Contribution of resolved and unresolved harmonic regions to brainstem speech-evoked responses in quiet and in background noise

M. Laroche,[1] H.R. Dajani,[1] A.M. Marcoux[2]

[1]School of Information Technology and Engineering, University of Ottawa, ON; [2]Audiology and Speech-Language Pathology Program, University of Ottawa, ON, Canada

Abstract

Speech auditory brainstem responses (speech ABR) reflect activity that is phase-locked to the harmonics of the fundamental frequency (F0) up to at least the first formant (F1). Recent evidence suggests that responses at F0 in the presence of noise are more robust than responses at F1, and are also dissociated in some learning-impaired children. Peripheral auditory processing can be broadly divided into resolved and unresolved harmonic regions. This study investigates the contribution of these two regions to the speech ABR, and their susceptibility to noise. We recorded, in quiet and in background white noise, evoked responses in twelve normal hearing adults in response to three variants of a synthetic vowel: i) Allformants, which contains all first three formants, ii) F1Only, which is dominated by resolved harmonics, and iii) F2&F3Only, which is dominated by unresolved harmonics. There were no statistically significant differences in the response at F0 due to the three variants of the stimulus in quiet, nor did the noise affect this response with the Allformants and F1Only variants. On the other hand, the response at F0 with the F2&F3Only variant was significantly weaker in noise than with the two other variants ($p<0.001$). With the response at F1, there was no difference with the Allformants and F1Only variants in quiet, but was expectedly weaker with the F2&F3Only variant ($p<0.01$). The addition of noise significantly weakened the response at F1 with the F1Only variant ($p<0.05$), but this weakening only tended towards significance with the Allformants variant ($p=0.07$). The results of this study indicate that resolved and unresolved harmonics are processed in different but interacting pathways that converge in the upper brainstem. The results also support earlier work on the differential susceptibility of responses at F0 and F1 to added noise.

Correspondence: Hilmi R. Dajani, School of Information Technology and Engineering, University of Ottawa, 161 Louis Pasteur, Ottawa, ON, Canada.
E-mail: hdajani@site.uottawa.ca

Key words: peripheral auditory processing, speech auditory brainstem responses.

Audiology Research 2011; 1:e7
doi:10.4081/audiores.2011.e7

Parts of this work were presented at the "AHS 2010 - International Conference on Adult Hearing Screening", Cernobbio (Italy), June 10-12, 2010.

Introduction

Speech auditory brainstem responses (speech ABR) can be recorded using scalp surface electrodes. With vowel stimuli, speech ABR is strikingly speech-like since it reflects neural activity that is phase-locked to the fundamental frequency (F0) and its harmonics at least up to and including the first formant frequency (F1). Recent evidence suggests that responses at F0 in the presence of interfering noise are more robust than responses at F1 (Russo *et al.*, 2004; Parbery-Clark *et al.*, 2009). There is also evidence that this dissociation between the two responses is present in some learning-impaired children, suggesting that speech ABR may be useful as a marker of central auditory processing disorders (Johnson *et al.*, 2005). Speech ABR may also be useful for the objective selecting and fitting of modern hearing aids, whose performance cannot be easily tested with simple non-speech sounds (Aiken and Picton, 2008).

Peripheral auditory processing can be broadly divided into a lower frequency region where the harmonics are resolved and a higher frequency region where they are not. Psychophysical and neurophysiological studies have determined that the pitch of a harmonic complex is more salient for resolved harmonics than for unresolved harmonics (e.g. Larsen *et al.*, 2008). Some studies have suggested different mechanisms for processing of the pitch of resolved and unresolved harmonics (Carlyon and Shackleton, 1994), but the existence of two mechanisms has also been questioned (Gockel and Carlyon, 2004). Because speech is a broadband signal, it is of interest to determine the contribution of the resolved and unresolved harmonic regions to the speech ABR, and whether the responses are differentially susceptible to added noise.

Methods

Twelve subjects (six females), 25-40 years old, participated in this study. None of the subjects had a history of hearing difficulty, and all had thresholds of 15 dB HL or less at 500, 1000, 2000, and 4000 Hz in their right ear. This study was carried out in accordance with the regulations of the University of Ottawa Research Ethics Board. We recorded speech ABRs using the BioMARK™ system (Biological Marker of Auditory Processing, Bio-logic Systems Corporation) with one measuring electrode at the vertex (Cz) and a reference electrode at the ipsilateral earlobe. The speech stimuli consisted of three variants of a 300 msec synthetic vowel /a/ generated using formant synthesis (F0=100 Hz, F1=700 Hz, F2=1220 Hz, F3=2600 Hz): i) Allformants which contains all first three formants, ii) F1Only whose content is dominated by resolved harmonics (based on the analysis in Micheyl and Oxenham, 2004), and iii) F2&F3Only whose content is dominated by unresolved harmonics. The sum of the F1Only and F2&F3Only stimuli is equal to the Allformants stimulus. Two noise conditions were also used: quiet and white noise added at an SNR of -5 dB relative to the Allformants stimulus. In the noisy condition, therefore, the noise level was the same with all three vowel variants. The order of the presentation of the six stimuli (3 vowel variants in 2 noise conditions) was randomized between subjects.

The stimulus sound levels were calibrated with the earphone inserted in a 2cc coupler attached to a Brüel & Kjaer Artificial Ear type 4152, and a Brüel & Kjaer Sound Level Meter type 2230. The level of each formant was kept the same in the stimulus variant in which it was present, leading to a maximum stimulus sound level of 80 dB SPL with the Allformants variant in noise. A control experiment, with the subject connected using the regular experimental procedure, but with the earphone inserted in an ear simulator that presents the same acoustic load to the transducer as when the insert earphone is in the ear confirmed that there was no electrical leakage from the sound generating equipment to the electrodes. Responses were synchronously averaged over 3000 stimulus repetitions that alternated in polarity. The responses at F0 were obtained by averaging responses to the two polarities, while the responses at F1 were obtained by averaging the difference between the responses to the two polarities (Aiken and Picton, 2008). Statistical analysis was done on the amplitudes at F0 and F1 in the response spectra, using repeated-measure t-tests with Bonferroni corrections for multiple comparisons.

Results

There were no statistically significant differences in the response at F0 due to the three variants of the stimulus in quiet, nor did the noise affect this response with the Allformants and F1Only variants (Figure 1). On the other hand, the response at F0 with the F2&F3Only variant was significantly weaker in noise than with the two other variants ($p<0.001$). In the quiet condition, the magnitude of the vector of the grand-averaged response at F0 to Allformants was less than half that of the vector sum of the grand-averaged responses to the F1Only and F2&F3Only stimulus variants, indicating that the contributions of the resolved and unresolved harmonic regions add up non-linearly when both are present in the stimulus (Figure 2). The latencies of the responses at F0 to the three variants in quiet ranged from 6.4 to 6.7 ms, excluding stimulus transmission time, indicating a common origin in the upper brainstem.

With the response at F1, there was no statistically significant difference with the Allformants and F1Only variants in quiet, but was expectedly weaker with the F2&F3Only variant ($p<0.01$) because the component at F1 is weak in this stimulus (Figure 3). The addition of noise significantly diminished the response at F1 with the F1Only variant ($p<0.05$), although this weakening only tended towards significance with the Allformants variant ($p=0.07$).

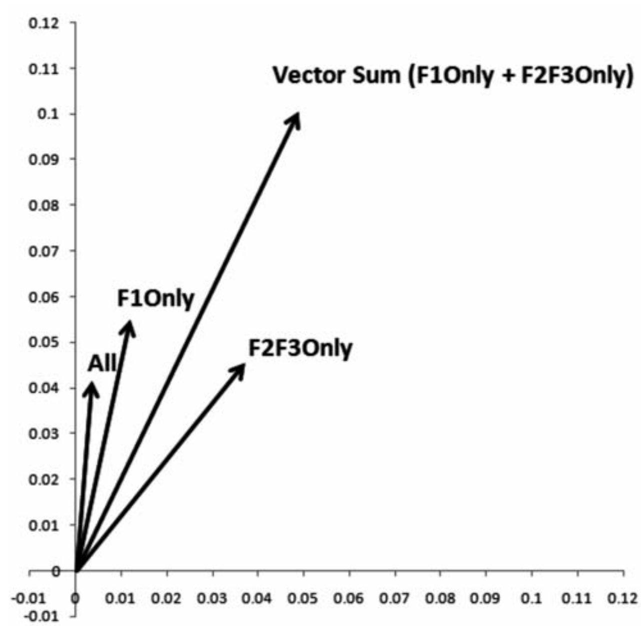

Figure 2. Vector diagram of grand-average responses at F0 in quiet. x and y axes are in micro-volts.

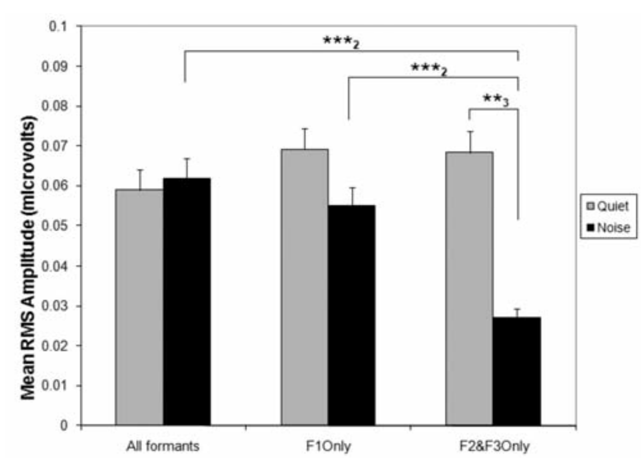

Figure 1. Comparison of response amplitudes (mean ± S.E.) at F0 with different stimulus variants in quiet and noise. Figure symbols: 1 Comparisons between responses to stimulus variants in quiet, 2 Comparisons between responses to stimulus variants in noise, 3 Comparisons between responses to one stimulus variant in quiet and noise. ***: P<0.001, **: P<0.01, *: P<0.05, +: P<0.1.

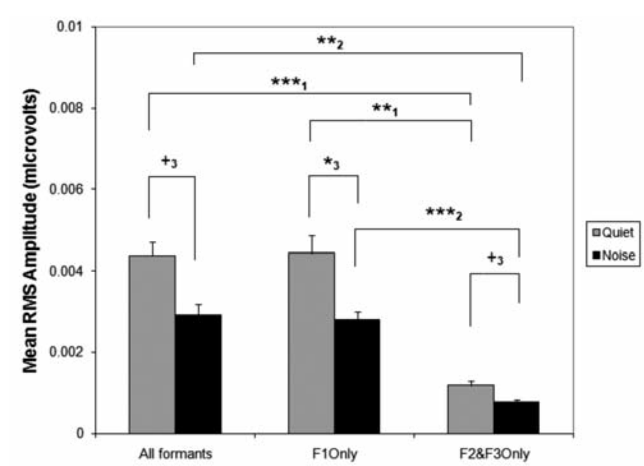

Figure 3. Comparison of response amplitudes (mean ± S.E.) at F1 with different stimulus variants in quiet and noise. Figure symbols are as described in the caption of Figure 1.

Discussion

The amplitude of the response at F0 generated with the stimulus dominated by resolved harmonics (F1Only) in quiet was not different from that generated with the stimulus dominated by unresolved harmonics (F2&F3Only), even though based on psychophysical studies the pitch of a harmonic complex is more salient perceptually for resolved harmonics than for unresolved harmonics (Shackleton and Carlyon, 1994). It is not clear why the strength of the evoked response does not follow perceptual saliency, but this may be because pitch salience depends on other forms of neural coding than the phase-locked activity that underlies the evoked response (e.g. rate-place coding), and on higher (cortical) levels of processing (Penagos *et al.*, 2004).

The responses at F0 obtained from the Allformants stimulus were not statistically different between the quiet and noisy conditions. This robustness to noise is consistent with the results of previous studies (Cunningham *et al.*, 2002; Parbery-Clark *et al.*, 2009), and is probably explained by the contribution of broad regions of the stimulus spectrum to the response. Even at the peripheral level, the fundamental periodicity is well represented in the interspike intervals recorded from the auditory nerve fibers with both low and high characteristic frequencies (Secker-Walker and Campbell, 1990; Cariani and Delgutte, 1996). The responses generated from the resolved vs. unresolved harmonic regions of the stimulus depicted a more complex picture. The response to the higher frequency unresolved harmonics region (i.e. with the F2&F3Only stimuli) was significantly affected by noise, although the differential susceptibility to noise with the unresolved harmonics versus the resolved harmonics may be due to the difference in the effective SNR level in each variant.

The latencies obtained between the stimuli and the responses in quiet are similar for all the stimulus variants, and are within the range of 5-10 ms that has been obtained in previous studies (Chandrasekaran and Kraus, 2010). Although multiple sites in the brainstem have been implicated in the generation of the speech ABR, the latencies we found fit with a main generating site in the upper brainstem (Akhoun *et al.*, 2008; Chandrasekaran and Kraus, 2010). The vector analysis of the responses at F0 depicts a more complex picture regarding the origin of the response. The response to Allformants (which contains both resolved and unresolved harmonics) in quiet is much smaller than the vector sum of the responses to F1Only and F2&F3Only, despite the Allformants stimulus being equal to the sum of the F1Only and F2&F3Only stimuli. This suggests that the responses at F0 due to the resolved and unresolved harmonics interact non-linearly either before reaching the upper brainstem or within it.

For the response at F1, the results showed it was weaker in quiet for the F2&F3Only stimulus compared to the Allformants and F1Only stimuli. This was expected because both Allformants and F1Only stimuli contain a strong F1 component, while it is weak in the F2&F3Only stimulus. The results also show that the response at F1 is more susceptible to noise than the response at F0, as reflected by the lower amplitudes obtained for the noisy condition, which is consistent with the findings of previous studies (Russo *et al.*, 2004; Johnson *et al.*, 2005). The basis behind the generation of the two responses may help explain the difference in robustness. A broad frequency range of the stimulus components contributes to the response at F0, while the response at F1 is due to a specific single frequency component.

Conclusions

Speech ABR has shown promise as a marker of central processing disorders in children (Johnson *et al.*, 2005), and is potentially useful in assessment of adult hearing impairment. There is a need, however, for a better characterization of the response and its basic underlying mechanisms in quiet and in noise. The results of this study indicate that the different pathways that process the pitch of resolved and unresolved harmonics converge in the upper brainstem. Moreover, along these pathways, there are non-linear interactions between the processing of resolved and unresolved harmonics, but the precise locus of these interactions is unknown. The results of the study also support earlier work on the differential susceptibility of speech ABR at F0 and F1 to added noise.

References

Aiken, S. J., Picton, T.W., 2008. Envelope and spectral frequency-following responses to vowel sounds, Hearing Res 245, 35-47.

Akhoun, I., Moulin, A., Jeanvoine, A., Ménard, M., Buret, F., Vollaire, C., Scorretti, R., Veuillet, E., Berger-Vachon, C., Collet, L., Thai-Van, H., 2008. Speech auditory brainstem response (speech ABR) characteristics depending on recording conditions, and hearing status: An experimental parametric study. J Neurosci Meth 172, 196-205.

Cariani, P.A., Delgutte, B., 1996. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. J Neurophysiol 76, 1698-1716.

Carlyon, R.P., Shackleton, T.M., 1994. Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms? J Acoust Soc Am 95, 3541-3554.

Chandrasekaran, B., Kraus, N., 2010. The scalp-recorded brainstem response to speech: Neural origins and plasticity. Psychophysiology 47, 236-246.

Cunningham, J., Nicol, G.T., King, C., Zecker, S.G., Kraus, N., 2002. Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. Hearing Res 169, 97-101.

Gockel, H., Carlyon, R.P., 2004. Across-frequency interference effects in fundamental frequency discrimination: Questioning evidence for two pitch mechanisms. J Acoust Soc Am 116, 1092-1104.

Johnson, K.L., Nicol, G.T., Kraus, N., 2005. Brain stem Response to speech: A biological marker of auditory processing. Ear Hearing 26, 424-434.

Larsen, E., Cedolin L., Delgutte B., 2008. Pitch representations in the auditory nerve: Two concurrent complex tones. J Neurophysiol 100, 1301-1319.

Micheyl, C., Oxenham, A.J., 2004. Sequential F0 comparisons between resolved and unresolved harmonics: No evidence for translation noise between two pitch mechanisms. J Acoust Soc Am 116, 3038-3050.

Parbery-Clark, A., Skoe, E., Kraus, N., 2009. Musical experience limits the degradative effects of background noise on the neural processing of sound. J Neurosci 29,14100-14107.

Penagos, H., Melcher, J.R., Oxenham, A.J., 2004. A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. J Neurosci 24, 6810-6815.

Russo, N., Nicol, T., Musacchia, G., Kraus, N., 2004. Brainstem responses to speech syllables. Clin Neurophysiol 115, 2021-2030.

Secker-Walker, H.E., Campbell, S.L., 1990. Time-domain analysis of auditory-nerve-fiber firing rates. J Acoust Soc Am 88, 1427-1436.

Shackleton, T.M., Carlyon, R.P., 1994. The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. J Acoust Soc Am 95, 3529-3540.