that our speech perception material did not sufficiently tap the repetition/novelty detection mechanism in the brain. In the previous studies, use of longer speech material may have activated repetition/novelty mechanisms, due to repetition of certain desired acoustic characteristics. The activation of these

repetition/novelty detection mechanisms might have then aided in improving speech perception in noise. However, this can be confirmed by systematically studying the relationship between contextual brainstem encoding and speech perception in noise using materials of differing complexity and duration.

Chapter 6

SUMMARY AND CONCLUSIONS

Human communication rarely occurs in optimal listening environments; rather we are often surrounded by background noise. Despite the frequent presence of noise, humans are able to separate target sounds from a complex soundscape. A key mechanism thought to underlie accurate perception in noise is the auditory system's ability to extract regularities from an ongoing acoustic signal. The aim of the present study was to test the context dependent encoding of speech in the presence of noise and to study the role of context dependent encoding in noise on speech perception in noise.

Fifteen normal hearing adults in the age range of 18 to 25 years were included in the study. FFRs were recorded in repetitive (only /da/) and variable (/da/ stimulus with /bu/, /bi/ and /gi/) paradigms, in quiet as well as noise conditions. Spectral analysis was carried out only for /da/ stimulus. FFRs in the repetitive and variable paradigms between quiet and noise conditions to derive context dependent brainstem encoding of speech. Fundamental frequency (H1-100Hz) and the higher harmonics were compared. The harmonics of interest were, harmonics of FFRs (H2, H3 & H4).

Paired sample t test was carried out to compare between repetitive and variable paradigm in quiet condition and in noise conditions separately. Then the FFRs in repetitive and variable paradigms were also compared separately between quiet and noise conditions. Further correlation between SNR-50 and context dependent effect was tested separately for quiet and noise conditions. Context effect was derived by subtracting the variable paradigm from repetitive paradigm and was tested for its correlation with SNR-50.

The results showed that variable paradigm yielded higher amplitudes compare to repetitive paradigm in both quiet and noise conditions. When repetitive paradigm was compared in quiet and noise conditions, there was no significant difference in H1, H2 and H4 amplitudes, but there was a higher amplitude seen in H3 in quiet condition. In variable paradigm on the other hand, there was no change in amplitude between the two conditions. Results of correlation showed that, there was no significant correlation between SNR-50 and context dependent encoding is evident H1,H2,H3 &H4 in both quiet and noise conditions.

From the present study it can be speculated that context-dependent effect may manifests as enhancement in the envelope encoding of repetitive stimuli and enhancement in the spectral encoding of the novel stimuli. The increase in the FFR amplitudes in the variable condition can be attributed to novelty detection mechanism in the auditory pathway. Such a novelty detection mechanism is possibly modulated by the corticofugal pathways innervating the brainstem. So it can be inferred that the novelty detection mechanism mediated by the cortico-fugal pathway may plays an important role in the stimulus encoding in the presence of noise. The preferential encoding of novel stimuli might thus help aid in better speech encoding.

In the present study, we used a strong research design and controlled the possible extraneous variables by adopting multiple measurements. Future studies can use the same research design having a large sample. Relationship between contextual brainstem encoding and speech perception in noise can be assessed using materials of differing complexity and duration. Also context dependent encoding can be assessed by changing the stimulus polarity.

REFERENCES

- Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, *245*(1–2), 35–47.
- Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in cognitive sciences*, *11*(11), 458-465.
- Banai, K., Nicol, T., Zecker, S. G., & Kraus, N. (2005). Brainstem timing: implications for cortical processing and literacy. *Journal of Neuroscience*, *25*(43), 9850-9857.
- Boutros, N. N., Gjini, K., Urbach, H., & Pflieger, M. E. (2011). Mapping Repetition Suppression of the N100 Evoked Response to the Human Cerebral Cortex. *Biological Psychiatry*, *69*(9), 883–889.
- Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press.
- Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*, *47*(2), 236-246.
- 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*(3), 311-319.
- 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*(4), 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(5), 758-767.
- Galbraith, G. C., Jhaveri, S. P., & Kuo, J. (1997). Speech-evoked brainstem frequency-following responses during verbal transformations due to word repetition. *Electroencephalography and clinical neurophysiology*, 102(1), 46-53.
- Galbraith, G. C., Threadgill, M. R., Hemsley, J., Salour, K., Songdej, N., Ton, J., & Cheung, L. (2000). Putative measure of peripheral and brainstem frequency-following in humans. *Neuroscience Letters*, 292(2), 123–127.
- 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(14), 8081-8086.
- Gnanateja, G. N., Ranjan, R., & Sandeep, M. (2012). Physiological bases of the encoding of speech evoked frequency following responses. *Journal of All India Institute of Speech and Hearing*, 31, 215-9.
- Gnanateja, G. N., Ranjan, R., Firdose, H., Sinha, S. K., & Maruthy, S. (2013). Acoustic basis of context dependent brainstem encoding of speech. *Hearing Research*, 304, 28–32.
- 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.
- Hanan Tonse & Sandeep, M. (2010). Effect of Context on Brainstem Encoding of Speech. Unpublished Masters Dissertation submitted to University of Mysore, Mysore.

- Keuroghlian, A. S., & Knudsen, E. I. (2007). Adaptive auditory plasticity in developing and adult animals. *Progress in neurobiology*, 82(3), 109-121.
- 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(2), 111-115.
- Kral, A., & Eggermont, J. 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(1), 259-269.
- Krishnan, A. (2002). Human frequency-following responses: representation of steady-state synthetic vowels. *Hearing research*, 166(1), 192-201.
- Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and language*, 110(3), 135-148.
- Krishnan, A., Gandour, J. T., Bidelman, G. M., & Swaminathan, J. (2009). Experience dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*, 20(4), 408.
- Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, 25(1), 161-168.
- Ma, X., & Suga, N. (2008). Corticofugal modulation of the paradoxical latency shifts of inferior collicular neurons. *Journal of neurophysiology*, 100(2), 1127-1134.
- Malmierca, M. S., Sanchez-Vives, M. V., Escera, C., & Bendixen, A. (2014). Neuronal adaptation, novelty detection and regularity encoding in audition. *Frontiers in Systems Neuroscience*, 8.

- 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*(4), 415–424.
- Maruthy, S., Kumar, U. A., & Gnanateja, G. N. (2017). Functional Interplay Between the Putative Measures of Rostral and Caudal Efferent Regulation of Speech Perception in Noise. *Journal of the Association for Research in Otolaryngology*, *18*(4), 635–648.
- 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 Sciences*, *104*(40), 15894-15898.
- Parbery-Clark, A., Strait, D. L., & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, *49*(12), 3338-3345.
- Ping, J., Li, N., Galbraith, G. C., Wu, X., & Li, L. (2008). Auditory frequency-following responses in rat ipsilateral inferior colliculus: *NeuroReport*, *19*(14), 1377–1380.
- Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural brain research*, *156*(1), 95-103.
- Russo, N. M., Skoe, E., Trommer, B., Nicol, T., Zecker, S., Bradlow, A., & Kraus, N. (2008). Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clinical Neurophysiology*, *119*(8), 1720-1731.
- Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology*, *115*(9), 2021-2030.

- Skoe, E., & Kraus, N. (2010). Auditory brainstem response to complex sounds: a tutorial. *Ear and hearing, 31*(3), 302.
- Skoe, E., & Kraus, N. (2010). Hearing It Again and Again: On-Line Subcortical Plasticity in Humans. *PLOS ONE, 5*(10), e13645.
- Slabu, L., Grimm, S., & Escera, C. (2012). Novelty Detection in the Human Auditory Brainstem. *The Journal of Neuroscience, 32*(4), 1447–1452.
- Song, J. H., Skoe, E., Wong, P. C., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of cognitive neuroscience, 20*(10), 1892-1902.
- Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency—effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience, 29*(3), 661-668.
- Strait, D. L., Hornickel, J., & Kraus, N. (2011). Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behavioral Brain Function, 7*(1), 44.
- Suga, N. (2008). Role of corticofugal feedback in hearing. *Journal of comparative physiology A, 194*(2), 169-183.
- Suga, N., Gao, E., Zhang, Y., Ma, X., & Olsen, J. F. (2000). The corticofugal system for hearing: recent progress. *Proceedings of the National Academy of Sciences, 97*(22), 11807–11814.
- Suga, N., Ji, W., Ma, X., Tang, J., Xiao, Z., & Yan, J. (2011). Corticofugal modulation and beyond for auditory signal processing and plasticity. In *Auditory and vestibular efferents*(pp. 313-352). Springer New York.
- Suga, N., Xiao, Z., Ma, X., & Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron, 36*(1), 9-18.

- Sumesh, K., & Animesh Barman. (2007). Brainstem Responses to Speech in Normal Hearing and Cochlear Hearing Loss Individuals. Unpublished Masters Dissertation submitted to University of Mysore, Mysore.
- Swaminathan, J., Krishnan, A., & Gandour, J. T. (2008). Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport*, *19*(11), 1163.
- Terreros, G., & Delano, P. H. (2015). Corticofugal modulation of peripheral auditory responses. *Frontiers in Systems Neuroscience*, *9*.
- 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*(3), 299-317.
- Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in cognitive sciences*, *13*(12), 532-540.
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature neuroscience*, *10*(4), 420.
- Yan, W., & Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nature Neuroscience*, *1*(1), 54–58.
- Zhang, Y., & Suga, N. (2005). Corticofugal feedback for collicular plasticity evoked by electric stimulation of the inferior colliculus. *Journal of Neurophysiology*, *94*(4), 2676-2682.