International Journal of Psychophysiology 95 (2015) 65–76

Contents lists available at ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

Infant cortical electrophysiology and perception of vowel contrasts

Barbara K. Cone *

Speech, Language and Hearing Sciences, University of Arizona, P.O. Box 210071, Tucson, AZ 85721, United States

ARTICLE INFO

Article history:
Received 14 October 2013
Received in revised form 31 May 2014
Accepted 3 June 2014
Available online 13 June 2014

Keywords:
Infant
Auditory evoked potential
Speech perception

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/s. 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/s 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.

1. 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 and Whitaker, 2013). It is known that children weight speech-feature cues differently than do adults (Nittrouer, 2004, 2007; Nittrouer and Lowenstein, 2007), but exactly how infants process temporal and spectral speech information is not well understood (Berg, 1991; Berg and Boswell, 1995).

The ability to perceive speech-features plays an important role in theories of infant speech perception and language acquisition (Jusczyk et al., 1998). Moreover, measures of such ability are critically important for fitting and fine-tuning hearing aids (McCreery and Stelmachowicz, 2011) and cochlear implants (Kirk and Choi, 2009). Speech-feature discrimination is essential to language development, in the segmenting of words and in assigning meaning to words (Stager and Werker, 1997; McMurray and Aslin, 2005). Classic studies of speech feature discrimination (Eilers et al., 1977) and categorical perception (Eimas, 1999; Eimas et al., 1971; Jusczyk et al., 1998; Kuhl, 1992, 2004; Trehub, 1979; Werker and 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 and 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-

---

* Tel.: +1 520 626 3710.
E-mail address: conewess@email.arizona.edu.

http://dx.doi.org/10.1016/j.ijpsycho.2014.06.002
0167-8760/© 2014 Elsevier B.V. All rights reserved.

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: /**u**d**u**/ vs. /**a**d**a**/, vowel place: /**u**d**u**/ vs. /**i**d**i**/, consonant voicing: /u**d**u/ vs. /u**t**u/, consonant continuance: /u**d**u/ vs. /u**z**u/, or consonant place: /u**d**u/ vs. /u**b**u/, or /u**b**u/ vs. /u**g**u/. 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.

1.1. 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 response (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 and 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 to 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 and 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 and 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 old 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 and Benasich, 2011; Molfese and 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., 2000). Yet, other investigators have demonstrated the inability to reliably obtain MMN in infants and young children (Morr et al., 2002), or for that matter, adults (Wunderlich and 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 and 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 and 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 et al. (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 old 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 old.

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 (ANSD) 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.