Published in final edited form as:

Int J Psychophysiol. 2015 February ; 95(2): 65–76. doi:10.1016/j.ijpsycho.2014.06.002.

Infant Cortical Electrophysiology and Perception of Vowel Contrasts

Barbara K Cone, Ph.D. [Professor]

Speech, Language and Hearing Sciences, University of Arizona, P.O. Box 210071, Tucson AZ 85721, conewess@email.arizona.edu, phone: 520-626-3710

Abstract

Cortical auditory evoked potentials (CAEPs) were obtained for vowel tokens presented in an oddball stimulus paradigm. Perceptual measures of vowel discrimination were obtained using a visually-reinforced head-turn paradigm. The hypothesis was that CAEP latencies and amplitudes would differ as a function of vowel type and be correlated with perceptual performance. Twenty normally hearing infants aged 4–12 months were evaluated. CAEP component amplitudes and latencies were measured in response to the standard, frequent token /a/ and for infrequent, deviant tokens /i/, /o/ and /u/, presented at rates of 1 and 2 tokens/second. The perceptual task required infants to make a behavioral response for trials that contained two different vowel tokens, and ignore those in which the tokens were the same. CAEP amplitudes were larger in response to the deviant tokens, when compared to the control condition in which /a/ served as both standard and deviant. This was also seen in waveforms derived by subtracting the response to standard /a/ from the responses to deviant tokens. CAEP component latencies in derived responses at 2/sec also demonstrated some sensitivity to vowel contrast type. The average hit rate for the perceptual task was 68.5%, with a 25.7% false alarm rate. There were modest correlations of CAEP amplitudes and latencies with perceptual performance. The CAEP amplitude differences for vowel contrasts could be used as an indicator of the underlying neural capacity to encode spectro-temporal differences in vowel sounds. This technique holds promise for translation to clinical methods for evaluating speech perception.

Introduction

Audibility is fundamental for the discrimination of speech features identifying consonants (e.g., place, manner and voicing) and vowels (e.g., formant positions). Recently published research has revealed an apparent discrepancy between infant tone detection threshold and speech threshold, calling into question whether audibility for one can be used to predict the other (Cone & Whitaker, 2013). It is known that children weight speech-feature cues differently than do adults (Nitttrouer 2004; 2007; Nitttrouer & Lowenstein, 2007), but exactly

© 2014 Elsevier B.V. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

how infants process temporal and spectral speech information is not well understood (Berg, 1991; Berg & Boswell, 1995).

The ability to perceive speech-features plays an important role in theories of infant speech perception and language acquisition (Juszyk, et al., 1998). Moreover, measures of such ability are critically important for fitting and fine-tuning hearing aids (McCreery & Stelmachowicz, 2011) and cochlear implants (Kirk & Choi, 2009). Speech-feature discrimination is essential to language development, in the segmenting of words and in assigning meaning to words (Stager & Werker, 1997; McMurray & Aslin, 2005). Classic studies of speech feature discrimination (Eilers et al., 1977;) and categorical perception (Eimas, 1999; Eimas et al., 1971; Juszyk, et al., 1998, Kuhl, 1992, 2004; Trehub, 1979; Werker & Tees, 1999) indicate that infants have the capacity to discriminate between many acoustic features of speech and that this capacity is shaped by experience during the first year of life. Exposure to the native language and its phonological contrasts appears to sharpen perceptual boundaries between acoustic features, while the boundaries for non-native language phonemes are diminished or become extinct (Werker et al., 1981; Werker & Tees 1984). Thus, infants with hearing loss will have impaired audibility with concomitant sensory deprivation leading to developmental delay in their perceptual abilities to distinguish between phonemes, even in their native language (Moeller et al., 2007).

Studies on infant speech feature detection and discrimination have employed habituation or visual reinforcement paradigms (e.g. Werker, et al., 1998). These behavioral methods have not been widely adopted in speech, language, and hearing clinics, which typically report speech detection thresholds without any measure of discrimination between speech sounds. The development of reliable psychophysical and physiological methods for evaluating speech-feature detection and discrimination would be of tremendous benefit for diagnostic and rehabilitative audiology and speech pathology. Such methods would have applications for assaying the perceptual abilities of infants with hearing loss as well as infants with normal hearing who are at risk for developmental communication disorders and language impairments. These methods could also be used to document the effects of treatment.

Eisenberg et al. (2004, 2007) have made efforts to translate research laboratory techniques for studying infant speech feature discrimination to methods used in the clinic. They developed a test known as “VRA-SPAC” (Visual Reinforcement Assessment of the Perception of Speech Pattern Contrasts), to test the abilities of young infants to discriminate between speech-features of vowel height, vowel place, consonant voicing, consonant continuance or manner, and consonant place. In this test, infants hear a constantly repeating token until “habituated” and then are taught to respond to a speech-feature contrast such as vowel height: /**udu**/ vs. /**ada**/, vowel place: /**udu**/ vs. /**idi**/, consonant voicing: /**udu** vs. /**utu**/, consonant continuance: /**udu**/ vs. /**uzu**/, or consonant place: /**udu**/ vs. /**ubu**/, or /**ubu**/ vs. /**ugu**/). Eisenberg et al. reported data for a small sample (N=11) of normally hearing infants in the age range of 7–15 months, and some older infants and toddlers (aged 9–21 months) with hearing loss. Infants and toddlers with hearing loss lagged in their discrimination abilities in comparison to younger, normal hearing infants. Some normally hearing infants could not learn the discrimination task, and this outcome has frustrated efforts to translate the method into clinical use.

Electrophysiologic measures for speech feature perception

Auditory evoked potentials from the brainstem and cortex may circumvent the problem presented by psychophysical measures of speech perception: that infants and observers must learn the detection task required, and maintain performance of this task at a criterion level. An initial step towards clinical use of auditory evoked potentials is to establish the relationship between perceptual and electrophysiologic results in the laboratory. During the past 35 years, much knowledge of infant auditory system development and sensory capacity has been obtained from the auditory brainstem responses (ABR). ABR thresholds for clicks and tonebursts, and the absolute and interpeak latencies of wave I – V components reflect increased capacity for neural synchrony and temporal processing that follows the time-course for brainstem myelination over the first 18 months of life (Hecox & Galambos, 1974). ABR thresholds for clicks and tonebursts suggest that adult-like sensitivity is obtained, at least in the mid-high frequencies, during the first year of life, which is well before perceptual thresholds approximate adult levels (Werner et al., 1993). Recently, ABRs evoked by consonant-vowel syllables have been used to document spectral and temporal encoding of speech features at the level of the brainstem (for review see Chandrasekaran & Kraus 2010), but data from infants has not yet been published using these stimuli.

Tones or noise modulated at rates greater than 60 Hz can be used to evoke a brainstem response known as the auditory steady-state response (ASSR). Cone & Garinis (2009) reported results of speech-feature discrimination in conjunction with auditory steady state responses (ASSR) evoked by multi-frequency mixed modulation stimuli that approximated the temporal-spectral complexity of speech. Twenty-eight infants under 1 year of age were tested on a speech token discrimination task, contrasting place (/ba/ vs. /da/) or place and manner (/ba/ vs. /sa/). These results showed that infant abilities to discriminate speech-features improved with stimulus level. Furthermore, speech-feature discrimination scores were correlated with the ASSR measures to complex stimuli, which were used to estimate the amount of acoustic speech information available to the listener. These results indicate that electrophysiological measures hold promise as a metric of speech-feature perception abilities in infants.

The obligatory (or exogenous) cortical auditory evoked potentials (CAEPs) can be used to understand the physiological processes and neural substrates underlying speech-feature perception in infants. Kurtzberg et al. (1984) found topographical differences in the scalp distribution of CAEPs of newborns that reflected place of articulation of consonants (/da/ vs. /ba/) and waveform morphology differences that reflected voice onset time (/ta/ vs. /da/ and /ba/). Novak et al. (1989) recorded CAEPs to formants extracted from synthesized CV syllables but found no systematic effect of formant center frequency on the responses recorded during the first 6 months of life. Wunderlich et al. (2006) also used speech tokens to evoke CAEPs in infants and young children. In newborns, the speech tokens evoked a much larger amplitude response than did tones, but this finding was not consistent in older infants (aged 13–41 months) or children (aged 4–6 years). None of these studies related the CAEP latencies and amplitudes to detection or discrimination of these stimuli in the same infants. Yet, other groups have shown that CAEPs recorded in newborns or during infancy

can be used to predict language outcomes in later childhood (for review, see Benasich et al., 2002; Choudhury & Benasich, 2011; Molfese & Molfese, 1997).

Another obligatory CAEP that has been applied to the study of speech perception is the mismatch negativity (MMN) or mismatch response (MMR). The MMN is revealed as the *difference* between the CAEP waveform for a frequently presented stimulus token and that for an infrequent, contrasting token. The onset latency of MMN seen in this difference or derived waveform is in the range of 150–200 ms, or somewhat prolonged relative to the negative trough latency for CAEP component N1. Stimulus contrasts used to evoke MMN can differ by one or more temporal or spectral parameters or by different speech features such as a difference in voicing (/ta/ vs. /da/ or place of articulation (/da/ vs. /ba/) or vowel type. MMN is present for speech token contrasts in pre-term newborns and is thought to be “developmentally stable” by some investigators (Cheour et al., 2001). Yet, other investigators have demonstrated inability to reliably obtain MMN in infants and young children (Morr et al, 2002), or for that matter, adults (Wunderlich & Cone-Wesson 2001) even for stimulus differences that are known to be perceptually salient.

An acoustic change complex (ACC) is apparent in the CAEP when the auditory system is stimulated with a steady-state stimulus that then has an abrupt change in one parameter, such as level or frequency or spectro-temporal complexity (Martin & Boothroyd, 1999; 2000). The ACC appears to be an onset response (P1-N1-P2) to the stimulus change. The ACC can be appreciated in response to speech tokens, such as if a steady state /s/ is followed by a vowel. In this case, there is an onset response for the consonant /s/ and also for the onset of the vowel. The latency of the onset response to the acoustic change from /s/ to /a/ is prolonged and the amplitude attenuated relative to that observed for the initial onset response (Martin et al, 2008). Small & Werker (2012) demonstrated that the ACC could be obtained in infants as young as 4 months in response to speech tokens that varied with respect to acoustic features differentiating Hindi vs. English consonant-vowel tokens.

Although the use of CAEPs for clinical audiologic or neurologic evaluation purposes was largely eclipsed by the ABR during the past 30 years, some recent clinical research results have re-invigorated their relevance. Sharma and colleagues (2002; 2005) have demonstrated that CAEPs are a reliable metric of cortical plasticity and development brought on by the use of cochlear implants. Their studies indicate that CAEP latency change in the first months of implant use is a “biomarker” of expected auditory maturation or plasticity following electrical stimulation of the auditory nerve. They have also shown, furthermore, that children implanted after 7 years of age do not demonstrate the CAEP latency shifts to age-appropriate values, irrespective of the duration of cochlear implant use. These findings are correlated with attenuated speech perception benefits from implantation in comparison to those who are implanted before 3.5 years of age.

Another clinically relevant study was completed by Rance et al. (2002), who measured CAEPs from a group of infants and young children (age range 6–92 months) diagnosed with auditory neuropathy spectrum disorder (ANS) and from an age-matched group of children with sensorineural hearing loss (SNHL). Although the stimuli were presented using an odd-ball paradigm, contrasting speech syllables /bad/ vs. /dad/ or pure tone samples that

contrasted tones that had a 10% frequency difference (e.g., 3.0 kHz vs. 3.3 kHz), only the P1-N1-P2 obligatory components for the standard stimulus were considered. They found that CAEPs for tones and speech tokens were present in over 85% of those with SNHL, but for only 60% of those with ANSD. Whereas absence of CAEPs in the SNHL group could be accounted for on the basis of severity of loss and stimulus output limits, this was not the case for those with ANSD. That is, CAEP presence/absence in response to a supra-threshold stimulus was not related to the severity of the pure tone hearing loss, nor was it attributable to age. Despite this incongruence, there was a strong positive correlation between the presence of CAEP and speech perception abilities, as measured by a standardized test of phoneme perception, in children with ANSD. The children with ANSD who had CAEPs had average speech perception scores that were, on average, 50% better than those for whom CAEP were absent. These findings suggested that CAEP could be used clinically as a prognostic measure in infants and young children with ANSD. More recently, other lab groups have replicated these findings. For example, Sharma et al. (2011) have shown a positive correlation between CAEP-P1 latency and auditory skill development (as measured by parental questionnaire) in children with ANSD.

More recently, the use of CAEPs in the verification of hearing aid fittings in infants has also been suggested as a clinical application (Carter et al., 2010). To this end, Golding et al. (2007) obtained CAEPs in response to brief speech sound tokens in infants and toddlers (aged 2–41 months) with hearing loss who were tested in the sound field with and without their hearing aids. The presence of a CAEP indicated that the hearing aids provided enough gain to evoke a response to speech sound tokens presented at a conversational level. The CAEP findings were positively correlated with responses to a parent questionnaire regarding functional hearing abilities demonstrated by infants when using their amplification.

Recently, Cone & Whitaker (2013) investigated the relationship between perceptual detection thresholds for tonebursts and speech-sound tokens and CAEPs in 36 infants between the ages of 4–12 months. First, CAEP amplitude and latency input-output functions were obtained for tonebursts and speech tokens. The tonal stimuli were 50 ms tonebursts at 0.5, 1.0, 2.0 and 4.0 kHz. The speech sound tokens, /a/, /i/, /o/, /u/, /m/, /s/, and /f/, were created from natural speech samples and were also 50 ms in duration. All CAEP tests were completed while the infants were *awake* and engaged in quiet play. Observer-based psychophysical methods were used to establish perceptual detection thresholds for the same speech sounds and tonebursts used to evoke CAEPs.

Not only were infant CAEP component latencies prolonged by 100–150 ms in comparison to adult responses, their CAEP latency-vs.-level input-output functions were steeper compared to adults. Yet, CAEP amplitude growth functions with respect to stimulus SPL were found to be adult-like in this age group. Despite the fact that the CAEP amplitude input-output functions were similar to those in adults, the infants' perceptual thresholds were elevated with respect to those found in adults. This suggested that, although CAEP latencies indicated immaturity in neural transmission at the level of the cortex, amplitude growth with respect to stimulus SPL was adult-like at this age, particularly for the P1-N1 component. The reasons for the discrepancy between electrophysiologic and perceptual threshold may be due to perceptual immaturity in temporal resolution abilities and the broadband listening

strategy employed by infants (Werner & Boike, 2001). Furthermore, the immaturity of CAEP latency suggests that there are developmental-dependencies in the encoding of stimulus level and/or spectral-temporal properties at the level of the cortex that would ultimately affect speech feature detection and discrimination.

These issues motivated a study of CAEPs evoked by a change in acoustic properties of vowel tokens, and an infant's ability to demonstrate perception of those acoustic differences. An additional objective was to significantly extend the data-base of CAEP measurement in awake infants using clinically relevant stimuli so that these methods could be used in pediatric audiological assessment.

By testing CAEPs in awake, alert infants, we aimed to show that these evoked potentials could be used as a metric of physiological development in the same way that ABRs are used at the level of the brainstem. That is, the motivation was to use CAEPs to understand underlying physiological processes and the neural substrates of perception. There is a need, furthermore, for studies that link infant speech perception to CAEP component latency and amplitude in order to establish speech-feature discrimination abilities, which would greatly supplement measurements of threshold sensitivity. Such research would also contribute to theories linking sensory capacity, perception, and neural encoding to speech perception development.

Knowledge of auditory development above the level of the brainstem is needed to explain the difference between sensory capacity (as indicated by measures of ABR and OAE) and perceptual abilities. It is critically important to use ecologically valid stimuli, such as speech sounds, to further our understanding of how the ear and brain develop the capacity to perceive acoustically complex stimuli. These issues motivate the aims of the current proposal and the long-term goal of providing more sensitive and specific measures of infant speech perception and to characterize the time course of such development in the first year of life.

The aim of the current study was to obtain perceptual and electrophysiologic responses to vowel tokens. CAEPs were obtained in awake infants for vowel segments that varied in vowel place and height and were presented in an odd-ball stimulus paradigm. Perceptual measures of vowel discrimination were obtained using a visually-reinforced head-turn paradigm. The broad hypothesis was that CAEP onset-response components P1-N1-P2 latencies and amplitudes would differ as a function of vowel type, such as observed in ACC stimulus paradigms, and be correlated with infant perceptual performance. It was also hypothesized that these CAEP component latencies and amplitudes would vary with stimulus presentation rate, specifically, that a 2/s stimulus rate would yield smaller amplitudes and prolonged latencies compared to responses obtained at 1/s owing to neural adaptation. It was further predicted that the difference waveforms for frequent vs. infrequent tokens, sometimes referred to as mismatch responses (MMR), would have larger amplitudes at the 2/s rate because of the neural adaptation to the frequently presented token.

Methods

Subjects

Twenty infants, aged 4.0–11.8 months participated in the experiment. The group included 6 females. The mean age at the time of the CAEP test was 5.4 months, while the mean age at the time of the behavioral test was 6.8 months. All infant participants were born at full-term and had passed their newborn hearing screens. None had risk factors for hearing or neurologic impairment. All infants had normal otoscopic and tympanometric test findings and passed a distortion product otoacoustic emission (DPOAE) screening test at the time of admission to the study. Otoscopy and tympanometry were repeated if a parent reported that the infant had experienced an upper respiratory or ear infection prior to the visit. Data were obtained only when otoscopy and tympanometry were normal. Infants were allowed to participate in up to 5 test sessions to obtain the electrophysiologic and perceptual data. The data set reported here is part of a larger study in which there were 32 infant participants tested with a large combination of vowel contrasts. In this report, /a/ (standard) vs. /i/, /o/ or /u/ (deviant) contrasts are reported, as they are representative of the primary findings for all vowel contrasts tested. Only those subjects that had CAEP control trials for the /a/ stimulus (described below) were included in this data set.

Stimuli

The stimuli were synthesized (Story, 2011) vowel tokens: /a/, /i/, /o/, /u/, each with a 500 ms total duration, and shaped with a 10 ms linear onset-offset ramp. The spectra of these tokens are shown in Figure 1. Synthesized, rather than natural speech tokens were used so that the fundamental and formant frequencies and their levels could be precisely controlled. The stimuli were presented at 70 dB SPL in the sound field, via a JBL Control 1× model speaker with a Crown D-75A amplifier. These stimuli were calibrated in the sound field using a Larsen Davis Model 824 sound pressure level meter using with a half-inch microphone suspended from the ceiling in the position approximating the position of the infant's head during testing. During calibration, the samples were presented in a steady-state fashion with an inter-stimulus interval of < 1.0 ms.

For CAEP tests, stimuli were presented in pairs with one standard (/a/) token and one deviant token using an oddball paradigm with a deviant probability of 25%. There was also a “control” condition of no vowel change in which the standard and deviant tokens were both /a/. CAEP tests were run at both a 1 Hz and 2 Hz stimulus rate. The 2 Hz rate resulted in a quasi-steady-state stimulus train, as the duration of each token was 500 ms.

For the perceptual discrimination test, two vowel tokens were concatenated for a total stimulus duration of 1000 ms. A change or “contrast” stimulus consisted of two tokens in which the second token was different from first (e.g., /a+/i/). The “no change” stimulus in this case consisted of two identical 500 ms tokens, /a+/a/). The rate of stimulus presentation for the vowel-change discrimination test was 0.4 Hz.

CAEP Vowel-Change Test

CAEPs were recorded using silver-silver chloride disposable electrodes placed at Cz (vertex, non-inverting), A2 (right mastoid, inverting), and A1 (left mastoid, ground) using electrode paste and paper tape, after cleansing each site with Nu-Prep. Electrode impedances were maintained at $<10\text{ k}\Omega$ and with $<3\text{ k}\Omega$ inter-electrode impedance. If an electrode became displaced during testing it was re-placed and electrode impedances checked prior to resuming recording. The EEG was amplified by a factor of 94 dB and filtered at 1–30 Hz and digitized at 0.5 kHz. Amplitude-based artifact rejection level was set at between 50–90 μV , levels that have proven efficacious in past research (Wunderlich et al., 2006; Cone & Whitaker, 2013).

All recordings were made using Intelligent Hearing Systems Smart-EP system. The MMN/P300 software module was used to present the stimuli in an oddball paradigm, and obtain separate waveforms in response to the standard (75% probability) and deviant (25% probability) tokens. Responses to the standard and deviant stimuli were averaged and stored in separate buffers. At least 50 artifact-free samples were obtained in response to the “deviant” stimulus for each contrast tested. The samples were averaged over a 500 ms epoch following the onset of each token. Each vowel contrast was tested using a rate of 1 Hz and 2 Hz, and control waveforms (no stimulus change) were also tested at both rates.

Infants were tested while awake and seated in a high chair or on their parent’s lap. A test assistant manipulated toys to keep the infants in a quiet but alert state conducive for CAEP tests.

Vowel-discrimination Test

Infants were trained and tested using visually-reinforced operant procedures with an observer-based psychophysical method as described previously (Cone & Garinis; 2009; Cone & Whitaker, 2013). The Intelligent Hearing Systems VRISD software module was used to present the stimuli and record the response.

Infants were brought into the test booth and seated in a high chair while the “control” stimulus was presented continuously; the stimulus used for this habituation procedure was the 1000 ms token created by concatenating two samples of the same token, e.g. /a/-/a/. Infants were trained to respond to a change in the speech stimulus. The training consisted of pairing the presentation of a change token (e.g. /a/-/i/) with visual reinforcement, which was an animated cartoon played on a video screen, placed at 45 degrees with respect to the infant’s head position. The pairing of the reinforcer with the change trial was used to teach the infant to emit a behavior that could be used during the testing phase. The behavior could be a head turn, an eye movement towards the reinforcer, a change or cessation of facial expression, or increase in body movement (Cone & Garinis; 2009; Cone & Whitaker, 2013). Training consisted of 5 pairings of the change trial with the reinforcer. Between change trials, the background, control stimulus (/a/-/a/-/a/-/a/) was played continuously. After 5 pairings, a probe trial was used in which the stimulus change was presented, but the reinforcer was withheld until the infant emitted a response. If the infant responded, the reinforcer was introduced. The infant had to respond correctly to two/three probe trials

before testing began. If the infant did not reach this criterion, 5 more training trials were given, followed by two probe trials. Up to 15 training trials were given per test session, but vowel discrimination testing was discontinued if the infant did not meet criterion.

The testing phase consisted of 12 trials for a given vowel contrast. During testing, the observer controlled when a trial was initiated, but, with all auditory monitors and feed-back turned off, was therefore masked as to whether the trial contained a vowel change token or not. Up to six trials per test were control trials. The order of trials was automatically randomized by the computer software. The observer voted on each trial by pressing a response button, that is, the observer had to determine whether the trial contained a vowel change based solely on observation of the infant's behavior. When the observer made a correct detection, the infant was given the visual reinforcer. When the observer made a correct rejection (no response on a control trial) there was no reinforcer. Likewise, there was no reinforcer for a miss (failure to detect a change trial) or a false alarm (voting that a change token had occurred on a control trial). The percentage of correct responses for both change and control trials was determined.

During all training and testing a test assistant manipulated toys at mid-line to keep the infants centered and in an appropriate response state for the vowel discrimination test.

Procedure

Infants could participate in the CAEP vowel-change test and the perceptual vowel-discrimination test on the same day. The order of test (CAEP or perceptual) was not controlled. Typically, the test session started with a perceptual test for one vowel contrast; those methods are described, above. Then, CAEPs were tested. For CAEP tests, the test order for contrasts was /i/, /o/ then /u/ for all subjects, but the order for presentation rate (1 Hz vs. 2 Hz) was randomized. The CAEP control trial blocks (no change in vowel token) were run early in the test session, typically after blocks to obtain waveforms for a vowel contrast condition had been run.

Data Analyses

CAEP Data Analyses—CAEP waveforms were analyzed using rule-based visual detection methods, by the principal investigator (BC) and two research assistants who had been trained by the principal investigator. The rules for visual detection of CAEP components included latency and amplitude criteria for each component, based upon values derived from the published literature on infant CAEP (Wunderlich & Cone-Wesson, 2006). CAEP component peaks P1, N1, P2, and N2 were marked with a cursor to determine peak latency, and amplitudes were calculated as the difference between the peak and succeeding trough, or trough-to peak, e.g., P1-N1 or N1-P2. Derived waveforms were created to normalize the differences in CAEPs obtained in response to vowel contrasts. The derived waveform consisted of subtracting the response to the standard token (/a/) from the response to the deviant token: /i/, /o/ or /u/ for the contrast trials, and /a/ in the control condition. Component latencies and amplitudes were measured for the derived waveforms.

Vowel Discrimination Test Analyses—The IHS Smart VRA software records observer responses for each trial. From the total number of trials, the percentage of hits (correct detection of stimulus change), misses (failure to detect stimulus change), false alarms (detection in the absence of stimulus change) and correct rejections (correct judgment of a control trial) were calculated. A group mean d' value was calculated following the methods described in Cone & Whitaker (2013).

Descriptive and inferential statistical analyses (analyses of variance) were completed using Statview (v5.0.2) software.

Results

CAEP

Responses from a representative infant are shown in Figure 2 for vowel change trials in which /a/ is the standard stimulus. Responses obtained at a stimulus rate of 1/s are shown in Frame A and those obtained at 2/s are in Frame B. The waveforms are averaged over 500 ms from vowel onset (in the case of the /a/ standards, presented at a probability of 75%) or 500 ms from vowel change onset (in the case of the deviant token presented with a 25% probability). The derived waveforms are shown in Frames C and D for the two presentation rates, respectively. In general, the P1-N1 complex was evident in response to standard and deviant tokens. The P2-N2 complex is less evident in responses obtained at 2/s compared to those obtained at 1/s. The derived waveforms of CAEPs obtained at 2/s show large differences in component amplitude when control vs. contrast trials are compared. These are not obvious at 1/s.

Table 1 summarizes the percentage of CAEP components evident in each test condition. Considering the results for the stimulus rate of 1/s, P1-N1 was the most frequently present component. P1-N1 was obtained in 92% of all conditions when the /a/ token was used as the standard and another vowel token was used as the deviant. For the control condition in which /a/ served as the standard and the deviant, the probability of obtaining a P1-N1 dropped by over 20% compared to the standard-contrast condition. For the contrast conditions, P1-N1 components were present for the /u/ deviants at a decreased rate compared to /i/ and /o/ deviants. This pattern of component presence was also found for N1-P2 and P2-N2 although overall presence for later components was reduced in comparison to P1-N1. The column labeled, “Any”, indicates the presence (in percent) of any response component for a given stimulus condition. The final column, all absent, indicates the percentage of trials in which no CAEP components were obtained. Responses to the /a/ were present in 96% of trials when it was used as the standard in the contrast condition (i.e. with /i/, /o/ or /u/ deviants). Overall, there were no significant differences in response rate (84%) for the /a/ control deviant condition compared to the response rates for /i/, /o/ or /u/ contrast deviant conditions at 86%, 86% and 80%, respectively.

When the stimulus was presented at 2/s, there were larger differences in response rates for individual or any CAEP components as a function of stimulus type. P1-N1 was the most frequently observed CAEP component across conditions, and the later latency components were less frequently observed. The /o/ deviant token evoked a P1-N1 in 90% of the trials,

whereas the /a/ deviant token in the control condition (/a/ as both standard and deviant) evoked the P1-N1 in only 63% of trials. Considering all deviant conditions and the presence of any CAEP component, it is clear that the /i/, /o/ and /u/ deviants had the highest response rates at 81,94 and 91%, respectively, whereas a CAEP was only present for the /a/ deviant in the control condition 69% of the time. The conditions with the highest percentage of all absent components were for the /a/ standard token, when contrasted with deviants /i/, /o/ or /u/ and for the /a/ token used as a deviant in the control condition, when /a/ was also the standard.

Figure 3 displays mean amplitudes for each CAEP component as a function of stimulus type and rate. When components were absent, the amplitude value was entered as 0. As can be seen in the figure, the CAEP component amplitudes were larger for contrasting deviants compared to the control /a/ deviant condition. Also, CAEP amplitudes for the /a/ standard condition were significantly reduced at the 2/s rate compared to those at 1/s, although the responses to deviant tokens at 2/s were the same amplitude or larger than at 1/s. The P1-N1, N1-P2 and P2-N2 component amplitudes were then averaged across components to form one dependent amplitude variable. Repeated measures analyses of variance were used to assess differences in component amplitude due to rate (1/s vs. 2/s) and stimulus condition (/a/ standard, and /a/, /i/, /o/ and /u/ as deviants). A significant interaction of rate and stimulus condition was found ($F_{1,4} = 13.049, p < .0001$) so separate analyses were completed for each rate. At a stimulus rate of 1/s the effect of stimulus type on CAEP amplitude approached significance ($F_{4,304} = 2.22, p = .07$). Table 2 summarizes the mean amplitude differences for each of the stimulus contrasts, with the associated critical differences and p-values from a post hoc test using Bonferroni-Dunn corrections. None of the contrasts reached statistical significance ($p < 0.005$ for the post hoc test). The largest differences were observed for the /a/ deviant and /i/ deviant contrasts. This means that the amplitude of the response in response to the /i/ deviant was larger than in the control condition when /a/ was used as both standard and deviant. The amplitude difference between the /a/ standard to the /i/ deviant contrast was also large, and trended towards significance.

At the 2/s rate, the effect of stimulus type was statistically significant, $F_{4,309} = 37.33, p < .0001$. Table 2 summarizes the results of the Bonferroni-Dunn post-hoc tests and these reveal that all contrast conditions with /i/, /o/ or /u/ deviants produced a significant difference in CAEP amplitude. Also, comparisons of the amplitude for /a/ used as deviant in the control condition, with all other contrast deviant tokens were significant. The control condition (/a/ standard and /a/ deviant) did not yield a significant difference for standard vs. deviant, in fact the mean amplitude difference for this condition was two orders of magnitude smaller than in any contrast condition.

Statistical analyses of the derived waveform data were also completed. The mean component amplitudes in the derived waveforms are shown in Figure 4; means were calculated using an amplitude value of 0 for missing components. The amplitudes for derived components P1-N1, N1-P2 and P2-N2 were averaged for each subject to form one dependent amplitude factor. The effect of stimulus type on derived waveform amplitude was determined separately for derived waveforms for 1/s and 2/s. The stimulus types were /a/-/a/ (control), /i/-/a/, /o/-/a/ and /u/-/a/, with the last 3 pairs designated as contrasts. There were

no statistically significant differences in amplitude owing to stimulus type for derived responses obtained at 1/s ($F_{3,90}=0.796$, $p=.50$). When derived waveform amplitudes at 2/s are considered, the effect of stimulus type was significant ($F_{3,86}=6.21$, $p=.0007$). Inspection of the amplitude means plotted in Figure 4 for the rate condition of 2/s indicates that in most cases the derived waveform amplitudes for the contrast stimuli are nearly twice the amplitude differences found for the control (/a-/a/) stimuli. Bonferroni-Dunn post hoc comparisons are shown in Table 2. The /i-/a/ and /u-/a/ difference waveforms were significantly larger than the control difference waveform (/a-/a/), however the /o-/a/ vs. control difference did not reach significance.

CAEP peak latencies were evaluated. The mean peak latencies for each component are plotted in Figure 5. Repeated measures analyses of variance were used to evaluate if there were significant differences in latency when responses to standard tokens /a/ were compared to those used as contrasts (/i/, /o/ and /u/). These analyses were completed separately for the two stimulus rates. It should be noted that “missing” cells for latency in the repeated measures analysis of variance skew the latency values to the latency of the present components. So, given that P1 and N1 were obtained more often than later peaks, the results would be skewed towards earlier latencies. The effect of stimulus type on peak latency was significant ($F_{3,107}=4.23$, $p=.0071$) at 1/s but not for 2/s. There were fewer observations available at 2/s rate because of missing data when responses were absent owing to neural adaptation of the CAEP for the standard stimulus. This likely affected the statistical power at this rate.

For the derived waveforms, significant differences in component latencies owing to stimulus type were seen only at the 2/s rate ($F_{3,48}=4.69$, $p=.0059$). Latencies were prolonged for the contrast conditions relative to the control conditions. These were most apparent for the later components, P2 and N2 (see Figure 6). Bonferroni-Dunn post-hoc tests revealed that the difference in latency for the /a-/a/ control difference wave form and /u-/a/ difference contrast was significant.

Perception of Vowel-Contrasts

Perceptual results indicate that, on average, there was a 68.5 % hit-rate for detecting a vowel change, and a 25.7% false alarm rate. The distribution of hit rates obtained for all of the vowel contrast tests ($N=57$) reveals that the hit rate was less than 50% for 15.8% of the tests. The distribution of false alarm rates indicated that 28.1% of the trials resulted in a false alarm rate of greater than 50%. Table 3 summarizes the hit and false alarm rates for the vowel contrast trials. The highest hit rates were found for the /a-/u/ contrasts at 77.1%, but the hit rate was only 58.8% for the /a-/o/ contrast. False alarm rates were also highest for the /a-/u/ contrast and lowest for the /a-/o/ contrast. An analysis of variance for hit rate as a function of contrast type reached statistical significance ($F_{2,54}=3.38$, $p<.05$), but the analysis of variance for false alarm rate did not reach significance. A mean d' value of 1.134 was calculated based upon the mean hit and false alarm rates across all vowel contrast tests and the 20 subjects tested.

The performance advantage for the /a-/u/ contrast in light of the significant differences found in the CAEP derived waveform latency (at the 2/s rate) for this contrast suggested that

there may be a correlation between the hit rate and CAEP response parameters. The correlation of CAEP component amplitudes in the derived waveforms with hit rate resulted in only two instances in which the correlation value exceeded 0.60, which were for P1-N1 amplitude for the /a/-i/ contrast ($r = 0.62$) and the P2-N2 amplitude for the /a/-u/ contrast ($r = 0.76$). Correlation of hit rate with CAEP latencies for the derived waveform have only 3 instances in which the r-value exceeded 0.6: N2 latency for the /a/-o/ contrast ($r = 0.81$), and P2 and N2 latencies for the /a/-u/ contrast in which the correlation was $r = -0.77$ for each component. These correlations suggest that the /a/-u/ contrast had more perceptual salience and resulted in CAEP difference waveforms with larger amplitudes and shorter latencies than other contrasts. Yet, caution must be exercised in interpretation of these results owing to the fact that they are based upon a limited data set of 8 instances in which hit rate, and CAEP data were available for all contrasts. CAEPs are obtained from infants and for vowel contrasts that do not produce a reliable behavioral response. This is evidenced by the high percentage of CAEPs present for the vowel contrasts (Table 1) whereas, the average detection rate (hit rate) for the vowel contrast was only 68.5%, and the mean d' value only 1.134.

Discussion

The immature cortex has the ability to encode a change in vowel spectrum and this is evident as an onset response. As can be seen from their frequency spectra (Figure 1), the fundamental frequency of the voice was held constant so only the change in formant energy would be the basis for a response that differed significantly from the control condition of no vowel change. In fact, an onset response to the deviant /a/ in the control condition (when /a/ was both standard and deviant), was observed in nearly all trials. This is likely due to the fact that all tokens had 10 ms linear onset-offset ramps; furthermore, they were presented in quasi-steady state manner such that the train of tokens was a complex periodic waveform with a triangular modulation envelope. Yet, the averaged and derived waveforms revealed much larger CAEP amplitudes for the conditions in which a deviant stimulus differed from the standard stimulus, at least when vowel tokens were presented at a rate of 2/s.

The question could be asked whether the P1-N1-P2 potential obtained for the vowel contrast is what Martin and others have labeled Acoustic Change Complex (ACC) (Martin & Boothroyd, 1999; 2000; Martin et al., 2008) or whether it could also be interpreted as mismatch negativity (MMN), owing to the use of an odd-ball stimulus paradigm (Naatanen et al., 2004) and the deviant-standard derived waveforms. The results obtained using a 1/s rate appear to be qualitatively similar to previous MMN investigations. That is, at 1/s, the derived waveforms for the contrast conditions were not consistently different from the control condition when there was no vowel change. This is consistent with the poor response-to-noise ratio for MMN, which makes it difficult to detect in individual subjects and thus cannot be relied upon as an indicator of neural capacity for perceiving stimulus change (Wunderlich & Cone-Wesson, 2001). Additionally, the infant MMN is reportedly smaller in amplitude, more diffuse, and not analogous to the adult MMN (Cheour et al., 2000). Only 50 artifact-free sweeps were obtained in response to the deviant stimulus, and this would further diminish the probability of defining the MMN. Furthermore, the mismatch responses in infants at this age are noted to exhibit a late positive peak rather than

a negativity and at prolonged latencies relative to what is observed in older children or adults (Choudhury & Benasich, 2011; Dehaene-Lambertz & Dehaene, 1994).

Conversely, at the fast rate of 2/s, large amplitude differences were found between the control and contrast conditions. At this rate, it was obvious there were very few definable peaks in response to the standard stimulus. The response to the standard appeared to be completely adapted. Thus, the introduction of a contrasting stimulus at an effective rate of 0.25 Hz was more likely to evoke a response to the deviant contrast, as adaptation for the improbable stimulus would be minimal at this rate. The fact that there were responses to the /a/ deviant in the control (no stimulus change) condition indicates the extent to which rate plays a role in the response to the infrequent stimulus. These rate effects on onset responses are particularly marked for CAEPs in infants (Wunderlich et al., 2006) and children (Gilley et al., 2005; Sussman et al., 2008). This type of recovery from neural refractoriness has also been discussed as a mechanism contributing to MMN (Morr et al., 2002). The difference waveforms for contrast conditions had significantly larger amplitudes than those found in the control condition, indicating that this was attributable to the change in vowel spectrum.

The results of the present study are similar to those of Hämäläinen et al. (2011), who obtained mismatch responses in a group of awake infants aged 6 months of age. The stimulus paradigm used by Hämäläinen et al. had standard tone pairs in which the frequencies of both tokens were the same (100 Hz), and deviant tone pairs in which the second token was 200 Hz higher (300 Hz) than the first token. In this case, the salient cue was a frequency difference between the tone pair. Because they tested the same age of infants as in the current study, and they were testing for response differences to a frequency change, their findings are relevant to the current experimental results. They found large positive-negative going components in the CAEP response to deviants that are similar in latency and amplitude to those found in the present study (see their Figure 2). They also performed source-analysis from their multi-channel recording and describe slightly different dipole sources for their positive and negative components of the waveform. They interpret the negative component as a pre-cursor of the adult mismatch negativity, “involved in the processing of feature-specific sensory information.”

On the other hand, Small & Werker (2012) obtained CAEPs from 4-month old infants using consonant vowel (CV) tokens (e.g., /da/) and also two syllable CV-CV tokens (e.g., /dada/). The single syllable tokens evoked a P1-N1-P2 onset response. When using two syllable tokens they also observed an onset response for the first token, although a second onset response for the second syllable was highly variable across different CVs. A P1-N1 complex for the second syllable onset was observed in all infants only for the /daba/ token, and inconsistently for other two-syllable token types. Some infants exhibited a broad positive or negative peak following the second syllable, but this was difficult to interpret as a response to stimulus change because this also occurred in some of the responses to single-syllable tokens. Small & Werker concluded that the ACC could be obtained in 4 month old infants, given their results for the /daba/ token.

It could be argued that the response to the vowel change obtained in the present experiment could contain components of mismatch negativity, given that an oddball stimulus paradigm was used, and these results are similar to others testing similarly aged infants with an oddball paradigm (Hämäläinen et al., 2011). Yet, the statistically significant results for responses to stimulus change were found for the quasi-steady state stimulus condition (2 tokens/s), which is similar to paradigms used by Dimitrijevic et al. (2009) to obtain ACC. More experimentation is needed to determine whether the responses obtained with the quasi-steady state paradigm employing standard and deviant tokens are those of the ACC, the MMN, or a combination of both potentials superimposed on one another. Varying the rate of stimulation, the probability of deviant tokens, and recording using dense-array EEG and subsequent source localization analysis may help to disambiguate the nature of this response. For the present, “acoustic mismatch potential” (AMP), seems to be a reasonable compromise in terms of what to label this response.

Infant Perceptual Performance and Cortical Electrophysiology

Previous findings with regard to infant speech feature discrimination performance and the presence of an evoked response to multiple components of a complex sound simulating speech features indicated a shared dependence upon stimulus level (Cone & Garinis, 2009). More recent work (Cone & Whitaker, 2013) revealed a discrepancy between infant perceptual thresholds for speech and tonal tokens and the CAEP amplitude input-output functions. CAEP “thresholds” (i.e., the presence of a physiological response) estimated from these input-output functions showed CAEPs obtained at levels below perceptual threshold for the same stimulus.

This discrepancy between perceptual and electrophysiologic results had previously been explored by Werner et al. (1993; 1994). They showed that ABR thresholds for tonal stimuli had obtained adult values by 6 months of age, whereas perceptual thresholds in infants were still elevated with respect to adults. They further found that the wave I – V inter-peak latencies, or central conduction times, were correlated with the perceptual threshold, and so they speculated that the attainment of threshold would be dependent upon maturation of the auditory brainstem and cortex. An initial hypothesis of the work comparing cortical electrophysiology to perceptual results in infants (Cone & Whitaker, 2013) was that there would be a close correspondence between them. This does not appear to be the case for either detection, or, based upon these data from the current experiment, discrimination tasks. CAEPs show clear evidence of neural encoding of acoustic differences between adjacent stimuli, but the infant behavioral response is not always detectable, even after training. Yet, there were some correlations that were evident between the two measures. Specifically, the hit rates were highest for discrimination of the /a/ vs. /u/ contrasts and the CAEP difference waveforms demonstrated the largest amplitudes and shortest latencies for this contrast. The CAEP differences can be explained acoustically: contrasts with larger acoustic differences yield MMNs with larger amplitudes and shorter latencies (Sams et al., 1985). Similarly, a shift in spectral content from high to low appears to evoke larger ACCs than in the opposite direction (Martin & Boothroyd, 2000). The vowel token /a/ used in this study had slightly higher mid-high frequency content than did the /u/, which had more energy in the lower part of the spectrum owing to the formant differences. It is well known that unit responses of the

auditory cortex also show sensitivity to direction of change for frequency-modulated stimuli (Mendelson et al., 1994). It may be that the acoustic salience of the /a/ vs. /u/ contrast, in the direction of high to low frequency change, was also an effective cue for better perceptual performance.

The CAEP results from this age group appear to reflect mainly sensory encoding of acoustic features of the stimulus. The links between the electrophysiologic indices of vowel feature discrimination and perceptual abilities are not strong, which is likely due to the immaturity of the primary and association auditory cortices. The refinement of perceptual abilities is certainly dependent upon the maturation of the primary auditory cortex and its association areas. Although CAEP components P1-N1-P2 are evident in infants under 1 year of age, they are quite prolonged in latency and the scalp distribution of these potentials is diffuse and immature (Wunderlich et al., 2006). Anatomical studies indicate that inter-layer connections of cortical auditory areas are not yet developed and auditory stimulation during this sensitive period of early infancy is needed to establish them fully (Kral & Eggermont, 2007). The auditory cortex can encode the acoustic differences in vowel stimuli, but further cortical maturation must be necessary in order for these differences to affect perception and behavioral responses. The immaturity of these cortical connections may be such that discrimination behavior cannot consistently be observed, even with short-term training.

These findings may have translation to clinical applications, despite the weak association between behavioral evidence of vowel contrast perception and cortical electrophysiology. Reliable electrophysiologic techniques for evaluating single subject speech-feature detection and discrimination would be of tremendous benefit for diagnostic and rehabilitative audiology and speech pathology. Such methods would have applications for assaying the sensory encoding abilities of infants with hearing loss and those with normal hearing but who are at risk for developmental communication disorder and language impairment (Choudhury & Benasich, 2011). These methods could also be used to document the effects of treatment, such as the provision of hearing aids or cochlear implants in the case of hearing loss and assistive devices to improve the signal to noise ratio of the acoustic signal in those with language based learning disorders (Hornickel et al. 2012). These preliminary results for the AMP, obtained with an innovative stimulus and recording paradigm, show promise for translation to a clinical procedure for documenting the cortical electrophysiologic processes and neural substrates underlying speech-feature perception.

Acknowledgements

Drs. Kristin Camerata and Jessie Liu Ross are gratefully acknowledged for their assistance in data collection and analysis. Spencer Smith, B.S., is acknowledged for his assistance with data analysis and design of the figures. Drs. Huanping Dai and Mark Borgstrom provided consultation regarding statistical analyses of the data.

This work was supported in part by NIH-NIDCD K24 DC 008826 to Barbara Cone.

Bibliography

Benasich AA, Thomas JJ, Choudhury N, Leppänen PHT. The importance of rapid auditory processing abilities to early language development: Evidence from converging methodologies. *Dev Psychobiol.* 2002; 40:278–292. [PubMed: 11891639]

- Berg KM. Auditory temporal summation in infants and adults: Effects of stimulus bandwidth and masking noise. *Percept Psychophys.* 1991; 50(4):324–320.
- Berg KM, Boswell AE. Temporal summation of 500-Hz tones and octave band noise bursts in infants and adults. *Percept and Psychophys.* 1995; 57(2):183–190.
- Carter L, Golding M, Dillon H, Seymour J. The detection of infant cortical auditory evoked potentials (CAEPs) using statistical and visual detection techniques. *J Am Acad Audiol.* 2010; 21(5):347–356. [PubMed: 20569668]
- Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiol.* 2010; 47(2):236–246.
- Cheour M, Leppänen PH, Kraus N. Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clin Neurophysiol.* 2000; 111(1):4–16. [PubMed: 10656505]
- Choudhury N, Benasich AA. Maturation of auditory evoked potentials from 6 to 48 months: Prediction to 3- and 4-year language and cognitive abilities. *Clin Neurophysiol.* 2011; 122:320–338. [PubMed: 20685161]
- Cone B, Garinis A. Auditory steady state responses and speech feature discrimination in infants. *J Am Acad Audiol.* 2009; 20(10):629–643. [PubMed: 20503801]
- Cone B, Whitaker R. Dynamics of infant cortical auditory evoked potentials (CAEPs) for tone and speech tokens. *Int J. Pediatr Otorhinolaryngol.* 2013; 77(7):1162–1173. [PubMed: 23722003]
- Dehaene-Lambertz G, Dehaene S. Speed and cerebral correlates of syllable discrimination in infants. *Nature.* 1994; 370(6487):292–295. [PubMed: 8035876]
- Dimitrijevic A, Lolli A, Michalewski HJ, Pratt H, Zeng F-G, Starr A. Intensity changes in a continuous tone: Auditory cortical potentials comparison with frequency changes. *Clin Neurophysiol.* 2009; 120:374–383. [PubMed: 19112047]
- Eilers RE, Wilson WR, Moore JM. Developmental changes in speech discrimination in infants. *J Speech Hear Res.* 1977; 20:766–779. [PubMed: 604689]
- Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. Speech perception in infants. *Science.* 1971; 171(968):303–306. [PubMed: 5538846]
- Eimas PD. Segmental and syllabic representations in the perception of speech by young infants. *J Acoust Soc Am.* 1999; 105(3):1901–1911. [PubMed: 10089612]
- Eisenberg, LS.; Martinez, AS.; Boothroyd, A. Perception of phonetic contrasts in infants: Development of the VRASPAC. In: Miyamoto, RT., editor. *Cochlear Implants. International Congress Series.* Vol. 1273. Amsterdam: Elsevier; 2004. p. 364-367.
- Eisenberg LS, Martinez AS, Boothroyd A. Assessing auditory capabilities in young children. *Int J Pediatr Otorhinolaryngol.* 2007; 71:1339–1350. [PubMed: 17604127]
- Gilley PM, Sharma A, Dorman M, Martin K. Developmental changes in refractoriness of the cortical auditory evoked potential. *Clin Neurophysiol.* 2005; 116:648–657. [PubMed: 15721079]
- Golding M, Pearce W, Seymour, Cooper J, King A, Ching T, Dillon H. The relationship between obligatory cortical auditory evoked potentials (CAEPs) and functional measures in young infants. *J Am Acad Audiol.* 2007; 18(2):117–125. [PubMed: 17402298]
- Hämäläinen JA, Ortiz-Mantilla S, Benasich AA. Source localization of event-related potentials to pitch change mapped onto age-appropriate MRIs at 6 months of age. *Neuroimage.* 2011; 54(3):1910–1918. 2011. [PubMed: 20951812]
- Hecox K, Galambos R. Brain stem auditory evoked responses in human infants and adults. *Archives of Otolaryngology.* 1974; 99:30–33. [PubMed: 4809084]
- Hornickel J, Zecker SG, Bradlow AR, Kraus N. Assistive listening devices drive plasticity in children with dyslexia. *Proceedings of the National Academy of Sciences of the United States of America.* 2012; 109(41):16731–16736. [PubMed: 22949632]
- Jusczyk, PW.; Houston, D.; Goodman, M. Speech perception during the first year. In: Slater, A., editor. *Perceptual Development: visual auditory and speech perception in infancy.* East Sussex, U.K: Psychology Press; 1998. p. 357-388.
- Kirk, KI.; Choi, S. Clinical investigations of cochlear implant performance. In: Niparko, John K., editor. *Cochlear Implants, Principles and Practices.* 2nd Edition. Philadelphia: Lippincott Williams and Wilkins; 2009.

- Kral A, Eggermont JJ. What's to lose and what's to learn: Development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Res Rev.* 2007; 56:259–269. [PubMed: 17950463]
- Kuhl, PK. Psychoacoustics and speech perception: internal standards, perceptual anchors and prototypes. In: Werner, LA.; Rubel, E., editors. *Developmental Psychoacoustics*. Washington, D.C: American Psychological Association; 1992. p. 293–332.
- Kuhl PK. Early language acquisition: Cracking the speech code. *Nature Rev Neurosci.* 2004; 5:831–843. [PubMed: 15496861]
- Kurtzberg D, Hilpert PL, Kreuzer JA, et al. Differential maturation of cortical auditory evoked potentials to speech sounds in normal full-term and very low-birth weight infants. *Dev Med Child Neurol.* 1984; 26:466–475. [PubMed: 6479466]
- Martin BA, Boothroyd A. Cortical, auditory, event-related potentials in response to periodic and aperiodic stimuli with the same spectral envelope. *Ear Hear.* 1999; 20:33–44. [PubMed: 10037064]
- Martin BA, Boothroyd A. Cortical, auditory, evoked potentials in response to changes of spectrum and amplitude. *J Acoust Soc Am.* 2000; 107:2155–2161. [PubMed: 10790041]
- Martin BA, Tremblay KL, Korczak P. Speech Evoked Potentials: From the laboratory to the clinic. *Ear Hear.* 2008; 29:285–313. [PubMed: 18453883]
- McCreery RW, Stelmachowicz PG. Audibility-based predictions of speech recognition for children and adults with normal hearing. *J Acoust Soc Am.* 2011; 130(6):4070–4081. [PubMed: 22225061]
- McMurray B, Aslin RN. Infants are sensitive to within-category variation in speech perception. *Cognition.* 2005; 95(2):B15–B26. [PubMed: 15694642]
- Mendelson JR, Schreiner CE, Sutter ML, Grasse KL. Functional topography of cat primary auditory cortex: responses to frequency modulated sweeps. *Exp Brain Res.* 1993; 94:65–87. [PubMed: 8335076]
- Moeller MP, Stelmachowicz PG, Hoover B, et al. Vocalizations of infants with hearing loss compared to infants with normal hearing. Part I—Phonetic development. *Ear Hear.* 2007; 28:605–627. [PubMed: 17804976]
- Molfese DL, Molfese VJ. Discrimination of language skills at five years of age using event-related potentials recorded at birth. *Dev Neuropsychol.* 1997; 13(2):135–156.
- Morr ML, Shafer VL, Kreuzer JA, Kurtzberg D. Maturation of mismatch negativity in typically developing infants and preschool children. *Ear & Hear.* 2002; 23(2):118–136.
- Naatanen R, Pakarinen S, Rinne T, et al. The mismatch negativity: towards the optimal paradigm. *Clin Neurophysiol.* 2004; 115:140–144. [PubMed: 14706481]
- Nittrouer S. The role of temporal and dynamic signal components in the perception of syllable-final stop voicing by children and adults. *J Acoust Soc Am.* 2004; 115(4):1777–1790. [PubMed: 15101656]
- Nittrouer, Susan. Dynamic spectral structure specifies vowels for children and adults. *J Acoust Soc Am.* 2007; 122(4):2328–2339. [PubMed: 17902868]
- Nittrouer S, Lowenstein JH. Children's weighting strategies for word-final stop voicing are not explained by auditory sensitivities. *J Speech Lang Hear Res.* 2007; 50(1):58–73. [PubMed: 17344548]
- Novak GP, Kurtzberg D, Kreuzer JA, Vaughan HF Jr. Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. *Electroencephalogr Clin Neurophysiol.* 1989; 73(4):295–305. [PubMed: 2477216]
- Rance G, Cone-Wesson B, Wunderlich J, Dowell R. Speech and cortical event related potentials in children with auditory neuropathy. *Ear Hear.* 2002; 23(3):239–253. [PubMed: 12072616]
- Sams P, Paavilainen P, Alho K, Näätänen R. Auditory frequency discrimination and event-related potentials. *Electroencephalogr and clin Neurophysiol.* 1985; 62:437–448.
- Sharma A, Cardon G, Henion K, Roland P. Cortical maturation and behavioral outcomes in children with auditory neuropathy spectrum disorder. *Int J Audiol.* 2011; 50(2):98–106. [PubMed: 21265637]

- Sharma A, Dorman MF, Kral A. The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. *Hear Res.* 2005; 203:134–143. [PubMed: 15855038]
- Sharma A, Dorman MF, Spahr AJ. A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear.* 2002; 23:532–539. [PubMed: 12476090]
- Skoe E, Kraus N. Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* 2010; 31:302–324. [PubMed: 20084007]
- Small SA, Werker JF. Does the ACC have potential as an index of early speech discrimination ability? A preliminary study in 4-month-old infants with normal hearing. *Ear Hear.* 2012; 33(6):e59–e69. [PubMed: 22785572]
- Stager CL, Werker JF. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature.* 1997; 388:381–382. [PubMed: 9237755]
- Story, B. TubeTalker: An airway modulation model of human sound production. In: Fels, Sidney; d'Allessandro, Nicolas, editors. *Proceedings of the First Annual Workshop on Performative Speech and Singing Synthesis.* Vancouver British Columbia, Canada: 2011 Mar. p. 14-15.
- Sussman E, Steinshneider M, Gumenyuk V, Grushko J, Lawson K. The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hear Res.* 2008; 236:61–79. [PubMed: 18207681]
- Trehub SE. Reflections on the development of speech perception. *Canadian J Psychol.* 1979; 33(4): 368–381.
- Werker JF, Gilbert JHV, Humphrey K, Tees RC. Developmental Aspects of Cross-Language Speech Perception. *Child Dev.* 1981; 52:349–355. [PubMed: 7238150]
- Werker JF, Tees. Cross-language speech perception: Evidence for perceptual re-organization during the first year of life. *Infant Behav. Dev.* 1984; 7:49–63.
- Werker JF, Tees RC. Influences on infant speech processing: toward a new synthesis. *Annu Rev Psychol.* 1999; 50:509–535. [PubMed: 10074686]
- Werker, JF.; Shi, R.; Desjardins, R.; Pegg, JE.; Polka, L.; Patterson, M. Three methods for testing infant speech perception. In: Slater, A., editor. *Perceptual Development: Visual, Auditory and Speech Perception in Infancy.* East Sussex, UK: Psychology Press Ltd; 1998. p. 389-420.
- Werner LA, Boike K. Infants' sensitivity to broad-band noise. *J Acoust Soc Am.* 2001; 109:2103–2111. [PubMed: 11386562]
- Werner LA, Folsom RC, Mancl LR. The relationship between auditory brainstem response latencies and behavioral thresholds in normal hearing infants and adults. *Hear Res.* 1994; 77:88–98. [PubMed: 7928741]
- Werner LA, Folsom RC, Mancl LR. The relationship between auditory brainstem response and behavioral thresholds in normal hearing infants and adults. *Hear Res.* 1993; 68:131–141. [PubMed: 8376210]
- Wunderlich JL, Cone-Wesson BK. Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential. *J Acoust Soc Am.* 2001; 109(4):1526–1537. [PubMed: 11325124]
- Wunderlich JL, Cone-Wesson BK. Maturation of the cortical auditory evoked potentials in infants and young children: A review. *Hear Res.* 2006; 212(1–2):212–223. [PubMed: 16480841]
- Wunderlich JL, Cone-Wesson BK, Shepherd R. Maturation of the cortical auditory evoked potential in infants and young children. *Hear Res.* 2006; 212(1–2):185–202. [PubMed: 16459037]

Highlights

- Infant CAEPs indicated sensory and neural capacity for discrimination of vowels.
- Perceptual measures of vowel discrimination capacity were less robust than CAEPs.
- CAEP and perceptual measures of vowel discrimination were modestly correlated.
- CAEP tests may translate clinically for evaluation of at-risk infants.

Vowel Spectra

Figure 1.

Vowel Spectra- Each of the tokens was recorded using an Etymotic ER-7c probe microphone inserted into an insert phone sound tube. The spectra were calculated using SpectraPlus v XYZ software. Note that /o/ and /u/ stimuli have a more limited bandwidth compared to the /a/ and /i/ stimuli, with limited energy for frequencies about 4 kHz.

Figure 2.

CAEP waveforms- The waveforms shown are from a representative infant participant. All waveforms were averaged over a 500 ms epoch initiated at the onset of the speech token. A) Rate = 1/s. The onset responses, P1-N1-P2, are shown for each vowel token. The token /a/ was the standard stimulus at a probability of 75% and the contrasting vowel was present with a probability of 25%. The /a/ deviant at 25% is the control condition. CAEP components P1, N1, P2, N2 are labeled for the response to the /i/ deviant token. B) Rate = 2/s. Responses to the standard and deviant /a/ tokens are of low amplitude owing to

adaptation. C) Derived waveforms, rate= 1/s. The waveforms were derived by subtracting the response to standard from the response to the deviant stimulus. The waveform for /a/-/a/ is the control condition. D) Derived waveforms, rate = 2/s. As in the original waveforms, the derived responses for the control condition, /a/-/a/, are of low amplitude, and individual components are not evident.

Figure 3.

CAEP component amplitudes. The amplitudes of CAEP components P1-N1, N1-P2, and P2-N2 are graphed for each vowel token, for Rate 1 = 1/s and Rate 2 = 2/s. Error bars indicate standard error of the mean. In general, amplitudes for contrasting vowels, /i/, /o/ and /u/ are larger than those measured in the control condition when /a/ was used for both standard and deviant.

Figure 4.

CAEP derived response component amplitudes. The amplitudes of CAEP components P1-N1, N1-P2, and P2-N2 are graphed for each derived wave, for Rate 1 = 1/s and Rate 2 = 2/s. Error bars indicate standard error. At 2/s, derived waveforms for contrast conditions are larger than those for the control condition.

Figure 5.

CAEP component latencies. The latencies of CAEP components P1, N1, P2, and N2 are graphed for each vowel token, for Rate 1 = 1/s and Rate 2 = 2/s. Latencies for each component do not vary significantly for vowel type nor for control vs. contrast conditions. Error bars indicate standard error.

Figure 6. derived response component latencies. The latencies of CAEP components P1, N1, P2, and N2 are graphed for each derived wave, for Rate 1 = 1/s and Rate 2 = 2/s. Latencies were prolonged for the contrast conditions relative to the control conditions, which were most apparent for the later components, P2 and N2. Error bars indicate standard error.

Table 1

Presence of CAEP components as a function of stimulus type and rate. Cells indicate percent of responses present for each condition.

	Percent Present				All Absent
	P1-N1	N1-P2	P2-N2	ANY	
/a/ standard (control)	78	49	47	82	18
/a/ deviant (control)	81	56	47	84	16
/a/ standard (contrast)	92	71	58	96	4
/i/ deviant	77	67	68	86	14
/o/ deviant	76	52	47	86	14
/u/ deviant	68	49	49	80	20
/a/ standard (control)	74	45	35	77	23
/a/ deviant (control)	63	31	9	69	31
/a/ standard (contrast)	59	29	20	62	38
/i/ deviant	75	66	55	81	19
/o/ deviant	90	67	49	94	6
/u/ deviant	89	72	72	91	9

Table 2

Summary of post hoc tests for CAEP component amplitude as a function of stimulus type and rate.

	Contrast		Mean difference	Critical Difference	P-value
Rate= 1/s	/a/ std, /a/ dev		0.51	3.19	0.65
	/a/ std, /i/ dev		2.39	2.59	0.01
	/a/ std, /o/ dev		0.03	2.58	0.97
	/a/ std, /u/ dev		0.41	2.57	0.65
	/a/ dev, /i/ dev		-2.9	3.48	0.02
	/a/ dev, /o/ dev		-0.54	3.47	0.66
	/a/ dev, /u/ dev		-0.92	3.46	0.45
	/i/ dev, /o/ dev		2.36	2.94	0.02
	/i/ dev, /u/ dev		1.98	2.92	0.06
	/o/ dev, /u/ dev		-0.38	2.91	0.72
Rate= 2/s					
	/a/ std, /a/ dev		0.04	2.75	0.97
	/a/ std, /i/ dev		-6.47	2.37	<.0001
	/a/ std, /o/ dev		-6.14	2.27	<.0001
	/a/ std, /u/ dev		-7.85	2.30	<.0001
	/a/ dev, /i/ dev		-6.51	3.07	<.0001
	/a/ dev, /o/ dev		-6.18	2.99	<.0001
	/a/ dev, /u/ dev		-7.89	3.02	<.0001
	/i/ dev, /o/ dev		0.32	2.65	0.73
	/i/ dev, /u/ dev		-1.387	2.68	0.14
	/o/ dev, /u/ dev		-1.71	2.59	0.06
	diff 1	diff 2			
	/a/-/a/	/i/-/a/	-5.38	4.84	0.0035
	/a/-/a/	/o/-/a/	-3.59	4.95	0.05
	/a/-/a/	/u/-/a/	-7.51	4.89	<.0001

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

		Mean difference	Critical Difference	P-value
	/o/-/a/	1.79	5.04	0.34
	/i/-/a/	-2.13	4.94	0.25
	/o/-/a/	-3.92	5.06	0.04

Table 3

Hit and false alarm rates for vowel change perception test.

	% Hits	% False Alarm
/a/-i/	70.2	31.9
/a/-o/	58.8	17.7
/a/-u/	77.1	27.0