

Intertrial auditory neural stability supports beat synchronization in preschoolers



Kali Woodruff Carr^{a,b}, Adam Tierney^{a,b,1}, Travis White-Schwoch^{a,b}, Nina Kraus^{a,b,c,d,*}

^a Auditory Neuroscience Laboratory, Northwestern University, 2240 Campus Drive, Evanston, IL 60208 USA

^b Department of Communication Sciences, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA

^c Department of Neurobiology & Physiology, Northwestern University, 2205 Tech Drive, Evanston, IL 60208, USA

^d Department of Otolaryngology, Northwestern University, 675 North St Clair, Chicago, IL, USA

ARTICLE INFO

Article history:

Received 4 April 2015

Received in revised form 17 October 2015

Accepted 3 December 2015

Available online 12 December 2015

Keywords:

Auditory processing

Sensorimotor beat synchronization

FFR

Speech processing

Children

ABSTRACT

The ability to synchronize motor movements along with an auditory beat places stringent demands on the temporal processing and sensorimotor integration capabilities of the nervous system. Links between millisecond-level precision of auditory processing and the consistency of sensorimotor beat synchronization implicate fine auditory neural timing as a mechanism for forming stable internal representations of, and behavioral reactions to, sound. Here, for the first time, we demonstrate a systematic relationship between consistency of beat synchronization and trial-by-trial stability of subcortical speech processing in preschoolers (ages 3 and 4 years old). We conclude that beat synchronization might provide a useful window into millisecond-level neural precision for encoding sound in early childhood, when speech processing is especially important for language acquisition and development.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Learning requires ongoing and repeated associations between stimuli and their implications (Hebb, 1949). Across modalities, stable perceptual representation of stimuli from one experience to the next allows for the emergence of coherent internal representations, while neural instability characterizes individuals with clinical disorders (e.g., autism, dyslexia, attention deficit, and schizophrenia; cf. Dinstejn et al., 2015). This neural stability comes into play when an individual interacts with sound; unstable processing in the auditory system has been observed in individuals with language impairments (Ahissar et al., 2000; Evans et al., 2009; Hornickel et al., 2009; Hornickel and Kraus, 2013). Stable neural processing of structured temporal patterns may be particularly crucial for language acquisition and development: anticipation and detection of the timing of auditory events allows a listener

to tune in to and predict important acoustic features (Large and Jones, 1999; McAuley et al., 2006) necessary for distinguishing and reproducing syllabic segments, prosodic cues, and the rapidly changing acoustic features that differentiate meaningful segments of speech (Baruch and Drake, 1997; Bertoncini and Mehler, 1981; Eimas et al., 1971; Ramus, 2000; Saffran et al., 1996; Tallal, 1980). Thus, stable neural coding of speech timing during early childhood – a period of intense, rapid learning and an age critical for mapping meaning to auditory input (Kuhl et al., 1992; Ruben, 1997) – could be acutely important for language learning.

Such precision and stability of speech processing in the human auditory system can be captured by examining the intertrial stability of the frequency following response (FFR) to a consonant–vowel speech syllable, a noninvasive measure of subcortical neural encoding, which records the summation of synchronous electrical activity originating from the auditory midbrain. The FFR reflects both temporal and spectral physiognomies of auditory stimuli with fine resolution (Skoe and Kraus, 2010). A high degree of intertrial stability of the FFR is associated with good reading ability in children, while intertrial variability has been observed to correlate with poorer reading skills (Hornickel and Kraus, 2013).

Beat synchronization, or entraining a motor movement to an auditory beat, has proved an intriguing tool for assessing sensorimotor timing (reviewed systematically in Repp, 2005; Repp and Su, 2013), and has been linked to the aforementioned intertrial

* Corresponding author at: Northwestern University, Auditory Neuroscience Laboratory, 2240 Campus Dr., Frances Searle Building Rm 2-233, Evanston, IL 60208, USA. Tel.: +1 847 491 3181.

E-mail addresses: kali@u.northwestern.edu (K. Woodruff Carr), adamtierney@gmail.com (A. Tierney), whiteschwoch@northwestern.edu (T. White-Schwoch), nkraus@northwestern.edu (N. Kraus).

¹ Present address: Department of Psychological Sciences, Birkbeck University of London, Malet Street, London WC1E 7HX, UK.

neural stability of the FFR to speech in adolescents (Tierney and Kraus, 2013a,b). Synchronizing to an external beat likely relies on temporal fidelity for auditory perceptual coding, motor production, and coupling between auditory and motor systems (Sowiński and Dalla Bella, 2013). The auditory midbrain appears to be particularly important for beat synchronization, as it is uniquely positioned to play an integrating role: inferior colliculus receives ascending connections from subcortical auditory structures and motor areas (e.g., basal ganglia; Coleman and Clerici, 1987; Kudo and Niimi, 1980) and descending input from cortex (Bajo et al., 2010), in addition to sending information to cerebellum (another area crucial for fine motor control) via dorsolateral pontine nuclei (Hashikawa, 1983; Mower et al., 1979; Saint Marie, 1996).

Examining links between sound processing in auditory midbrain and beat synchronization could inform our knowledge of the biology responsible for transformation of perceived periodicity in auditory stimuli to motor output. Tierney and Kraus (2013a,b) have established a systematic relationship between intertrial stability of subcortical speech encoding and the consistency of beat synchronization in adolescents, proposing auditory system stability as a biological mechanism common to speech processing and beat-keeping. In young children, the ability to synchronize to a beat relates to precision of subcortical speech-envelope tracking, as well as pre-literacy skills thought to predict future reading skills such as phonological awareness and auditory short-term memory (Woodruff Carr et al., 2014).

Here, we expand upon previous work (Woodruff Carr et al., 2014) to explore the neurophysiology underlying individual differences in preschoolers who are able to synchronize motor movements to isochronous beats at prosodic stress rates. We predicted more consistent auditory-motor timing, as revealed through beat synchronization, would relate to higher levels of intertrial neural stability for processing speech syllables. Furthermore, our previous work identifying links between beat synchronization and neural envelope tracking precision led us to hypothesize that stability of low-frequency encoding in particular would relate to beat synchronization, because the envelope measure is filtered to capture low-frequency modulations. Our findings suggest that stability of auditory neural encoding may be an important foundation for sensorimotor integration in preschoolers. Furthermore, beat synchronization may serve as a useful behavioral tool for assessing developmental auditory neural function in young children.

2. Methods

2.1. Participants

Twenty-five children (15 females), ages three and four years old ($M=4.34$, $SD=0.56$), were recruited from the Chicago area. No child had a history of a neurologic condition, a diagnosis of autism spectrum disorder, a family history of language learning disorders, or second language exposure. All children had normal age-adjusted scaled scores for both verbal ($M=13.48$, $SD=3.24$) and nonverbal ($M=13.52$, $SD=2.84$) intelligence estimated with the Wechsler Preschool and Primary Scale of Intelligence, third edition (WPPSI; Pearson/PsychCorp, San Antonio, TX), passed a screening of peripheral auditory function (normal otoscopy, Type A tympanograms, and distortion product otoacoustic emissions at least 6 dB above the noise floor from 0.5 to 4 kHz) and had normal click-evoked auditory brainstem responses (identifiable wave V latency of <5.8 ms). Informed consent and assent was obtained from legal guardians and children, respectively, in accordance with procedures approved by the Northwestern University Institutional Review Board and children were monetarily compensated for their participation.

2.2. Beat synchronization

Our beat synchronization task was based on Kirshner and Tomasello's (2009) social drumming entrainment paradigm for preschoolers. The experimenter sat across from the child with two conga drums between them, one for the experimenter and one for the participant. Each conga had a Pulse Percussion DR-1 drum trigger attached to the underside of its drumhead to record the drum hits and convert vibrations into voltage in real time with no delay. The experimenter covertly listened and drummed to an isochronous beat presented through an in-ear headphone and encouraged the child to imitate and drum along with the experimenter. Auditory stimuli and drum hits of both the experimenter and participant were recorded as two separate two-channel recordings in Audacity version 2.0.5. Four trials were performed: two trials at 2.5 Hz followed by two trials at 1.67 Hz. Each trial was 20 s in duration, resulting in 50 isochronous drum hits for the 2.5 Hz trials and 33 drum hits for the 1.67 Hz trials. The use of two rates allowed for the assessment of general synchronization ability as opposed to synchronization to a specific rate, reducing the potential bias of an individual's preferred tempo.

2.2.1. Data processing

Synchronization data were processed using software developed in house in MATLAB (Mathworks, Inc., Natick, MA). Due to the high intersubject variability in intensity and rapidity of drumming, drum hits for the experimenter and participant were detected by setting an amplitude threshold and a refractory period on a participant-by-participant basis. The first point at which the signal exceeded the amplitude threshold was marked as a hit, immediately followed by a refractory period during which the program did not mark peaks (to ensure multiple points were not marked for each hit). Accuracy of automated hit detection was checked manually to ensure onsets were correctly marked for each hit.

2.2.2. Data analysis

Beat synchronization ability was assessed using circular statistics (Fisher, 1993), a useful tool for assessing sensorimotor synchronization when there is not one-to-one correspondence of hits and pacing stimuli (Kirshner and Tomasello, 2009; Sowiński and Dalla Bella, 2013; Fujii and Schlaug, 2013), as is the case with this dataset: children frequently missed hits or did not synchronize continuously over a session. Each drum hit was assigned a relative phase angle (θ or "accuracy") in degrees by subtracting the hit time from the nearest experimenter's hit, dividing the result by the ISI, and multiplying by 360. The mean of all vectors resulted in R , a measurement of the extent to which participants tended to maintain a constant temporal relationship between their drum hits and the experimenter's. We define beat synchronization "consistency" as the average vector length across each of the two trials and across both rates. These two measures seem largely independent (correlation between consistency and accuracy: $r_{(25)} = -0.275$, $p = 0.183$). Recent work has shown the ability to synchronize to an external beat is still developing during this age (Kirshner and Tomasello, 2009; Woodruff Carr et al., 2014). Therefore, Rayleigh's test was applied to the set of all vectors produced in the two trials for a given rate to determine whether a participant was successfully synchronizing (the null hypothesis of this test is that the distribution of data points occur randomly in time near or away from the pacing stimuli onsets, indicative of chance performance; $p > 0.05$). The two trials at each rate were combined to compute a Rayleigh's p -value for each rate. If a child's Rayleigh's test resulted in a $p < 0.05$ at both rates, the child was included in analyses. Our previous work (Woodruff Carr et al., 2014) investigated group differences in neural processing between children who could ($p < 0.05$) and could not ($p > 0.05$) synchronize; the current investigation expands upon this

work by investigating neural correlates of synchronization ability within an expanded group of successful synchronizers.

2.3. Neurophysiology

2.3.1. Stimuli

Frequency following responses (FFRs) were elicited to 170 ms six-formant stop consonant–vowel speech syllables [ba], [da], and [ga] at 80 dB SPL at a 4.35 Hz sampling rate. Syllables were synthesized at 20 kHz with voicing onset at 5 ms, a 50 ms formant transition, and a 120 ms steady state vowel using a Klatt-based formant synthesizer (Klatt, 1980). Stimuli differed only in onset frequency of the second formant (F_2 : [ba] = 900 Hz, [da] = 1700 Hz, [ga] = 2480 Hz), shifting to 1240 Hz for the steady-state vowel. Over the transition period for all stimuli the first and third formants were dynamic ($F_1 = 400$ –720 Hz, $F_3 = 2580$ –2500 Hz) with the fundamental frequency, fourth, fifth, and sixth formants constant ($F_0 = 100$, $F_4 = 3300$, $F_5 = 3750$, and $F_6 = 4900$ Hz). All stimuli were presented in alternating polarities (stimulus waveform was inverted 180°) with an interstimulus interval of 81 ms controlled by E-Prime version 2.0 (Psychology Software Tools, Inc., Sharpsburg, PA). Each stimulus was presented 4200 times, with presentation order randomized for each participant.

Additionally, auditory brainstem responses (ABRs) were collected to a 100 μ s square wave click stimulus presented in rarefaction at 80 dB SPL at a 31.3 Hz sampling rate. 2000 sweeps were presented.

2.3.2. Recording parameters

Stimuli were presented monaurally to the right ear through an insert earphone (ER-3, Etymotic Research, Elk Grove Village, IL) while the participant sat watching a movie of their choice in a sound-attenuated booth (IAC Acoustics, Bronx, NY). The left ear remained unblocked so that the soundtrack of the movie (<40 dB SPL) was audible but not loud enough to mask presented stimuli. FFRs were collected using BioSEMI Active2 with an ActiABR module recorded in LabView 2.0 (National Instruments, Austin, TX). Responses were digitized at 16.384 kHz and collected with an online bandpass filter from 100 to 3000 Hz (20 dB/decade roll-off). The active electrode was placed at the vertex (Cz), with references on each earlobe. Grounding electrodes CMS and DRL were placed on the forehead at Fp1 and Fp2, respectively. Only ipsilateral-referenced (Cz–Right earlobe) responses were used in analyses. Offset voltage was <50 mV for all electrodes.

2.3.3. Data reduction and processing

FFRs to speech stimuli were offline amplified in the frequency domain 20 dB per decade for 3 decades below 100 Hz, bandpass filtered from 70 to 2000 Hz (12 dB/octave roll-off), segmented into epochs with an interval of –40 to 210 ms (in relation to the stimulus onset), and baseline-corrected to the pre-stimulus period. Responses exceeding $\pm 35 \mu$ V were rejected as artifacts and remaining sweeps were averaged. Final responses to each syllable comprised 2000 artifact-free sweeps of each polarity, and responses to the two polarities were added to accentuate the response to the speech envelope (Aiken and Picton, 2008) and limit the influence of cochlear microphonic and stimulus artifact (Campbell et al., 2012). Data reduction occurred in MATLAB using custom scripts.

2.3.4. Data analysis

The FFR faithfully reproduces spectrotemporal stimulus features due to the inferior colliculus' ability to encode fine timing information (Liu et al., 2006; Warrier et al., 2011). Intertrial stability of the ABR was assessed using a procedure previously reported

(Hornickel and Kraus, 2013; Tierney and Kraus, 2013a,b). To calculate the stability of a participant's response to the speech stimuli, 2000 of 4000 trials were randomly selected and averaged. The remaining 2000 trials were also averaged. The two sub-averaged waveforms were then correlated over the 0 to 170 ms range to determine their similarity. These steps were repeated 300 different times, each with different random samplings of trials, and the 300 correlation values were averaged to generate a final measure of intertrial neural response stability. Neural stability for encoding the click stimulus was assessed over 0 to 11.95 ms of the response using the same methodology: the two sub-averaged waveforms (1000 randomly selected, averaged trials) were correlated to determine similarity 100 times, each with different random samplings of trials, and the correlation values were averaged. To specifically examine the timing variability in frequency encoding of the 170 ms speech-evoked FFR, intertrial phase-locking (Tierney and Kraus, 2013a,b) was calculated in 20 Hz windows surrounding the fundamental frequency of the stimulus (100 Hz) and its harmonics up to 1000 Hz. Time–frequency spectrum was calculated using a short-time fast Fourier transform that resulted in a matrix containing two measures for each time x frequency point: a vector length (the extent to which each frequency is encoded in the FFR) and phase (the timing of that frequency). To specifically analyze the timing variability, each vector was transformed into a unit vector. For each frequency, the 4000 vectors were averaged and the length of the resulting vector was calculated as a measure of the consistency of phase across trials. Low frequency phase-locking was captured by averaging across the vectors for 100, 200, 300, and 400 Hz, while high frequency phase-locking was computed as a mean of the vectors for harmonics 500–1000 Hz. Intertrial neural stability and phase-locking correlation values were Fisher and log transformed to conform to the expectations of a linear model (normality and sphericity). Data analysis occurred in MATLAB.

2.4. Statistical analysis

Pearson correlations were used to compare intertrial neural stability and phase-locking across stimuli (mean of [ba] + [da] + [ga]) to beat synchronization consistency. Phase-locking was averaged across low-frequency harmonics (100–400 Hz) and high-frequency harmonics (500–1000 Hz). Hierarchical two-step linear regressions were employed to determine how neural stability and phase-locking predicted variance in beat synchronization over and above demographic factors. Statistics were computed using SPSS (SPSS, Inc., Chicago, IL).

3. Results

3.1. Intertrial neural stability

We found a systematic relationship between beat synchronization consistency (a measure of the extent to which participants were able to maintain a constant temporal relationship between their drum hits and the pacing stimulus events) and intertrial neural stability, a measure of trial-by-trial variability in auditory midbrain. Those who more consistently synchronized had higher intertrial neural stability (composite of [ba], [da], and [ga]: $r_{(25)} = 0.554$, $p = 0.004$; Fig. 1a). There was no relationship between beat synchronization accuracy and intertrial neural stability ($r_{(25)} = 0.092$, $p = 0.663$). This relationship was specific to periodic stimuli such as speech: intertrial stability of the click-evoked auditory brainstem response (ABR) did not correlate with beat synchronization ($r_{(25)} = 0.048$, $p = 0.819$).

To determine the unique predictability of beat synchronization from neural response stability, we performed a hierarchical

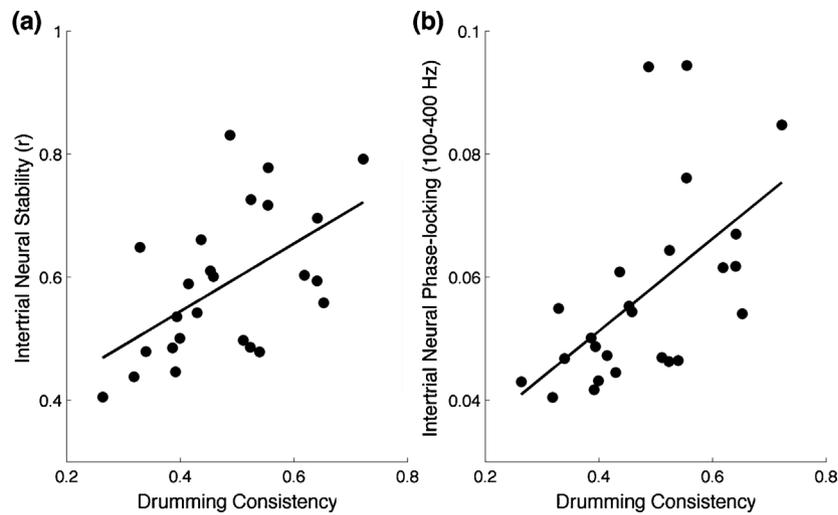


Fig. 1. The ability to consistently time motor movements to an auditory beat relates to (a) intertrial neural stability ($r_{(25)} = 0.554, p = 0.004$) and (b) low-frequency (100–400 Hz) intertrial phase-locking ($r_{(25)} = 0.609, p = 0.001$) of neural responses to sound.

Table 1

Hierarchical two-step linear regression results: (A) demographics alone do not significantly explain variability in beat synchronization, but the addition of intertrial neural stability significantly improves the model, explaining 23.3% ($p = 0.010$) of beat synchronization variance over and above age, sex, and intelligence. Combined with demographic measures, this model predicts 45.2% of variance in consistency of beat entrainment ($p = 0.032$). (B) The addition of neural phase-locking significantly improves the model, explaining 26.8% ($p = 0.005$) of beat synchronization variance over and above age, sex, and intelligence. Combined with demographic measures, this model predicts 48.7% of variance in consistency of beat entrainment ($p = 0.018$). * $p < 0.05$, ** $p < 0.01$.

Predictor	Regression (A)		Regression (B)	
	ΔR^2	β	ΔR^2	β
Step 1	0.219		0.219	
Age		-0.082		-0.082
Sex		0.017		0.017
Verbal intelligence		0.486*		0.486*
Nonverbal intelligence		-0.175		-0.175
Step 2	0.233**		0.268**	
Age		-0.105		-0.041
Sex		0.139		0.155
Verbal intelligence		0.371		0.304
Nonverbal intelligence		-0.131		-0.162
Intertrial neural stability		0.512*		-
Neural phase-locking (100–400 Hz)		-		0.568**
Total R^2	0.452*		0.487*	

linear regression. On the first step the independent variables sex, age, and verbal and non-verbal intelligence scores failed to predict unique variance in beat synchronization ($R^2 = 0.219, F_{(4,20)} = 1.339, p = 0.270$). On the second step we added the independent variable of intertrial neural stability. This step improved the model, explaining an additional 23.3% of beat synchronization variance ($F_{(1,19)} = 8.067, p = 0.010$), over and above age, sex, IQ, and vocabulary. Our overall model accounts for 45.2% ($F_{(5,19)} = 3.129, p = 0.032$) of variance in beat synchronization consistency (see Table 1A for full regression results). Assuming an alpha level of $p < 0.05$, a post hoc power analysis revealed a large effect size (Cohen's $f^2 = 0.425$) and sufficient power (0.816).

3.2. Intertrial neural phase-locking

To investigate intertrial stability at specific frequencies, a measure of intertrial phase-locking to the fundamental frequency (F_0)

and its harmonics was computed. We again discovered a systematic relationship between consistency of beat synchronization and phase-locking, specifically at lower frequencies: F_0 and its first three harmonics (composite of [ba], [da], and [ga] at 100–400 Hz: $r_{(25)} = 0.609, p = 0.001$; Fig. 1b), as was reported for adolescents in previous work (Tierney and Kraus, 2013a,b). Fig. 2 further illustrates this relationship between intertrial neural phase-locking and beat synchronization consistency. Investigation of higher frequency phase-locking (500–1000 Hz) revealed no link to beat synchronization consistency ($r_{(25)} = 0.330, p = 0.108$; see Table 2 for correlations between phase-locking to each frequency and beat synchronization consistency), suggesting a frequency specificity of this effect. These relationships were not observed for synchronization accuracy (100–400 Hz: $r_{(25)} = 0.039, p = 0.853$; 500–1000 Hz: $r_{(25)} = -0.009, p = 0.965$).

Again, hierarchical linear regression modeling was performed to control for demographic factors. On the first step the independent variables sex, age, and verbal and non-verbal intelligence did not predict unique variance of beat synchronization, but with the addition of the independent variable of low-frequency phase-locking the model was improved, explaining an additional 26.8% of beat synchronization consistency variance ($F_{(1,19)} = 9.927, p = 0.005$), over and above demographic factors. The overall model accounts for 48.7% ($F_{(5,19)} = 3.605, p = 0.018$) of variance in consistency of beat synchronization. (See Table 1B for full regression results.) Assuming an alpha level of $p < 0.05$, a post hoc power analysis revealed a large effect size (Cohen's $f^2 = 0.522$) and sufficient power (0.887). This effect was specific to lower frequencies (100–400 Hz): incorporating high frequency (500–1000 Hz) phase-locking into the regression model in a third step did not improve its fit ($\Delta R^2 < 0.001, p = 0.943$).

4. Discussion

These results suggest that successful beat synchronization in young children relies in part on stable temporal encoding in the auditory system. Less variability when encoding sound may allow for more regularly-timed motor reactions. We propose trial-by-trial neural stability supports the developmental process of coordinating auditory-motor beat synchronization in young children. The present findings are in line with previous studies linking the ability to tap consistently to a beat and stability of subcortical sound processing (Tierney and Kraus, 2013a,b), and provide a unique developmental perspective for a neural timing metric that

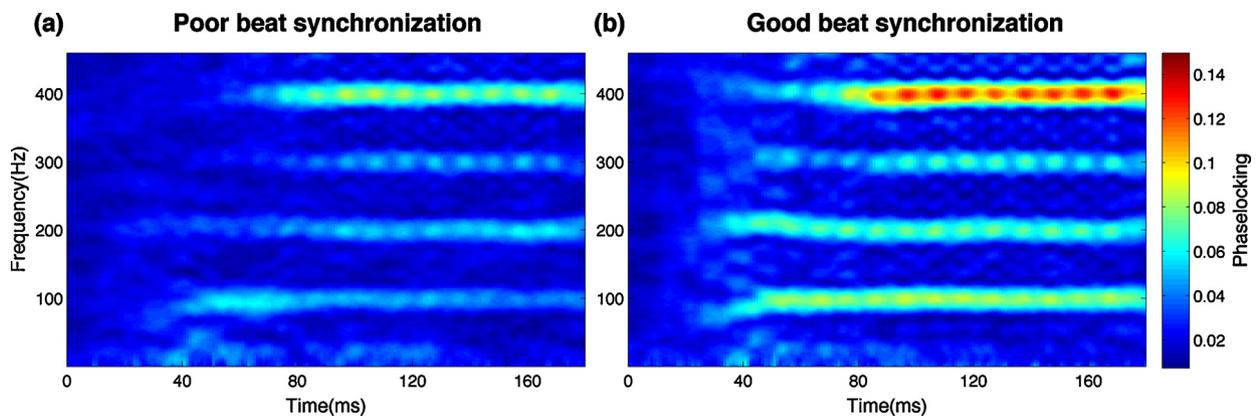


Fig. 2. To further illustrate the robust relationship between intertrial neural phase-locking and beat synchronization, participants were dichotomized as relatively (a) poor ($N = 13$) or (b) good ($N = 12$) synchronizers based on a median split according to their beat synchronization consistency. The good beat synchronization group's phase-locking power to the stimulus [da] is more robust for the fundamental frequency (100 Hz) and its harmonics (at 200, 300, and 400 Hz; $F_{(1,23)} = 12.967$, $p = 0.002$).

Table 2

Pearson correlation r -values for intertrial phase-locking at the fundamental frequency (F_0) and its subsequent harmonics (H_2 – H_{10}) with beat synchronization consistency. * $p < 0.05$, ** $p < 0.01$.

Phase-locking frequency	Beat synchronization consistency
F_0	0.552**
H_2	0.421*
H_3	0.510**
H_4	0.462*
H_5	0.241
H_6	0.232
H_7	0.318
H_8	0.227
H_9	0.345
H_{10}	0.273

underlies literacy skills (Hornickel and Kraus, 2013; Tierney and Kraus, 2013a,b). This mechanism appears to be in place at an early age, during the development of many important language skills, and prior to explicit reading instruction.

During beat synchronization, perception and production systems must interact synchronously for optimal performance: the brain extracts and estimates stimulus periodicity as well as assesses discrepancies between the stimulus and one's own motor output (Rauschecker, 2011). The subcortical nervous system works to detect sub-second differences in intervals (basal ganglia) and integrate this performance feedback across modalities (through connections from the dorsal cochlear nucleus to the cerebellum) to make subtle timing adjustments (Merchant et al., 2008), resulting in error-correction of asynchronies that does not always necessitate conscious effort (Ito, 2008; Repp, 2000; Schwartz and Kotz, 2013). Although the influence of motor variability during beat synchronization cannot be ruled out, we believe this is not a main factor influencing our results. Other work with children this age reports no correlation between motor variability in spontaneous tapping and synchronization tasks (Drake et al., 2000), suggesting that variability in synchronization performance in young children is primarily driven by factors other than motor variability. We suggest our neural stability measure captures auditory-motor integration, and future work is needed to parse the influence of these separate, but connected, systems.

Our results suggest coherence of temporal encoding across timescales. In particular, we demonstrate rapid intertrial neural stability for encoding frequency information from 100 to 400 Hz (for speech stimuli presented at 4.35 Hz) relates to consistency of beat synchronization (to rates approximating speech syllables, at 1.67 and 2.5 Hz; see Table 3). This connection between millisecond-level timing in the auditory midbrain and coordination of motor

movements to synchronize at much slower rates may be a function of hierarchical temporal scaffolding, with incredibly fast neural fidelity (i.e., intertrial stability of the FFR for dynamic formant transitions and periodic vowels) acting as temporal subdivisions to support sensorimotor synchrony (i.e., beat synchronization consistency) at slower rates. This finding is coherent with previous work demonstrating concomitance between beat synchronization and low-frequency temporal encoding precision: correlations were observed between beat synchronization consistency and subcortical envelope tracking, but not for broadband stimulus encoding (Woodruff Carr et al., 2014). Together, we suggest the ability to tune in to and exploit slow modulations of spectral information emerges first developmentally, supporting more stable trial-by-trial neural encoding.

The FFR is generated by a summation of simultaneous, synchronously-firing neurons throughout subcortical auditory nuclei; therefore intertrial variability of an FFR may result from a number of circumstances: a failure of eighth nerves to synchronize (e.g., auditory neuropathy), greater receptor adaptation or fatigue, and/or slower recovery from firing (i.e., prolonged refractory periods) (Don et al., 1977; Starr et al., 2003; Schaette et al., 2005). It is difficult to pinpoint the cause of this jitter, but future work using intracranial recordings is necessary to determine its source. If reliable animal models for beat synchronization are discovered (Cook et al., 2013; Hasegawa et al., 2011; Hattori et al., 2013; Large and Gray, 2015; Patel et al., 2009), it may be possible to explore local temporal jitter within inferior colliculus, a primary generator of the FFR, and how this relates to beat synchronization abilities.

The ability to consistently perceive and anticipate time intervals in sound streams may explain previously-observed links between auditory-motor synchronization and phonological processing: if input to the auditory system is not coherent from one experience to the next, this could hinder the developmental of a refined phonemic inventory. Increased neural variability would make the process of learning the correct probabilities and statistics of acoustic events challenging, and individuals with poor neural stability could exhibit difficulties in predicting their environment. In the case of autism, individuals with greater neural noise also exhibit heightened sensitivity to details at the consequence of an impaired ability to integrate details into gestalt percepts (Dinstein et al., 2015). Neural instability might be responsible for some of the deficits exhibited by children with language difficulties who struggle to process timing information in speech, through the process of stochastic resonance (McDonnell and Abbott, 2009). Stochastic resonance is a phenomenon where a signal normally too weak to be detected is boosted by noise. This may bias children with

Table 3

Pearson correlation r -values for beat synchronization consistency at each rate and the average of the two rates with neural stability measures. * $p < 0.05$, ** $p < 0.01$.

	Beat synchronization consistency		
	2.5 Hz	1.67 Hz	Average
Intertrial neural stability	0.444*	0.425*	0.544**
Neural phase-locking (100–400 Hz)	0.495*	0.460*	0.609**
Neural phase-locking (500–1000 Hz)	0.207	0.310	0.330

autism to focus on details rather than attempt to integrate them, and could also explain a pattern observed in the auditory domain for children with dyslexia. These children with auditory-based learning disorders exhibit an allophonic mode of speech perception, demonstrating higher sensitivity to irrelevant phonemic distinctions (Serniclaes et al., 2004). Supporting this idea, greater variability in auditory-neurophysiological responses elicited by speech have been reported in poor readers (Hornickel and Kraus, 2013; White-Schwoch et al., 2015) and animal models of dyslexia (Centanni et al., 2013).

In light of our current results, we suggest more stable trial-by-trial encoding of low-frequency (100 to 400 Hz) spectrotemporal acoustic features supports stable internal representations of sounds imperative for language learning. This stability of sound encoding might eventually bootstrap phonological development through cognitive systems that are engaged during listening and learning (cognitive-sensory coupling) such as attention and working memory (cf. Kraus and White-Schwoch, 2015), subsequently facilitating reading acquisition. Additionally, individuals diagnosed with speech and language impairments such as specific language impairment and dyslexia are less accurate than age- and language-matched controls at synchronizing to prosodic stress-rate tempi (Corriveau and Goswami, 2009; Thomson et al., 2006; Thomson and Goswami, 2008), suggesting that unstable neural responses may contribute to poor auditory processing important for both beat synchronization and development of literacy skills.

Although the present work does not explicitly relate these metrics to language proficiency, it does reveal a relationship between subcortical speech processing and auditory-motor synchronization at a prosodic rate, both metrics that independently relate to language competency (Hornickel and Kraus, 2013; Tierney and Kraus, 2013a,b; Thomson and Goswami, 2008; White-Schwoch et al., 2015; Woodruff Carr et al., 2014). Moreover, we observed this relationship between beat synchronization consistency and auditory neural stability only in response to periodic speech-like sounds and not acoustically-simple click stimuli (cf. Hornickel and Kraus, 2013), which we believe provides further evidence that the degree of neural fidelity with which an individual is able to resolve dynamic stimuli such as speech is intimately tied to beat synchronization.

5. Conclusion

In summary, we provide biological evidence of relationships between auditory-motor beat synchronization and intertrial neural stability for encoding speech sounds, establishing these links for the first time in emergent readers. We suggest that stable neural responses to sound may be integral to the emergence of sensorimotor synchronization skills. These findings illuminate a potential underlying neural mechanism that links the ability to synchronize and development of phonological processing during preschool years, as observed in previous work (Woodruff Carr et al., 2014). Encouragingly, neural synchrony can be improved through auditory training (Hornickel et al., 2012; Russo et al., 2005; Song et al., 2012). In addition, musical training has been shown to improve beat synchronization (Slater et al., 2013). Given

relationships observed in the present work, it seems possible that a rhythm-based music training program might prove successful for refining temporal processing and consequential language and literacy skills (Bhide et al., 2013; Kuhl, 2007). Future work exploring the potential therapeutic benefits of an auditory-motor integration task, such as drumming in time to music or speech, could inform preemptive treatment for children with substandard auditory processing before behavioral struggles manifest.

Author contributions

KWC, AT, and NK designed research; KWC performed research; AT contributed analytic techniques; KWC, AT, and TW-S analyzed data; KWC, AT, TW-S, and NK wrote the paper.

Acknowledgements

This work was supported by the National Institutes of Health (R01 HD069414 to NK, T32 DC009399 to KWC) and the Hugh Knowles Hearing Center of Northwestern University (to NK). We thank members of the Auditory Neuroscience Laboratory for their assistance with data collection as well as Trent Nicol, Jennifer Krizman, and Evan C. Davies for comments on an earlier draft of the manuscript.

References

- Ahissar, M., Protopapas, A., Reid, M., Merzenich, M.M., 2000. Auditory processing parallels reading abilities in adults. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6832–6837. <http://dx.doi.org/10.1073/pnas.97.12.6832>.
- Aiken, S.J., Picton, T.W., 2008. Envelope and spectral frequency-following responses to vowel sounds. *Hear. Res.* 245, 35–47. <http://dx.doi.org/10.1016/j.heares.2008.08.004>.
- Bajo, V.M., Nodal, F.R., Moore, D.R., King, A.J., 2010. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat. Neurosci.* 13, 253–260. <http://dx.doi.org/10.1038/nn.2466>.
- Baruch, C., Drake, C., 1997. Tempo discrimination in infants. *Infant Behav. Dev.* 20, 573–577. [http://dx.doi.org/10.1016/S0163-6383\(97\)90049-7](http://dx.doi.org/10.1016/S0163-6383(97)90049-7).
- Bertoncini, J., Mehler, J., 1981. Syllables as units in infant speech perception. *Infant Behav. Dev.* 4, 247–260. [http://dx.doi.org/10.1016/S0163-6383\(81\)80027-6](http://dx.doi.org/10.1016/S0163-6383(81)80027-6).
- Bhide, A., Power, A., Goswami, U., 2013. A rhythmic musical intervention for poor readers: a comparison of efficacy with a letter-based intervention: musical intervention for poor readers. *Mind Brain Educ.* 7, 113–123. <http://dx.doi.org/10.1111/mbe.12016>.
- Campbell, T., Kerlin, J.R., Bishop, C.W., Miller, L.M., 2012. Methods to eliminate stimulus transduction artifact from insert earphones during electroencephalography. *Ear Hear.* 33, 144–150. <http://dx.doi.org/10.1097/AUD.0b013e3182280353>.
- Centanni, T.M., Booker, A.B., Sloan, A.M., Chen, F., Maher, B.J., Carraway, R.S., Khodaparast, N., Rennaker, R., Loturco, J.J., Kilgard, M.P., 2013. Knockdown of the dyslexia-associated gene *Kiaa0319* impairs temporal responses to speech stimuli in rat primary auditory cortex. *Cereb. Cortex*. <http://dx.doi.org/10.1093/cercor/bht028>.
- Coleman, J.R., Clerici, W.J., 1987. Sources of projections to subdivisions of the inferior colliculus in the rat. *J. Comp. Neurol.* 262, 215–226. <http://dx.doi.org/10.1002/cne.902620204>.
- Cook, P., Rouse, A., Wilson, M., Reichmuth, C., 2013. A California sea lion (*Zalophus californianus*) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *J. Comp. Psychol.* 127, 412–427. <http://dx.doi.org/10.1037/a0032345>.
- Corriveau, K.H., Goswami, U., 2009. Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat. *Cortex* 45, 119–130. <http://dx.doi.org/10.1016/j.cortex.2007.09.008>.
- Dinstein, I., Heeger, D.J., Behrmann, M., 2015. Neural variability: friend or foe? *Trends Cogn. Sci.* 19, 322–328. <http://dx.doi.org/10.1016/j.tics.2015.04.005>.
- Don, M., Allen, A.R., Starr, A., 1977. Effect of click rate on the latency of auditory brain stem responses in humans. *Ann. Otol. Rhinol. Laryngol.* 86, 186–195.
- Drake, C., Jones, M.R., Baruch, C., 2000. The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition* 77, 251–288.
- Eimas, P.D., Siqueland, E.R., Jusczyk, P., Vigorito, J., 1971. Speech perception in Infants. *Science* 171, 303–306. <http://dx.doi.org/10.1126/science.171.3968.303>.
- Evans, J.L., Saffran, J.R., Robe-Torres, K., 2009. Statistical learning in children with specific language impairment. *J. Speech Lang. Hear. Res.*: JSLHR 52, 321–335. [http://dx.doi.org/10.1044/1092-4388\(2009\)07-0189](http://dx.doi.org/10.1044/1092-4388(2009)07-0189).
- Fisher, N.I., 1993. *Statistical Analysis of Circular Data*. Cambridge University Press, Cambridge.

- Fujii, S., Schlaug, G., 2013. The Harvard Beat Assessment Test (H-BAT): a battery for assessing beat perception and production and their dissociation. *Front. Hum. Neurosci.* 7, <http://dx.doi.org/10.3389/fnhum.2013.00771>.
- Hasegawa, A., Okanoya, K., Hasegawa, T., Seki, Y., 2011. Rhythmic synchronization tapping to an audio-visual metronome in budgerigars. *Sci. Rep.* 1, <http://dx.doi.org/10.1038/srep00120>.
- Hashikawa, T., 1983. The inferior colliculopontine neurons of the cat in relation to other collicular descending neurons. *J. Comp. Neurol.* 219, 241–249, <http://dx.doi.org/10.1002/cne.902190209>.
- Hattori, Y., Tomonaga, M., Matsuzawa, T., 2013. Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Sci. Rep.* 3, 1566, <http://dx.doi.org/10.1038/srep01566>.
- Hebb, D.O., 1949. *The Organization of Behavior*. Wiley, New York, NY.
- Hornickel, J., Kraus, N., 2013. Unstable representation of sound: a biological marker of dyslexia. *J. Neurosci.* 33, 3500–3504, <http://dx.doi.org/10.1523/JNEUROSCI.4205-12.2013>.
- Hornickel, J., Skoe, E., Nicol, T., Zecker, S., Kraus, N., 2009. Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13022–13027, <http://dx.doi.org/10.1073/pnas.0901123106>.
- Hornickel, J., Zecker, S.G., Bradlow, A.R., Kraus, N., 2012. Assistive listening devices drive neuroplasticity in children with dyslexia. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16731–16736, <http://dx.doi.org/10.1073/pnas.1206628109>.
- Ito, M., 2008. Control of mental activities by internal models in the cerebellum. *Nat. Rev. Neurosci.* 9, 304–313, <http://dx.doi.org/10.1038/nrn2332>.
- Kirschner, S., Tomasello, M., 2009. Joint drumming: social context facilitates synchronization in preschool children. *J. Exp. Child Psychol.* 102, 299–314, <http://dx.doi.org/10.1016/j.jecp.2008.07.005>.
- Klatt, D., 1980. Software for a cascade/parallel formant synthesizer. *J. Acoust. Soc. Am.* 67, 971–995.
- Kraus, N., White-Schwach, T., 2015. Unraveling the biology of auditory learning: a cognitive-sensorimotor-reward framework. *Trends Cogn. Sci.*, <http://dx.doi.org/10.1016/j.tics.2015.08.017>.
- Kudo, M., Niimi, K., 1980. Ascending projections of the inferior colliculus in the cat: an autoradiographic study. *J. Comp. Neurol.* 191, 545–556, <http://dx.doi.org/10.1002/cne.901910403>.
- Kuhl, P.K., 2007. Is speech learning “gated” by the social brain? *Dev. Sci.* 10, 110–120, <http://dx.doi.org/10.1111/j.1467-7687.2007.00572.x>.
- Kuhl, P., Williams, K., Lacerda, F., Stevens, K., Lindblom, B., 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608, <http://dx.doi.org/10.1126/science.1736364>.
- Large, E.W., Gray, P.M., 2015. Spontaneous tempo and rhythmic entrainment in a Bonobo (*Pan paniscus*). *J. Comp. Psychol.*, <http://dx.doi.org/10.1037/com000011>.
- Large, E.W., Jones, M.R., 1999. The dynamics of attending: how people track time-varying events. *Psychol. Rev.* 106, 119–159, <http://dx.doi.org/10.1037/0033-295X.106.1.119>.
- Liu, L.-F., Palmer, A.R., Wallace, M.N., 2006. Phase-locked responses to pure tones in the inferior colliculus. *J. Neurophysiol.* 95, 1926–1935, <http://dx.doi.org/10.1152/jn.00497.2005>.
- McAuley, J.D., Jones, M.R., Holub, S., Johnston, H.M., Miller, N.S., 2006. The time of our lives: life span development of timing and event tracking. *J. Exp. Psychol. Gen.* 135, 348–367, <http://dx.doi.org/10.1037/0096-3445.135.3.348>.
- McDonnell, M.D., Abbott, D., 2009. What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology. *PLoS Comput. Biol.* 5, e1000348, <http://dx.doi.org/10.1371/journal.pcbi.1000348>.
- Merchant, H., Zarco, W., Prado, L., 2008. Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *J. Neurophysiol.* 99, 939–949, <http://dx.doi.org/10.1152/jn.01225.2007>.
- Mower, G., Gibson, A., Glickstein, M., 1979. Tectopontine pathway in the cat: laminar distribution of cells of origin and visual properties of target cells in dorsolateral pontine nucleus. *J. Neurophysiol.* 42, 1–15.
- Patel, A.D., Iversen, J.R., Bregman, M.R., Schulz, I., 2009. Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* 19, 827–830, <http://dx.doi.org/10.1016/j.cub.2009.03.038>.
- Ramus, F., 2000. Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288, 349–351, <http://dx.doi.org/10.1126/science.288.5464.349>.
- Rauschecker, J.P., 2011. An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hear. Res.* 271, 16–25, <http://dx.doi.org/10.1016/j.heares.2010.09.001>.
- Repp, B.H., 2005. *Sensorimotor synchronization: a review of the tapping literature*. *Psychon. Bull. Rev.* 12, 969–992.
- Repp, B.H., 2000. Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychol. Res. Psychol. Forsch.* 63, 106–128, <http://dx.doi.org/10.1007/PL00008170>.
- Repp, B.H., Su, Y.-H., 2013. Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon. Bull. Rev.* 20, 403–452, <http://dx.doi.org/10.3758/s13423-012-0371-2>.
- Ruben, R.J., 1997. A time frame of critical/sensitive periods of language development. *Acta Otolaryngol. (Stockh.)* 117, 202–205.
- Russo, N.M., Nicol, T.G., Zecker, S.G., Hayes, E.A., Kraus, N., 2005. Auditory training improves neural timing in the human brainstem. *Behav. Brain Res.* 156, 95–103, <http://dx.doi.org/10.1016/j.bbr.2004.05.012>.
- Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. *Statistical learning by 8-month-old infants*. *Science* 274, 1926–1928.
- Saint Marie, R.L., 1996. Glutamatergic connections of the auditory midbrain: selective uptake and axonal transport of D-[³H]aspartate. *J. Comp. Neurol.* 373, 255–270, [http://dx.doi.org/10.1002/\(SICI\)1096-9861\(19960916\)373:2<AID-CNE8>3.0.CO;2-2](http://dx.doi.org/10.1002/(SICI)1096-9861(19960916)373:2<AID-CNE8>3.0.CO;2-2).
- Schaette, R., Gollisch, T., Herz, A.V.M., 2005. Spike-train variability of auditory neurons in vivo: dynamic responses follow predictions from constant stimuli. *J. Neurophysiol.* 93, 3270–3281, <http://dx.doi.org/10.1152/jn.00758.2004>.
- Schwartz, M., Kotz, S.A., 2013. A dual-pathway neural architecture for specific temporal prediction. *Neurosci. Biobehav. Rev.* 37, 2587–2596, <http://dx.doi.org/10.1016/j.neubiorev.2013.08.005>.
- Serniclaes, W., Heghe, S.V., Mousty, P., Carré, R., Sprenger-Charolles, L., 2004. Allophonic mode of speech perception in dyslexia. *J. Exp. Child Psychol.* 87, 336–361, <http://dx.doi.org/10.1016/j.jecp.2004.02.001>.
- Skoe, E., Kraus, N., 2010. Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* 31, 302–324, <http://dx.doi.org/10.1097/AUD.0b013e3181c8b272>.
- Slater, J., Tierney, A., Kraus, N., 2013. At-risk elementary school children with one year of classroom music instruction are better at keeping a beat. *PLoS ONE* 8, e77250, <http://dx.doi.org/10.1371/journal.pone.0077250>.
- Song, J.H., Skoe, E., Banai, K., Kraus, N., 2012. Training to improve hearing speech in noise: biological mechanisms. *Cereb. Cortex* 22, 1180–1190, <http://dx.doi.org/10.1093/cercor/bhr196> (N. Y. N 1991).
- Sowiński, J., Dalla Bella, S., 2013. Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia* 51, 1952–1963, <http://dx.doi.org/10.1016/j.neuropsychologia.2013.06.027>.
- Starr, A., Michalewski, H.J., Zeng, F.-G., Fujikawa-Brooks, S., Linthicum, F., Kim, C.S., Winnier, D., Keats, B., 2003. Pathology and physiology of auditory neuropathy with a novel mutation in the MPZ gene (Tyr145->Ser). *Brain J. Neurol.* 126, 1604–1619, <http://dx.doi.org/10.1093/brain/awg156>.
- Tallal, P., 1980. Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang.* 9, 182–198.
- Thomson, J.M., Fryer, B., Maltby, J., Goswami, U., 2006. Auditory and motor rhythm awareness in adults with dyslexia. *J. Res. Read.* 29, 334–348, <http://dx.doi.org/10.1111/j.1467-9817.2006.00312.x>.
- Thomson, J.M., Goswami, U., 2008. Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. *J. Physiol. (Paris)* 102, 120–129, <http://dx.doi.org/10.1016/j.jphysparis.2008.03.007>.
- Tierney, A., Kraus, N., 2013a. The ability to move to a beat is linked to the consistency of neural responses to sound. *J. Neurosci.* 33, 14981–14988, <http://dx.doi.org/10.1523/JNEUROSCI.0612-13.2013>.
- Tierney, A.T., Kraus, N., 2013b. The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain Lang.* 124, 225–231, <http://dx.doi.org/10.1016/j.bandl.2012.12.014>.
- Warrier, C.M., Abrams, D.A., Nicol, T.G., Kraus, N., 2011. Inferior colliculus contributions to phase encoding of stop consonants in an animal model. *Hear. Res.* 282, 108–118, <http://dx.doi.org/10.1016/j.heares.2011.09.001>.
- White-Schwach, T., Woodruff Carr, K., Thompson, E.C., Anderson, S., Nicol, T., Bradlow, A.R., Zecker, S.G., Kraus, N., 2015. Auditory processing in noise: a preschool biomarker for literacy. *PLOS Biol.* 13, e1002196, <http://dx.doi.org/10.1371/journal.pbio.1002196>.
- Woodruff Carr, K., White-Schwach, T., Tierney, A.T., Strait, D.L., Kraus, N., 2014. Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14559–14564, <http://dx.doi.org/10.1073/pnas.1406219111>.