



PAPER

Brainstem transcription of speech is disrupted in children with autism spectrum disorders

Nicole Russo,^{1,2} Trent Nicol,¹ Barbara Trommer,^{2,3} Steve Zecker¹ and Nina Kraus^{1,2,4}

1. The Auditory Neuroscience Laboratory, Department of Communication Sciences, Northwestern University, USA

2. Northwestern University Interdepartmental Neuroscience Program, USA

3. Departments of Pediatrics and Neurology, Feinberg School of Medicine, Northwestern University and Evanston Northwestern Healthcare Center for Neurodevelopmental Disabilities, USA

4. Department of Neurobiology and Physiology and Department of Otolaryngology, Northwestern University, USA

Abstract

Language impairment is a hallmark of autism spectrum disorders (ASD). The origin of the deficit is poorly understood although deficiencies in auditory processing have been detected in both perception and cortical encoding of speech sounds. Little is known about the processing and transcription of speech sounds at earlier (brainstem) levels or about how background noise may impact this transcription process. Unlike cortical encoding of sounds, brainstem representation preserves stimulus features with a degree of fidelity that enables a direct link between acoustic components of the speech syllable (e.g. onsets) to specific aspects of neural encoding (e.g. waves V and A). We measured brainstem responses to the syllable /da/, in quiet and background noise, in children with and without ASD. Children with ASD exhibited deficits in both the neural synchrony (timing) and phase locking (frequency encoding) of speech sounds, despite normal click-evoked brainstem responses. They also exhibited reduced magnitude and fidelity of speech-evoked responses and inordinate degradation of responses by background noise in comparison to typically developing controls. Neural synchrony in noise was significantly related to measures of core and receptive language ability. These data support the idea that abnormalities in the brainstem processing of speech contribute to the language impairment in ASD. Because it is both passively elicited and malleable, the speech-evoked brainstem response may serve as a clinical tool to assess auditory processing as well as the effects of auditory training in the ASD population.

Introduction

Autism spectrum disorders (ASD) are a cluster of disorders that includes autism, Asperger syndrome, and Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS) (Siegal & Blades, 2003; Tager-Flusberg & Caronna, 2007). Impairment in the social and communicative use of language is a hallmark of ASD. In severe cases, children with ASD are non-verbal. When speech is present, it is often slow to develop, echolalic, stereotypic, emotionless or excessively literal (Boucher, 2003; Shriberg, Paul, McSweeney, Klin, Cohen & Volkmar, 2001; Siegal & Blades, 2003). An individual with ASD will usually not engage in typical reciprocal communication (Rapin & Dunn, 2003; Tager-Flusberg & Caronna, 2007). Both receptive and expressive deficits can occur and both have been attributed, at least in part, to abnormalities of auditory processing (Siegal & Blades, 2003). Further, children with ASD have particular difficulties processing speech in background noise, as demonstrated by high speech perception thresholds and poor temporal resolution

and frequency selectivity (Alcantara, Weisblatt, Moore & Bolton, 2004).

The source-filter model of speech

Successful communication relies on being able to both produce and process speech sounds in a meaningful manner. The literature on speech production provides a useful dichotomy to describe the acoustics of speech (Fant, 1960; Ladefoged, 2001; Titze, 1994), the source-filter model. In this model, the vibration of the vocal folds represents the sound source. The lowest frequency of a periodic signal, such as speech sounds, is known as the fundamental frequency (F_0) and it is the rate of the vocal fold vibrations. Everything else – vocal tract, oral cavity, tongue, lips and jaw – comprises the filter. The vocal tract resonates and the resulting resonant frequencies are known as formants. The formants are conveyed through manipulations of the filter and provide cues about onsets and offsets of sounds. Broadly speaking, linguistic content – vowels and consonants – is transmitted

Address for correspondence: Nina Kraus, Department of Communication Sciences, Northwestern University, 2240 Campus Dr., Evanston, IL 60208, USA; e-mail: nkraus@northwestern.edu

by particular filter shapes, whereas nonlinguistic information, such as pitch and voice intonation, relies largely on characteristics of the source. Although source-filter cues are simultaneously conveyed in the acoustic stream of speech, remarkably, they can be readily transcribed as both discrete components and as a whole by the auditory brainstem (Kraus & Nicol, 2005).

Transcription in the auditory brainstem

The brainstem response has the capacity to reveal auditory pathway deficits in a non-invasive and passive manner, which has engendered its long history of clinical use even in difficult-to-test populations. Timing and periodicity of an evoking stimulus are preserved in this response, enabling it to reflect processing deficits that arise from the peripheral auditory system and ascending auditory pathway. Soundwaves propagate through the auditory nerve and lower brainstem structures including the cochlear nucleus, superior olivary complex, lateral lemniscus and inferior colliculus (Buchwald & Huang, 1975; Hood, 1998; Moller & Jannetta, 1985). Precision is such that timing delays on the order of fractions of milliseconds are diagnostically significant and the latency of responses lends insight into where in the pathway the anomalies occur (Hood, 1998; Jacobson, 1985).

Click stimuli, which are typically used to assess hearing, evoke short latency auditory brainstem responses (≤ 10 ms) that provide information limited to timing and amplitude. The brainstem frequency-following response (FFR), that cannot be elicited by a click, phase locks to the periodic components of a stimulus (in the case of speech, largely source information as well as formant frequency harmonics (filter)) (Galbraith, 1994; Galbraith, Amaya, de Rivera, Donan, Duong, Hsu, Tran & Tsang, 2004; Galbraith, Philippart & Stephen, 1996; Worden & Marsh, 1968) and is thought to originate in the auditory midbrain lateral lemniscus (Galbraith, 1994) and inferior colliculus (J.C. Smith, Marsh & Brown, 1975). In one study of the FFR to phrase-speech presented forward and backward, the far-field brainstem response was enhanced (increased signal-to-noise ratio) for forward speech (Galbraith *et al.*, 2004), suggesting that the auditory brainstem may respond preferentially to stimuli with environmental significance or with which people have greater experience.

The speech-evoked auditory brainstem response in quiet and background noise

Speech is a complex stimulus that, unlike a click, has environmental relevance. The speech-evoked brainstem response lends itself to the extraction of information about encoding of syllable onset, offset and periodicity (pitch and formant spectra). The brainstem response to the syllable /da/ reflects both phonetic/filter (transient) and prosodic/source (periodic) acoustic features with remarkable fidelity (Johnson, Nicol & Kraus, 2005;

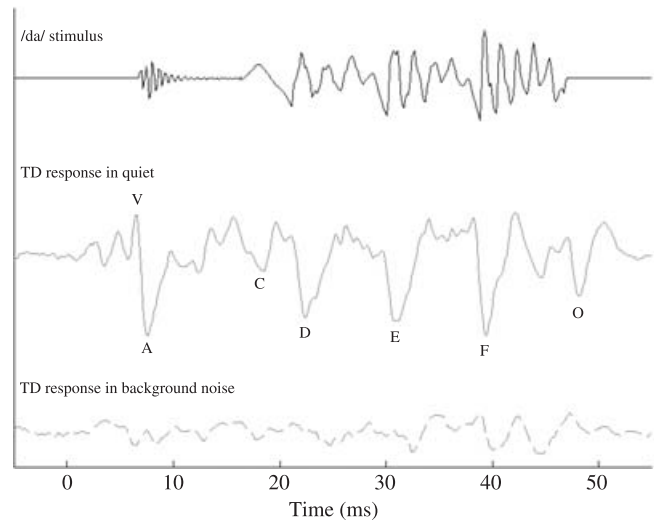


Figure 1 Speech stimulus /da/ and typically developing grand average response in quiet and background noise conditions. The /da/ has an onset burst followed by a transition to the periodic vowel portion. The stimulus waveform is shifted ~ 7 ms to compensate for neural lag in the response. Brainstem responses to /da/ were robust in quiet, reflecting stimulus features with great precision. Waves V and A reflect the onset of the /da/ stimulus, wave C represents the transition to the periodic portion, waves D, E, and F comprise the frequency-following response, and wave O signals the offset of the response. The wavelengths between waves D–E and E–F correspond to the fundamental frequency (F_0 , pitch) of the stimulus, while F_1 and higher frequency components are encoded in the smaller peaks between the dominant F_0 waves. In background noise (dashed line), many of the transient response peaks are abolished, while sustained activity (frequency-following response) and waves F and O persisted.

King, Warrier, Hayes & Kraus, 2002; Kraus & Nicol, 2005; Russo, Nicol, Musacchia & Kraus, 2004; Johnson, Nicol, Zecker, Bradlow, Skoe & Kraus, 2008) (see Figure 1). The timing of these neural responses relays information about neural synchrony. Specifically, four transient responses (waves V, A, C, and O) and phase locking in the range of the first formant (F_1) convey the filter aspects of the speech syllable. The response to the onset of the syllable is indicated by a wave V and its negative trough, wave A. Wave C represents the transition to the periodic, voiced portion of the stimulus that corresponds to the vowel. Wave O corresponds to stimulus offset. The F_0 , which transmits pitch cues, is reflected in the time domain by FFR waves D, E and F and phase locking to that frequency. Source cues are conveyed by the F_0 in both the frequency and time domains. Moreover, the brainstem response to speech presented in background noise provides an index of auditory pathway function in challenging listening situations. Even in the normal system, transient onset synchrony and phase locking to frequencies in the range of F_1 are significantly diminished, whereas the encoding of F_0 remains robust.

Clinical correlations and utility of the auditory brainstem response

Auditory brainstem function has been linked to language impairment (Banai, Nicol, Zecker & Kraus, 2005; Cunningham, Nicol, Zecker, Bradlow & Kraus, 2001; Johnson, Nicol, Zecker & Kraus, 2007; King *et al.*, 2002; Wible, Nicol & Kraus, 2004, 2005) and also to auditory expertise, such that speech-evoked brainstem responses have been shown to be shaped and enhanced by lifelong linguistic (Krishnan, Xu, Gandour & Cariani, 2004, 2005; Xu, Krishnan & Gandour, 2006) and musical experience (Musacchia, Sams, Skoe & Kraus, 2007; Wong, Skoe, Russo, Dees & Kraus, 2007), possibly through corticofugal feedback to subcortical sensory circuitry (Ahissar & Hochstein, 2004; Kraus & Banai, 2007). Although prior studies have investigated cortical-evoked responses to speech and their relationship to language in individuals with ASD (Boddaert, Belin, Chabane, Poline, Barthelemy, Mouren-Simeoni, Brunelle, Samson & Zilbovicius, 2003; Boddaert, Chabane, Belin, Bourgeois, Royer, Barthelemy, Mouren-Simeoni, Philippe, Brunelle, Samson & Zilbovicius, 2004; Ceponiene, Lepistö, Shestakova, Vanhala, Alku, Näätänen & Yaguchi, 2003; Jansson-Verkasalo, Ceponiene, Kielinen, Suominen, Jäntti, Linna, Moilanen & Näätänen, 2003; Kasai, Hashimoto, Kawakubo, Yumoto, Kamio, Itoh, Koshida, Iwanami, Nakagome, Fukuda, Yamasue, Yamada, Abe, Aoki & Kato, 2005; Kuhl, Coffey-Corina, Padden & Dawson, 2005; Lepistö, Kujala, Vanhala, Alku, Huotilainen & Näätänen, 2005; Lepistö, Silokallio, Nieminen-von Wendt, Alku, Näätänen & Kujala, 2006), the majority of studies of the auditory brainstem have focused on responses to non-speech stimuli (i.e. clicks or pulses) (reviewed in Klin, 1993; Rapin & Dunn, 2003). Such studies yielded mixed results, reporting abnormal, increased conduction times (Maziade, Merette, Cayer, Roy, Szatmari, Cote & Thivierge, 2000; McClelland, Eyre, Watson, Calvert & Sherrard, 1992), normal responses (Courchesne, Courchesne, Hicks & Lincoln, 1985; Rumsey, Grimes, Pikus, Duara & Ismond, 1984; Tharpe, Bess, Sladen, Schissel, Couch & Schery, 2006) or deficits in only a subset of children (Rosenhall, Nordin, Brantberg & Gillberg, 2003). These varying results may reflect the diagnostic heterogeneity in ASD. In the only known study of speech-evoked brainstem responses in ASD, Russo and colleagues (Russo, Skoe, Trommer, Nicol, Zecker, Bradlow & Kraus, 2008) reported deficient processing of pitch cues in speech in a subset of children with ASD.

In the present study, we utilized the speech-evoked auditory brainstem response to investigate the language deficiencies characteristic of individuals with ASD. Our hypotheses were that (1) children with ASD would demonstrate deficits in auditory brainstem function relative to their typically developing (TD) counterparts and (2) that abnormalities in the neural processing of speech in the ASD group would correlate with their language impairment. Children participating in this study were required to have normal auditory brainstem responses to

clicks, thereby ruling out middle ear or cochlear dysfunction (Hood, 1998; Jacobson, 1985; Rosenhall *et al.*, 2003) and so that any intergroup differences would be attributable only to difficulties processing speech.

Methods

The Institutional Review Board of Northwestern University approved all research and consent and assent were obtained both from the parent(s) or legal guardian(s) and the child.

Children were acclimated to the testing location and equipment prior to experimental data being collected. They were allowed to visit the laboratory and interact with the tester on multiple occasions. Some children brought electrodes home with them to better familiarize themselves with the neurophysiological procedure.

Participants

Of the 45 children originally recruited for this study, six (all children with ASD) were excluded for the following reasons: abnormal click-evoked brainstem responses ($n = 2$), mental ability below inclusion cutoff ($n = 1$), non-compliance resulting in inability to test ($n = 1$), parental decision to discontinue due to the required time commitment of the study ($n = 1$) and relocation ($n = 1$). Participants included 21 verbal children with ASD ($n = 19$ boys, 2 girls) and 18 typically developing children (TD, $n = 10$ boys, 8 girls). Age range was 7–13 years old and mean age (SD) did not differ between groups (TD, $M = 9.7$, $SD = 1.934$ vs. ASD, $M = 9.90$, $SD = 1.921$; independent two-tailed t -test; $t(37) = .295$, $p = .77$).

Study participants were recruited from community and internet-based organizations for families of children with ASD. Children in the diagnostic group were required to have a formal diagnosis of one of the ASDs made by a child neurologist or psychologist and to be actively monitored by their physicians and school professionals at regular intervals. Parents were asked to supply the names of the examining professionals, their credentials, office location, date of initial evaluation and the specific diagnosis made. Diagnoses included autism ($n = 1$), Asperger Disorder ($n = 7$), PDD-NOS ($n = 1$), and a combined diagnosis (e.g. Asperger Disorder/PDD-NOS; $n = 12$). The diagnosis of ASD was supplemented by observations during testing such that included subjects were noted to have some or all of the following: reduced eye contact, lack of social or emotional reciprocity; perseverative behavior; restricted range of interests in spontaneous and directed conversation; repetitive use of language or idiosyncratic language; abnormal pitch, volume, and intonation; echolalia or scripted speech; and stereotyped body and hand movements. Diagnosis was also supplemented by an internal questionnaire that provided developmental history, a description of current symptoms, and functional level at time of entry into the study.

Table 1 Mental and language ability scores. Mental ability was assessed using the Wechsler Abbreviated Scale of Intelligence (WASI); it provided full-scale, verbal and performance intelligence scores. Although significantly lower than TD children on measures of full-scale (overall) and verbal measures, the children with ASD demonstrated scores well within the normal range on all three components. Language ability was assessed using the Clinical Evaluation of Language Fundamentals – 4th edn. (CELF); this test provided indices of core, expressive, and receptive language abilities. For all indices of language ability, the children with ASD scored significantly lower than TD children, but within the normal range. Mean scores (standard deviations) are reported

Test battery	Subtest	TD (<i>n</i> = 18) Mean (SD)	ASD (<i>n</i> = 21) Mean (SD)
WASI (mental ability)	Full scale	117.78 (11.909)	107.33 (14.62)
	Verbal	115.89 (2.683)	101.95 (15.045)
	Performance	115.39 (11.62)	110.9 (13.707)
CELF (language ability)	Core	112.39 (9.388)	98.71 (20.219)
	Expressive	113.22 (12.638)	102.57 (22.171)
	Receptive	113.06 (7.712)	96.71 (19.233)

Further inclusion criteria for both TD and ASD groups were (1) the absence of a confounding neurological diagnosis (e.g. active seizure disorder, cerebral palsy), (2) normal peripheral hearing as measured by air threshold pure-tone audiogram and click-evoked auditory brainstem responses and (3) a full-scale mental ability score whose confidence interval included a value ≥ 80 .

Hearing screening

On the first day of testing, children underwent a hearing threshold audiogram for bilateral peripheral hearing (≤ 20 dB HL) for octaves between 250 and 8000 Hz via an air conduction threshold audiogram on a Grason Stadler model GSI 61. Children wore insert earphones in each ear and were instructed to press a response button every time they heard a beep. At each subsequent test session, children were required to pass a follow-up hearing screening at 20 dB HL for octaves between 125 and 4000 Hz. These screenings were conducted using a Beltone audiometer and headphones.

Cognitive and academic testing

All cognitive and academic testing took place in a quiet office with the child seated across a table from the test administrator. Full-scale mental ability was assessed by four subtests of the Wechsler Abbreviated Scale of Intelligence (WASI, four subtests) (Woerner & Overstreet, 1999). This testing was supplemented by scores of performance (Block Design and Matrix Reasoning subtests) and verbal (Vocabulary and Similarities) mental ability which were not part of inclusion criteria (Table 1; mean and standard deviations). The Clinical Evaluation of Language Fundamentals-4 (CELF) (Semel, Wiig & Secord, 2003) was administered to provide indices of core, expressive and receptive language abilities (Table 1). However, performance on the CELF was not used as a study inclusion criterion.

Stimuli and data collection

All neurophysiological recordings took place in a sound attenuated chamber. During testing, the child sat

comfortably in a recliner chair and watched a movie (DVD or VHS) of his or her choice. The movie soundtrack was presented in free field with the sound level set to < 40 dB SPL, allowing the child to hear the soundtrack via the unoccluded, non-test ear. To enhance compliance, children were accompanied by their parent(s) in the chamber. Children were permitted breaks during testing as needed.

Auditory brainstem responses were collected via the Navigator Pro (Bio-logic Systems Corp., a Natus Company, Mundelein, IL) using BioMARK software (<http://www.brainvolts.northwestern.edu/projects/clinicaltechnology.php>). All auditory stimuli were presented monaurally into the right ear through insert earphones (ER-3, Etymotic Research, Elk Grove Village, IL, USA). Responses were recorded via three Ag-AgCl electrodes, with contact impedance of ≤ 5 k Ω , located centrally on the scalp (Cz), with the ipsilateral earlobe as reference and forehead as ground. For all data, trials with artifacts exceeding 23.8 μ V were rejected online.

Click-evoked responses

Click-evoked wave V auditory brainstem response latencies were used to assess hearing, as well as for confirmation that the position of the ear insert did not change during the test session. The click stimulus, a 100 μ s duration broadband square wave, was presented at a rate of 13/sec. Click-evoked responses were sampled at 24 kHz and were online bandpass filtered from 100 to 1500 Hz, 12 dB/octave. Two blocks of 1000 sweeps each were collected at the onset of the session and an additional block of 1000 sweeps was collected at the conclusion to confirm that ear insert placement did not change. A criterion of being within one cursor-click on the recording window (~ 0.04 ms) was implemented. In most cases there was no change and in no subject was there more than the allowable difference between initial and final click replications.

Speech-evoked responses

Auditory brainstem responses were recorded in response to a 40 ms speech syllable /da/ which was synthesized in Klatt (Klatt, 1980) (Figure 1). In this syllable, the voicing

begins at 5 ms and the first 10 ms are bursted. The frequency components are as follows: F_0 : 103–125 Hz, F_1 : 220–720 Hz, F_2 : 1700–1240 Hz, F_3 : 2580–2500 Hz, F_4 : 3600 (constant), F_5 : 4500 (constant).

The speech-evoked brainstem responses were collected in two different conditions, at conversational speech level in quiet (80 dB SPL) and at the same intensity with simultaneous white background noise (75 dB). The /da/ stimuli were presented with alternating polarity in order to minimize stimulus artifact and cochlear microphonic (Gorga, Abbas & Worthington, 1985). For each condition, three blocks of 2000 sweeps were collected at a rate of 10.9/sec. A 75 ms recording window (including a 15 ms pre-stimulus period and 20 ms period after stimulus offset) was used. The three response blocks were then averaged together for a final waveform to be used in all analyses. Responses were sampled at 6856 Hz and bandpass filtered on-line from 100 to 2000 Hz, 12 dB/octave to isolate the frequencies that are most robustly encoded at the level of the brainstem.

Data analysis

Standardized tests of cognitive (WASI) and language (CELF) abilities were scored according to each test's age appropriate normative range.

The speech-evoked brainstem response was characterized by measures of pre-stimulus neural activity, peak latency and amplitude, the onset response complex (waves V and A) interpeak duration, slope and amplitude; and FFR amplitude (root-mean-square; RMS, fast Fourier transform amplitude; FFT), stimulus-to-response and quiet-to-noise inter-response correlations. RMS and FFT measures were performed over the 22–40 ms range, isolating the main peaks of the FFR. Stimulus-to-response correlations were performed in both quiet and background noise conditions by shifting the 13–34 ms range of the stimulus in time until the best correlation with the response was found. This maximum typically occurred at 8.5 ms and aligned the FFR peaks with those of the stimulus waveform. Quiet-to-noise inter-response correlations were performed over the entire response as well as the 11–40 ms FFR range, in each case allowing for up to 2 ms of lag. These measures have been previously described (Russo *et al.*, 2004). A subset of six of the original 21 children with ASD were re-evaluated at a second session to establish response reliability over time in children with ASD and demonstrated that the responses reported below were stable over time in these children.

For the statistical analysis, multivariate analyses of variance (MANOVA) tests were conducted, followed by independent two-tailed Student's *t*-tests. Due to limitations inherent in the interpretation of a MANOVA (Tabachnick & Fidell, 2007), two-tailed independent *t*-tests and effect sizes were calculated to describe diagnostic group differences (*p*-values $\leq .05$ were considered significant and Cohen's *d* effect sizes > 0.50 were deemed meaningful). A preliminary multivariate analysis of

variance ruled out age, sex and mental ability as potential co-factors, with the exception of sex on speech-evoked wave V latency. Reported *p*-values are those of the *t*-tests, except for wave V latency. Due to the sex effect, an analysis of co-variance with sex as a co-variate was conducted for wave V latency; therefore, *F*-values are reported. Levene's Test for Equality of Variances was applied to each statistical analysis and, when relevant, the reported *p*-values reflect corrections based on unequal variances. Pearson's correlations were conducted to establish relationships between neural processing of sound and cognitive/academic measures; the correlation analyses were collapsed across (irrespective of) diagnostic category. Correlations were considered significant if they were moderate to strong ($r \geq 0.35$) with $p \leq .05$.

Results

Cognitive and academic testing (Table 1)

Although in the normal range, both full-scale and verbal mental ability scores from the WASI were significantly lower in the children with ASD ($t(37) = 2.42$, $p = .021$, $d = .82$ and $t(37) = 3.11$, $p = .004$, $d = 1.01$, respectively), whereas performance mental ability did not differ between groups ($t(37) = 1.09$, $p = .282$, $d = .35$). On the CELF test, children with ASD scored significantly lower on indices of core and receptive language ability ($t(37) = 2.77$, $p = .01$, $d = .87$ and $t(37) = 3.58$, $p = .001$, $d = 1.12$, respectively), but not expressive language ability ($t(37) = 1.80$, $p = .08$, $d = .59$).

Click-evoked brainstem responses

Wave V latency ranged from 5.15 to 5.90 ms ($M = 5.56$ ms, $SD = .178$), consistent with the previously reported normal range (Gorga *et al.*, 1985; Hood, 1998; Jacobson, 1985). Wave V latency did not differ between groups ($t(37) = 1.46$, $p = .149$, $d = .46$).

Speech-evoked response fidelity

The brainstem response to /da/ (Figure 1) consists of seven transient response peaks (V-O). In the quiet condition, waves were identifiable 100% of the time in the TD group; however, six children with ASD were missing a wave, specifically in the FFR region, although not always the same wave. Because onset response components are known to become abolished or diminished by background noise even in normal subjects (Russo *et al.*, 2004), these measures were omitted from analyses. Analyses of responses in background noise were restricted to those waves (F and O) which were reliably present in background noise (~90% typical responses). Nevertheless, all sustained response measures (RMS, FFT, and correlations) were evaluated. Despite normal click-evoked responses, the ASD group showed pervasive deficits

Table 2 Significant speech-evoked auditory brainstem response measures. Significant differences ($p < .05$) between TD children and children with ASD were identified for several speech-evoked auditory brainstem response measures. Mean values (standard deviations) are reported

	TD ($n = 18$) Mean (SD)	ASD ($n = 21$) Mean (SD)
Quiet response measures:		
Wave V latency (ms)	6.54 (0.174)	6.73 (0.267)
Wave A latency (ms)	7.48 (0.232)	7.85 (0.412)
Onset response VA duration (ms)	0.94 (0.168)	1.13 (0.323)
Wave D latency (ms)	22.38 (0.433)	22.77 (0.549)
Wave F latency (ms)	39.25 (0.27)	39.54 (0.4)
Noise response measures:		
Wave F amplitude (μV)	-0.11 (0.079)	-0.05 (0.096)
Stimulus-to-response-in-background-noise lag (ms)	8.85 (0.961)	8.17 (1.112)
Stimulus-to-response-in-background-noise correlation (r -value)	0.21 (0.088)	0.14 (0.108)
Quiet-to-noise inter-response correlation (entire response) (r -value)	0.4 (0.185)	0.27 (0.166)
Quiet-to-noise inter-response correlation (11–40 ms range) (r -value)	0.49 (0.206)	0.36 (0.23)
Composite response measures:		
Onset synchrony in quiet (z -score)	0.00 (2.359)	-3.84 (4.192)
Transient responses in quiet (z -score)	0.00 (3.49)	-5.84 (5.714)
Phase locking in quiet (z -score)	0.00 (1.423)	-1.99 (2.349)
Neural synchrony in noise (z -score)	0.00 (2.603)	-3.17 (3.422)

transcribing speech in quiet and background noise, as was evident in both the onset and FFR portions of the response. Significant group differences are described below. Due to the large number of dependent variables, only the means and standard deviations of measures that were significantly different between groups are reported in Table 2.

Transcription of phonetic (filter) aspects of speech

Wave V and A latencies were significantly delayed in children with ASD (wave V: $F(1, 36) = 4.45$, $p = .042$, $d = .87$; wave A: $t(37) = 3.45$, $p = .001$, $d = 1.18$) (Figure 2). Onset response duration was also significantly prolonged in the ASD group ($F(1, 36) = 4.67$, $p = .037$, $d = .75$).

Transcription of fundamental frequency (source) aspects of speech

Quiet response wave D and F latencies were significantly delayed in children with ASD (wave D: $t(37) = 2.47$, $p = .018$, $d = .81$; wave F: $t(37) = 2.62$, $p = .013$, $d = .87$), and wave F amplitude showed a trend toward being reduced ($t(37) = 1.91$, $p = .064$, $d = .61$). Wave F amplitude in background noise was smaller in children with ASD ($t(37) = 2.14$, $p = .039$, $d = .70$). These differences are shown in Figure 3. It should be noted that the timing of waves D, E and F also reflect spectrotemporal formant frequency (filter) components of the stimulus (Johnson *et al.*, 2008a). In the frequency domain, there were no between-group differences in F_0 magnitude.

Neural synchrony in background noise: the sustained response

Stimulus-to-response-in-background-noise correlations over the FFR (13–34 ms) were lower ($t(37) = 2.41$, $p = .021$,

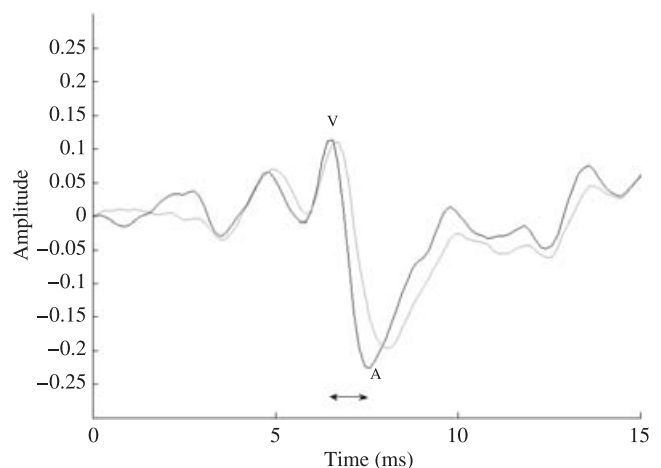


Figure 2 Comparison of grand average onset responses to /da/ in quiet in TD children ($n = 18$; black line) and children with ASD ($n = 21$; gray line). The neural response to the onset of speech sounds was less synchronous in children with ASD (gray) as compared to TD children (black). Notably the onset response in children with ASD showed significant delays in waves V and A, and also a longer interpeak interval (horizontal arrow).

$d = .78$) and the maximum correlation occurred at a shorter lag ($t(37) = 2.03$, $p = .050$, $d = .66$) in the ASD group. Quiet-to-noise inter-response correlations over the entire response (0–49 ms) and restricted to the FFR range (11–40 ms) were also significantly lower in the ASD group ($t(37) > 2.47$, $p < .018$, $d > .80$, both ranges), indicating poorer response fidelity in the ASD group. These findings are consistent with excessive response degradation by background noise in the ASD group relative to the TD controls.

Overall, subcortical transcription of sound was pervasively disrupted in ASD children as was evidenced by

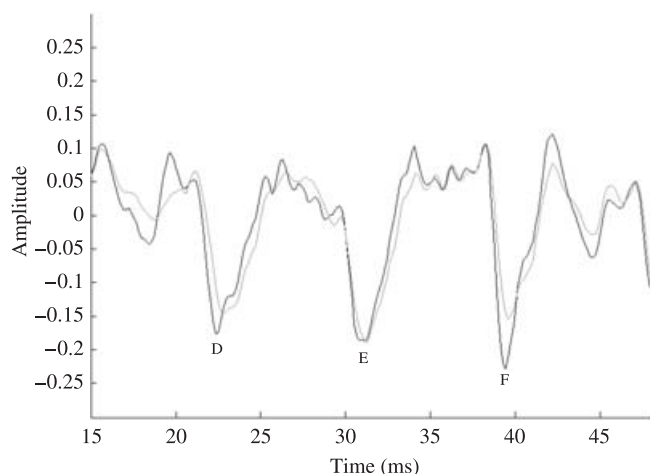


Figure 3 Comparison of grand average frequency-following responses to /da/ in quiet in TD children (black line) and children with ASD (gray line). The frequency-following response in children with ASD showed significant delays in peaks D and F; peak F was also reduced in amplitude, demonstrating reduced phase locking in brainstems of children with ASD.

both phonetic/filter (delayed onset) and F_0 /source deficits (delayed FFR waves D and F and smaller amplitude of wave F).

Relationship between neurophysiology and behavior

Individual measures

Quiet. Performance mental ability was related to wave C amplitude ($r(37) = -.38, p = .018$), while all scores of mental ability (performance, verbal and full-scale) were also related to offset wave O amplitude (performance and verbal: $r(37) = -.35, p = .028$; full-scale: $r(37) = -.38, p = .018$), such that better scores on tests of mental ability were indicative of larger amplitudes of these transient responses.

Noise. Higher performance and full-scale mental ability scores were associated with earlier wave F latency in background noise ($r(37) = -.35, p = .04$ and $r(37) = -.38, p = .022$, respectively). Higher core and receptive language indices were associated with greater quiet-to-noise inter-response correlations (11–40 ms range; $r(37) = .36, p = .025$ and $r(37) = .35, p = .028$, respectively).

Composite analyses

Based on the variables that differed significantly between groups, we computed four composite scores: onset synchrony in quiet (waves V and A latencies and VA duration), transient responses in quiet (waves V, A, D and F latencies, and VA duration), phase locking in quiet (waves D and F latencies), and neural synchrony in noise (wave F amplitude, stimulus-to-response-in-background-noise correlations and lags, and quiet-to-noise inter-response

correlations). Composite z-scores were significantly worse in children with ASD compared to TD children on all measures (onset synchrony in quiet: $t(37) = 3.59, p = .001, d = 1.13$; transient responses in quiet: $t(37) = 3.92, p < 0.001, d = 1.23$; phase locking in quiet: $t(37) = 3.26, p = .003, d = 1.03$; neural synchrony in noise: $t(37) = 3.21, p = .003, d = 1.04$).

Of the composite scores, neural synchrony in noise was the only variable that correlated significantly with behavior. Higher core and receptive language indices were indicative of better resilience in background noise (i.e. greater neural synchrony score) ($r(37) = .53, p < .001$ and $r(37) = .36, p = .02$, respectively).

Discussion

Summary

To our knowledge, this is the first study to investigate the sensory transcription of both filter and source aspects of speech in ASD, as well as to investigate effects of background noise on brainstem processing in this population. The major finding is that children with ASD demonstrate reduced neural synchrony and phase locking to speech cues in both quiet and background noise at the level of the brainstem, despite normal click-evoked responses. Thus, the sensory transcription of speech is disrupted due to an inability to accurately process both filter cues, which help to distinguish between consonants and vowels, and source cues, which help to determine speaker identity and intent. Similar to recent findings of correlations between behavioral mental ability and cortical evoked potentials in children with ASD (Salmond, Vargha-Khadem, Gadian, de Haan & Baldeweg, 2007), our results indicate relationships between speech-evoked brainstem responses and behavior. In quiet, more robust encoding of cues that signify formant transitions and offset (amplitudes of transient waves C and O) were related to better cognitive abilities. In background noise, wave F latency was correlated with higher cognitive ability. Finally, neural resilience to background noise (quiet-to-noise inter-response correlations; wave F amplitude) was strongly related to better core and receptive language abilities. The transcription of speech in noise remaining consistent with transcription of speech in quiet indicates less of a deleterious effect of background noise and may account for better speech perception in noise. Taken together, these findings support the notion that subcortical biological deficits may underlie social communication problems in ASD.

Click versus speech

Although there were differences between the control and ASD groups with respect to brainstem processing of speech stimuli, all subjects exhibited normal processing of non-speech (click) stimuli. This finding is consistent

with data from cortical-evoked potentials in ASD which indicate that the 'speech-ness' of a stimulus predisposes it to abnormal processing (Boddaert *et al.*, 2003; Boddaert *et al.*, 2004; Ceponiene *et al.*, 2003; Jansson-Verkasalo *et al.*, 2003; Kasai *et al.*, 2005; Lepistö *et al.*, 2005, 2006).

The different subcortical encoding of click versus speech stimuli probably reflects the distinctive acoustic and environmental characteristics of the stimuli themselves, as has been the case with cortical responses (Binder, Frost, Hammeke, Bellgowan, Springer, Kaufman & Possing, 2000; Liebenthal, Binder, Piorkowski & Remez, 2003). For example, clicks are brief in duration, have a rapid onset and flat broadband spectral components, whereas speech syllables, with their consonant-vowel (CV) combinations, are longer in duration, have ramped onset, and complex, time-varying spectral content. Backward masking (masking of the preceding consonant by an after-coming vowel) is characteristic of CV syllables but not clicks. Whereas the response to a click is limited to the onset, the response to speech syllables includes the FFR which is thought to activate different response mechanisms from onset responses (Akhoun, Gallégo, Moulin, Ménard, Veuillet, Berger-Vachon, Collet & Thai-Van, 2008; Hoormann, Falkenstein, Hohnsbein & Blanke, 1992; Khaladkar, Kartik & Vanaja, 2005; Song, Banai, Russo & Kraus, 2006).

It is also possible that the speech-click response discrepancy reflects different exposure that people have to the two stimuli. Lifelong linguistic and musical experience have been shown to influence brainstem transcription of sounds (Krishnan *et al.*, 2005; Krishnan, Gandour & Cariani, 2004; Musacchia *et al.*, 2007; Wong *et al.*, 2007; Xu *et al.*, 2006). Moreover, the developmental time-course of the speech-evoked brainstem response differs from the click response (Johnson, Nicol, Zecker & Kraus, 2008). Thus, speech-like sounds would be expected to invoke experience-related shaping of afferent sensory circuitry unlike clicks, which are laboratory-based and artificial. Taken together, these qualities make speech sounds a better stimulus for investigating language-related auditory processing deficits at the brainstem.

Relationship to other language-impaired populations

Both genetic (Bartlett, Flax, Logue, Smith, Vieland, Tallal & Brzustowicz, 2004; Herbert & Kenet, 2007; Smith, 2007) and behavioral evidence suggest that the language impairment in ASD has features in common with other developmental language and language-based learning disorders, including specific language impairment (SLI) and dyslexia (Herbert & Kenet, 2007; Oram Cardy, Flagg, Roberts, Brian & Roberts, 2005; Rapin & Dunn, 2003; Smith, 2007). Similar behavioral characteristics include both delayed or abnormal language development and problems with pragmatics of language. Physiologically, children with ASD show some similarities in subcortical transcription of sound with a subgroup of children with other language-based learning disorders (e.g. poor readers).

Common response traits between the groups are a deficit in onset synchrony and reduced fidelity of the response in background noise with respect to the stimulus (Cunningham *et al.*, 2001; Johnson *et al.*, 2007; King *et al.*, 2002; Russo, Nicol, Zecker, Hayes & Kraus, 2005; Wible *et al.*, 2004, 2005; see Banai & Kraus, 2008, for review). Consistent with known deficits in phonologic awareness, children with other language impairments (such as dyslexia) also demonstrate deficits in offset responses and phase locking to higher frequencies (F_1 range), while F_0 -related (pitch-related) response features are unimpaired. In contrast, children with ASD demonstrate more severe degradation of the response to speech in background noise (reduced quiet-to-noise inter-response correlations), amidst robust encoding of frequencies in the F_1 range and normal offset encoding. Although as a group, children with ASD did not exhibit pitch-related deficits (F_0 amplitude, FFR interpeak latencies), isolation of the subset of children with ASD ($n = 5$) previously identified as having poor brainstem pitch tracking in response to fully voiced speech syllables (Russo *et al.*, 2008) reveals pitch processing impairments to the syllable /da/. This subset of children did show prolonged interpeak latencies and reduced amplitude of F_0 . Because of the relationship with source cues and prosody, the source-related brainstem deficits in ASD may be associated, to some extent, with behavioral difficulties with prosody. Taken together, these data both provide overlap as well as isolate a different basis for speech transcription interference in brainstem responses of children with ASD compared to individuals with other language impairments. The data indicate a more pervasive impairment in speech transcription in ASD – in processes important for extracting timing, pitch, and harmonic information from phonemes. These results invite further investigation of speech-related brainstem processing deficits and the language impairment in ASD.

Implications

These data suggest that subcortical auditory pathway dysfunction may contribute to the social communication impairment in ASD. The mechanisms regulating the relationship between brainstem transcription and language acquisition, perception, and production likely involve corticofugal modulation of afferent sensory function. The malleability of this subcortical transcription process by short-term (Russo *et al.*, 2005; Song, Skoe, Wong & Kraus, 2007) and life-long (Krishnan *et al.*, 2004, 2005; Musacchia *et al.*, 2007; Wong *et al.*, 2007; Xu *et al.*, 2006) experience with sound further supports corticofugal modulation. Additional studies are warranted to further investigate reciprocal (top-down and bottom-up) influences. Consistent with prior work demonstrating deficient neural correlates of pitch in response to a fully voiced speech syllable in the auditory brainstem of children with ASD (Russo *et al.*, 2008), this study indicates that the brainstem deficit in ASD affects the transcription of various acoustic cues relevant for speech perception.

An advantage of using speech-evoked brainstem responses in the assessment of auditory function in the ASD population is that responses are objective and can be collected passively and non-invasively. Moreover, children with abnormal speech-evoked brainstem responses represent the best candidates for auditory training (Hayes, Warrier, Nicol, Zecker & Kraus, 2003; King *et al.*, 2002; Russo *et al.*, 2005). Consequently, brainstem responses may eventually have a place as a diagnostic and outcome measure of the efficacy of auditory training programs in the ASD population.

Acknowledgements

We would like to thank the children who participated in this study and their families. We would also like to thank Erika Skoe for her contribution to this research, particularly in software development. This research was supported by NIH R01 DC01510. The authors declare that they have no competing financial interests.

References

- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, **8** (10), 457–464.
- Akhoun, I., Gallégo, S., Moulin, A., Ménard, M., Veuillet, E., Berger-Vachon, C., Collet, L., & Thai-Van, H. (2008). The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. *Clinical Neurophysiology*, **119** (4), 922–933.
- Alcantara, J.I., Weisblatt, E.J., Moore, B.C., & Bolton, P.F. (2004). Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, **45** (6), 1107–1114.
- Banai, K., & Kraus, N. (2008). The dynamic brainstem: implications for APD. In D. McFarland & A. Cacece (Eds.), *Current controversies in central auditory processing disorder* (pp. 269–289). San Diego, CA: Plural Publishing Inc.
- Banai, K., Nicol, T., Zecker, S.G., & Kraus, N. (2005). Brainstem timing: implications for cortical processing and literacy. *Journal of Neuroscience*, **25** (43), 9850–9857.
- Bartlett, C.W., Flax, J.F., Logue, M.W., Smith, B.J., Vieland, V.J., Tallal, P., & Brzustowicz, L.M. (2004). Examination of potential overlap in autism and language loci on chromosomes 2, 7, and 13 in two independent samples ascertained for specific language impairment. *Human Heredity*, **57** (1), 10–20.
- Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., & Possing, E.T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, **10** (5), 512–528.
- Boddaert, N., Belin, P., Chabane, N., Poline, J.B., Barthelemy, C., Mouren-Simeoni, M.C., Brunelle, F., Samson, Y., & Zilbovicius, M. (2003). Perception of complex sounds: abnormal pattern of cortical activation in autism. *American Journal of Psychiatry*, **160** (11), 2057–2060.
- Boddaert, N., Chabane, N., Belin, P., Bourgeois, M., Royer, V., Barthelemy, C., Mouren-Simeoni, M.C., Philippe, A., Brunelle, F., Samson, Y., & Zilbovicius, M. (2004). Perception of complex sounds in autism: abnormal auditory cortical processing in children. *American Journal of Psychiatry*, **161** (11), 2117–2120.
- Boucher, J. (2003). Language development in autism. *International Journal of Pediatric Otorhinolaryngology*, **67** Suppl. 1, S159–S163.
- Buchwald, J.S., & Huang, C. (1975). Far-field acoustic response: origins in the cat. *Science*, **189** (4200), 382–384.
- Ceponiene, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., & Yaguchi, K. (2003). Speech-sound-selective auditory impairment in children with autism: they can perceive but do not attend. *Proceedings of the National Academy of Sciences of the United States of America*, **100** (9), 5567–5572.
- Courchesne, E., Courchesne, R.Y., Hicks, G., & Lincoln, A.J. (1985). Functioning of the brain-stem auditory pathway in non-retarded autistic individuals. *Electroencephalography and Clinical Neurophysiology*, **61** (6), 491–501.
- Cunningham, J., Nicol, T., Zecker, S.G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. *Clinical Neurophysiology*, **112** (5), 758–767.
- Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.
- Galbraith, G.C. (1994). Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. *Electroencephalography and Clinical Neurophysiology*, **92** (4), 321–330.
- Galbraith, G.C., Amaya, E.M., de Rivera, J.M., Donan, N.M., Duong, M.T., Hsu, J.N., Tran, K., & Tsang, L.P. (2004). Brain stem evoked response to forward and reversed speech in humans. *NeuroReport*, **15** (13), 2057–2060.
- Galbraith, G.C., Philippart, M., & Stephen, L.M. (1996). Brainstem frequency-following responses in Rett syndrome. *Pediatric Neurology*, **15** (1), 26–31.
- Gorga, M.P., Abbas, P.J., & Worthington, D.W. (1985). Stimulus calibrations in ABR measurements. In J.T. Jacobson (Ed.), *The auditory brainstem response* (pp. 49–62). San Diego, CA: College-Hill Press.
- Hayes, E.A., Warrier, C.M., Nicol, T.G., Zecker, S.G., & Kraus, N. (2003). Neural plasticity following auditory training in children with learning problems. *Clinical Neurophysiology*, **114** (4), 673–684.
- Herbert, M.R., & Kenet, T. (2007). Brain abnormalities in language disorders and in autism. *Pediatric Clinics of North America*, **54** (3), 563–583.
- Hood, L. (1998). *Clinical applications of the auditory brainstem response*. San Diego, CA: Singular Publishing Group.
- Hoormann, J., Falkenstein, M., Hohnsbein, J., & Blanke, L. (1992). The human frequency-following response (FFR): normal variability and relation to the click-evoked brainstem response. *Hearing Research*, **59** (2), 179–188.
- Jacobson, J.T. (Ed.) (1985). *The auditory brainstem response*. San Diego, CA: College-Hill Press.
- Jansson-Verkasalo, E., Ceponiene, R., Kielinen, M., Suominen, K., Jäntti, V., Linna, S.L., Moilanen, I., & Näätänen, R. (2003). Deficient auditory processing in children with Asperger syndrome, as indexed by event-related potentials. *Neuroscience Letters*, **338** (3), 197–200.

- Johnson, K.L., Nicol, T.G., & Kraus, N. (2005). Brain stem response to speech: a biological marker of auditory processing. *Ear and Hearing*, **26** (5), 424–434.
- Johnson, K.L., Nicol, T.G., Zecker, S.G., & Kraus, N. (2007). Auditory brainstem correlates of perceptual timing deficits. *Journal of Cognitive Neuroscience*, **19** (3), 376–385.
- Johnson, K.L., Nicol, T.G., Zecker, S.G., Bradlow, A.R., Skoe, E., & Kraus, N. (2008). Brainstem encoding of voiced consonant-vowel stop syllables. *Clinical Neurophysiology*, **119** (11), 2623–2635.
- Johnson, K.L., Nicol, T.G., Zecker, S.G., & Kraus, N. (2008). Developmental plasticity in the human auditory brainstem. *Journal of Neuroscience*, **28** (15), 4000–4007.
- Kasai, K., Hashimoto, O., Kawakubo, Y., Yumoto, M., Kamio, S., Itoh, K., Koshida, I., Iwanami, A., Nakagome, K., Fukuda, M., Yamasue, H., Yamada, H., Abe, O., Aoki, S., & Kato, N. (2005). Delayed automatic detection of change in speech sounds in adults with autism: a magnetoencephalographic study. *Clinical Neurophysiology*, **116** (7), 1655–1664.
- Khaladkar, A., Kartik, N., & Vanaja, C. (2005). Speech burst and click evoked ABR. *Audiology Online*.
- King, C., Warrier, C.M., Hayes, E., & Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neuroscience Letters*, **319** (2), 111–115.
- Klatt, D. (1980). Software for cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, **67**, 971–975.
- Klin, A. (1993). Auditory brainstem responses in autism: brainstem dysfunction or peripheral hearing loss? *Journal of Autism and Developmental Disorders*, **23** (1), 15–35.
- Kraus, N., & Banai, K. (2007). Auditory-processing malleability. *Current Directions in Psychological Science*, **16** (2), 105–110.
- Kraus, N., & Nicol, T. (2005). Brainstem origins for cortical ‘what’ and ‘where’ pathways in the auditory system. *Trends in Neurosciences*, **28** (4), 176–181.
- Krishnan, A., Xu, Y., Gandour, J.T., & Cariani, P.A. (2004). Human frequency-following response: representation of pitch contours in chinese tones. *Hearing Research*, **189** (1–2), 1–12.
- Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, **25** (1), 161–168.
- Kuhl, P.K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Developmental Science*, **8** (1), F1–F12.
- Ladefoged, P. (2001). *A course in phonetics*. Fort Worth, TX: Harcourt College Publishers.
- Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., & Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Research*, **1066** (1–2), 147–157.
- Lepistö, T., Silokallio, S., Nieminen-von Wendt, T., Alku, P., Näätänen, R., & Kujala, T. (2006). Auditory perception and attention as reflected by the brain event-related potentials in children with Asperger syndrome. *Clinical Neurophysiology*, **117** (10), 2161–2171.
- Liebenthal, E., Binder, J.R., Piorkowski, R.L., & Remez, R.E. (2003). Short-term reorganization of auditory analysis induced by phonetic experience. *Journal of Cognitive Neuroscience*, **15** (4), 549–558.
- McClelland, R.J., Eyre, D.G., Watson, D., Calvert, G.J., & Sherrard, E. (1992). Central conduction time in childhood autism. *British Journal of Psychiatry*, **160**, 659–663.
- Maziade, M., Merette, C., Cayer, M., Roy, M.A., Szatmari, P., Cote, R., & Thivierge, J. (2000). Prolongation of brainstem auditory-evoked responses in autistic probands and their unaffected relatives. *Archives of General Psychiatry*, **57** (11), 1077–1083.
- Moller, A., & Jannetta, P. (1985). Neural generators of the auditory brainstem response. In J.T. Jacobson (Ed.), *The auditory brainstem response* (pp. 13–31). San Diego, CT: College-Hill Press.
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America*, **104** (40), 15894–15898.
- Oram Cardy, J.E., Flagg, E.J., Roberts, W., Brian, J., & Roberts, T.P. (2005). Magnetoencephalography identifies rapid temporal processing deficit in autism and language impairment. *NeuroReport*, **16** (4), 329–332.
- Rapin, I., & Dunn, M. (2003). Update on the language disorders of individuals on the autistic spectrum. *Brain and Development*, **25** (3), 166–172.
- Rosenhall, U., Nordin, V., Brantberg, K., & Gillberg, C. (2003). Autism and auditory brain stem responses. *Ear and Hearing*, **24** (3), 206–214.
- Rumsey, J.M., Grimes, A.M., Pikus, A.M., Duara, R., & Ismond, D.R. (1984). Auditory brainstem responses in pervasive developmental disorders. *Biological Psychiatry*, **19** (10), 1403–1418.
- Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology*, **115** (9), 2021–2030.
- Russo, N., Nicol, T., Zecker, S., Hayes, E., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, **156** (1), 95–103.
- Russo, N., Skoe, E., Trommer, B., Nicol, T., Zecker, S., Bradlow, A., & Kraus, N. (2008). Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clinical Neurophysiology*, **119**, 1720–1731.
- Salmond, C.H., Vargha-Khadem, F., Gadian, D.G., de Haan, M., & Baldeweg, T. (2007). Heterogeneity in the patterns of neural abnormality in autistic spectrum disorders: evidence from ERP and MRI. *Cortex*, **43**, 686–699.
- Semel, E., Wiig, E.H., & Secord, W.A. (2003). *Clinical evaluation of language fundamentals* (4th edn.; CELF). San Antonio, TX: Harcourt Assessment.
- Shriberg, L.D., Paul, R., McSweeney, J.L., Klin, A.M., Cohen, D.J., & Volkmar, F.R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research*, **44** (5), 1097–1115.
- Siegel, M., & Blakes, M. (2003). Language and auditory processing in autism. *Trends in Cognitive Sciences*, **7** (9), 378–380.
- Smith, J.C., Marsh, J.T., & Brown, W.S. (1975). Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology*, **39** (5), 465–472.
- Smith, S.D. (2007). Genes, language development, and language disorders. *Mental Retardation and Developmental Disabilities Research Reviews*, **13** (1), 96–105.

- Song, J., Banai, K., Russo, N.M., & Kraus, N. (2006). On the relationship between speech- and nonspeech-evoked auditory brainstem responses. *Audiology and Neuro-Otology*, **11** (4), 233–241.
- Song, J., Skoe, E., Wong, P.C.M., & Kraus, N. (2007). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Human Brain Mapping Conference*. Chicago, IL.
- Tabachnick, B., & Fidell, L. (2007). Multivariate analysis of variance and covariance. In S. Hartman (Ed.), *Using multivariate statistics* (4th edn., pp. 243–310). Boston, MA: Allyn & Bacon.
- Tager-Flusberg, H., & Caronna, E. (2007). Language disorders: autism and other pervasive developmental disorders. *Pediatric Clinics of North America*, **54** (3), 469–481.
- Tharpe, A.M., Bess, F.H., Sladen, D.P., Schissel, H., Couch, S., & Schery, T. (2006). Auditory characteristics of children with autism. *Ear and Hearing*, **27** (4), 430–441.
- Titze, I.R. (1994). Fluctuations and perturbations in vocal output. In *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.
- Wible, B., Nicol, T., & Kraus, N. (2004). Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. *Biological Psychology*, **67** (3), 299–317.
- Wible, B., Nicol, T., & Kraus, N. (2005). Correlation between brainstem and cortical auditory processes in normal and language-impaired children. *Brain*, **128** (Pt 2), 417–423.
- Woerner, C., & Overstreet, K. (Eds.) (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: The Psychological Corporation.
- Wong, P.C., Skoe, E., Russo, N.M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, **10** (4), 420–422.
- Worden, F.G., & Marsh, J.T. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography and Clinical Neurophysiology*, **25** (1), 42–52.
- Xu, Y., Krishnan, A., & Gandour, J.T. (2006). Specificity of experience-dependent pitch representation in the brainstem. *NeuroReport*, **17** (15), 1601–1605.

Received: 31 January 2008

Accepted: 21 May 2008