Effect of Context on Brainstem Encoding of Speech

[1]Hanan Tonse & [2]Sandeep M.

Abstract

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. It is not well understood 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 and 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 suggest a possibility of online plasticity at the level of brainstem regulated by the corticofugal network.

Key words: *Speech evoked brainstem responses, plasticity, context and corticofugal pathway.*

Introduction

It was believed that the processing operations that occur in the relay nuclei of the brainstem are not specific to speech sounds (Scott & Johnsrude, 2003). However, experimenters in the past decade have measured the neural transcription of speech from the auditory brainstem (Johnson, Nicol, Zecker & Kraus, 2008a; Hornickel, Skoe, Nicol, Zecker & Kraus, 2009; Tzounopoulos & Kraus, 2009). The presence of 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 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 allows for a means to examine how plasticity occurs at the level of brainstem in humans.

Two models are currently being debated to account for the effect of experience on the representation of speech at the brainstem. One is the local reorganization model, which states that the brainstem function is modified over a long time scale (Krishnan & Gandour, 2009; Krishnan, Swaminathan & Gandour, 2009). The other model 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 & Yan, 1997).

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 and medial geniculate body. The auditory cortex and the corticofugal pathway is reported to evoke small short term changes in the subcortical nuclei, in response to a sound that is repeatedly delivered (Suga, Xiao, Ma & Ji, 2002). It is reported that, when the sound becomes relevant by associative learning, its responses are enhanced by the activation of the corticofugal pathway. This 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 the corticofugal 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 was a need to study the brainstems sensitivity to ongoing contextual demands.

Chandrasekaran, Hornickel, Skoe, Nicol and Kraus (2009) elicited brainstem response to speech syllable

[1]E-mail: hanantonse@gmail.com; [2]Lecturer in Audiology, E-mail: msandeepa@gmail.com.

/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 (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.

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 stimulus). Here, both speeches 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 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 have help in better understanding the complex mechanisms in the brainstem. Hence the present study was taken up with an objective to study the effect of context on the brainstem responses elicited by stimulus /da/.

Method

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 (<15 dBHL) at octave frequencies between 250 Hz and 8 kHz, which was measured using a calibrated GSI-61 clinical audiometer. They had type 'A' tympanogram with the presence of ipsilateral and contralateral reflexes, measured by calibrated GSI-Tympstar Middle Ear Analyzer which ensured normal middle ear functioning. 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 0 dB SNR. All the subjects were meritorious students from different parts of the country, pursuing their bachelors and masters degree in Speech and Hearing.

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/ of 40 ms duration (core stimulus) was obtained from Professor Nina Kraus, Principal investigator at the Auditory Neuroscience lab, Northwestern University, Chicago. 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 16000 Hz 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 16000 Hz. The recorded syllables were then analyzed using the Formant based analysis by synthesis of the ACOPHON 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 1 and 2 respectively.

The sampling frequency of both the synthetically generated syllables (/ba/ and /ta/) was converted from 16000 Hz to 48000 Hz, 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 48000 Hz 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 40 ms across all four stimuli.

All the four stimuli were individually normalized and then group normalized to obtain equal average RMS power of 93.4 dBSPL. They were then loaded into the personal computer with the Bio-Logic Navigator Pro AEP Software (Version 7.0). The synthetic speech

Table 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 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 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 dBSPL. 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 approximated mean behavioral threshold was 25 dBSPL for the syllable /da/ and 30 dBSPL for the syllables /ba/ and /ta/, and white noise respectively.

Procedure

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 5 kOhms 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. In condition 1, 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/P_{300} 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%. 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 order of recording was randomized to avoid the order effect. The Figure 1 represents the different stimulus conditions.

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.

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, and F 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, and F.

The response amplitude between 0 and 1 ms was approximately 0µV across all the recordings. Hence there was no baseline amplitude correction used and the peak amplitude of waves B, C, D, E, and F were obtained directly from the instrument by placing the cursor on the respective waves.

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 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; (103-121 Hz), (454-719 Hz) and (721-1155 Hz). 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, 2008b). 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.

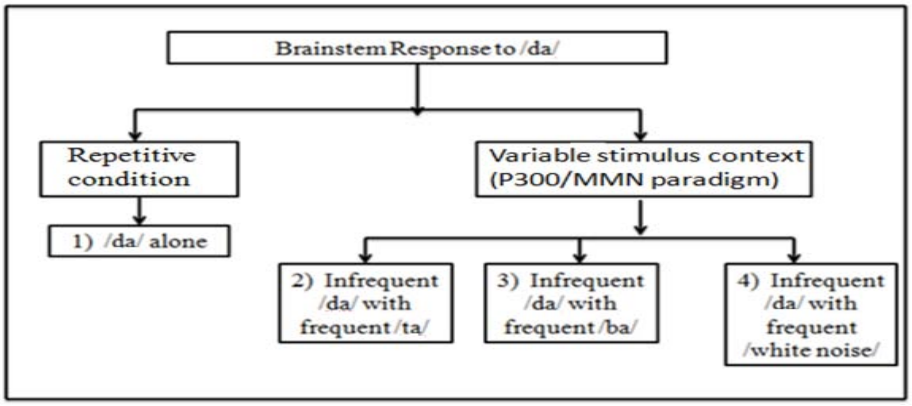

Figure 1: Representation of different stimulus conditions used.

Table 3: Stimulus and acquisition parameters used to elicit the Brainstem Responses

S. No	Stimulus Parameters	
1	Stimuli	Repetitive context: /da/ Variable context: Frequent stimuli: /ba/ or /ta/ or noise
2	Frequent to infrequent ratio	2:1
3	Ear	Right/left
4	Duration of stimulus	40 ms
5	Intensity	70 dBnHL
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	Amplification	100000
4	Artifact rejection	+ or – 23.8 µV
5	Filter setting	100-2000 Hz

The data thus obtained was used for the comparison between the responses obtained in repetitive stimulus condition and in different stimulus contexts and also for the comparison across the responses obtained in different contexts.

Results

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).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, and F, 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.

Results of subjective analysis: The responses were subjectively analyzed to identify the waves V, A, B C, D, E, and F. It was observed in the analysis that not all the recordings had all the aforementioned waves. Waves V, A, D, E, and F had higher percentage (>83%) compared to waves B and C (<70%) across all four conditions. Hence waves B and C were not analyzed on Repeated measures ANOVA.

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 2. The mean and standard deviation of peak latency of wave V and A, and V to A peak to peak amplitude is given in Table 4.

It is evident from the table that 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 was 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. Results of the Sidak Post-Hoc test showed that 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. Also, 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. Furthermore, 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.

Results of sustained response: 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 3. The mean and standard deviation of peak latency and amplitude obtained from the subjects across the four test conditions are given in Table 5.

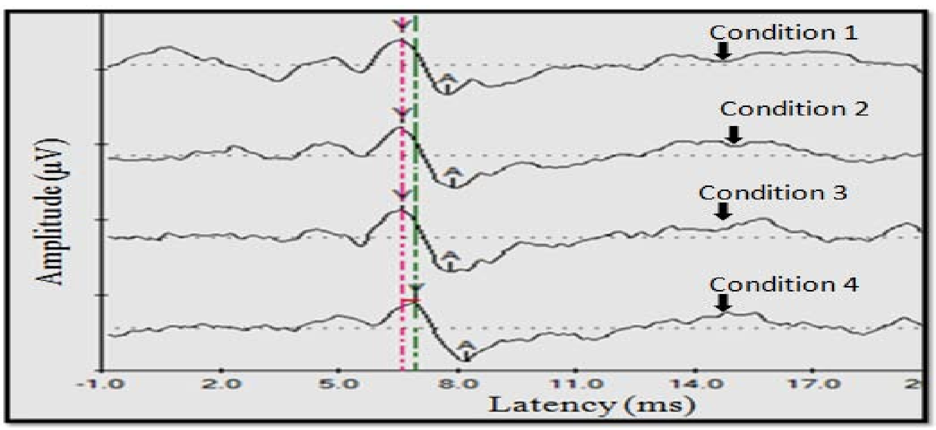

Figure 2: Onset responses elicited by syllable /da/ across 4 conditions in a representative subject.

Table 4: 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

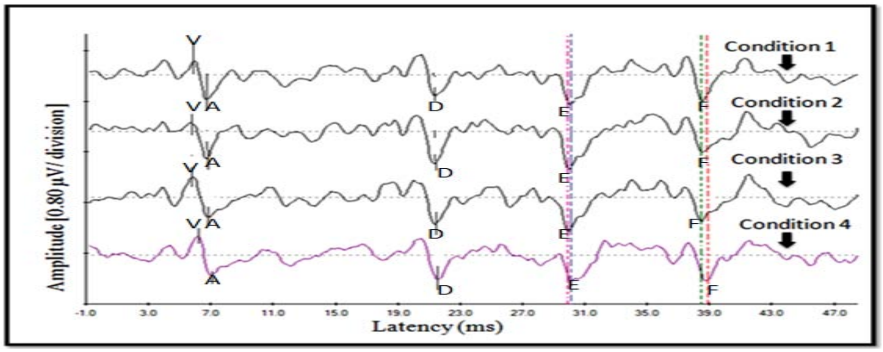

Figure 3: Sustained response elicited by syllable /da/ across 4 conditions in a representative subject.

Table 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 (µV)		
	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 5 that the mean peak latency of waves D, E and F in condition 1 differed from that in conditions 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 post-hoc test results showed that 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. Also, 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. Furthermore, there was no significant difference in the mean peak latency of wave F across the conditions 1, 2 and 3.

Results of objective analysis

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 6.

Table 6: The mean and standard deviation (SD) of amplitude of fundamental frequency (F0), first formant (F1) and higher harmonics (HF)

Condition	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 6 that the mean amplitudes of F0, F1 and HF were 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. 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.

Discussion

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, and F 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 white noise (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 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 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, et al., (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, et al., 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/, with respect to the burst spectrum and F2 transition. The burst spectra of /ba/ has primary concentration of energy at the low frequencies (500-1500 Hz) while /da/ has a relatively flat or high frequency (>4000 Hz) 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 brainstems 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, et al., (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.

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.

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 (721-1155 Hz) 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, Hornickelet al., (2009), wherein the peak amplitudes corresponding to lower harmonics H2 and H4 (200 Hz and 400 Hz) 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.

Conclusions

From the findings of the present study it can be concluded 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.

References

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.

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.

Halle, M., Huges, G. W., & Radley, J. P. (1957). Acoustic properties of consonants. *Journal of the Acoustical Society of America, 29*(1), 107-116.

Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Subcortical differentiation of voiced stop consonants: relationships to reading and speech in noise perception. *Proceedings of the National Academy of Science, 106*(31), 13022–13027.

Howard, D. M., & Angus, J. A. S (.2009). *Acoustics and Psychoacoustics.* (4th edition), UK; Linacre House, Jordan Hill.

Johnson, K. L., Nicol, T., Zecker, S. G., & Kraus, N. (2008a). Developmental plasticity in the human auditory brainstem. *Journal of Neuroscience, 28*(15), 4000-4007.

Johnson, K, L., Nicol, T, G., Zecker, S. G, Bradlow, A., Skoe, E., & Kraus, N. (2008b). Brainstem encoding of voiced consonant-vowel stop syllables. *Clinical Neurophysiology, 119,* 2623-2635.

Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem inprocessing 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.

Luo, F., Wang, Q., Kashini, A., & Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *Journal of Neuroscience, 28*, 11615-11621.

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.

Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology, 115*, 2021-2030.

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. (2010). Hearing it again and again: on-line subcortical plasticity in humans. *PlosONE, 5*, (10), e13645.

Suga, N., Xiao, Z., Ma, X., and Ji, W. (2002). Plasticity and corticofugal modulation or 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.

Zhang, Y., Suga, N., & Yan, J. (1997). Corticofugal modulation of frequency processing in bat auditory system. *Nature, 387*, 990-903.