

THE AUDITORY BRAINSTEM IS A BAROMETER OF RAPID AUDITORY LEARNING

E. SKOE,^{a,c,g} J. KRIZMAN,^{a,b,c,g} E. SPITZER^{a,c,g} AND N. KRAUS^{a,c,d,e,f,g*}

^a Auditory Neuroscience Laboratory, Evanston, IL, USA

^b Bilingualism and Psycholinguistics Laboratory, Evanston, IL, USA

^c Department of Communication Sciences, Evanston, IL, USA

^d Institute for Neuroscience, Evanston, IL, USA

^e Department of Neurobiology and Physiology, Evanston, IL, USA

^f Department of Otolaryngology, Evanston, IL, USA

^g Northwestern University, Evanston, IL, USA

Abstract—To capture patterns in the environment, neurons in the auditory brainstem rapidly alter their firing based on the statistical properties of the soundscape. How this neural sensitivity relates to behavior is unclear. We tackled this question by combining neural and behavioral measures of statistical learning, a general-purpose learning mechanism governing many complex behaviors including language acquisition. We recorded complex auditory brainstem responses (cABRs) while human adults implicitly learned to segment patterns embedded in an uninterrupted sound sequence based on their statistical characteristics. The brainstem's sensitivity to statistical structure was measured as the change in the cABR between a patterned and a pseudo-randomized sequence composed from the same set of sounds but differing in their sound-to-sound probabilities. Using this methodology, we provide the first demonstration that behavioral-indices of rapid learning relate to individual differences in brainstem physiology. We found that neural sensitivity to statistical structure manifested along a continuum, from adaptation to enhancement, where cABR enhancement (patterned > pseudo-random) tracked with greater rapid statistical learning than adaptation. Short- and long-term auditory experiences (days to years) are known to promote brainstem plasticity and here we provide a conceptual advance by showing that the brainstem is also integral to rapid learning occurring over minutes. © 2013 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: auditory, brainstem, statistical learning.

*Correspondence to: N. Kraus, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA. Tel: +1-847-491-3181; fax: +1-847-491-2523.

E-mail address: nkraus@northwestern.edu (N. Kraus).

URL: <http://www.brainvolts.northwestern.edu> (N. Kraus).

Abbreviations: cABR, auditory brainstem responses to complex sounds; F_0 , fundamental frequency; FFT, fast Fourier transform; HL, hearing level; IC, inferior colliculus; MOC, medial olivocochlear; RMANOVA, repeated measures analysis of variance; SSA, stimulus-specific adaptation; SNR, signal-to-noise ratio; SPL, sound pressure level.

INTRODUCTION

The nervous system is continuously flooded with sensory information. Statistical learning facilitates the discovery of patterns in the environment by piggybacking on the nervous system's fundamental ability to compute statistical relationships between objects separated in time (Saffran, 2003). These statistical calculations, which generally occur rapidly (within minutes) and pre-consciously (Abla et al., 2008; Cunillera et al., 2009), can be observed in the auditory, visual, and tactile modalities (Conway and Christiansen, 2005; Turk-Browne et al., 2009; Mitchel and Weiss, 2011) and across different species (Toro and Trobalon, 2005; Saffran et al., 2008). In humans, this basic learning mechanism is believed to bootstrap language acquisition and promote the development of general cognitive abilities (Saffran et al., 2001; Graf Estes et al., 2007; Conway et al., 2009; Evans et al., 2009).

Behavioral-indices of auditory-based statistical learning have been traced to neuroplasticity in the human auditory cortex (Bonte et al., 2005; Abla et al., 2008; Francois and Schon, 2011). However, statistical learning is likely not the sole purview of the cortex. As evidence for this, there is a body of literature demonstrating that neurons throughout the auditory system, including subcortical centers such as the auditory nerve and inferior colliculus, change their firing patterns according to the statistical features of the soundscape (Ulanovsky et al., 2003, 2004; Nelken, 2004; Dean et al., 2005; Perez-Gonzalez et al., 2005; Malmierca et al., 2009; Wen et al., 2009; Antunes et al., 2010). Consistent with this work in animal models, sensitivity to simple stimulus statistics is also observed in far-field responses from the human auditory brainstem (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011; Slabu et al., 2012), whose cellular generators include the inferior colliculus (Smith et al., 1975; Moller et al., 1994; Chandrasekaran and Kraus, 2010). Human studies have also revealed that the auditory brainstem undergoes learning-related plasticity in response to short-term (de Boer and Thornton, 2008; Song et al., 2008, 2012; Carcagno and Plack, 2011; Chandrasekaran et al., 2012; Hornickel et al., 2012b) and long-term auditory experiences (Krishnan et al., 2005; Musacchia et al., 2007; Krizman et al., 2012), as well as rapid within-session plasticity (Skoee and Kraus, 2010b), suggesting that the brainstem subserves learning over both abbreviated and protracted time scales. Based on this collection of evidence, we hypothesized that the auditory brainstem is part of the circuitry mediating rapid statistical learning.

While human brainstem activity is known to be sensitive to salient statistics, such as when a single pattern is repeated without interruption or an individual sound is repeated with different probabilities, we are the first to address (1) how this neural sensitivity to stimulus statistics relates to rapid auditory learning of those patterns, and (2) whether this sensitivity generalizes to more ecologically valid conditions in which multiple sound patterns are embedded within a single uninterrupted sequence. Based on past work showing that multisession auditory training triggers enhancements in brainstem function (de Boer and Thornton, 2008; Song et al., 2008, 2012; Carcagno and Plack, 2011; Chandrasekaran et al., 2012; Hornickel et al., 2012b; Anderson et al., 2013), we predicted that rapid statistical learning would correlate with brainstem physiology, with enhanced neural responses to patterned sequences being associated with greater behavioral learning (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011). By coupling non-invasive neurophysiological and behavioral techniques, we take a new approach to investigating brainstem responses to complex sounds (cABRs). While cABRs have been instrumental in revealing that short-term auditory experiences (lasting days or weeks) promote learning-related brainstem plasticity (Kraus and Chandrasekaran, 2010; Skoe and Kraus, 2010a), until now, cABRs have only ever been measured before or after learning occurred, not during the learning process itself.

EXPERIMENTAL PROCEDURES

Experimental design

We utilized a well-established method for examining rapid statistical learning whereby subjects hear a sound sequence formed from an arbitrary concatenation of patterns and are tested afterward on their ability to recognize each pattern in isolation (Saffran et al., 1999; Ablat et al., 2008). In most statistical learning paradigms, sequences are created from a small inventory of patterns composed of speech syllables, tones, or other sounds and these patterns are strung together without immediate repeats or overt grouping cues to demarcate pattern boundaries (Saffran et al., 1999; Cunillera et al., 2008; Schon and Francois, 2011). In such paradigms, patterns can only be discovered by calculating the transitional probability of specific sound combinations within the sequence. For example, in an arbitrary sequence created from “paku”, “dibe”, “togi”, “dubo” (e.g., pakudibetogidibedubopakutogidubo) every time a “pa” is heard, there is a 100% probability that the next sound will be “ku”. Such statistical learning paradigms lead to implicit learning (Perruchet and Pacton, 2006) within a matter of minutes, as observed in adults (Saffran et al., 1999), children (Graf Estes et al., 2007), infants (Saffran et al., 1996), neonates (Teinonen et al., 2009) and non-human species (Toro and Trobalon, 2005; Saffran et al., 2008). Moreover, in normal populations performance is similar for linguistic and non-linguistic sequences, suggesting that statistical learning is a general-purpose learning mechanism (Saffran et al., 1999).

The brainstem’s role in rapid statistical learning was probed by recording cABRs, far-field responses to auditory input that reflect the synchronous activity of neural populations primarily in the lateral lemniscus and inferior colliculus (Smith et al., 1975; Moller et al., 1994; Chandrasekaran and Kraus, 2010). Recordings were made from young adults while they passively

listened to eight complex tones that were sequenced into two novel sound sequences composed from the same sounds but with different tone-to-tone statistics (i.e., transitional probabilities). The sound sequences contained either no underlying structure (pseudo-random) or structure in the form of four reoccurring two-tone (doublet) patterns. Within the patterned sequence the doublets can be considered “words” or meaningful units; in contrast, there were no meaningful sound combinations within the pseudo-random sequence. Subjects heard the pseudo-random condition first, followed by the patterned condition. After listening to the patterned sequence for 15 min., subjects were tested on how well they implicitly learned the patterns (Figs. 1 and 2).

Subjects

28 adults participated in this study (18 females, age range: 18.20–29.04 years, mean = 21.11). Written informed consent was obtained from all participants and all experimental protocols were reviewed and approved by Northwestern University’s Institutional Review Board. All participants had normal bilateral hearing (air conduction thresholds < 20 dB HL for 0.125, 0.250, 0.5, 1, 2, 3, 4, 6, and 8 kHz and ≤ 5 dB difference between left-ear and right-ear pure tone averages), normal click-evoked ABR wave V latencies based on lab-internal norms (30.1 Hz, 80 dB SPL), as well as normal neurological function (self-report). Subjects had a mean non-verbal IQ of 115.18 ± 10.227 (as measured by the Wechsler Abbreviate Scale of Intelligence Matrix Design subtest or the Test of Nonverbal Intelligence (TONI-3) and a range of formal musical (instrumental) instruction, from 0 to 13 years with an average of 4.21 ± 3.947 years.

Stimuli

The sound sequences were formed from eight 333-ms triangle waves (Fig. 3). The fundamental frequencies (F_0) of the individual complex tones were 262, 294, 330, 350, 370, 393, 416, and 440 Hz with each tone mapping onto a specific musical note (C4, D4, E4, F4, F#4, G4, G#4, A4, respectively). Sound stimuli were created in Adobe Audition (Adobe System Corp., San Jose, CA) with a 50-ms ramp (triangular window) applied to the onset and offset of the stimulus in the MATLAB programming environment (The Mathworks, Natrick, NJ). Ramping was applied to prevent clipping when the stimuli were sequenced. The complex tones contained only odd harmonics of the F_0 and each successive harmonic diminished in amplitude by $1/H^2$, where H = harmonic number. Triangle waves were chosen because they have a natural sound quality, with a timbre akin to a clarinet.

Tone sequences were generated with algorithms in MATLAB, resulting in one pseudo-random sequence and one patterned sequence (Audio clips 1–2). In the patterned and pseudo-random conditions, each tone had an equal probability of occurrence (1 of 8, or 12.5%) but the local neighborhoods, including the first-order transitional probabilities, were different (Table 1). To create the pseudo-random sequence, the eight tones were sorted pseudo-randomly within the sequence, with the proviso that no tone was repeated in immediate succession. For the patterned sequence the local neighborhood of each sound was constrained, and therefore quite predictable. In this sequence, the eight tones were grouped into four doublets (two-tone combinations): EC, F#F, DG, G#A that each occurred with a 25% probability (Fig. 2). As a first step to creating the patterned sequences, a deep structure was formed by stringing together the numbers 1–4 in pseudo-random order, with each number appearing 120 times with no immediate repeats (e.g., 1-2-3-1-3-2-4...). The individual doublets were then mapped onto the deep structure, with numbers 1–4 being replaced with EC, F#F, DG, G#A, respectively, to create the stimulus sequence.

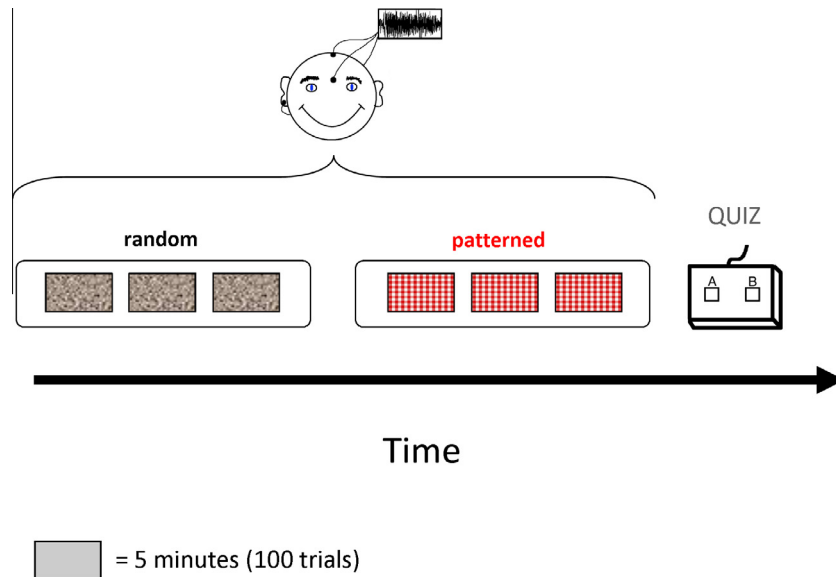


Fig. 1. Experimental design. Auditory brainstem responses to complex sounds (cABRs) were obtained using scalp-electrodes while subjects listened to continuous sequences of complex tones that formed either a pseudo-random sequence (black) or a patterned sequence composed of four recurring doublets (red) (Fig. 2). Electrodes were placed on the central vertex, forehead, and right earlobe. All subjects heard two conditions, with the pseudo-random sequence presented first. Each 5-min sequence (small rectangles) was presented three times, with intervening breaks. Within each 5-min block, each tone was heard approximately 100 times. After hearing the third block of the patterned condition, subjects were given a two-alternative forced choice quiz that tested their ability to distinguish the doublets from foils, two-tone combinations that never occurred in the patterned sequence. During the experiment, subjects sat in a comfortable reclining chair in a soundproof, electrically shielded booth. Subjects were instructed to stay awake while the sounds were presented and to keep their gaze on the nature images appearing on the screen in from them. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

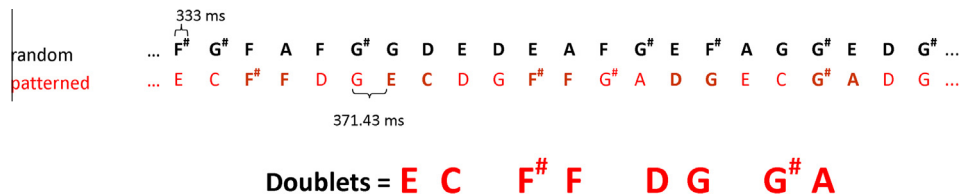


Fig. 2. Stimulus characteristics. Snapshots of the pseudo-random (black) and patterned (red) sequences are depicted here to illustrate their defining characteristics. Each sample sequence represents 8.17 s of the respective condition. The sequences were composed of eight 333-ms complex tones, with each mapping onto a different musical note. Within the sequences, the global statistics of the individual sounds were matched, such that each tone played with a 12.5% probability, while varying the local context of the sound (Table 1). In the pseudo-random sequence, no tone was repeated in immediate succession, but the sequence otherwise had no predictable structure. The patterned sequence was created from a set of four two-tone patterns (EC, F#F, DG, G#A) that were concatenated pseudo randomly without conspicuous pattern breaks. Each pattern occurred with a probability of 25% within the sequence but no pattern was played twice in a row (Table 1 and online Audio clips 1–2). For illustrative purposes, the patterns are plotted in alternating shades of light and dark red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Each sequence was presented three times (approximately 5 min/sequence), with intervening breaks between blocks (Fig. 1). After the subject reached the requisite number of artifact-free trials per block (100/tone) (see next section), the stimulus sequence was immediately stopped and the subject was given a short break. To allow for a small percentage of motion or other myogenic artifacts, an extra 20 trials per tone was buffered into the sequence; however, in a majority of the cases, the subjects reached the target number of artifact-free trials (100/tone) before reaching the end of the sequence.

Electrophysiology

Stimulus presentation. Sounds were delivered binaurally using Stim2 (Gentask module; Compumedics, Inc., Charlotte, NC) at 70 dB SPL via ER-3A ear insert tubephones (Etymotic Laboratories, Elk Grove Village, IL) with an inter-tone interval of 38.43 ms.

Recording and data processing procedure. cABRs were recorded with an analog-to-digital rate of 20 kHz using scalp-electrodes and a PC-based hardware/software system (SynAmps 2 amplifier, Neuroscan Acquire, Compumedics, Inc.). Three Ag–AgCl electrodes were placed on the scalp in a vertical montage (Hood, 1998) (the active electrode at the central vertex (Cz), reference electrode on the right earlobe, and the ground electrode on the forehead; Fig. 1). Contact impedance was kept <5 k Ω . Recordings were made in continuous (non-averaged) mode with an online filter of 0.5–3000 Hz.

As part of offline processing, recordings were filtered from 30 to 2000 Hz (12 dB/octave) and then each tone was epoched with a window of –10 to 350 ms (Neuroscan Edit). After baseline correcting each response to the mean voltage of the noise floor (–10 to 0 ms), trials containing myogenic artifact were discarded, using an automated procedure that flagged trials with activity exceeding the range of $\pm 35 \mu\text{V}$. On average, 3.56% of trials were discarded due to artifact. The rejection rate

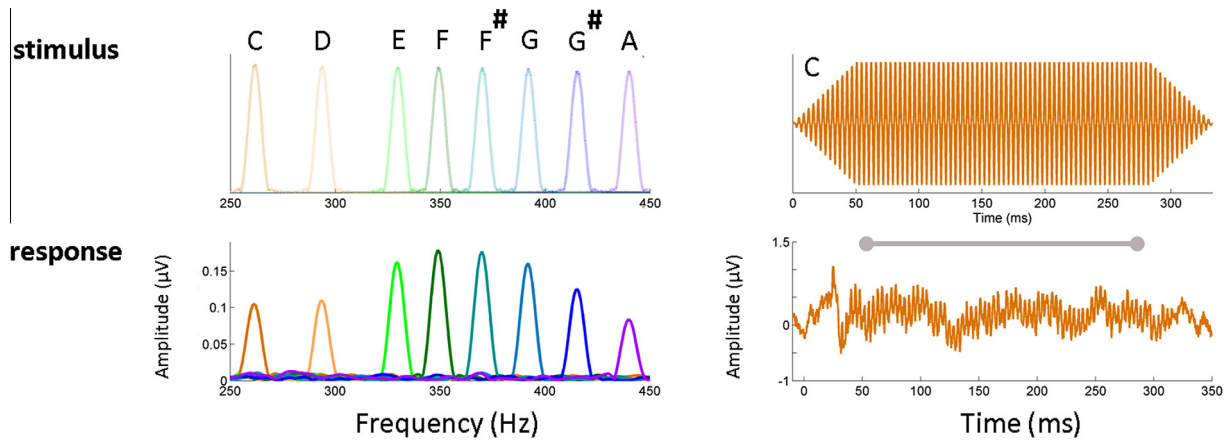


Fig. 3. The fundamental frequency of the stimulus is preserved in the complex auditory brainstem response (cABR) via phaselocking. cABRs were recorded to sound sequences composed of eight complex tones ranging in their fundamental frequency (F_0) from 262–440 Hz. To illustrate the fidelity of brainstem phaselocking, frequency domain waveforms are plotted for each tone (top left) and their respective responses (bottom left). Each color represents a different stimulus. The time domain waveform for the stimulus with the lowest F_0 (middle C on a piano) is plotted on the top right, with the response plotted below it. The stimulus amplitude is ramped during the initial and final 50 ms. For each response, the phaselocked component corresponding to unramped portion of the stimulus (55–278 ms, marked by gray line) was converted to the frequency domain and the amplitude at the stimulus F_0 was calculated. A representative subject is plotted.

Table 1. First-order transitional probabilities for the pseudo-random (A) and patterned sequences (B). First-order transitional probabilities, defined as the probability of two sounds being successive within the sequence, were calculated post hoc after the sequences were generated. (A) For the pseudo-random sequence, all sound combinations occur, except that no sound follows itself. Because the sequences were created with this “sample without replacement approach”, the average transitional probability is roughly equal to 14.3% or 1/7. In addition, because the sequence was created with a pseudo-random number generator, and the sequence was finite, the probabilities are matched but not identical. (B) For the patterned sequence, composed of four recurring doublets, the first-order transitional probabilities are more constrained such that only certain sound combinations occur. Doublets are defined as sound combinations with a transitional probability of 100% (bold)

Given...	...probability it is followed by							
	C	D	E	F	F#	G	G#	A
<i>A. Pseudo-random condition:</i>								
C		11.11	17.95	11.97	11.11	20.51	18.80	8.55
D	20.83		15.00	11.67	15.83	14.30	13.33	10.83
E	13.22	19.83		13.22	14.05	11.57	8.26	19.83
F	17.80	13.56	12.71		11.02	12.71	15.25	16.10
F#	16.24	7.69	19.66	13.68		11.11	15.38	15.38
G	10.92	14.29	16.81	10.08	10.08		21.01	16.81
G#	11.90	17.46	10.32	14.29	20.63	13.49		11.90
A	6.72	15.97	9.24	23.53	12.61	17.65	14.29	
<i>B. Patterned condition:</i>								
C		34.40			33.60		32.00	
D						100.00		
E	100.00							
F		31.09	34.45				34.45	
F#				100.00				
G			35.59		33.90		29.66	
G#								100.00
A		32.76	35.34		31.90			

was stable across blocks as evaluated by a six-way repeated measures analysis of variance (RMANOVA) ($F(5, 135) = 0.968$, $p = 0.440$). The remaining pool of artifact-free trials was averaged, with the mean number of trials being 298.47 ± 2.376 per tone for the pseudo-random condition and 299.64 ± 1.617 per tone for the patterned condition. The number of trials was matched between the two conditions ($F(1, 27) = 2.086$, $p = 0.160$).

Although cABRs are traditionally evoked to a single stimulus (e.g., speech syllable) that is presented repetitively over many thousands of trials (Skoe and Kraus, 2010a), this study follows

a recent trend to record a smaller number of sweeps to a greater variety of stimuli (Chandrasekaran et al., 2009; Skoe and Kraus, 2010b, 2012; Parbery-Clark et al., 2012).

Extracting the response to the fundamental frequency (F_0). The response to the F_0 served as the primary dependent variable because brainstem nuclei demonstrate robust phaselocking to the F_0 of tonal stimuli below 2 kHz (Moushegian et al., 1973) and because the brainstem response to the F_0 can be modified by short- and long-term auditory learning (Krishnan et al., 2005;

Wong et al., 2007; Song et al., 2008, 2012; Kraus and Chandrasekaran, 2010; Carcagno and Plack, 2011; Jeng et al., 2011; Skoe and Kraus, 2012).

The phaselocked component (55–278 ms) of each cABR subaverage (Fig. 3) was analyzed by applying a fast Fourier transform (FFT) with zero padding (Moushegian et al., 1973; Skoe and Kraus, 2010a), with the resultant response spectrum having a 1-Hz resolution. This 55–278-ms time window was chosen because it reflects when the stimulus amplitude is unchanging (50–273 ms), after accounting for the roughly 5-ms delay between when the stimulus enters the ear canal and when the inferior colliculus responds (Hood, 1998). For each tone, the amplitude of the response to the F_0 was obtained for each subject by finding the peak in the response spectrum nearest the F_0 of the stimulus (262, 294, 330, 350, 370, 393, 416, and 440 Hz, respectively) (Skoe and Kraus, 2010b). For the pseudo-random condition, the mean frequencies of the response were 261.54, 293.89, 329.71, 349.72, 369.89, 392.29, 415.50, and 440.04 Hz, respectively (SD = 1.835, 1.197, 0.763, 0.772, 0.875, 1.512, 0.923, 0.508). For the patterned condition, the mean frequencies were 261.29, 294.07, 329.46, 349.57, 370.25, 392.46, 415.43 and 439.75 Hz, respectively (SD = 1.697, 2.227, 1.138, 1.731, 1.735, 1.710, 1.260, 2.119). The frequencies of the F_0 peaks did not differ between the patterned and pseudo-random conditions ($F(1,27) = 0.015$, $p = 0.902$, as assessed by a Condition \times Tone RMANOVA). For each F_0 peak, a frequency domain signal-to-noise ratio (SNR) was calculated by dividing the amplitude of the F_0 peak by the average energy of the noise floor. We defined the noise floor as the average spectral energy 30 Hz above and 30 Hz below the F_0 peak.

Measuring the extent of pattern enhancement. To gauge how the cABR to the patterned condition (Patt) differed from the pseudo-random (Rand) condition, a percent change was calculated:

$$\text{Percent change (\%)} = ((\text{Patt} - \text{Rand})/\text{Rand}) * 100$$

where Patt and Rand represent the average F_0 amplitude to the eight tones for the patterned and pseudo-random conditions, respectively. A positive percent change indicates that the response was larger (i.e., enhanced) in the patterned condition relative to the pseudo-random one.

Experimental instructions and setting

At the outset of each block, subjects heard the following pre-recorded instructions: “You will now hear a series of tones. Listen carefully to the sounds because later on you will be asked some questions to gauge how well you remembered the sounds. Please keep your eyes open and focus your gaze on the image on the screen. Try to sit as relaxed as possible. This section will last 15 min – you will get a break every 5 min or so.” To facilitate alertness while minimizing muscle movement, subjects were shown a slideshow of 60 nature photos. The 1280 \times 857 pixel images were played from standard DVD player and projected into the testing chamber onto a large projector screen in front of the participant. Each photo was presented for 1 min with a four second fade between each photo. Because statistical learning can be interrupted by a concurrent task that is attention demanding (Toro et al., 2005), subjects did not perform a photo-related task nor other secondary task.

Behavioral assessment of learning

Immediately after the patterned condition, subjects were tested on how well they learned the individual doublets comprising the patterned sequence. In a two-alternative forced-choice test,

each doublet was paired with a foil pair, two sounds that were never played together in the patterned sequence (Saffran et al., 1999; Ablá et al., 2008). Stimuli were presented via Stim2. Subjects chose the more familiar sounding doublet in the pair by pressing either the “A” or “B” button on a response box: “A” for the first doublet and “B” for the second. Each doublet was paired once with one of four foils (EF, F#C, DA, G#G), creating 16 comparisons. As with the “real” doublets, the foils began with E, F#, D, G# (respectively) but differed in which tone followed; this manipulation allowed us to test how the subjects learned the statistical structure of the patterned sequence. Behavioral scores were converted to percent correct, with 50% representing chance performance. For two subjects, behavioral scores fell outside two standard deviations of the group mean (at 2.72 and –2.53, respectively). Following common practice for reducing the influence of outlying points, these scores were corrected to be at +2 or –2 (respectively) standard deviations of the mean (Tabachnick and Fidell, 2007; Field and Miles, 2010).

RESULTS

Neurophysiological data

To confirm that the responses were sufficiently robust, frequency domain SNRs were computed. When collapsing across notes, the average SNR was 3.84 ± 1.229 (range: 1.76–6.05) for the pseudo-random condition and 3.78 ± 1.308 (range: 1.18–5.87) for the patterned condition. The SNR did not differ between conditions nor did the energy of the noise floor ($t(27) = 0.554$, $p = 0.584$; $t(27) = 0.293$, $p = 0.771$, respectively). Based on these analyses, we concluded that the responses were robust enough to analyze further.

As a first analysis, we wanted to establish that the cABR is different between the patterned and pseudo-random conditions on the group level. A Condition \times Tone RMANOVA revealed a main effect of condition ($F(1,27) = 5.548$, $p = 0.026$), with the response to the patterned condition being smaller than the pseudo-random condition when considering the subjects as a group (Fig. 4).

However, individual subjects demonstrated a spectrum of physiological outcomes that ranged from adaptation to enhancement, with the percent change following a normal distribution (Shapiro Wilk’s Test of Normality: $W(28) = 0.944$, $p = 0.139$) that extended from –30.64% to 10.42% (average percent change: $-3.52 \pm 9.827\%$). 35.71% of the subjects showed an enhanced cABR to the patterned condition.

Behavioral index of statistical learning

Across the group of subjects, behavioral performance exceeded chance ($t(27) = 5.658$, $p < 0.005$), suggesting that on the group level the subjects learned, at least to some degree, the structure of the patterned sequence. Scores ranged from 37.50% to 85.15% and followed a normal distribution (Shapiro Wilk’s Test of Normality: $W(28) = 0.934$, $p = 0.078$), with the average performance being $61.27 \pm 10.542\%$. This range and distribution of individual differences in statistical learning is on par with what has been reported previously (Saffran et al., 1999; Ablá et al., 2008).

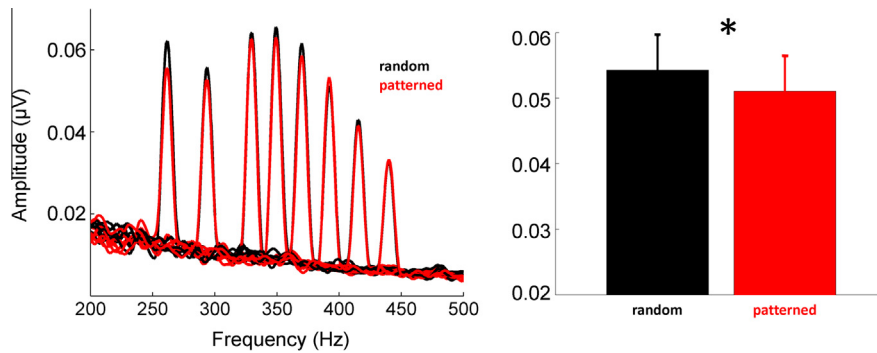


Fig. 4. Context-dependent brainstem encoding. On the group level, responses are smaller in the patterned condition (red) relative to the pseudo-random condition (black) ($F(1,27) = 5.548, p = 0.026$). Frequency domain waveforms for each note are overlaid (left), with the mean amplitude for each condition plotted as a bar graph (right). (+1 standard error of the mean). To capture the differences between conditions, the responses to the individual tones in the pseudo-random condition are all plotted in black and responses to the individual tones in the patterned condition are all plotted in red. This group average of brainstem function cloaks individual differences in brainstem function that can be seen in Fig. 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Relationship between neurophysiology and behavior

As predicted, statistical learning performance correlated with brainstem physiology, such that higher behavioral scores mapped to greater neural enhancements in the patterned condition relative to the pseudo-random one (Pearson's $r = 0.560, p = 0.002$) (Fig. 5). It was further confirmed that this brain-behavior relationship reflected the physiological differences between conditions and not the amplitude of the individual conditions. That is, behavioral performance did not correlate with the response to either the patterned or pseudo-random conditions, the two neural measures used to derive the percent change (behavior vs. pseudo-random condition: Pearson's $r = -0.238, p = 0.223$; behavior vs. patterned condition: Pearson's $r = -0.032, p = 0.873$).

The effect of musical training

Musicians have been shown to be better statistical learners (Schon and Francois, 2011; Shook et al., 2012; Francois et al., 2013), in addition to showing brainstem enhancements to statistically predictable stimulus conditions that correlate with years of musical practice (Parbery-Clark et al., 2011). Based on this, we predicted that novice musicians would show behavioral and neural enhancements in our statistical learning paradigm. As predicted, subjects with musical training ($n = 19$) outperformed those without musical training ($n = 9$) on the behavioral test of statistical learning ($63.98 \pm 9.180\%$ vs. $55.56 \pm 11.46\%$; $t(26) = -2.096, p = 0.046$); however, as a group, the musically trained subjects did not show larger physiological enhancements than the untrained subjects ($t(26) = -1.292, p = 0.225$). This is because the neural enhancements were observed in both trained and untrained subjects. Yet, when considering only those subjects with musical training, more years of training did correlate with greater physiological enhancements (Pearson's $r = 0.457, p = 0.043$).

Controlling for time- and state-dependent effects

In a separate control study, we confirmed that the response to the pseudo-random condition is stable when presented for a longer period. This test-retest

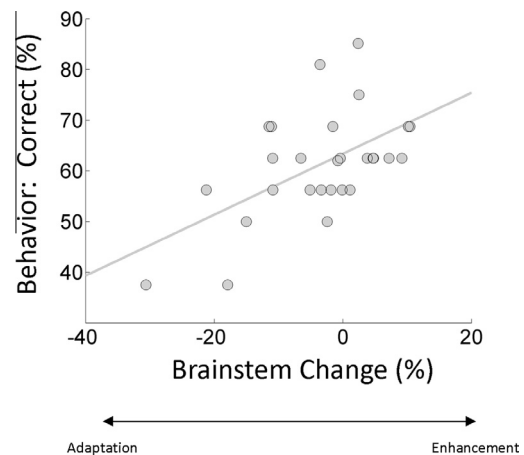


Fig. 5. Brainstem physiology predicts behavioral outcomes. Percent correct on the behavioral test is plotted as a function of percent change in the complex auditory brainstem response (cABR). Percent change is defined as: $((\text{Patt} - \text{Rand})/\text{Rand}) * 100$, where Patt and Rand represent the average cABR to the patterned and pseudo-random conditions, respectively. Greater neural enhancements to the pattern condition relate to greater behavioral learning (Pearson's $r = 0.560, p = 0.002$).

confirmation was performed on a different set of subjects ($n = 18, 14$ females), who had similar demographics to those tested on the patterned condition (age: $18.95\text{--}27.26$ years, non-verbal IQ: 119.70 ± 7.174 , years of formal music instruction: 7.56 ± 5.00 years). In this case, when the pseudo-random condition was presented a second time (six blocks total), there was no main effect of condition ($F(1,27) = 1.156, p = 0.297$, as assessed by a Condition \times Tone RMANOVA). Consistent with this, the change from the first to the second presentation of the pseudo-random condition was not statistically different from 0% (one sample t -test, $t(17) = -0.580, p = 0.569$); that is $((\text{Rand}_2 - \text{Rand}_1)/\text{Rand}_1) * 100 = 0\%$. This finding is consistent with literature demonstrating that the auditory brainstem response has a high test-retest reliability (Chiappa et al., 1979; Song et al., 2011a; Hornickel et al., 2012a). Because the

experimental setting, instruction, and duration were identical to the central Statistical Learning experiment, the results of this control study help to rule out the potential influence of task or state-dependent effects (Baas et al., 2006; Hairston et al., 2013).

DISCUSSION

For humans, statistical learning provides a cornerstone for language and music skills by underpinning the process of grouping sounds into meaningful units (Saffran, 2003; Graf Estes et al., 2007). Here we examined the subcortical correlates of this fundamental learning mechanism by recording cABRs during the learning process, something that has been never done before. We found that the amount of statistical learning that takes place over 15 min relates to the strength with which statistically probable sound sequences are indexed in the brainstem. Although our stimuli did not contain the full complexities of natural sound systems, the ability to segment our relatively simple sequence into constituent patterns draws on the same mechanisms needed to parse ongoing speech and music (Saffran et al., 1999; Romberg and Saffran, 2010). Thus, with this statistical learning paradigm, we believe we are tapping into the neural networks that subserves human's ability to learn complex sound systems such as language and music.

The brainstem is an active sound processor: Extending previous findings

As demonstrated previously in human and animal models, the auditory brainstem is an active sound processor that is sensitive to statistical features of auditory input, including changes in predictability and immediate context (Dean et al., 2005, 2008; Perez-Gonzalez et al., 2005; Chandrasekaran et al., 2009; Malmierca et al., 2009; Skoe and Kraus, 2010b; Parbery-Clark et al., 2011; Strait et al., 2011; Slabu et al., 2012). However, in humans this sensitivity is not absolute and does not emerge under all stimulus conditions (Slabu et al., 2012) or for all populations (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011; Strait et al., 2011). For example, in children with language-based learning impairments the cABR does not change when the statistics of the stimulus train are altered. This is in contrast to typically developing children whose cABRs are enhanced when the stimulus train contains a single (predictable) sound compared to a pseudo-randomized mix of sounds in which the target sound is presented both infrequently and unpredictably (Chandrasekaran et al., 2009). Similar findings have emerged for highly trained expert musicians compared to non-musicians using the same stimulus paradigm (Parbery-Clark et al., 2011). In this case, musicians, unlike non-musicians, demonstrate larger cABRs to the statistically predictable condition versus the unpredictable one. Notably, these previous cABR findings dovetail with behavioral evidence that musicians outperform non-musicians on auditory statistical learning tasks (Francois and Schon,

2011; Shook et al., 2012; Francois et al., 2013) and language-impaired children perform at chance on such tasks (Evans et al., 2009).

While this forerunning work hinted that the auditory brainstem might be involved in statistical learning, and might account for individual differences in human learning, conclusive statements could not be drawn for a variety of reasons. First, and foremost, unlike the current study, which recorded cABRs as part of a learning paradigm, in previous studies, neurophysiological assessments did not incorporate behavioral measures of statistical learning. The stimulus parameters used previously were also not optimal for establishing a link between brainstem physiology and statistical learning. Due to how the stimulus paradigms were constructed, with some sounds occurring more frequently than others, it was unclear whether neural enhancements in humans reflected the brainstem's ability to track statistical relationships between adjacent sounds, a requisite for learning linguistic and musical structure, or merely its ability to track how frequently an individual sound occurs within a sound sequence (Chandrasekaran et al., 2009; Skoe and Kraus, 2010b; Parbery-Clark et al., 2011; Strait et al., 2011; Slabu et al., 2012).

To tease out the role that the brainstem plays in statistical learning, and also to assess the computational sensitivity of the brainstem, we incorporated a behavioral test of learning and compared cABRs between patterned and pseudo-random sound sequences where the frequency of occurrence for each sound was equated between sequences. In addition, to make the sound sequences more naturalistic, we modeled the patterned condition after those used previously in the statistical learning literature (Saffran et al., 1999; Graf Estes et al., 2007). The patterned sequence contained four interknit patterns, with the only grouping cue being the statistical dependencies between sounds. Additionally, although the sound patterns repeated within the sequence, there were no local repeats, with each pattern always separated by at least one other. Thus, to discover the patterns within the continuous stream of sounds, inter-sound statistics, not the individual sound probability, needed to be tracked over time.

By recording cABRs as part of a statistical learning paradigm and controlling for the frequency of occurrence, our results reinforce and extend the view that the brainstem is an active sound processor that adjusts dynamically based on the immediate behavioral context (Perez-Gonzalez et al., 2005; Tzounopoulos and Kraus, 2009; Kraus and Chandrasekaran, 2010; Skoe and Kraus, 2010b). We speculate that if the brainstem were only sensitive to the frequency of occurrence but not other statistical cues imparting structure (e.g., the likelihood of two sounds being temporally adjacent), then the two conditions should have produced indistinguishable cABRs because the frequency of occurrence was matched. The fact that the cABR is different between the pseudo-random and patterned conditions, and the fact that the response to the

pseudo-random condition is stable over an extended time period as revealed in the control study, leads us to more definitively conclude that the brainstem *can* index statistical relationships among sounds, in addition to tracking the overall probability of individual sounds.

Notably, however, we observed individual differences in how stimulus statistics were indexed in the cABR. Within our population of normal young adults, the brainstem's sensitivity to stimulus statistics manifested along a continuum from adaptation to enhancement. Similar to what has been shown for cortical-evoked potentials (Abla et al., 2008), we found that the more enhanced the subcortical response to the patterned (predictable) condition was relative to the pseudo-random condition, the greater the learning outcome. Moreover, we found that behavioral performance correlated with the *change* in physiology between the patterned and pseudo-random conditions, but it did not correlate with the amplitude of the response to the individual conditions. This is interesting because it suggests that individual differences in auditory learning are not simply driven by baseline differences in auditory brainstem function *but instead reflect individual differences in how the brainstem adjusts to different stimulus conditions.*

Taken together, our findings suggest that differences in auditory brainstem physiology may, at least partially, account for individual differences in statistical learning that are observed across normal and impaired populations (Saffran et al., 1999; Abla et al., 2008; Evans et al., 2009; Scott-Van Zeeland et al., 2010; Misyak and Christiansen, 2012; Peretz et al., 2012). In this regard, the auditory brainstem provides a barometer of individual differences in behavioral outcomes. We suggest that these physiological differences may reflect genetic and/or experiential differences among subjects (Parbery-Clark et al., 2009; Soros and Stanton, 2012; Hairston et al., 2013; Hornickel and Kraus, 2013). Although the continuum of function that we observed could not be explained by differences in musical history across the whole group, within the musically trained subjects, more years of practice tracked with greater physiological enhancements. This result supports the idea that experience can shape these processes, but that it alone cannot account for the range of inter-individual function we observed.

Neural mechanisms of statistical learning: The interaction of local and top-down mechanisms within the inferior colliculus

Multiple lines of evidence from animal and human data trace the neural origins of the cABR to the rostral brainstem, with a major contributor being the inferior colliculus (IC) (Smith et al., 1975; Moller et al., 1994; Chandrasekaran and Kraus, 2010; Warrier et al., 2011). The IC, one of the most metabolically active centers of the brain (Sokoloff, 1977), serves as a relay center for ascending and descending auditory information (Winer, 2006). Evidence from animal models points to the IC as having a kind of “primitive intelligence” that can track complex stimulus statistics over time (Perez-Gonzalez et al., 2005). We now expand upon our understanding of the IC, and the functional role it plays in processing and

learning stimulus statistics by concluding that the IC is part of the neural circuitry mediating rapid auditory learning.

While studies of IC function in humans and laboratory animals have revealed that neurons in the IC adjust their responses based on the statistics of the sound input, the effect manifested in different forms. Direct recordings of the IC from anesthetized animals indicate that single neurons and neuronal populations adapt to commonly occurring sound combinations and features (Perez-Gonzalez et al., 2005; Malmierca et al., 2009). Yet, in contrast, for far-field IC recordings from awake humans statistically predictable stimulus conditions produce either no change or enhanced cABRs (Chandrasekaran et al., 2009; Skoe and Kraus, 2010b; Parbery-Clark et al., 2011; Slabu et al., 2012). Our data help to unify these seemingly contradictory results in the animal and human literatures by revealing a physiological spectrum in a healthy young adult population, with some individuals showing enhanced and others showing attenuated responses to patterned sound sequences.

To provide a unified account of the IC and its sensitivity to statistical regularities, we theorize that adaptation represents a default mode of indexing statistical regularities within the IC. This default mode may, however, be overridden by a top-down (exogenous) mechanism(s) that accentuates the transmission of sounds or sound features, leading to an enhanced response under certain stimulus conditions and/or in certain individuals. Thus, we raise the possibility that neural enhancements seen in the cABR may reflect learning-related neuroplastic changes occurring in higher-order centers of the brain that alter the functional state of the auditory brainstem. Such top-down initiated changes could arise by activating the corticofugal system, a communication link between the auditory cortex and subcortical structures. Through this system of efferents the cortex is able to rapidly alter subcortical sensory processing to improve, in an online fashion, the input the cortex receives (Suga et al., 2002). By our argument, activation of the corticofugal system may induce a temporary change in the balance between excitatory and inhibitory mechanisms in the IC, and this change may override adaptive mechanisms that are endogenous to subcortical structures producing a response enhancement (Yan et al., 2005; Nelken and Ulanovsky, 2007; Bauerle et al., 2011). In this regard, learning-related neural plastic changes need not necessarily have occurred at the level of the IC but may reflect instructive feedback from the cortex that is relayed during the learning process.

There are several layers of evidence from the animal literature to support our argument. First, there is evidence that the corticofugal pathway is necessary for auditory learning, given that learning is disrupted when this efferent pathway is ablated (Bajo et al., 2010). Second, IC neurons that adapt to a repeated stimulus are found in a region of the midbrain that is rich with corticofugal connections (Herbert et al., 1991; Perez-Gonzalez et al., 2005). Third, because animal physiology is typically recorded under anesthetized conditions, and because anesthesia mutes efferent activity (Boyev et al., 2002), direct recordings from the

IC in animal models likely reflect the default mode (i.e., adaptation) of processing stimulus statistics. However, the very existence of this mass of top-down connections in the IC suggests that these neurons, which are sensitive to stimulus statistics, might be under the influence of the auditory cortex for more active listening conditions (Perez-Gonzalez et al., 2005; Nelken and Ulanovsky, 2007).

Although the current methodology for recording cABRs in humans cannot provide direct evidence that top-down modulation has occurred, or isolate the points at which top-down modulation might be initiated and terminated, such techniques have been developed in humans for studying medial olivocochlear (MOC) efferents, the lowest branch of the corticofugal pathway that originates in the superior olive and innervates the Organ of Corti (Zhao and Dhar, 2011). From this work on the human MOC, we know that auditory learning can alter efferent activity (de Boer and Thornton, 2008) and that efferent function falls along a continuum with, for example, highly trained musicians having stronger efferent function (Perrot et al., 1999) and children with auditory processing disorders having reduced function (Muchnik et al., 2004).

Moreover, top-down processes, such as selective attention, can affect even the most peripheral levels of the auditory system as seen in both human and animal models (Delano et al., 2007; Srinivasan et al., 2012). While attentional demands were not experimentally manipulated in the current study, we cannot rule out that attention did not play a role in our findings. Subjects were not given explicit instructions as to what to attend to in the stimuli but were instead instructed to “listen carefully to the sounds” irrespective of the condition. However, in line with what has been proposed by Baker et al. (2004) for visual statistical learning, the structured nature of the patterned sound sequence may have made it inherently more attention grabbing than the unpatterned sequence. As such, involuntary attention to the stimuli may have differed between the patterned and unpatterned conditions (Baker et al., 2004). Given known differences in efferent function among individuals, combined with inter-individual differences in brainstem physiology (Galbraith et al., 2000; Banai et al., 2009; Chandrasekaran et al., 2009; Parbery-Clark et al., 2011; Song et al., 2011b; Soros and Stanton, 2012; Hairston et al., 2013) and attentional abilities (Soveri et al., 2011; Krizman et al., 2012), we speculate that the ability to learn the statistics of a soundscape hinges on individual differences in brainstem physiology, both in terms of how inter-sound statistical relationships are calculated and how these calculations are modified by individual differences in top-down processes.

Which came first, brainstem enhancements or learning?

While we have established a link between brainstem physiology and learning occurring over approximately 15 min, this link may be partially obfuscated by the “noisiness” of behavioral data. Although cABRs may

reveal the physiological conditions promoting behavioral learning, unlike neuroelectric measurements, behavioral tests may underestimate the extent of learning because they can be tainted by the subject’s motivation, task understanding, etc. (Francois and Schon, 2010). Moreover, while cABR enhancements are associated with better behavioral learning, our methodology cannot dissect whether brainstem enhancements are the instigator or the result of greater behavioral outcomes, or when in the chain of events that underlie learning the brainstem enhancements emerge. Thus, the relationship between cABR enhancements and improved behavioral performance remains a “chicken or egg” question. Developing non-invasive methodology to pinpoint when top-down modulation is initiated in humans, isolating the roles of excitatory and inhibitory pathways, and tracking behavioral learning over time, may unlock the sequence of neural events that underlie learning in addition to shedding light on the source(s) of the individual differences we observed.

CONCLUSION

There is mounting evidence that the auditory brainstem is tuned by learning-related experiences lasting days to years (Kraus and Chandrasekaran, 2010; Skoe and Kraus, 2010a). We now provide a significant scientific advance by showing that the auditory brainstem is also integral to learning occurring over minutes. Future studies should test the limits of the brainstem’s computational abilities and its involvement in online learning by using more complex stimulus statistics that approximate natural language.

CONFLICT OF INTEREST

The authors declare no competing financial interests.

Acknowledgments—Supported by NSF-0842376 and the Knowles Hearing Center at Northwestern University. We extend our thanks to Beverly Wright for her discussion of the data and Scott Miller for sharing his stimuli. We also thank Trent Nicol, Travis White-Schwoch, Samira Anderson, and Kali Woodruff for their comments on an earlier draft of the manuscript.

REFERENCES

- Abla D, Katahira K, Okanoya K (2008) On-line assessment of statistical learning by event-related potentials. *J Cogn Neurosci* 20:952–964.
- Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N (2013) Reversal of age-related neural timing delays with training. *Proc Natl Acad Sci USA* 110:4357–4362.
- Antunes FM, Nelken I, Covey E, Malmierca MS (2010) Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *PLoS One* 5:e14071.
- Baas JM, Milstein J, Donlevy M, Grillon C (2006) Brainstem correlates of defensive states in humans. *Biol Psychiatry* 59:588–593.
- Bajo VM, Nodal FR, Moore DR, King AJ (2010) The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci* 13:253–260.

- Baker CI, Olson CR, Behrmann M (2004) Role of attention and perceptual grouping in visual statistical learning. *Psychol Sci* 15:460–466.
- Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N (2009) Reading and subcortical auditory function. *Cereb Cortex* 19:2699–2707.
- Bauerle P, von der Behrens W, Kossl M, Gaese BH (2011) Stimulus-specific adaptation in the gerbil primary auditory thalamus is the result of a fast frequency-specific habituation and is regulated by the corticofugal system. *J Neurosci* 31:9708–9722.
- Bonte ML, Mitterer H, Zellagui N, Poelmans H, Blomert L (2005) Auditory cortical tuning to statistical regularities in phonology. *Clin Neurophysiol* 116:2765–2774.
- Boyev KP, Liberman MC, Brown MC (2002) Effects of anesthesia on efferent-mediated adaptation of the DPOAE. *J Assoc Res Otolaryngol* 3:362–373.
- Carcagno S, Plack CJ (2011) Subcortical plasticity following perceptual learning in a pitch discrimination task. *J Assoc Res Otolaryngol* 12:89–100.
- Chandrasekaran B, Kraus N (2010) The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47:236–246.
- Chandrasekaran B, Hornickel J, Skoe E, Nicol T, Kraus N (2009) Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron* 64:311–319.
- Chandrasekaran B, Kraus N, Wong PC (2012) Human inferior colliculus activity relates to individual differences in spoken language learning. *J Neurophysiol* 107:1325–1336.
- Chiappa KH, Gladstone KJ, Young RR (1979) Brain stem auditory evoked responses: studies of waveform variations in 50 normal human subjects. *Arch Neurol* 36:81–87.
- Conway CM, Christiansen MH (2005) Modality-constrained statistical learning of tactile, visual, and auditory sequences. *J Exp Psychol Learn Mem Cogn* 31:24–39.
- Conway CM, Pisoni DB, Kronenberger WG (2009) The importance of sound for cognitive sequencing abilities: the auditory scaffolding hypothesis. *Curr Dir Psychol Sci* 18:275–279.
- Cunillera T, Gomila A, Rodriguez-Fornells A (2008) Beneficial effects of word final stress in segmenting a new language: evidence from ERPs. *BMC Neurosci* 9:23.
- Cunillera T, Camara E, Toro JM, Marco-Pallares J, Sebastian-Galles N, Ortiz H, Pujol J, Rodriguez-Fornells A (2009) Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage* 48:541–553.
- de Boer J, Thornton AR (2008) Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *J Neurosci* 28:4929–4937.
- Dean I, Harper NS, McAlpine D (2005) Neural population coding of sound level adapts to stimulus statistics. *Nat Neurosci* 8:1684–1689.
- Dean I, Robinson BL, Harper NS, McAlpine D (2008) Rapid neural adaptation to sound level statistics. *J Neurosci* 28:6430–6438.
- Delano PH, Elgueda D, Hamame CM, Robles L (2007) Selective attention to visual stimuli reduces cochlear sensitivity in chinchillas. *J Neurosci* 27:4146–4153.
- Evans JL, Saffran JR, Robe-Torres K (2009) Statistical learning in children with specific language impairment. *J Speech Lang Hear Res* 52:321–335.
- Field AP, Miles J (2010) *Discovering statistics using SAS: (and sex and drugs and rock 'n' roll)*. Los Angeles: SAGE.
- Francois C, Schon D (2010) Learning of musical and linguistic structures: Comparing event-related potentials and behavior. *Neuroreport* 21:928–932.
- Francois C, Schon D (2011) Musical expertise boosts implicit learning of both musical and linguistic structures. *Cereb Cortex* 21:2357–2365.
- Francois C, Chobert J, Besson M, Schon D (2013) Music training for the development of speech segmentation. *Cereb Cortex*. <http://dx.doi.org/10.1093/cercor/bhs180>.
- Galbraith GC, Buranahirun CE, Kang J, Ramos OV, Lunde SE (2000) Individual differences in autonomic activity affects brainstem auditory frequency-following response amplitude in humans. *Neurosci Lett* 283:201–204.
- Graf Estes K, Evans JL, Alibali MW, Saffran JR (2007) Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychol Sci* 18:254–260.
- Hairston WD, Letowski TR, McDowell K (2013) Task-related suppression of the brainstem frequency following response. *PLoS One* 8:e55215.
- Herbert H, Aschoff A, Ostwald J (1991) Topography of projections from the auditory cortex to the inferior colliculus in the rat. *J Comp Neurol* 304:103–122.
- Hood LJ (1998) *Clinical applications of the auditory brainstem response*. San Diego: Singular Pub. Group.
- Hornickel J, Kraus N (2013) Unstable representation of sound: a biological marker of dyslexia. *J Neurosci* 33:3500–3504.
- Hornickel J, Knowles E, Kraus N (2012a) Test-retest consistency of speech-evoked auditory brainstem responses in typically-developing children. *Hear Res* 284:52–58.
- Hornickel J, Zecker SG, Bradlow AR, Kraus N (2012b) Assistive listening devices drive neuroplasticity in children with dyslexia. *Proc Natl Acad Sci USA* 109:16731–16736.
- Jeng FC, Hu J, Dickman B, Montgomery-Reagan K, Tong M, Wu G, Lin CD (2011) Cross-linguistic comparison of frequency-following responses to voice pitch in American and Chinese neonates and adults. *Ear Hear* 32:699–707.
- Kraus N, Chandrasekaran B (2010) Music training for the development of auditory skills. *Nat Rev Neurosci* 11:599–605.
- Krishnan A, Xu Y, Gandour J, Cariani P (2005) Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res* 25:161–168.
- Krizman J, Marian V, Shook A, Skoe E, Kraus N (2012) Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proc Natl Acad Sci USA* 109:7877–7881.
- Malmierca MS, Cristaudo S, Perez-Gonzalez D, Covey E (2009) Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J Neurosci* 29:5483–5493.
- Misyak JB, Christiansen MH (2012) Statistical learning and language: an individual differences study. *Lang Learn* 62:302–331.
- Mitchel AD, Weiss DJ (2011) Learning across senses: cross-modal effects in multisensory statistical learning. *J Exp Psychol Learn Mem Cogn* 37:1081–1091.
- Moller AR, Jannetta PJ, Jho HD (1994) Click-evoked responses from the cochlear nucleus: a study in human. *Electroencephalogr Clin Neurophysiol* 92:215–224.
- Moushegian G, Rupert AL, Stillman RD (1973) Laboratory note. Scalp-recorded early responses in man to frequencies in the speech range. *Electroencephalogr Clin Neurophysiol* 35:665–667.
- Muchnik C, Ari-Even Roth D, Othman-Jebara R, Putter-Katz H, Shabtai EL, Hildesheimer M (2004) Reduced medial olivocochlear bundle system function in children with auditory processing disorders. *Audiol Neurootol* 9:107–114.
- Musacchia G, Sams M, Skoe E, Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc Natl Acad Sci USA* 104:15894–15898.
- Nelken I (2004) Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol* 14:474–480.
- Nelken I, Ulanovsky N (2007) Mismatch negativity and stimulus-specific adaptation in animal models. *J Psychophysiol* 21:214–223.
- 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.
- Parbery-Clark A, Strait DL, Kraus N (2011) Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia* 49:3338–3345.
- Parbery-Clark A, Tierney A, Strait DL, Kraus N (2012) Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* 219:111–119.

- Peretz I, Saffran J, Schon D, Gosselin N (2012) Statistical learning of speech, not music, in congenital amusia. *Ann N Y Acad Sci* 1252:361–367.
- Perez-Gonzalez D, Malmierca MS, Covey E (2005) Novelty detector neurons in the mammalian auditory midbrain. *Eur J Neurosci* 22:2879–2885.
- Perrot X, Micheyl C, Khalfa S, Collet L (1999) Stronger bilateral efferent influences on cochlear biomechanical activity in musicians than in non-musicians. *Neurosci Lett* 262:167–170.
- Perruchet P, Pacton S (2006) Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cogn Sci* 10:233–238.
- Romberg AR, Saffran JR (2010) Statistical learning and language acquisition. *WIREs Cogn Sci* 1:906–914.
- Saffran JR (2003) Musical learning and language development. *Ann N Y Acad Sci* 999:397–401.
- Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–1928.
- Saffran JR, Johnson EK, Aslin RN, Newport EL (1999) Statistical learning of tone sequences by human infants and adults. *Cognition* 70:27–52.
- Saffran JR, Senghas A, Trueswell JC (2001) The acquisition of language by children. *Proc Natl Acad Sci USA* 98:12874–12875.
- Saffran JR, Hauser M, Seibel R, Kapfhammer J, Tsao F, Cushman F (2008) Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* 107:479–500.
- Schon D, Francois C (2011) Musical expertise and statistical learning of musical and linguistic structures. *Front Psychol* 2:167.
- Scott-Van Zeeland AA, McNealy K, Wang AT, Sigman M, Bookheimer SY, Dapretto M (2010) No neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders. *Biol Psychiatry* 15:345–351.
- Shook A, Marian V, Bartolotti J, Schroeder SR (2013) Musical experience influences statistical learning of a novel language. *Am J Psychol* 126:95–104.
- Skoe E, Kraus N (2010a) Auditory brain stem response to complex sounds: a tutorial. *Ear Hear* 31:302–324.
- Skoe E, Kraus N (2010b) Hearing it again and again: on-line subcortical plasticity in humans. *PLoS One* 5:e13645.
- Skoe E, Kraus N (2012) A little goes a long way: how the adult brain is shaped by musical training in childhood. *J Neurosci* 32:11507–11510.
- Slabu L, Grimm S, Escera C (2012) Novelty detection in the human auditory brainstem. *J Neurosci* 32:1447–1452.
- Smith JC, Marsh JT, Brown WS (1975) Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalogr Clin Neurophysiol* 39:465–472.
- Sokoloff L (1977) Relation between physiological function and energy metabolism in the central nervous system. *J Neurochem* 29:13–26.
- Song JH, Skoe E, Wong PC, Kraus N (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. *J Cogn Neurosci* 20:1892–1902.
- Song JH, Nicol T, Kraus N (2011a) Test-retest reliability of the speech-evoked auditory brainstem response. *Clin Neurophysiol* 122:346–355.
- Song JH, Skoe E, Banai K, Kraus N (2011b) Perception of speech in noise: neural correlates. *J Cogn Neurosci* 23:2268–2279.
- Song JH, Skoe E, Banai K, Kraus N (2012) Training to improve hearing speech in noise: biological mechanisms. *Cereb Cortex* 22:1180–1190.
- Soros P, Stanton SG (2012) On variability and genes: inter-individual differences in auditory brain function. *Front Hum Neurosci* 6:150.
- Soveri A, Rodriguez-Fornells A, Laine M (2011) Is there a relationship between language switching and executive functions in bilingualism? Introducing a within group analysis approach. *Front Psychol* 2:183.
- Srinivasan S, Keil A, Stratis K, Woodruff Carr KL, Smith DW (2012) Effects of cross-modal selective attention on the sensory periphery: cochlear sensitivity is altered by selective attention. *Neuroscience* 223:325–332.
- Strait DL, Hornickel J, Kraus N (2011) Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behav Brain Funct* 7:44.
- Suga N, Xiao Z, Ma X, Ji W (2002) Plasticity and corticofugal modulation for hearing in adult animals. *Neuron* 36:9–18.
- Tabachnick BG, Fidell LS (2007) Using multivariate statistics. Boston: Pearson/Allyn & Bacon.
- Teinonen T, Fellman V, Naatanen R, Alku P, Huotilainen M (2009) Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci* 10:21.
- Toro JM, Trobalon JB (2005) Statistical computations over a speech stream in a rodent. *Percept Psychophys* 67:867–875.
- Toro JM, Sinnott S, Soto-Faraco S (2005) Speech segmentation by statistical learning depends on attention. *Cognition* 97:B25–B34.
- Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK (2009) Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J Cogn Neurosci* 21:1934–1945.
- Tzounopoulos T, Kraus N (2009) Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron* 62:463–469.
- Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6:391–398.
- Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. *J Neurosci* 24:10440–10453.
- Warrier CM, Abrams DA, Nicol TG, Kraus N (2011) Inferior colliculus contributions to phase encoding of stop consonants in an animal model. *Hear Res* 282:108–118.
- Wen B, Wang GI, Dean I, Delgutte B (2009) Dynamic range adaptation to sound level statistics in the auditory nerve. *J Neurosci* 29:13797–13808.
- Winer JA (2006) Decoding the auditory corticofugal systems. *Hear Res* 212:1–8.
- Wong PC, Skoe E, Russo NM, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* 10:420–422.
- Yan J, Zhang Y, Ehret G (2005) Corticofugal shaping of frequency tuning curves in the central nucleus of the inferior colliculus of mice. *J Neurophysiol* 93:71–83.
- Zhao W, Dhar S (2011) Fast and slow effects of medial olivocochlear efferent activity in humans. *PLoS One* 6:e18725.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.neuroscience.2013.03.009>.

(Accepted 12 March 2013)
(Available online 19 March 2013)