Frequency-dependent Effects of Background Noise on Subcortical Response Timing

A. Tierney[1], **A. Parbery-Clark**[1,2], **E Skoe**[1,2], and **N Kraus**[1,2,3,4,5,*]

[1]Auditory Neuroscience Laboratory, Northwestern University, Evanston, Illinois, http://www.brainvolts.northwestern.edu

[2]Communication Sciences, Northwestern University, Evanston, Illinois

[3]Institute for Neuroscience, Northwestern University, Evanston, Illinois

[4]Neurobiology and Physiology, Northwestern University, Evanston, Illinois

[5]Otolaryngology, Northwestern University, Evanston, Illinois

Abstract

The addition of background noise to an auditory signal delays brainstem response timing. This effect has been extensively documented using manual peak selection. Peak picking, however, is impractical for large-scale studies of spectrotemporally complex stimuli, and leaves open the question of whether noise-induced delays are frequency-dependent or occur across the frequency spectrum. Here we use an automated, objective method to examine phase shifts between auditory brainstem responses to a speech sound (/da/) presented with and without background noise. We predicted that shifts in neural response timing would also be reflected in frequency-specific phase shifts. Our results indicate that the addition of background noise causes phase shifts across the subcortical response spectrum (70 – 1000Hz). However, this noise-induced delay is not uniform such that some frequency bands show greater shifts than others: low-frequency phase shifts (300–500 Hz) are largest during the response to the consonant-vowel formant transition (/d/), while high-frequency shifts (720–1000 Hz) predominate during the response to the steady-state vowel (/a/). Most importantly, phase shifts occurring in specific frequency bands correlate strongly with shifts in the latencies of the predominant peaks in the auditory brainstem response, while other frequency bands do not correlate with latency shifts. This finding confirms the validity of phase shift detection as an objective measure of timing differences and reveals that this method detects noise-induced shifts in timing that may not be captured by traditional peak latency measurements.

Keywords

auditory; brainstem; phase; timing; noise

*Corresponding author: Nina Kraus, Northwestern University, 2240 Campus Drive, Evanston IL, 60208, USA, nkraus@northwestern.edu, Phone: 847-491-3181, Fax: 847-491-2523.

1. Introduction

The precise representation of the temporal characteristics of sound is crucial for the successful comprehension of speech. In real-world environments, speech is often accompanied by background noise that obscures its rapidly changing acoustic features (e.g., formant transitions). As a result, speech sounds featuring rapid spectrotemporal changes are particularly difficult to perceive when presented in noise (Miller and Nicely 1954; Brandt and Rosen 1980; Nishi et al. 2010). Speech perception may, therefore, depend in part on the extent to which a listener can maintain a robust neural representation of the temporal features of a signal when that signal is embedded in noise or otherwise degraded (Parbery-Clark et al. 2009; Anderson et al. 2010; Bidelman and Krishnan 2010; Hornickel et al. 2011).

The presence of background noise causes delays in auditory brainstem activity that are larger when stimuli are presented at lower signal to noise ratios. These delays occur for simple stimuli such as pure tones and clicks (Ananthanarayan and Durrant 1992; Burkard and Sims 2002) as well as more complex stimuli such as speech syllables (Cunningham et al. 2001; Russo et al. 2004; Parbery-Clark et al. 2009; Anderson and Kraus 2010; Anderson et al. 2010 Song et al. 2010). Background noise also affects cortical processing resulting in delayed cortical responses, including P1, N1, P2, N2, MMN, and P3 (Whiting et al. 1998; Martin et al. 1999; Warrier et al. 2004; Billings et al. 2009; Parbery-Clark et al. 2011). Thus, the presence of background noise systemically affects the auditory system. Latency shifts may reflect a decrease in neural synchronization and/or a decrease in the number of neurons firing, causing responses to be both smaller in amplitude and delayed (Don et al. 1977; Burkard and Sims 2002). As presenting a signal in noise is effectively similar to presenting the same signal at a lower intensity in quiet, these results are in line with previous work showing a relationship between latency and stimulus intensity (Akhoun et al. 2008). These latency shifts are behaviorally relevant, as the extent of noise-induced neural response delays has been linked to performance on tests of speech in noise perception (Parbery-Clark et al. 2009; Anderson et al. (2010)).

Historically, noise-induced shifts in neural timing have been studied via manual picking of peak latencies. This methodology has proven to be a reliable method of documenting the effects of background noise on the timing of the neural response to both simple stimuli such as clicks (Thümmler et al. 1981; Burkard and Hecox 1983; Gott and Hughes 1989; Burkard and Sims 2002) and complex sounds such as synthesized speech syllables (Cunningham et al. 2001; Russo et al. 2004, 2009; Parbery-Clark et al. 2009; Anderson et al. 2010). Picking peaks is, however, time- and resource-intensive, especially when stimuli are complex, resulting in a large number of peaks that need to be selected across multiple conditions. Another potential difficulty in using peak latencies to study noise's effect on the auditory brainstem response (ABR) is that the presence of background noise decreases the amplitude of responses and degrades response morphology (Yamada et al. 1979; Thümmler et al. 1981; Burkard and Hecox 1983; Gott and Hughes 1989; Cunningham et al. 2001; Burkard and Sims 2002; Russo et al. 2004; Parbery-Clark et al. 2009; Song et al. 2010; Li and Jeng 2011). This reduction in response amplitude can bring a response so close to the noise floor that peaks are difficult to identify (Anderson et al. 2010), often necessitating large number of trials or, in some cases, making peak-picking impossible even after thousands of sweeps have been collected. A more automated and objective measure of noise-induced latency shifts would make the use of complex stimuli more practical for both clinicians and scientists.

An objective method of calculating shifts in the timing of neural responses is to measure the frequency-specific phase shift between responses (John and Picton 2000; Skoe et al. 2011).

Latency delays in any periodic signal translate to lags in the phase of the frequencies that make up that signal. While this method (the "cross-phaseogram") has been used to replicate previous results based on manual peak latency measurements (Skoe et al. 2011), a direct comparison between peak latency shifts and phase shifts in different frequency bands has not been performed. In this study, which we consider a natural extension of our lab's previous work, frequency-dependent phase shifts were measured between ABRs to the speech sound /da/ presented in quiet and in a background of six-talker babble. This speech sound consists of a transition period (the consonant /d/), in which there are rapid changes in frequency, and a steady-state period (the vowel /a/), in which the frequency content remains constant. Phase shifts in the response to the transition and steady-state were compared to latency shifts of manually-selected response peaks. The selected peaks were the largest, most predominant peaks in the response, and occurred at a period of roughly ~10 ms.

In addition to delineating the relationship between latency shifts and phase shifts, a further aim of the present study was to better understand how background noise impacts ABRs to speech. The cross-phaseogram can reveal frequency-specific information about noise-induced timing delays, and therefore it can be used to answer questions that cannot be addressed using traditional latency measurements. For example, it remains unknown whether the addition of background noise to an acoustic signal leads to equivalent neural delays across the frequency spectrum, or whether delays are more severe in particular frequency bands. By examining phase shifts, it is possible to investigate the presence of noise-induced timing shifts, their predominance in different frequency bands, and differences between the transition and the steady-state components of the response.

Methods

2.1 Stimuli

The 170-ms speech syllable /da/ was synthesized using a Klatt synthesizer (Klatt 1980). The fundamental frequency (F_0) and fourth through sixth formants of the syllable (F_4–F_6) remained constant throughout the /da/ (170ms) at 100, 3300, 3750, and 4900 Hz, respectively. During the consonant-vowel formant transition period (0–60 ms), F_1 rose from 400 to 720 Hz, F_2 fell from 1700 to 1240 Hz, and F_3 fell from 2850 to 2500 Hz. For the remainder of the sound, the steady-state vowel portion, formant positions remained constant. Background noise consisted of 45 s of multi-talker babble spoken by six different speakers (two males and four females) that looped without pauses throughout the electrophysiological recording. Sentences included in the babble were grammatically correct but semantically anomalous; for further information about the acoustical characteristics of the babble, see Smiljanic and Bradlow (2005).

2.2 Subjects

Forty-eight right-handed young adults, 33 female and 15 male, participated in the study. Participants' ages ranged from 18 to 32 (mean = 22.8 ±3.36). All subjects had hearing thresholds ≤ 20 dB hearing level (HL) from 0.125 to 8 kHz. Subjects reported no history of learning disabilities or neurological deficits, and all subjects had IQ scores in the normal range as measured by the TONI test (Brown et al. 1997). All subjects had normal wave V ABR latencies evoked by a 100 μs click presented at 80 dB sound pressure level (SPL) and presented at a rate of 31.1 Hz.

2.3 Procedure

The speech syllable /da/ was presented binaurally using Stim[2] (Compumedics, Inc.) through insert ear phones (ER-3; Etymotic Research) in alternating polarities at 80 dB SPL with an interstimulus interval of 83 ms. In the noise condition, the speech syllable /da/ and

background noise were presented simultaneously, with the /da/ presented at a +10 dB signal-to-noise ratio.

ABRs to the /da/ sound in the quiet and noise conditions were collected with four Ag-AgCl scalp electrodes using Neuroscan Acquire 4.3 (Compumedics, Inc.) at a 20 kHz analog to digital conversion rate. Brainstem responses were recorded with a vertical montage (active electrode at Cz, linked reference electrodes on both earlobes, ground electrode at Fpz). Electrode impedances were maintained below 5 kOhms. During data collection, to ensure that the subjects remained still but awake, subjects watched a subtitled movie of their choice while seated comfortably in a darkened room. Each electrophysiological test session lasted approximately 50 minutes.

2.4 Data processing

The continuous EEG was filtered offline through a 70–2000 Hz bandpass filter using Neuroscan Edit (Compumedics, Inc.), then epoched from −40 to 210 ms, relative to stimulus onset at 0 ms. Epochs were baseline corrected using the prestimulus portion of the response (−40 to 0 ms). Any sweep exceeding ± 35 μV was excluded due to movement artifact. After artifact rejection, separate 3000-sweep averages to the two stimulus polarities were created and then added offline to minimize the contribution of stimulus artifact and cochlear microphonic to the response (Gorga et al. 1985; Aiken and Picton 2008).

A "cross-phaseogram" was constructed following Skoe et al. (2011), using custom routines coded in MATLAB (The MathWorks Inc.): phase shifts were calculated on 40-ms overlapping windows of the response; the midpoint of the first window started at −20 ms, with each subsequent window shifted by 1 ms, and the final window centered on 160 ms. First, each of these windows was baseline-corrected, then ramped on and off using a Hanning window. Next, the cross-frequency spectrum of each window was calculated and converted to phase angles using the cross-power spectral density function. Jumps between successive blocks of greater than π were corrected to their 2π complement. The resulting cross-phaseogram plot (see example in Fig 2) is a three-dimensional (3D) image, with the degree of shift mapped to different values on the red-green-blue color spectrum. Specifically, for regions colored in red, responses to /da/ in noise were delayed relative to responses to /da/ in quiet; for regions colored in blue, responses to /da/ in noise were earlier than responses to /da/ in quiet. Regions colored in green indicate that there was no effect of noise on the phase of responses.

To calculate peak latency shifts, four peaks in the transition portion of the response (33, 43, and 53 ms) and 10 peaks in the steady-state portion of the response (63, 73, 83, 93, 103, 113, 123, 133, 143, 153, and 163 ms) were identified. (See Fig 1 for an example response waveform and the locations of the analyzed peaks.) The latency of each peak in each subject's response was independently verified by three highly trained peak-pickers. To calculate noise-induced peak latency shifts, the latency of each peak in the response to the quiet condition was subtracted from the latency of the corresponding peak in the response to the noise condition.

2.5 Data analysis

Analyses were conducted separately on the response to the consonant-vowel transition (13–63 ms) and response to the steady-state vowel (63–183 ms), see Fig. 1. First, to determine whether or not there was a noise-induced shift in peak latencies, latency shifts for all transition peaks (33, 43, and 53 ms) were averaged. Next, a t-test was used to determine whether this composite transition latency shift score was different from zero ms, i.e. whether the response in noise was different in latency from the response in quiet. Next, latency shifts

for all steady-state peaks (63, 73, 83, 93, 103, 113, 123, 133, 143, 153, and 163 ms) were averaged, and a t-test determined whether this composite steady-state latency shift score was different from zero. Similarly, to determine whether or not there was a noise-induced shift in phase, phase shifts over the response spectrum (70–1000 Hz) were calculated for the transition and steady-state portions of the response, and a t-test determined whether they differed from zero radians. A phase shift of zero radians indicates no effect of noise on the timing on the brainstem response, a positive phase shift indicates that the response in noise is delayed relative to the response in quiet whereas a negative phase shift indicates that the response in noise is earlier than the response in quiet. Next, phase shifts in four different frequency ranges were analyzed(Fig 2): 70–300 Hz, 300–500 Hz, 500–720 Hz, and 720–1000 Hz. The selection of these frequency ranges was motivated by qualitative examination of the cross-phaseogram (Fig 2B), which suggested differences in noise-induced phase shifts between these frequency ranges. Above 1000 Hz, the brainstem response rapidly falls into the noise floor (Kuwada et al. 1984), and phase calculations on frequencies above this threshold are therefore not meaningful. In both the transition and the steady-state of the response, the phase shifts in these four frequency bands were compared using a one-way ANOVA to determine whether the degree of phase shift was uniform or differed across frequencies.

Correlations between peak latencies and phase shifts were calculated, using MATLAB, over five different frequency ranges: 70–1000 Hz, 70–300 Hz, 300–500 Hz, 500–720 Hz, and 720–1000 Hz. The average of the latency shifts of transition peaks (at 33, 43, and 53 ms) was correlated with phase shifts during the formant transition (13–63 ms), and correspondingly the average of latency shifts of steady-state peaks (at 63, 73, 83, 93, 103, 113, 123, 133, 143, 153, and 163 ms) was correlated with phase shifts during the steady-state vowel (63–183 ms).

3. Results

The effects of background noise on the brainstem response to speech were analyzed by analyzing both shifts in peak latencies and shifts in phase. There was a significant phase shift between the quiet and the noise condition in the 70–1000 Hz frequency range, as shown in Fig 2; the phase of the noise condition lagged compared to the quiet condition. This shift was present in both the transition period (0.54 ± 0.27 radians, $t(47) = 14.1$, $p < 1.0 \times 10^{-15}$) and the steady-state period (0.45 ± 0.30 radians, $t(47) = 10.4$, $p < 1 \times 10^{-13}$). Similarly, the addition of noise led to an increase in peak latencies. This timing shift was present for both the transition peaks (mean latency shift = 0.55 ± 0.34 ms, $t(47) = 11.2$, $p < 1 \times 10^{-14}$) and the steady-state peaks (mean latency shift = 0.18 ± 0.13 ms, $t(47) = 9.29$, $p < 1 \times 10^{-11}$).

Next, we examined whether the phase shifts were uniform across the frequency spectrum or whether they were frequency-dependent. The phase shift in the formant transition was not uniform across frequencies, but differed significantly between the four frequency ranges analyzed ($F(3,188) = 4.92$, $p < 0.01$). The average phase shifts (in radians) for the frequency ranges were 70–300 Hz: 0.49 ± 0.24, 300–500 Hz: 0.66 ± 0.35, 500–720 Hz: 0.37 ± 0.54, and 720–1000 Hz: 0.63 ± 0.49. Similarly, the phase shift in the steady-state period differed significantly between the four frequency ranges ($F(3,188) = 4.25$, $p < 0.01$). Average steady-state phase shifts for the four frequency ranges were 70–300 Hz: 0.35 ± 0.24, 300–500 Hz: 0.37 ± 0.30, 500–720 Hz: 0.41 ± 0.43, and 720–1000 Hz: 0.61 ± 0.59.

Finally, we investigated whether phase shifts and peak latency shifts were correlated, and if so, whether this relationship depended on the frequency range in which phase shifts were analyzed. Peak latency shifts were highly correlated with phase shifts in both the transition

and the steady-state portions of the response (see Table 1 and Fig 3). In the transition, latency shifts correlated with phase shifts between 70 and 300 Hz ($p < 1 \times 10^{-4}$), but were not significantly correlated with phase shifts in any other frequency bands ($p > 0.3$). In the steady-state, however, peak latency shifts were correlated with phase shifts between 70 and 300 Hz ($p < 0.01$), 300 and 500 Hz ($p < 1 \times 10^{-10}$), and 500 and 720 Hz ($p < 1 \times 10^{-6}$), but were not correlated with phase shifts between 720 and 1000 Hz.

4. Discussion

The effects of background noise on auditory brainstem response timing were investigated by analyzing phase shifts and latency shifts between the response to a speech sound presented with and without background noise. The addition of noise resulted in both shifts in response peak timing and shifts in the phase of the response spectrum. Consistent with our predictions, analyzing the phase shift brought about by the addition of noise is a valid and objective way to investigate response timing. Specifically, peak latency shifts were strongly correlated with phase shifts; however, peak latency shifts correlated best with phase shifts occurring in a restricted range of frequencies.

The strong correlations between latency and phase shifts confirm that the cross-phaseogram can be used as an objective measure of noise-induced shifts in timing. This method could, therefore, help both clinicians and researchers rapidly and automatically assess the impact of noise on ABRs without having to manually identify peaks, making the use of complex stimuli more practical. Given that the extent of noise-induced latency shifts in ABRs to speech stimuli correlates with speech-in-noise performance (Parbery-Clark 2009; Anderson et al. 2010; Hornickel et al. 2011), the analysis of noise-induced phase shifts in brainstem responses could automate and render objective this potentially diagnostic information about the source of a patient's difficulties hearing speech in noise.

Our results suggest that phase shifts are also present in frequencies that are not reflected in peak latencies. For example, in both the transition and the steady-state, although the largest quiet-to-noise phase shifts occurred at relatively high frequencies, it was smaller, lower-frequency phase shifts that best correlated with delays in manually-selected peaks. Therefore, by quantifying phase shifts, it is possible to document the effect of noise on the response across a wide range of frequencies, including high-frequency components of the response, which may be reflected in other aspects of the response not associated with the large periodic peaks commonly used for the quantification of the effects of noise on ABR response timing (Parbery-Clark et al. 2009, Anderson et al. 2010).

Phase shifts resulting from the addition of noise were not uniform across frequency nor time, but were larger in certain frequency bands and certain portions of the response. The source of this frequency and time dependence, however, is unclear. One possibility is that there is a greater amount of energy in the background noise at certain frequency bands, because our background noise consisted of multi-talker babble which, unlike white noise, is not uniform across frequencies. This explanation, however, is not consistent with our finding that these bands show larger phase shifts only during certain portions of the response.

That larger mid-frequency (300–500 Hz) phase shifts are only present during the transition is likely a consequence of the rapid spectrotemporal changes occurring within this frequency band during this portion of the stimulus. This result reveals that the spectrotemporal complexity of the information being represented determines the extent to which the response is susceptible to neural desynchronization in noise. This dissociation between the formant transition and the steady-state is consistent with previous research showing effects of

background noise on peak latency only during the transition period of the response, and no such effects during the steady-state (Parbery-Clark et al. 2009; Anderson et al. 2010).

Significant steady-state noise-induced shifts occurred only above 700 Hz. However, these phrase shifts were not present in the first 25 ms of the response. This somewhat paradoxical pattern of results can be explained by the fact that these frequencies are at the edge of the range of frequencies to which the brainstem can phase-lock. As the characteristic frequency of a neuron increases above 500 Hz, the proportion of units that are capable of phase-locking declines (Liu et al. 2006, Akhoun et al. 2008b). This weak phase-locking is even further degraded in the presence of background noise, which decreases the amplitude of components of neural responses (Thümmler et al. 1981; Burkard and Hecox 1983; Gott and Hughes 1989; Whiting et al. 1998; Cunningham et al. 2002; Russo et al. 2004; Warrier et al. 2004; Billings et al. 2009; Song et al. 2010). These frequencies are, therefore, only weakly represented in the ABR and may, consequently, be more susceptible to desynchronization than other frequencies. This susceptibility would explain why phase shifts are largest above 700 Hz during the steady state, but why, then, are high-frequency phase shifts not present in the beginning of the response? In order for phase shifts to be detected in a particular frequency band, there must be a measurable neural response in these frequencies (Skoe et al. 2011). The higher frequencies, however, are very weakly represented during the transition (see Fig 4), which may prevent detection of phase shifts. In summary, for phase shifts to be measurable, noise-induced neural desynchronization must occur, but must not be so severe that phase-locking is eliminated entirely.

Latency shifts correlated with phase shifts in a wider range of frequencies in the steady state, as compared to the transition. This may stem from the fact that higher harmonics are more strongly encoded in the steady-state as indicated in Fig 4. The large periodic peaks occurring every 10 ms are driven by both the 100 Hz component of the response and the higher harmonics, which combine to determine the shape and latency of the peaks. The degree to which timing shifts at a particular frequency affect the latency of a peak may depend on the amplitude of that frequency. This would explain why timing shifts at harmonics above 300 Hz have little effect on peak latencies during the transition, as they are only weakly represented in the response (see Fig 4). Similarly, phase shifts above 720 Hz do not correlate with latency shifts in either the steady-state or the transition, most likely because, as discussed above, those frequencies are not strongly represented throughout the entire response.

In summary, phase shifts provide a surrogate for noise-induced shifts in brainstem response timing. Moreover, phase shifts in certain frequency ranges correlate very highly with latency shifts measured via manual peak-picking methods. However, phase shifts also occur at frequencies that do not correlate with latency shifts. Taken together, these results confirm the cross-phaseogram's usefulness as a measure of noise-induced changes in neural response timing and demonstrate its capacity to reveal the complex, frequency-dependent effects of noise on the timing of the auditory brainstem response. As a result, the cross-phaseogram may be a useful tool for both clinicians and scientists studying the impact of noise on the auditory system. Future work could examine links between phase shifts in different frequency bands and perceptual speech-in-noise ability, to determine the behavioral relevance of frequency-dependent shifts in phase in various clinical and expert populations.

Acknowledgments

References

Aiken S, Picton T. Envelope and spectral frequency-following responses to vowel sounds. Hearing Research. 2008; 245:35–47. [PubMed: 18765275]

Akhoun I, Gallégo S, Moulin A, Ménard M, Veuillet E, Berger-Vachon C, Collet L, Thai-Van H. The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. Clinical Neurophysiology. 2008; 119:922–33. [PubMed: 18291717]

Akhoun I, Moulin A, Jeanvoine A, Menard M, Buret F, Vollaire C, Scorretti R, Veuillet E, Berger-Vachon C, Collet L, Thai-Van H. Speech auditory brainstem response (speech ABR) characteristics depending on recording conditions, and hearing status: an experimental parametric study. Journal of Neuroscience Methods. 2008b; 175:196–205. [PubMed: 18789971]

Ananthanarayan A, Durrant J. The frequency-following response and the onset response: evaluation of frequency specificity using a forward-masking paradigm. Ear and Hearing. 1992; 13:228–232. [PubMed: 1397764]

Anderson S, Kraus N. Objective neural indices of speech-in-noise perception. Trends in Amplification. 2010; 14:73–83. [PubMed: 20724355]

Anderson S, Skoe E, Chandrasekaran B, Kraus N. Neural timing is linked to speech perception in noise. Journal of Neuroscience. 2010; 30:4922–4926. [PubMed: 20371812]

Bidelman G, Krishnan A. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. Brain Research. 2010; 1355:112–125. [PubMed: 20691672]

Billings CJ, Tremblay KL, Stecker GC, Tolin WM. Human evoked cortical activity to signal-to-noise ratio and absolute signal level. Hearing Research. 2009; 254:15–24. [PubMed: 19364526]

Brandt J, Rosen J. Auditory phonemic perception in dyslexia: categorical identification and discrimination of stop consonants. Brain and Language. 1980; 9:324–337. [PubMed: 7363076]

Brown, L.; Sherbenou, RJ.; Johnsen, SK. Test of Nonverbal Intelligence (TONI-3). Austin, TX: Pro-ed, Inc; 1997.

Burkard R, Hecox K. The effect of broadband noise on the human brainstem auditory evoked response. I. Rate and intensity effects. Journal of the Acoustical Society of America. 1983; 74:1204–1213. [PubMed: 6643843]

Burkard R, Sims D. A comparison of the effects of broadband masking noise on the auditory brainstem response in young and older adults. American Journal of Audiology. 2002; 11:13–22. [PubMed: 12227352]

Cunningham J, Nicol TG, Zecker SG, Bradlow AR, Kraus N. Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. Clinical Neurophysiology. 2001; 112:758–767. [PubMed: 11336890]

Don M, Allen A, Starr A. Effect of click rate on the latency of auditory brain stem responses in humans. Annals of Otology. 1977; 86:186–195.

Gorga, M.; Abbas, P.; Worthington, D. Stimulus calibration in ABR measurements. In: Jacobsen, J., editor. The auditory brainstem response. San Diego: College-Hill; 1985. p. 49-62.

Gott P, Hughes E. Effect of noise masking on the brain-stem and middle-latency auditory evoked potentials: central and peripheral components. Electroencephalography and Clinical Neurophysiology. 1989; 74:131–138. [PubMed: 2465888]

Hornickel J, Chandrasekaran B, Zecker S, Kraus N. Auditory brainstem measures predict reading and speech-in-noise perception in school-aged children. Behavioral Brain Research. 2011; 216:597–605.

John MS, Picton TW. Human auditory steady-state responses to amplitude-modulated tones: phase and latency measurements. Hearing Research. 2000; 141:57–79. [PubMed: 10713496]

Klatt D. Software for a cascade/parallel formant synthesizer. Journal of the Acoustical Society of America. 1980; 67:13–33.

Kuwada S, Yin T, Syka J, Buunen T, Wickesberg R. Binaural interaction in low-frequency neurons in inferior colliculus of the cat. IV. Comparison of monaural and binaural response properties. Journal of Neurophysiology. 1984; 51:1306–1325. [PubMed: 6737032]

Li X, Jeng F. Noise tolerance in human frequency-following responses to voice pitch. Journal of the Acoustical Society of America. 2011; 129:EL21–EL26. [PubMed: 21302977]

Liu L, Palmer A, Wallace M. Phase-locked responses to pure tones in the inferior colliculus. Journal of Neurophysiology. 2006; 95:1926–1935. [PubMed: 16339005]

Martin B, Kurtzberg D, Stapells D. The effects of decreased audibility produced by high-pass noise masking on N1 and the mismatch negativity to speech sounds /ba/ and /da/. Journal of Speech and Language Hearing Research. 1999; 42:271–286.

Miller G, Nicely P. An analysis of perceptual confusions among some english consonants. Journal of the Acoustical Society of America. 1954; 27:338–352.

Nishi K, Lewis DE, Hoover BM, Choi S, Stelmachowicz PG. Children's recognition of American English consonants in noise. The Journal of the Acoustical Society of America. 2010; 127:3177–3188. [PubMed: 21117766]

Parbery-Clark A, Skoe E, Kraus N. Musical experience limits the degradative effects of background noise on the neural processing of sound. Journal of Neuroscience. 2009; 29:14100–14107. [PubMed: 19906958]

Parbery-Clark A, Marmel F, Bair J, Kraus N. What subcortical-cortical relationships tell us about processing speech in noise. European Journal of Neuroscience. 2011; 33:549–557. [PubMed: 21255123]

Russo N, Nicol T, Musacchia G, Kraus N. Brainstem responses to speech syllables. Clinical Neurophysiology. 2004; 115:2021–2030. [PubMed: 15294204]

Russo N, Nicol T, Trommer B, Zecker S, Kraus N. Brainstem transcription of speech is disrupted in children with autism spectrum disorders. Developmental Science. 2009; 12:557–567. [PubMed: 19635083]

Skoe E, Nicol T, Kraus N. Cross-phaseogram: Objective neural index of speech sound differentiation. Journal of Neuroscience Methods. 2011; 196:308–317. [PubMed: 21277896]

Smiljanic R, Bradlow A. Production and perception of clear speech in Croatian and English. Journal of the Acoustical Society of America. 2005; 118:1677–1688. [PubMed: 16240826]

Song J, Skoe E, Banai K, Kraus N. Perception of speech in noise: neural correlates. Journal of Cognitive Neuroscience. 2010 e-publication ahead of print. 10.1162/jocn.2010.21556

Thümmler I, Tietze G, Matkei P. Brain-stem responses when masking with wide-band and high-pass filtered noise. Scandiavian Audiology. 1981; 10:255–259.

Warrier CM, Johnson KL, Hayes EA, Nicol T, Kraus N. Learning impaired children exhibit timing deficits and training-related improvements in auditory cortical responses to speech in noise. Experimental Brain Research. 2004; 157:431–441.

Whiting KA, Martin BA, Stapells DR. The effects of broadband noise masking on cortical event-related potentials to speech sounds /ba/ and /da. Ear and Hearing. 1998; 19:218–231. [PubMed: 9657596]

Yamada O, Kodera K, Hink R, Suzuki J. Cochlear distribution of frequency-following response initiation. A high-pass masking noise study. Audiology. 1979; 18:381–387. [PubMed: 496720]

- We examined noise-induced delays in the auditory brainstem response.

- We measured these delays using peak latency picking and detection of phase shifts.

- Phase shifts in certain frequency bands correlated very highly with latency shifts.

- Thus, phase shift detection is a viable way to objectively measure noise-induced timing shifts.

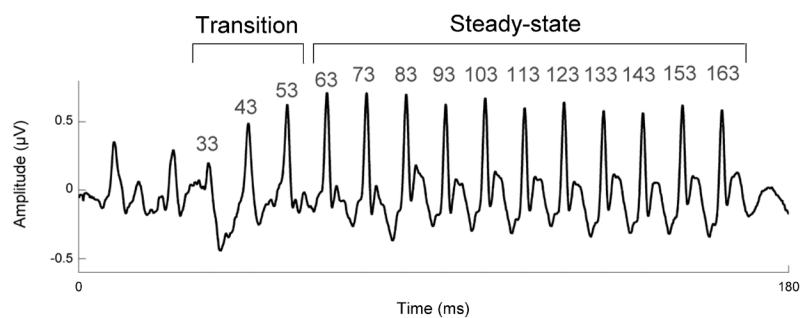- We also find phase shifts at other frequencies that do not correlate with peak latency shifts.

Fig. 1.
Group average brainstem response to the speech sound /da/ in quiet. Neural response peaks corresponding to the transition occur at 33, 43, and 53 ms. Neural response peaks corresponding to the steady-state occur every 10 ms from 63–163ms.
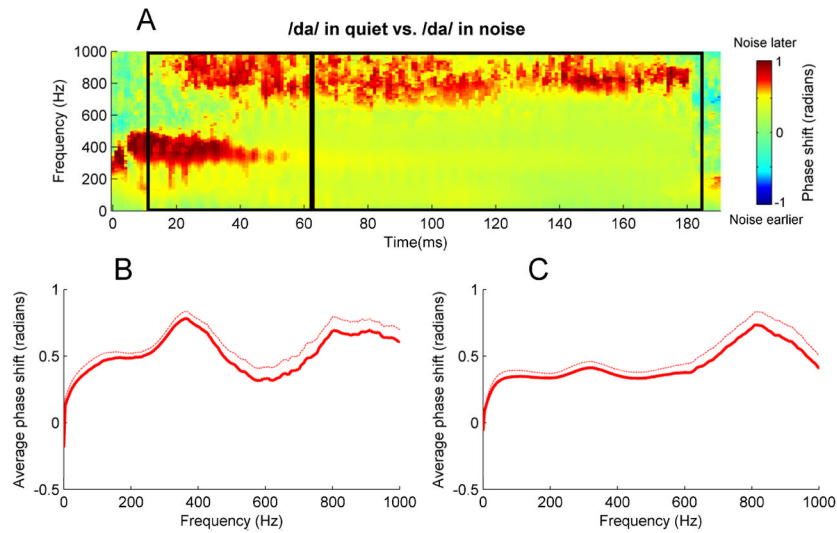
Fig. 2.
(A) Cross-phaseogram displaying noise-induced phase shifts. Formant transition and steady-state regions are demarcated by black boxes. Regions in red indicate times and frequencies for which the response was later in phase when the stimulus was presented in noise. Regions in blue correspond to times and frequencies for which the response was later in phase when the stimulus was presented in quiet. (B) Average neural phase shifts for during the transition (13–63 ms). Thick and thin lines correspond to means and standard errors, respectively. (C) Average phase shifts for the steady-state portion of the response (63–180 ms).
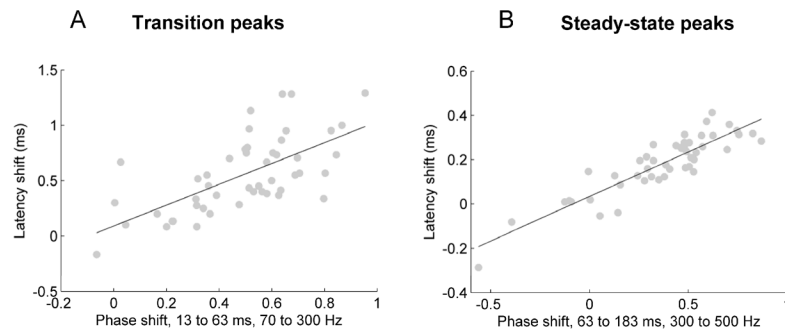
Fig. 3.
(A) Scatterplot displaying the relationship between the average latency shift of transition peaks at 33, 43, and 53 ms and the phase shift between 13 and 63 ms and between 70 and 300 Hz ($p < 1 \times 10^{-4}$). (B) Scatterplot displaying the relationship between the average latency shift of steady-state peaks 63, 73, 83, 93, 103, 113, 123, 133, 143, 153, and 163 ms and the phase shift between 63 and 183 ms and between 300 and 500 Hz ($p < 1 \times 10^{-15}$).
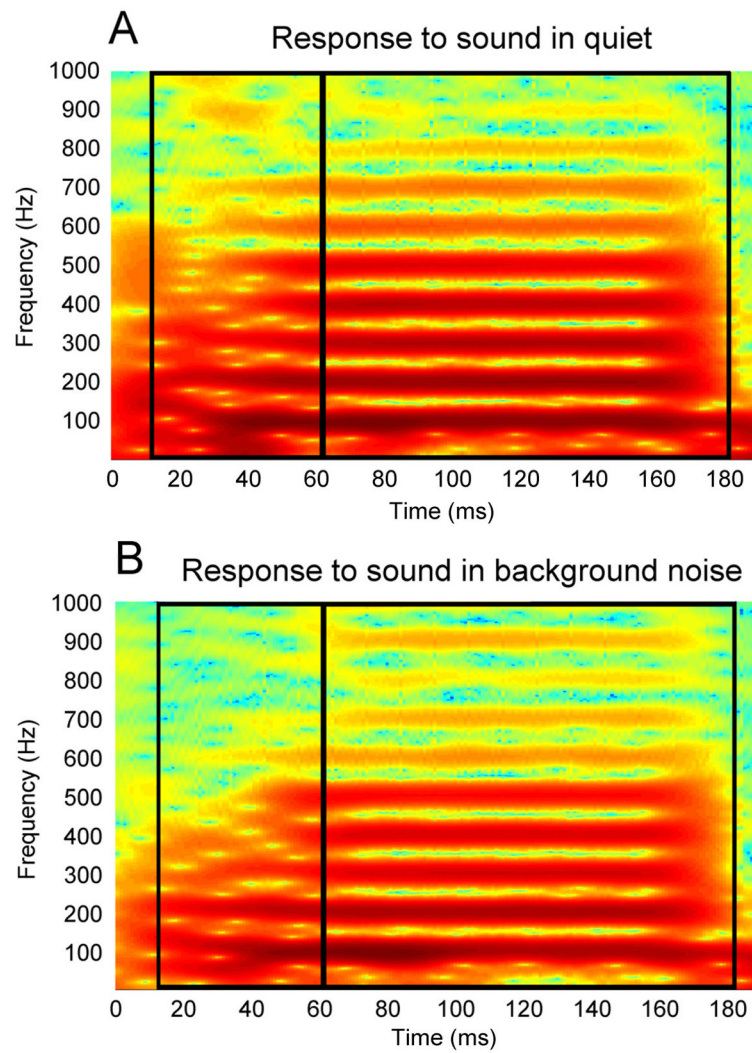
Fig. 4.
(A) Spectrogram of the average of the responses to the stimulus presented in quiet. (B) Spectrogram of the average of the responses to the stimulus presented in noise. Formant transition and steady-state regions are demarcated by black boxes.

Table 1

Correlations (r-values and p-valies) between phase shifts and peak latency shifts. Bolded correlations are significant (p < 0.05).

	Phase shift, 70–1000 Hz		Phase shift, 70–300 Hz		Phase shift, 300–500 Hz		Phase shift, 500–720 Hz		Phase shift, 720–1000 Hz	
	r	**p**	**r**	**p**	**r**	**P**	**r**	**p**	**r**	**p**
Transition latency shift	0.13	0.37	**0.65**	$\mathbf{4.4 \times 10^{-5}}$	−0.07	0.64	−0.08	0.58	0.08	0.57
Steady-state latency shift	**0.66**	$\mathbf{3.4 \times 10^{-5}}$	**0.46**	$\mathbf{1.2 \times 10^{-3}}$	**0.89**	$\mathbf{<1 \times 10^{-10}}$	**0.72**	$\mathbf{6.0 \times 10^{-7}}$	0.22	0.12