International Journal of Pediatric Otorhinolaryngology 79 (2015) 1055–1062

Contents lists available at ScienceDirect

International Journal of Pediatric Otorhinolaryngology

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

Objective measurement of high-level auditory cortical function in children

Aseel Almeqbel [a,*], Catherine McMahon [b,c]

[a] Department of Hearing and Speech Sciences, Faculty of Allied Health Sciences, Health Sciences Center, Kuwait University, Kuwait City, Kuwait
[b] Linguistics Department, Faculty of Human Sciences, Macquarie University, Sydney, NSW, Australia
[c] The HEARing Cooperative Research Centre (CRC), Melbourne, VIC, Australia

ARTICLE INFO

Article history:
Received 18 January 2015
Received in revised form 16 April 2015
Accepted 17 April 2015
Available online 7 May 2015

Keywords:
Temporal processing
Cortical auditory evoked potentials (CAEP)
Voice-onset-time (VOT)
Speech-in-noise
Amplitude-modulated broadband noise (AM-BBN)
N2 latency

ABSTRACT

Objective: This study examined whether the N2 latency of the cortical auditory evoked potential (CAEP) could be used as an objective indicator of temporal processing ability in normally hearing children.
Methods: The N2 latency was evoked using three temporal processing paradigms: (1) differences in voice-onset-times (VOTs); (2) speech-in-noise using the CV/da/embedded in broadband noise (BBN) with varying signal-to-noise ratios (SNRs); and (3) 16 Hz amplitude-modulated (AM) BBN presented (i) alone and (ii) following an unmodulated BBN, using four modulation depths. Thirty-four school-aged children with normal hearing, speech, language and reading were stratified into two groups: 5–7 years ($n = 13$) and 8–12 years ($n = 21$).
Results: The N2 latency shifted significantly and systematically with differences in VOT and SNR, and was significantly different in the two AM-BBN conditions.
Conclusions: For children without an N1 peak in the cortical waveform, the N2 peak can be used as a sensitive measure of temporal processing for these stimuli.
Significance: N2 latency of the CAEP can be used as an objective measure of temporal processing ability in a paediatric population with temporal processing disorder who are difficult to assess via behavioural response.

1. Introduction

Auditory temporal processing ability plays a significant role in developing speech, language and reading skills in children with normally developed hearing [1]. Temporal processing refers to the ability of the auditory system to receive and analyse temporal (timing) cues within the speech signal in the time domain such as voice-onset-time (VOT), duration, speech-in-noise and amplitude modulation. There is considerable evidence in the literature suggesting that auditory temporal processing deficits can be associated with speech, language and reading disorders [1–13]. Research has shown that a sub-group of children with specific language impairment (SLI) and/or developmental dyslexia (specific reading disability), exhibit poor speech perception abilities which are further impaired in the presence of competing noise [14–16]. In addition, poor speech understanding in noise is a consistently reported feature of children with auditory neuropathy spectrum disorder (ANSD) [17–21], where significantly higher signal-to-noise ratios (9 dB) are required when compared with normally hearing peers in order to correctly identify speech stimuli in background noise [22]. This pattern has also been observed in adults with ANSD. Hence, Narne and Vanaja [23] used behavioural measures in order to evaluate speech perception at 3 dB SNR (+10 dB, +5 dB and 0 dB) in adult participants with ANSD. The results showed that the adults with ANSD had greater reduction in speech scores in noise, compared with normally hearing adults. Moreover, a significant relationship has been found between temporal processing measures, such as temporal modulation transfer function (TMTF) and speech perception in noise, which suggests that disrupted temporal processing ability in ANSD may account for extreme difficulties in noise [23].

Early detection and diagnosis of an auditory temporal processing deficit is important to ensure early habilitation in young children with disrupted temporal processing abilities. One of the main limitations in detecting temporal processing deficits in infants and young children, however, is that they are unable to provide reliable behavioural responses to measures that assess temporal processing. In this regard, studies in normally hearing

---

[*] Corresponding author. Tel.: +965 249833412.
E-mail address: aseel.m@hsc.edu.kw (A. Almeqbel).

http://dx.doi.org/10.1016/j.ijporl.2015.04.026
0165-5876/© 2015 Elsevier Ireland Ltd. All rights reserved.

young adults have demonstrated that the objective cortical auditory evoked potential (CAEP), in particular the N1 peak latency, can be used as a sensitive measure of temporal changes in sound stimuli at the level of auditory cortex [24–26]. The N1 is also a reliable indicator for investigating neural timing of speech processing [27–29]. Therefore N1 latency in adults has been suggested as a clinically useful technique for measuring temporal processing [24,25,28]. However, due to the complexity of the development and maturation of the N1 component from infancy to adulthood [30,31], the N1 component of the CAEPs is either inconsistent or undeveloped in school-aged children, especially those younger than 8 years of age [32–35].

In general, the dominant features of the cortical response in young children are the P1, which varies in latency as a function of age, and the N2, a negativity following the P1 at about 200 or 250 ms [36,37]. N1, however, is age dependent [36–38] and typically emerges in an inconsistent pattern from 3 to 8 years. The presence of the N1 becomes more reliable around 9 to 11 years of age [36,39]. Some authors identify the negativity at 200–250 ms as the N1 [40], while others refer to it as the N2, regarding it as different from the adult auditory N1 wave [41], which is thought to emerge only at the age of 9 to 11 years [42]. Erring on the side of caution, in this paper we refer to the negativity at 200–250 ms as N2.

Few studies have measured temporal processing ability objectively in young children and most of these studies used a single paradigm, such as voice-onset-time [34] or speech-in-noise [43]. A discriminating feature of the current study, therefore, is that N2 latency was investigated as a function of three temporal processing paradigms (representing different magnitudes of temporal processing) in school-aged children with normal hearing, speech, language and reading skills: (1) differences in voice-onset-times (VOTs) using naturally produced stop consonant-vowel syllables /da/–/ta/; (2) speech-in-noise using the CV/da/embedded in broadband noise (BBN) with varying signal-to-noise ratios (SNRs); and (3) 16 Hz amplitude-modulated (AM) BBN presented in two conditions: alone and following an unmodulated BBN (representing a temporal change in the stimulus) using four modulation depths.

2. Method

2.1. Participants

Thirty-four school-aged children with normal hearing (16 males, 18 females) aged 5–12 years (mean 8.85, SD 2.16) were recruited. All children had English as their first language and all had pure-tone air conduction thresholds $\leq$15 dB HL at octave frequencies from 250 Hz to 8 kHz with type (A) tympanograms. They had no history of hearing or speech problems, or significant noise exposure, and no reported previous history of reading or learning problems. All participants had speech perception scores $\geq$90% using the Manchester Joiner Word List (MJWL) (open-set-speech perception) and had normal intelligence, operationalised as scores >85 on the Wechsler Nonverbal Scale of Ability (WNV) IQ test [44]. We divided the children into subgroups based on their age, 5–7 years ($n = 13$) and 8–12 years ($n = 21$), because of the known morphological differences in the cortical auditory evoked potential (CAEP) with age [39,45–47], and then into two groups of 17, to prevent participants becoming fatigued by a long testing session. Group I ($n = 17$ of which 7 were in the younger age range) participated in voice-onset-time, speech-in-noise, speech, language and reading tests, and group II ($n = 17$, of which 6 were in the younger age range) participated in the objective amplitude modulation detection test. Prior to testing, the purpose of the study was explained to each participant and informed consent was obtained from the children and their parents.

2.2. Stimuli

2.2.1. Voice-onset-time (VOT)

Two naturally produced stop consonant-vowel-syllables /da/–/ta/ were recorded by an Australian female speaker. These speech stimuli differed in VOT, for example the VOT for /da/ = 0.03 ms and for /ta/ = 0.95 ms. Speech stimuli /da/ and /ta/ were recorded using an AKG C535 condenser microphone connected to a Mackie sound mixer, with the microphone positioned 150 mm in front and at 45 degrees to the speaker's mouth to avoid air turbulence and popping. The mixer output was connected via an M-Audio Delta 66 USB sound device to a Windows computer running Cool Edit audio recording software and captured at 44.1 kHz 16 bit wave format. All speech stimuli were collected in a single session to maintain consistency of voice quality. Speech stimuli were modified after selection and recording using Cool Edit 2000 software. All speech stimuli of 200 ms duration were ramped with 20 ms rise and fall time to prevent any audible click arising from the rapid onset or offset of the waveform. The inter-stimulus interval (ISI), calculated from the onset of the preceding stimulus to the onset of the next stimulus was 1207 ms, as it has been shown that a slower stimulation rate results in more robust CAEP waveforms in immature auditory nervous systems [36].

2.2.2. Speech-in-noise

The speech stimulus /da/ was presented with varying signal-to-noise ratios. The speech stimulus /da/ was naturally recorded by an Australian female speaker. The speech sound was recorded using an AKG C535 condenser microphone connected to a Mackie sound mixer, with the microphone positioned 150 mm in front and at 45 degrees to the speaker's mouth to avoid air turbulence and popping. The mixer output was connected via an M-Audio Delta 66 USB sound device to a Windows computer running Cool Edit audio recording software and captured at 44.1 kHz 16 bit wave format. The speech stimulus was 60 ms in duration and ramped with 20 ms rise and fall time to prevent any audible click arising from the rapid onset or offset of the waveform. After a speech stimulus was selected and recorded, a BBN of 600 ms was generated using Praat software and the signal-to-noise ratio was altered using Matlab software with respect to the 65 dB SPL /da/ sound. This was then combined to create the speech signal /da/ embedded in different noise levels. Noise levels were 45, 65 and 75 dB SPL, which were chosen to create three signal-to-noise ratios (SNRs): +20 dB, 0 dB and −10 dB. The inter-stimulus interval (ISI), calculated from the onset of the preceding stimulus to the onset of the next stimulus was 1667 ms.

2.2.3. Amplitude modulation (AM) of broadband noise stimulus

Two conditions were used in this paradigm: (i) an amplitude-modulated BBN alone (300 ms), representing a temporally varying stimulus; and (ii) a BBN (600 ms) followed by an amplitude-modulated BBN (300 ms), representing a temporal change in the stimulus. These were used to determine whether changes in the N2 latency were a result of the amplitude-modulated signal per se or a temporal change in the signal from an unmodulated to a modulated signal. The modulation frequency was 16 Hz, and overall duration was 900 ms with 20 ms rise and fall time. The stimuli were generated using a 16-bit digital-to-analog converter with a sampling frequency of 44.1 kHz and low pass filter with a cut-off frequency of 20 kHz. The depth of the modulation was controlled by varying the amplitude of the modulating sine wave. Five modulation depths were used (100%, 75%, 50%, 25%, 0%). The inter-stimulus interval (ISI), calculated from the onset of the preceding stimulus to the onset of the next stimulus was 1307 ms for the first condition and 1907 ms for the second condition to ensure that there was a consistent period of silence between the stimuli.