

Context-Dependent Encoding in the Human Auditory Brainstem Relates to Hearing Speech in Noise: Implications for Developmental Dyslexia

Bharath Chandrasekaran,^{1,2,3,7} Jane Hornickel,^{1,2,7} Erika Skoe,^{1,2} Trent Nicol,^{1,2} and Nina Kraus^{1,2,4,5,6,*}

¹Auditory Neuroscience Laboratory

²Roxelyn and Richard Pepper Department of Communication Sciences

³Communication Neural Systems Research Group

⁴Institute for Neuroscience

⁵Neurobiology and Physiology

⁶Otolaryngology

Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA

⁷These authors contributed equally to this work

*Correspondence: nkraus@northwestern.edu

DOI 10.1016/j.neuron.2009.10.006

SUMMARY

We examined context-dependent encoding of speech in children with and without developmental dyslexia by measuring auditory brainstem responses to a speech syllable presented in a repetitive or variable context. Typically developing children showed enhanced brainstem representation of features related to voice pitch in the repetitive context, relative to the variable context. In contrast, children with developmental dyslexia exhibited impairment in their ability to modify representation in predictable contexts. From a functional perspective, we found that the extent of context-dependent encoding in the auditory brainstem correlated positively with behavioral indices of speech perception in noise. The ability to sharpen representation of repeating elements is crucial to speech perception in noise, since it allows superior “tagging” of voice pitch, an important cue for segregating sound streams in background noise. The disruption of this mechanism contributes to a critical deficit in noise-exclusion, a hallmark symptom in developmental dyslexia.

INTRODUCTION

Verbal communication often occurs in noisy backgrounds. Imagine a conversation with a friend in a noisy restaurant. To effectively converse with your friend you need to extract the information that he/she conveys from the irrelevant background noise. This task is particularly challenging because the competing noise (other talkers) has acoustic properties that overlap with the target signal (your friend’s voice). Yet, for the most part, communication is unimpeded even under such challenging conditions. This remarkable feat relies on a highly adaptive auditory system that continually modulates its activity based on

contextual demands. Successful completion of this complex task, extracting the speech signal, takes advantage of a predictable, repeating element (the pitch of your friend’s voice) amid the random, fluctuating background of many voices. The ability to “tag” the predictable elements in the environment (e.g., voice pitch) provides significant benefits to perception under adverse listening conditions (Bregman, 1994; Brox and Nooteboom, 1982; Sayles and Winter, 2008). How the nervous system functionally adapts and fine-tunes the representation of predictable auditory elements in the environment is currently unknown.

Understanding the relationship between the adaptive auditory system and perception of speech in noise is clinically relevant because recent studies have demonstrated that children with developmental dyslexia are particularly vulnerable to the deleterious effects of background noise (Sperling et al., 2005, 2006; Ziegler et al., 2009). Developmental dyslexia is a neurological disorder affecting reading and spelling skills in approximately 5%–10% of school-aged children (Demonet et al., 2004). A “core deficit” identified in these children is noise-exclusion, i.e., an inability to exclude noise from ongoing perceptual dynamics (Ahissar, 2007; Ahissar et al., 2006; Ramus and Szenkovits, 2008; Sperling et al., 2005; Ziegler et al., 2009). Behavioral studies have posited that noise-exclusion issues may be attributed to a neural impairment in extracting regularities (e.g., extracting a speaker’s voice in the midst of background noise) from the incoming sensory stream (Ahissar et al., 2006). Although the neural bases of such context-dependent encoding are unknown, it has been argued that lower perceptual structures play an important role in automatically fine-tuning responses to repeating elements in the incoming sensory stream (Ahissar, 2007).

Recent studies in animal models have argued that lower perceptual structures (i.e., auditory brainstem) are crucial for processing auditory signals in noisy environments (Luo et al., 2008). Auditory processing in lower perceptual structures involves an interplay between sensory and cognitive systems mediated by feedforward and feedback pathways (Tzounopoulos and Kraus, 2009). The massive efferent connections from the cortex to subcortical structures form the basis for such

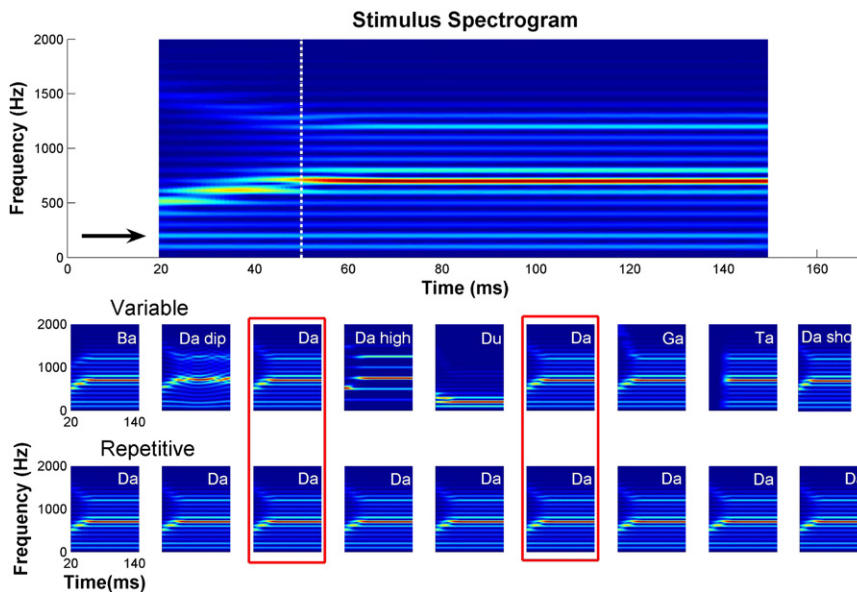


Figure 1. Stimulus Characteristics and Experimental Design

(Top) The spectrogram of the stimulus /da/. The boundary of the consonant-vowel formant transition and the steady-state vowel portion of the syllable is marked by a dashed white line. The spectrogram was generated via frequency analyses over 40 ms bins starting at time 0, and the midpoint of each bin is plotted. The stimulus /da/ is presented in variable (middle) and repetitive (bottom) contexts. As seen in the spectrograms, the stimuli in the variable context differed from /da/ in a number of spectral and temporal features. Responses to /da/ are event-matched between the two conditions to control for presentation order. The second harmonic of the stimulus, a dominant spectral element in the results, is marked with an arrow.

feedback-related top-down control (Winer, 2005). Although the functional role of these efferent connections is currently unknown, a recent study has hypothesized that corticofugal feedback may provide significant benefits in noisy environments by selectively amplifying relevant information in the signal, and inhibiting irrelevant information at the earliest stages of auditory processing (Luo et al., 2008).

In humans, the neural transcription of complex auditory stimuli such as speech can be measured noninvasively from lower levels of the central nervous system such as the auditory brainstem (Johnson et al., 2008; Hornickel et al., 2009; Tzounopoulos and Kraus, 2009). The auditory brainstem response faithfully preserves the complex harmonic characteristics of speech (Kraus and Nicol, 2005). Most speech sounds have a complex harmonic structure that relates to the source (vocal fold vibration) and filter (vocal tract characteristics). For example, in producing a vowel, a speaker causes his/her vocal folds (the source) to vibrate. This causes a glottal pulse, a periodic buzz-like sound made up of a fundamental frequency (F0) and integer multiples of that fundamental frequency (harmonics). The glottal pulses are shaped by the vocal tract (e.g., the tongue position in the oral cavity), and depending on the vowel, certain harmonics are boosted, resulting in a signature spectrum. These boosted harmonics are referred to as “formants” (e.g., F1, F2, F3, etc.). The fundamental frequency (F0) and the lower-numbered harmonics strongly relate to voice pitch (e.g., is the speaker male or female?), while the formant structure relates to speech identification (e.g., is the vowel /i/ or /a/?). Neural representation of both components (voice pitch, formant structure) is necessary for speech-in-noise perception. The voice pitch allows tagging of the speaker in noise; the formant structure is needed to discern the content of speech.

The frequency following response (FFR), a component of the auditory brainstem response, reflects neural phase-locking to F0 and its harmonics (Chandrasekaran and Kraus, 2009). The FFR closely mimics the incoming signal; when the FFR waveform,

recorded in response to words, is played back, subjects can identify the words with greater-than-chance accuracy (Galbraith et al., 1995). Recent studies have demonstrated that the FFR can serve as an index of long-term and training-related plasticity. Native speakers of a tone language, in which changes to voice pitch alone can change word meaning, represent voice pitch more robustly than nonnative speakers (Krishnan and Gandour, 2009; Krishnan et al., 2005). Similarly, musicians, who have long-term experience with musical pitch, show superior representation of voice pitch at the level of the brainstem, suggesting that plasticity is not specific to the domain of expertise (Musacchia et al., 2007; Wong et al., 2007). Typically, studies examining neural plasticity at the level of the brainstem have used two groups, a proficient group (e.g., musicians) and a control group (e.g., Krishnan et al., 2005, 2009; Musacchia et al., 2007; Wong et al., 2007). The general conclusions from these studies have been that processing in the human auditory brainstem is dynamic in nature. While it is generally agreed that the auditory brainstem is sensitive to auditory experience, the exact mechanism by which auditory experience modulates activity is as yet undetermined. An important issue is the extent to which plasticity is operational online (i.e., shows sensitivity to ongoing contextual demands) or reflects long-term structural and functional reorganization. Do we continuously fine-tune or shape our representation with repetition? Or does plasticity reflect a longer timescale that requires an extensive local reorganization of circuitry to better encode biologically relevant sounds?

In the current study we examine whether auditory brainstem responses can indeed be modulated online by context. To test this proposal we created a brainstem recording procedure that averages across responses to repetitive auditory stimulation. In Experiment 1 we examine context-dependent encoding by comparing auditory brainstem responses to a synthesized speech syllable, /da/, elicited in two different contexts: a predictable context versus a highly variable context. By matching trials between the two contexts (see Figure 1), we are able to examine

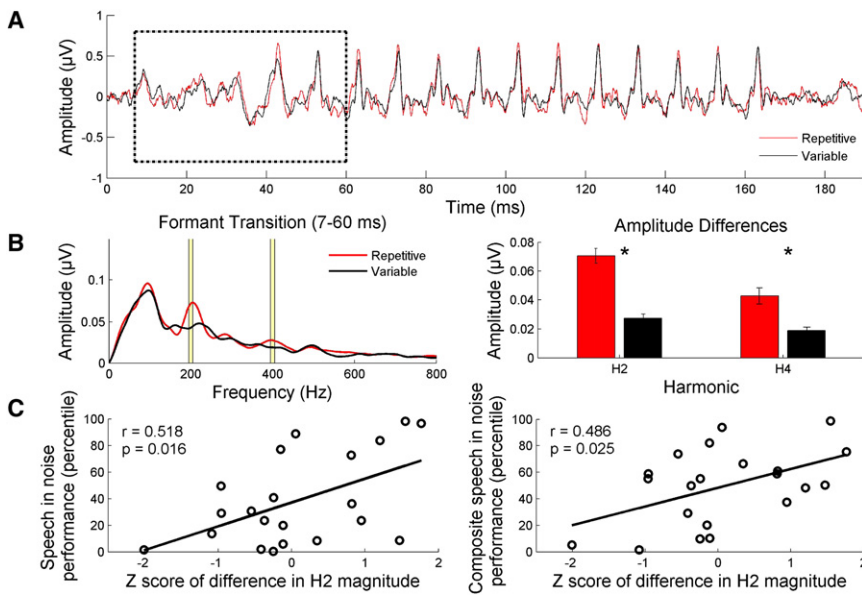


Figure 2. Experiment 1: Human Auditory Brainstem Responses Are Sensitive to Stimulus Context

(A) The grand averages of the time-amplitude responses in the repetitive (red) and variable (black) conditions are overlaid. The black box demarcates formant transition period (7–60 ms). Context did not affect measures of peak latency or response amplitude. (B) Grand-average spectra for the repetitive (red) and variable (black) conditions show enhanced encoding of the second (H2) and the fourth (H4) harmonics in the repetitive condition (left). Mean spectral amplitudes of the second (H2) and fourth (H4) harmonics are shown in the repetitive (red) and variable (black) conditions (right). Error bars represent 1 SEM. (C) The normalized difference in H2 magnitude between the two conditions (repetitive minus variable) is related to speech-in-noise perception measures (HINT-RIGHT, left; HINT-COMPOSITE, right).

differences in brainstem responses to the same stimulus under two different contexts, without a presentation order confound. Next, we examine whether the ability to fine-tune or sharpen brainstem responses to speech features online is functionally related to speech-in-noise perception in children. In Experiment 2, we examine context-dependent brainstem modulation in children with developmental dyslexia, a clinical group that is shown to have global deficits in repetition-induced sensory fine-tuning (Ahissar, 2007) as well as noise-exclusion (Sperling et al., 2005).

RESULTS

Context-Dependent Modulation of Brainstem Responses to Speech (Experiment 1)

In Experiment 1, we examined context-dependent brainstem encoding of speech in 21 typically developing children. Context-dependent effects were observed in the 7–60 ms time range of the response, which encompasses the response to the sound onset and the consonant-vowel formant transition period (Figure 2A), but not in the 60–180 ms time range that encompasses the response to the steady-state vowel. Spectral amplitudes of the lower harmonics (H2, H4), which lead to the perception of pitch, were enhanced in the repetitive context relative to the variable context (Figure 2B). No significant context effects were found for any of the latency (Figure S1 and Table S4 available online) or amplitude (Figure 2A) measures, suggesting that stimulus context does not modulate these measures. Within the spectral domain, multivariate repeated-measures ANOVAs (RmANOVAs) conducted on the average response magnitudes of the F0 and subsequent five harmonics yielded significant differences between the repetitive context and variable context conditions for the second (H2) and fourth (H4) harmonics during the formant transition region only (7–60 ms). Relative to the variable context, H2 and H4 amplitudes were significantly larger in the repetitive context [$F(1,18) = 13.952$, $p = 0.002$; $F(1,18) = 4.758$, $p = 0.043$, respectively]. Figure 2B

shows the grand averaged response spectrum for the 7–60 ms range and these differences are highlighted as bar charts. There was no significant effect of context for F0 or any harmonic amplitude for the steady-state vowel portion (60–180 ms), indicating that context-dependent effects only occur in response to the complex, time-varying portion of the stimulus, which is crucial for distinguishing speech sounds. Additionally, over the frequency region of interest included in the first formant range (400–720 Hz), the repetitive context elicited stronger spectral representations between 530–590 Hz than the variable context in the 7–60 ms time region [$t(20) = 4.217$, $p < 0.001$; see Figure S3], but not in the 60–180 ms time region [$t(20) = 0.428$, $p = 0.673$; see Figure S3].

Correlations between Speech-in-Noise Perception and Context Effects on Neural Encoding

To investigate the relationship between the extent of context-dependent brainstem encoding and behavioral indices of speech-in-noise perception, a series of Pearson's correlations were calculated. We evaluated the degree of brainstem dynamicity by computing the difference in spectral amplitudes (H2, H4) between the two conditions (repetitive context minus variable context) for each participant. These values were then converted to z-scores with larger positive values indicating enhanced encoding in the repetitive context relative to the variable context. The z-scored H2 and H4 spectral amplitude differences were correlated with behavioral performance in Hearing in Noise Test (HINT), a standardized test of speech perception in noise (see Figure 2C) administered to the children.

H2 difference scores correlated positively with HINT-RIGHT (noise source located to the right of the listener) percentile score ($r = 0.518$, $p = 0.016$) and HINT-COMPOSITE (composite of three noise conditions) ($r = 0.486$, $p = 0.025$), and to a lesser degree with HINT-FRONT (noise source located in front of the listener) ($r = 0.407$, $p = .067$), but not with HINT-LEFT ($r = 0.066$, $p = 0.777$). No significant correlations were found between H4

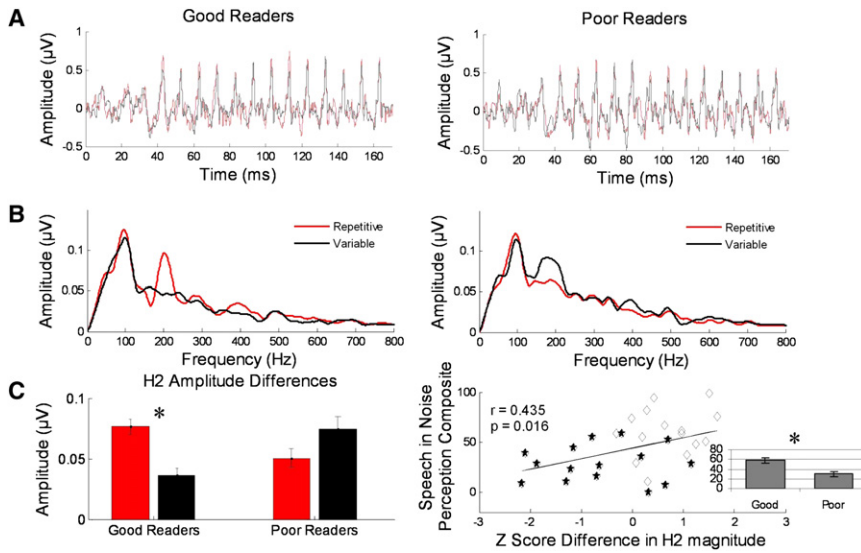


Figure 3. Experiment 2: Context Effects Are Seen for Good Readers, but Not for Poor Readers

(A) The grand averages of the responses in the repetitive (red) and variable (black) conditions are overlaid for the good (left) and poor (right) readers. (B) Grand average spectra over the formant transition period for the good (left) and poor (right) readers show enhanced harmonic encoding in the good readers in the repetitive (red) condition and enhanced encoding of harmonics in the poor readers in the variable (black) condition. (C) (Left) Bar plots of H2 amplitude support the response spectra, with greater H2 amplitude in good (left) as compared with poor (right) in the repetitive (red) condition, with the opposite effect in the variable (black) condition. (C) (Right) The normalized difference in H2 magnitude between the two conditions is again related to speech-in-noise perception measures (HINT-COMPOSITE) across the whole group. Good readers are plotted as open diamonds, and poor readers, as filled stars. Overall, poor readers show inferior speech-in-noise perception relative to good readers [$t(28) = -4.00$, $p < 0.001$; see inset]. Error bars represent standard error.

difference and speech-in-noise perception (see Supplemental Results section).

Context-Dependent Encoding in Children with Dyslexia (Experiment 2)

In Experiment 2, we examined context-dependent brainstem encoding of speech in children with good and poor reading skills ($n = 15$, both groups). Children with poor reading skills differed in the extent and nature of context-dependent spectral encoding within the 7–60 ms time period corresponding to the stimulus formant transition, but not during the 60–180 ms time period corresponding to the steady-state vowel. Multivariate RMANOVAs revealed significant interactions between context (repetitive, variable) and group (good readers, poor readers) for H2 amplitude [$F(1,28) = 17.099$, $p < 0.001$], H4 amplitude [$F(1,28) = 11.649$, $p = 0.002$] (Figure 3), and the F1 range [$F(1,28) = 6.827$,

$p = 0.014$; see Figure S4] in the formant transition region only (7–60 ms). Consistent with Experiment 1, post hoc paired t tests showed larger H2 and H4 amplitudes in the repetitive context than the variable context for good readers [$t(14) = 5.156$, $p < 0.001$; $t(14) = 2.805$, $p < 0.05$, respectively, Figures 3B and 4]. Also, in the F1 range, good readers showed larger spectral amplitude in the repetitive context relative to the variable context [$t(14) = 3.749$, $p = 0.002$, Figure S2]. In contrast, poor readers showed no significant differences between the two conditions (Figures 3B and 4), although a trend for H2 and H4 amplitudes to be greater for the variable context relative to the repetitive context was present [H2: $t(14) = -1.773$, $p = 0.098$; H4: $t(14) = -2.095$, $p = 0.055$]. Additionally, for poor readers, no significant differences were observed between the two contexts in the spectral amplitude within the F1 range (Figure S4). Additional post hoc independent t tests revealed larger H2 spectral

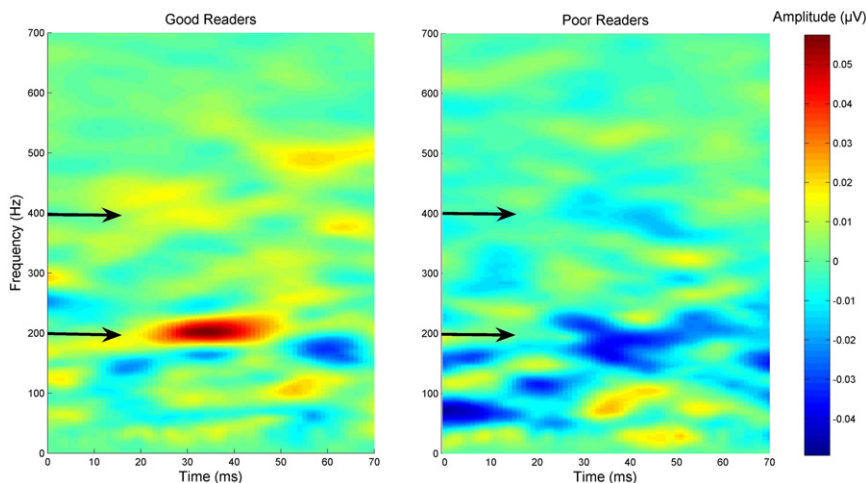


Figure 4. Experiment 2: Good Readers Show Superior Encoding in the Repetitive Context While the Opposite Is Seen in Poor Readers

Good readers (left) have greater H2 and H4 amplitudes (200 and 400 Hz, respectively) in the repetitive condition than they do in the variable condition (signified by warm colors), while the poor readers (right) show greater H2 and H4 amplitudes in the variable condition (signified by cool colors). Figures were created by first generating response spectrograms for both conditions, and then subtracting spectral amplitudes elicited in the variable context condition from those elicited in the repetitive condition for each group.

amplitude for good readers relative to poor readers in the repetitive context [$t(28) = -2.643$, $p = 0.013$]. In contrast, in the variable context, poor readers showed greater H2 spectral amplitude relative to good readers [$t(28) = 3.116$, $p = 0.004$]. Consistent with Experiment 1, there were no main effects of context, main effects of group, or interactions between group and context within the steady-state vowel portion (60–180 ms).

Correlations between Speech-in-Noise Perception and Context Effects in Good and Poor Readers

Similar to Experiment 1, which was restricted to normal readers, the difference in H2 encoding between the two contexts in Experiment 2 was also correlated with speech-in-noise perception in this broader group containing both good and poor readers. H2 difference was correlated with HINT-RIGHT ($r = 0.349$, $p = 0.058$, Figure 3D), HINT-COMPOSITE ($r = 0.436$, $p = 0.016$, Figure 3C), and HINT-FRONT ($r = 0.365$, $p = 0.048$), but not HINT-LEFT ($r = 0.082$, $p = 0.666$). These effects are maintained when controlling for verbal IQ ($r = 0.320$, $p = 0.091$; $r = 0.344$, $p = 0.067$; $r = 0.419$, $p = 0.024$, respectively). In addition, H2 difference scores were also significantly correlated with performance on a number of behavioral indices of reading ability (Table S2), although these effects were not maintained when verbal IQ was partialled out.

DISCUSSION

Our electrophysiological results provide the first evidence that the human auditory brainstem is sensitive to ongoing stimulus context. Stimulus repetition induces online plasticity that leads to an automatic sharpening of brainstem representation of speech cues related to voice pitch. This repetition-induced neural fine-tuning is strongly associated with perception of speech in noise, suggesting that this type of plasticity is indeed functional. The ability to modulate or sharpen the neural representation of voice pitch is crucial to speech perception in noise. This is because voice pitch is a critical cue in speaker identification and allows enhanced tagging of the speaker's voice, an important mechanism for segregating sound streams in background noise (Bregman, 1994; Brokx and Nootboom, 1982; Sayles and Winter, 2008). In a second experiment comparing good and poor readers, we find that brainstem encoding among poor readers is impaired and does not adapt as well to the repeating elements of the auditory signal. Poor readers also show a deficit in perceiving speech in noise, confirming previous studies that report noise-exclusion dysfunction in other sensory domains (Ziegler et al., 2009; Sperling et al., 2005, 2006). We elaborate on each of these findings separately in the following sections.

Human Auditory Brainstem Is Sensitive to Stimulus Context

In Experiment 1 we examined the effect of stimulus context on the auditory brainstem response to speech in typically developing children. Our data demonstrate that in the predictable context (relative to the variable context), representation of harmonic stimulus features that contribute to encoding voice pitch was enhanced within the time-varying period of the

response (7–60 ms), a period corresponding to the transition from the consonant to the vowel. This suggests that the human auditory brainstem is indeed modulated by short-term stimulus history.

How do these findings relate to current knowledge about the functioning of the human auditory brainstem? Studies have demonstrated experience-dependent modulation of the encoding of voice pitch (Krishnan et al., 2005; Musacchia et al., 2007; Wong et al., 2007; Song et al., 2008). Long-term experience with a tone language can improve the representation of native pitch contours (Krishnan and Gandour, 2009). While these studies have demonstrated that the auditory brainstem encoding is dynamic in nature, and reflects long-term auditory experience, the neurobiological mechanism that contributes to this plasticity has remained elusive. Two hypotheses on the nature of experience-dependent brainstem plasticity are being debated (Krishnan and Gandour, 2009). The corticofugal model states that top-down feedback via the corticofugal efferent network modifies brainstem function (Suga, 2008; Suga et al., 2002; Zhang et al., 1997). The local reorganization model posits that brainstem function is modulated over a longer timescale, i.e., the brainstem is reorganized to promote the encoding of frequently encountered sounds (Krishnan and Gandour, 2009; Krishnan et al., 2009). Both models require top-down modulation and are not mutually exclusive. The corticofugal model predicts moment-to-moment changes in brain function as a result of top-down feedback. The local reorganization model predicts top-down modulation of brainstem circuitry during learning, after which top-down feedback is no longer required. Thus, both models predict plasticity in relevant feature representation, but the timescales are vastly different. In the current experiment, the stimulus (/da) is exactly the same in both variable and repetitive conditions. Yet, online context determines the robustness of brainstem representation. These results can be explained within the framework of a corticofugal model of plasticity that argues that neural representation is continuously shaped online. In animal models, cortical neurons have been shown to rapidly adapt to improve signal quality in challenging environments (Atiani et al., 2009; Elhilali et al., 2009). The auditory cortex is also capable of improving signal quality by modulating response properties of brainstem neurons via the corticofugal pathways (Gao and Suga, 1998, 2000; Suga, et al., 2000; Suga and Ma, 2003; Yan and Suga, 1999; Zhang and Suga, 1997). Corticofugal modulation sharpens representation at the auditory brainstem by enhancing the response properties of physiologically matched subcortical neurons, while subduing the activity of unmatched subcortical neurons (Luo et al. 2008). This constant, automatic, top-down search to increase the signal-to-noise ratio (SNR) has been argued to provide significant benefits under adverse signal-to-noise conditions (Nahum et al., 2008).

Speech-in-Noise Perception Is Associated with Context-Dependent Brainstem Plasticity

The ability to tag the repeating elements in the auditory environment is important in determining success at accurately perceiving speech in noise (Ahissar et al., 2006). Here we show that repetition induces improved neural representation of cues that are relevant for perceiving voice pitch, an important cue

for segregating sound sources in noisy environments. Importantly, repetition-induced plasticity in representation of voice pitch was strongly associated with behavioral performance on speech-in-noise tests. This result suggests that the ability to fine-tune brainstem encoding of repeating elements in the auditory environment is important for speech-in-noise perception. Hypothesizing about the role of the corticofugal network in real-world situations, a recent animal study suggested that top-down selective processing is beneficial for perception in noisy environments (Luo et al., 2008). In the context of the current study, corticofugal modulation likely improves signal quality at the auditory periphery by selectively amplifying relevant features of the signal (e.g., voice pitch) based on top-down feedback.

Context-Dependent Brainstem Encoding of Speech Features Is Disrupted in Poor Readers

In Experiment 2, we examined the hypothesis that children with developmental dyslexia show a disruption in context-dependent brainstem encoding of speech features that may contribute to their generally reported noise-exclusion deficit. We found differences between children with good and poor reading skills in their brainstem representation of cues related to voice pitch and formant structure of the stop consonant /da/. Only good readers showed context-dependent brainstem encoding of speech features (i.e., their representation in the repetitive condition is enhanced compared to the variable condition). No significant effects of context were elicited from poor readers. This result demonstrates a deficit in fine-tuning auditory representation with experience in poor readers. This provides support for the anchor-deficit hypothesis (Ahissar, 2007; Ahissar et al., 2006), which posits that children with developmental dyslexia, unlike typically developing children, do not reap benefits from stimulus repetition. This suggests that their encoding deficits are not just related to the acoustics of the stimulus, but are also context dependent. Indeed, it has been argued that a general impairment in the ability to use top-down predictive cues to shape early sensory processing can explain noise-exclusion deficits experienced by dyslexics (Ramus, 2001; Ramus et al., 2003; Ramus and Szenkovits, 2008). Consistent with this hypothesis, our results demonstrate a speech-in-noise perception deficit in poor readers that is associated with the inability to modulate encoding of voice pitch based on context. Poorer sensory representation of regularities in the auditory environment may impair the ability to use voice pitch as a tag, thereby causing a deficit in noise-exclusion.

Previous studies in children with developmental language disorders have demonstrated that these children have particular difficulty processing stop consonants (Elliott et al., 1989; Tallal, 1975). It has been hypothesized that this difficulty may be due to a global deficit in encoding fast temporal events (Tallal, 1980). In the current study, repetition-induced enhancement in the representation of harmonic structure for good readers was restricted to the fast changing, time-varying formant transition portion of the signal. We found no context-dependent effects in the response corresponding to the vowel. These data are consistent with previous studies that report that the greatest neuroplasticity in brainstem responses occurs to the most

acoustically complex aspects of the stimuli (Krishnan et al., 2009; Song et al., 2008; Strait et al., 2009; Wong et al., 2007). Importantly, our data suggest that an auditory encoding deficit in dyslexia is not entirely related to stimulus parameters per se. Instead, we argue that auditory encoding deficits are context-dependent. In predictable contexts, children with dyslexia, relative to good readers, show an impairment in the ability to continuously fine-tune sensory representation. In contrast, no such deficit was found in the variable context, a context in which presentation is random. These data are thus consistent with a recent proposal that children with dyslexia are unable to benefit from prior exposure to auditory stimuli (Ahissar, 2007; Ahissar et al., 2006).

Impaired Context-Dependent Brainstem Encoding: Clinical Implications

Our discovery that children with dyslexia show deficits in context-dependent brainstem encoding of speech features is consistent with the proposal that a cogent explanation for the broad sensory deficit in dyslexia is a failure of top-down expectancy-based processes that enhance lower-level processing (Ramus et al., 2003). These top-down processes are particularly important for noise-exclusion (enhancing relevant aspects of the signal, while excluding irrelevant details; Luo et al., 2008). In typically developing children, we argue, repetitive auditory presentation induces expectancy-based enhancement of relevant features in the signal (e.g., voice pitch) via the corticofugal network. In contrast, poor readers appear to be unable to modulate their current lower-level representation as a result of top-down, expectancy-based fine-tuning. Interestingly, in the current study, dyslexic children showed enhanced brainstem representation of lower harmonics in the variable condition compared to good readers. The functional basis of enhanced spectral representation in a highly unpredictable auditory environment is unclear. Since ongoing representations are not influenced by prior experience, dyslexic children may be able to represent their sensory environment in a broader and arguably more creative manner (Everatt et al., 1999). However, stronger representation in a highly variable listening environment may also come at the cost of the ability to exclude irrelevant details (e.g., noise) from ongoing perceptual dynamics. We do find that individuals who show better representation in the variable context also demonstrate poorer speech-in-noise perception (see Figures 2 and 3).

From the perspective of the neural bases of speech perception, our findings demonstrate that speech encoding is a dynamic process that involves constant updating of current representation based on prior exposure. Indeed, these expectancy-based processes are crucial for speech perception in challenging listening environments. When SNR is seriously compromised, top-down context-dependent cues are critical, which explains the strong association between behavioral performance on speech-in-noise tests and context-dependent lower-level encoding of speech features.

From a clinical perspective, our results yield an objective neural index that can directly benefit assessment of children with reading problems. Noise-exclusion deficits are a hallmark clinical symptom in children with reading difficulties. In addition

to conventional intervention (phonological intervention/auditory training) strategies, children who show a context-dependent encoding deficit at the lower-level sensory stages may benefit from speech-in-noise training and/or use of augmentative communication (e.g., FM systems which eliminate background noise and provide an excellent SNR, thereby improving source segregation).

Conclusion

The current study demonstrates context-dependent modulation in the human auditory brainstem. Human auditory brainstem encoding is determined by both the acoustics of the incoming stimulus and the context in which the stimulus occurs. Such plasticity occurs more rapidly than previously thought, and may function to improve perception in challenging listening backgrounds. In children with developmental dyslexia, a broad deficit in the extraction of stimulus regularities may contribute to a critical deficit in noise-exclusion.

EXPERIMENTAL PROCEDURES

Participants

To be included in the study, children were required to have hearing thresholds ≤ 20 dB Hearing Level (HL) for octaves from 250 Hz to 8000 Hz and no air-bone conduction gap greater than 10 dB. Inclusionary criteria also included clinically normal auditory brainstem response latencies to click stimuli (100 μ s clicks presented at 80 dB Sound Pressure Level (SPL) at 31.1 Hz; see Table S4) and an estimate of intelligence of greater than 85 ($M = 123.4$, $SD = 16.5$) on the verbal subscore of the Wechsler Abbreviated Scale of Intelligence (WASI; The Psychological Corporation, 1999). Informed consent was obtained from all children and their legal guardians. The Internal Review Board at Northwestern University approved all procedures involved in this experiment.

Experiment 1

Participants were 21 right-handed children (12 male, age 8–13 years, $M = 10.4$; $SD = 1.6$) with no history of learning or neurological impairments.

Experiment 2

Participants in Experiment 2 were grouped into “poor readers” ($n = 15$) or “good readers” ($n = 15$), based on their performance on the Test of Word Reading Efficiency (Torgesen et al., 1999), a standardized test of reading ability. Only children with scores below 85 were included in the poor reading group. Additionally, poor readers carried an external diagnosis of reading or learning impairment made by professional clinicians, and attended a private school for the learning disabled. For the good reading group, we included children from Experiment 1 who had a reading score of >110 on the Test of Word Reading Efficiency. Also, all children in Experiment 2 underwent standardized tests of reading and spelling ability (Supplemental Experimental Procedures). Test results are summarized in Table S2. The good and poor reading groups ($n = 15$ for both) did not differ in age [$t(28) = -0.972$, $p = 0.339$] but did differ on verbal IQ [$t(28) = -3.673$, $p = 0.001$], which can be assumed given the dependence of this measure on short-term verbal working memory, which is known to be impaired in individuals with dyslexia. However, we took a conservative statistical approach and partialled out the contribution of verbal IQ in all correlations between physiological measures and behavioral indices (i.e., speech-in-noise perception).

Behavioral Procedures: Reading and Speech-in-Noise Measures

Behavioral indices of reading and speech-in-noise perception were collected. Reading ability was assessed with the Test of Word Reading Efficiency, which requires children to read a list of real words (Sight subtest) and nonsense words (Phoneme subtest) while timed (Torgesen et al., 1999). These subset scores are combined to form a Total score that was used to differentiate the good and poor readers in the present study.

Speech-in-noise perception was evaluated with the HINT (Bio-logic Systems Corp., Mundelein, IL). Sentence stimuli were presented in speech-shaped noise at varying SNRs in an adaptive paradigm in three different noise conditions: noise presented from the front, from the left, and from the right. In all conditions, the target sentences came from the front. A final threshold SNR value is calculated for each condition, yielding four measures (HINT-FRONT, HINT-RIGHT, HINT-LEFT, and HINT-COMPOSITE). Only age-normalized percentiles were used in the present analysis. In addition, for Experiment 2, the children underwent a number of cognitive tests. See the Supplemental Experimental Procedures and Table S2 for test descriptions and group differences.

Stimuli and Design

Stimulus and design for Experiment 1 and 2 were identical. Brainstem responses were elicited in response to the syllable /da/ presented to the right ear while the children watched a video of their choice. The /da/ stimulus was a 6 formant speech syllable synthesized in Klatt (1980). The stimulus was 170 ms long with a 5 ms voice onset time, a level fundamental frequency (F_0 : 100 Hz), and dynamic first, second, and third formants (F_1 : 400–720 Hz, F_2 : 1700–1240 Hz, F_3 : 2580–2500 Hz, respectively) during the first 50 ms. The fourth, fifth, and sixth formants were constant over the duration of the stimulus (F_4 : 3300 Hz, F_5 : 3750 Hz, F_6 : 4900 Hz, respectively). Brainstem responses to /da/ were collected from the scalp (at Cz) using Scan 4.3 Acquire (Compumedics) with Ag–AgCl scalp electrodes in a vertical, ipsilateral montage under two different conditions. In one session, 6300 sweeps of /da/ were presented with a probability of 100% (repetitive context). In a second session (variable context), 2100 sweeps of /da/ were presented randomly in the context of seven other speech sounds at a probability of 12.5%. The seven speech sounds varied in a number of acoustic features including formant structure (/ba/, /ga/, /du/), duration (a 163 ms /da/), voice-onset time (/ta/), and fundamental frequency (high pitch /da/, /da/ with a dipping pitch contour). For a detailed description of these stimuli, see Table S1. We then compared the brainstem responses to /da/ from the variable context condition to trial-matched /da/ responses in the repetitive context condition, resulting in 700 trials in each condition (see Figure 1). Importantly, by matching trials between the two conditions, we are able to examine differences in processing responses to the same stimuli under two different contexts without the confound of presentation order. Responses were offline bandpass filtered from 70 to 2000 Hz with a 12 dB roll-off, epoched from -40 to 190 ms (40 ms stimulus onset at time 0), and baseline corrected. The low-pass cutoff of 70 Hz was used to reduce cortical contribution. All stimuli were presented in alternating polarities via insert earphones at 80.3 dB SPL at a rate of 4.35 Hz and responses were digitized at 20,000 Hz. The fast presentation rate ensured that cortical contributions were minimized, since cortical neurons are unable to phase-lock at such fast rates (Chandrasekaran and Kraus, 2009). In addition to serving as a hearing screening, responses to 100 μ s clicks were collected before each auditory session (see Supplemental Experimental Procedures). Click-evoked wave V latencies were consistent across sessions for all participants in Experiment 1 and 2, ensuring that no differences existed in recording parameters across sessions [paired t test: $t(35) = 0.867$, $p = 0.392$].

Analyses

Events with amplitude greater than ± 35 μ V were rejected. Responses in the repetitive context condition were averaged according to their occurrence relative to the order of presentation in the variable context condition (Figure 1). Overall, an average of 700 trials were compared across the two conditions from each child.

In the current study, the responses were broken into two time ranges for analysis: 7–60 ms, which includes the response to the sound onset and the response to the formant transition, and 60–180 ms, which includes the response to the steady-state vowel (see Figure 2, top). Responses were examined in the time and frequency domains (Banai et al., 2009; Musacchia et al., 2007). To examine the strength of spectral encoding, average response magnitudes were calculated for 10 Hz wide bins surrounding the F_0 and subsequent five harmonics (100 Hz [F_0], 200 Hz [H_2], 300 Hz [H_3], 400 Hz [H_4], 500 Hz [H_5], and 600 Hz [H_6]). Since the F_1 sweeps from 400–720 Hz in the signal, an additional region of interest within the first formant trajectory

(400–720 Hz) was identified by comparing spectral encoding of responses to the repetitive context and variable context across 10 Hz wide bins for each participant in Experiment 1. The two conditions differed significantly (on point-to-point *t* tests) across 530 to 590 Hz and consequently, spectral amplitude averaged over that range was calculated for each child across the two conditions. Onset response latencies (peak and trough) were identified for each child and compared across both contexts to determine if context affected the conduction speed of the responses. Also, rectified mean amplitude (RMA) was calculated over both time ranges as a measure of overall response magnitude. SNR (RMA of prestimulus baseline/RMA of response) was calculated for both conditions and no significant differences were found [Experiment 1: variable mean: 1.40, repetitive mean: 1.59; paired *t* test: $t(20) = 0.568$, $p = 0.576$; Experiment 2: variable mean: 1.43, repetitive mean: 1.22; paired *t* test: $t(30) = -1.568$, $p = 0.0697$].

Statistical Analyses

For both time regions the mean spectral amplitudes for F0, H2–H6, and the F1 range were compared for the two conditions within each child using RmANOVAs and followup *t* tests. In Experiment 2, the 2(context) × 2(group) multivariate RmANOVAs were limited to H2, H4, and the F1 range (based on the results of Experiment 1). The differences in spectral amplitude of H2 and H4 in the 7–60 ms range between the two conditions (repetitive context minus variable context) were calculated for each child and normalized to the group mean by converting to a *z*-score. The *z*-scores were then correlated with the HINT measures in Experiment 1 and 2 and all other behavioral measures in Experiment 2 using Pearson's correlations.

SUPPLEMENTAL DATA

Supplemental data for this article include Supplemental Experimental Procedures, Supplemental Results, 4 tables, and 4 figures and can be found at [http://www.cell.com/neuron/supplemental/S0896-6273\(09\)00807-1](http://www.cell.com/neuron/supplemental/S0896-6273(09)00807-1).

ACKNOWLEDGMENTS

We would like to thank Steve Zecker, Brad Wible, Patrick Wong, and Ed Rubel for their comments on earlier versions of the manuscript. We would like to thank the members of the Auditory Neuroscience Laboratory who aided in data collection, and the children and their families who participated in the study. This work was supported by Grants NIH/NIDCD RO1-01510, F32DC008052, and NSF BCS-544846, and by the Hugh Knowles Center at Northwestern University.

Accepted: September 3, 2009

Published: November 11, 2009

REFERENCES

- Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends Cogn. Sci.* 11, 458–465.
- Ahissar, M., Lubin, Y., Putter-Katz, H., and Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nat. Neurosci.* 9, 1558–1564.
- Atiani, S., Elhilali, M., David, S.V., Fritz, J.B., and Shamma, S.A. (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron* 61, 467–480.
- Banai, K., Hornickel, J.M., Skoe, E., Nicol, T., Zecker, S., and Kraus, N. (2009). Reading and Subcortical Auditory Function. *Cerebral Cortex* 19, 2699–2707.
- Bregman, A.S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound* (Cambridge, MA: The MIT Press).
- Brokx, J.P.L., and Nootboom, S.G. (1982). Intonation and the perceptual separation of simultaneous voices. *J. Phonetics* 10, 23–36.
- Chandrasekaran, B., and Kraus, N. (2009). The scalp-recorded auditory brainstem response to speech: neural origins and plasticity. *Psychophysiology*, in press. Published online October 12, 2009. 10.1111/j.1469-8986.2009.00928.x.
- Demonet, J.F., Taylor, M.J., and Chaix, Y. (2004). Developmental dyslexia. *Lancet* 363, 1451–1460.
- Elhilali, M., Xiang, J., Shamma, S.A., and Simon, J.Z. (2009). Interaction between attention and bottom-up saliency mediates the representation of foreground and background in an auditory scene. *PLoS Biol.* 7, e1000129.
- Elliott, L.L., Hammer, M.A., and Scholl, M.E. (1989). Fine-Grained Auditory Discrimination in Normal Children and Children with Language-Learning Problems. *J. Speech Hear. Res.* 32, 112–119.
- Everatt, J., Steffert, B., and Smythe, I. (1999). An eye for the unusual: creative thinking in dyslexics. *Dyslexia* 5, 28–46.
- Galbraith, G.C., Arbagey, P.W., Branski, R., Comerchi, N., and Rector, P.M. (1995). Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport* 6, 2363–2367.
- Gao, E., and Suga, N. (1998). Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proc. Natl. Acad. Sci. USA* 95, 12663–12670.
- Gao, E., and Suga, N. (2000). Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: role of the corticofugal system. *Proc. Natl. Acad. Sci. USA* 97, 8081–8086.
- Hornickel, J., Skoe, E., Nicol, T., Zecker, S., and Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc. Natl. Acad. Sci. USA* 106, 13022–13027.
- Johnson, K.L., Nicol, T., Zecker, S.G., and Kraus, N. (2008). Developmental plasticity in the human auditory brainstem. *J. Neurosci.* 28, 4000–4007.
- Klatt, D.H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America* 67, 971–995.
- Kraus, N., and Nicol, T. (2005). Brainstem origins for cortical “what” and “where” pathways in the auditory system. *Trends in Neurosci.* 28, 176–181.
- Krishnan, A., and Gandour, J.T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain Lang.* 110, 135–148.
- Krishnan, A., Xu, Y., Gandour, J., and Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cogn. Brain Res.* 25, 161–168.
- Krishnan, A., Swaminathan, J., and Gandour, J.T. (2009). Experience-dependent Enhancement of Linguistic Pitch Representation in the Brainstem Is Not Specific to a Speech Context. *J. Cogn. Neurosci.* 21, 1092–1105.
- Luo, F., Wang, Q., Kashani, A., and Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *J. Neurosci.* 28, 11615–11621.
- Musacchia, G., Sams, M., Skoe, E., and Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. USA* 104, 15894–15898.
- Nahum, M., Nelken, I., and Ahissar, M. (2008). Low-level information and high-level perception: the case of speech in noise. *PLoS Biol.* 6, e126.
- Ramus, F. (2001). Dyslexia. Talk of two theories. *Nature* 412, 393–395.
- Ramus, F., and Szenkovits, G. (2008). What phonological deficit? *Q. J. Exp. Psychol.* 61, 129–141.
- Ramus, F., Rosen, S., Dakin, S.C., Day, B.L., Castellote, J.M., White, S., and Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain* 126, 841–865.
- Sayles, M., and Winter, I.M. (2008). Reverberation challenges the temporal representation of the pitch of complex sounds. *Neuron* 58, 789–801.
- Song, J.H., Skoe, E., Wong, P.C.M., and Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *J. Cogn. Neurosci.* 20, 1892–1902.
- Sperling, A.J., Lu, Z.L., Manis, F.R., and Seidenberg, M.S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nat. Neurosci.* 8, 862–863.
- Sperling, A.J., Lu, Z.L., Manis, F.R., and Seidenberg, M.S. (2006). Motion-Perception Deficits and Reading Impairment. *Psychol. Sci.* 17, 1047–1053.
- Strait, D.L., Kraus, N., Skoe, E., and Ashley, R. (2009). Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. *Eur. J. Neurosci.* 29, 661–668.

- Suga, N. (2008). Role of corticofugal feedback in hearing. *J. Comp. Physiol. A. Neuroethol. Sens. Neural. Behav. Physiol.* 194, 169–183.
- Suga, N., and Ma, X. (2003). Multiparametric corticofugal modulation and plasticity in the auditory system. *Nat. Rev. Neurosci.* 4, 783–794.
- Suga, N., Gao, E., Zhang, Y., Ma, X., and Olsen, J.F. (2000). The corticofugal system for hearing: recent progress. *Proc. Natl. Acad. Sci. USA* 97, 11807–11814.
- Suga, N., Xiao, Z., Ma, X., and Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron* 36, 9–18.
- Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang.* 9, 182–198.
- Tallal, P. (1975). Perceptual and linguistic factors in the language impairment of developmental dysphasics: an experimental investigation with the token test. *Cortex* 11, 196–205.
- The Psychological Corporation (1999). Wechsler Abbreviated Scale of Intelligence (WASI), San Antonio, TX.
- Torgesen, J.K., Wagner, R.K., and Rashotte, C.A. (1999). Test of Word Reading Efficiency (Austin, TX: Pro-Ed).
- Tzounopoulos, T., and Kraus, N. (2009). Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron* 62, 463–469.
- Winer, J.A. (2005). Decoding the auditory corticofugal systems. *Hear. Res.* 207, 1–9.
- Wong, P.C., Skoe, E., Russo, N.M., Dees, T., and Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* 10, 420–422.
- Yan, J., and Suga, N. (1999). Corticofugal amplification of facilitative auditory responses of subcortical combination-sensitive neurons in the mustached bat. *J. Neurophysiol.* 81, 817–824.
- Zhang, Y., and Suga, N. (1997). Corticofugal amplification of subcortical responses to single tone stimuli in the mustached bat. *J. Neurophysiol.* 78, 3489–3492.
- Zhang, Y., Suga, N., and Yan, J. (1997). Corticofugal modulation of frequency processing in bat auditory system. *Nature* 387, 900–903.
- Ziegler, J.C., Pech-George, C., George, F., and Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Sci.* 12, 732–745.