Hearing Research 328 (2015) 34–47

Contents lists available at ScienceDirect

Hearing Research

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

Research paper

Auditory-neurophysiological responses to speech during early childhood: Effects of background noise

Travis White-Schwoch [a], Evan C. Davies [a], Elaine C. Thompson [a], Kali Woodruff Carr [a], Trent Nicol [a], Ann R. Bradlow [b], Nina Kraus [a, c, *]

[a] *Auditory Neuroscience Laboratory & Department of Communication Sciences, Northwestern University, Evanston, IL, USA*[1]
[b] *Department of Linguistics, Northwestern University, Evanston, IL, USA*
[c] *Department of Neurobiology & Physiology, Northwestern University, Evanston, IL and Department of Otolaryngology, Northwestern University, Chicago, IL, USA*

ARTICLE INFO

Article history:
Received 21 March 2015
Received in revised form
30 May 2015
Accepted 12 June 2015
Available online 23 June 2015

Keywords:
Speech recognition in noise
Auditory development
Electrophysiology
FFR
Auditory processing
cABR

ABSTRACT

Early childhood is a critical period of auditory learning, during which children are constantly mapping sounds to meaning. But this auditory learning rarely occurs in ideal listening conditions—children are forced to listen against a relentless din. This background noise degrades the neural coding of these critical sounds, in turn interfering with auditory learning. Despite the importance of robust and reliable auditory processing during early childhood, little is known about the neurophysiology underlying speech processing in children so young. To better understand the physiological constraints these adverse listening scenarios impose on speech sound coding during early childhood, auditory-neurophysiological responses were elicited to a consonant-vowel syllable in quiet and background noise in a cohort of typically-developing preschoolers (ages 3–5 yr). Overall, responses were degraded in noise: they were smaller, less stable across trials, slower, and there was poorer coding of spectral content and the temporal envelope. These effects were exacerbated in response to the consonant transition relative to the vowel, suggesting that the neural coding of spectrotemporally-dynamic speech features is more tenuous in noise than the coding of static features—even in children this young. Neural coding of speech temporal fine structure, however, was more resilient to the addition of background noise than coding of temporal envelope information. Taken together, these results demonstrate that noise places a neurophysiological constraint on speech processing during early childhood by causing a breakdown in neural processing of speech acoustics. These results may explain why some listeners have inordinate difficulties understanding speech in noise. Speech-elicited auditory-neurophysiological responses offer objective insight into listening skills during early childhood by reflecting the integrity of neural coding in quiet and noise; this paper documents typical response properties in this age group. These normative metrics may be useful clinically to evaluate auditory processing difficulties during early childhood.

1. Introduction

The world is inherently noisy, forcing talkers and listeners to compete with a bedlam of environmental and industrial sounds, additional voices, and more. This acoustic turbulence presents a challenge during early childhood, when children are attempting to make sense of the soundscape by forming precise representations of speech sounds to develop a rich and diverse lexicon. Due to the confluence of auditory and cognitive factors contributing to speech recognition in adverse listening conditions, and the heterogeneous development of central auditory processing, younger children are especially susceptible to the effects of background noise on speech

---

Abbreviations: FFR, frequency-following response; ABR, auditory brainstem response; F_0, fundamental frequency; LP, learning problems; CV, consonant vowel; M, mean; SD, standard deviation; SPL, sound pressure level; CMS/DRL, common mode sense/driven right leg; ENV, envelope; TFS, temporal fine structure; SNR, signal-to-noise ratio; RMS, root mean squared; FFT, fast Fourier transformation; RMANOVA, repeated measures analysis of variance; SEM, standard error of the mean

* Corresponding author. 2240 Campus Dr. Evanston, IL, 60208, USA.
E-mail address: nkraus@northwestern.edu (N. Kraus).

[1] www.brainvolts.northwestern.edu.

http://dx.doi.org/10.1016/j.heares.2015.06.009
0378-5955/

understanding (Hall III et al., 2002; Leibold and Buss, 2013; Wightman and Kistler, 2005). Although this susceptibility abates as children mature, a consequence of this protracted development is that most critical auditory mapping experiences occur before children have achieved adult-like speech recognition in these listening conditions. Success during this early childhood learning process has lifelong implications for auditory perception and cognition, and communication skills more broadly. Poor auditory processing under adverse listening conditions, in turn, has been linked to childhood learning problems (Bradlow et al., 2003; Cunningham et al., 2001; Ziegler et al., 2009; but see Messaoud-Galusi et al., 2011).

The auditory frequency-following response (FFR) is the product of synchronous firing of midbrain nuclei and reflects neural activity necessary for auditory perception in noise (Kraus et al., 2000; Zeng et al., 1999).[2] Even subtle dyssynchronies are linked to poor auditory processing in noisy and reverberant listening environments (Anderson et al., 2013b; Fujihira and Shiraishi, 2014; Ruggles et al., 2012) whereas enhancements in subcortical neural synchrony are associated with superior perception in challenging listening scenarios (Anderson et al., 2013c; Bidelman and Krishnan, 2010; Song et al., 2012). We believe, therefore, that the FFR provides a means to explore the neurophysiology contributing to auditory processing in noise.

1.1. *The speech-evoked FFR: a snapshot of auditory processing*

Auditory-neurophysiological responses reflect neural coding of multiple complex sound features, including the transient and periodic acoustic events found in speech. These response properties are collectively the product of an integrative auditory-cognitive system that is shaped through life experience (Kraus and Nicol, 2014). In fact, the FFR's neural generators are extensively connected to, and modified by, sensory, limbic, and cognitive circuits. Therefore, our view is that the response reflects experience with sound (for better or worse) and that a thorough evaluation of the response within an individual provides a unique window into auditory processing.

In particular, the FFR to speech can simultaneously quantify the midbrain coding of multiple acoustic properties of speech sounds. Depending on the stimulus and recording paradigm, this may include cues that contribute information about the talker (such as pitch-related information), cues that provide information about what was said (such as formant cues that convey phonemic identity), and temporal cues (such as the envelope and temporal fine structure). This biological mosaic reflects minute aspects of auditory processing with extreme granularity. Interestingly, these elements of neural coding are not necessarily strongly intercorrelated within an individual. Consequently, different populations of listeners have distinct "signature" patterns of response properties that may include relative strengths and weaknesses in the neural processes important for everyday communication (Kraus and Nicol, 2014). Therefore, analyzing multiple aspects of the response within an individual, or group of individuals, can offer converging information about the quality of speech sound coding and, potentially, auditory processing at large. Importantly, a single response can provide divergent information about the neural processing of orthogonal acoustic cues (White-Schwoch et al., in press).

1.2. *Consonants and vowels in quiet and noise*

Perceptual evidence from children and adults has shown that consonants are more difficult than vowels to recognize in adverse listening environments (Johnson, 2000). Compared to vowels, consonants comprise acoustic transients (the onset burst) and fast-changing spectral content with relatively low amplitude (the transition to or from the adjacent phoneme); these acoustic properties make consonants more susceptible to masking. Vowels, on the other hand, typically are of longer duration, are higher in intensity, and have relatively stable spectral content.

Speech recognition in noise is more challenging for preschoolers than for older children (Hall III et al., 2002; Leibold and Buss, 2013). Moreover, there is evidence from school-aged children that the neural coding of transient and dynamic speech cues is tenuous in noise relative to quiet, placing a neurophysiological constraint on consonant processing (Cunningham et al., 2001). This consonant liability in noise has been observed in auditory midbrain, thalamus, and cortex using near-field multiunit recordings in an animal model (Cunningham et al., 2002). It is unclear, however, whether and how this susceptibility to masking manifests during early childhood. It is important to understand the course of typical development to lay the groundwork to explore and identify deviations. Children with listening difficulties can exhibit poor auditory-temporal processing that may be characterized as a developmental delay (Wright and Zecker, 2004), and these children exhibit multimodal deficits parsing signals in noise (Sperling et al., 2005; Ziegler et al., 2009; but see Messaoud-Galusi et al., 2011). Here, our strategy is to examine neural coding of consonants in noise in typically developing children, with the aim of providing a neurophysiological framework to explore development, deviations, and individual differences.

At the same time, there is some evidence that listening in noise may carry benefits. For one, there is the phenomenon of "stochastic resonance," which demonstrates that background noise can improve perceptual thresholds (Douglass et al., 1993), including in the auditory system (Morse and Evans, 1996; Zeng et al., 2000). However, it is important to point out that there is a difference between perceptual *acuity* and extracting *meaning* from a signal (e.g., Anderson et al., 2013b; Souza et al., 2007). With regards to learning, Moucha and colleagues (2005) reported that exposing rats to background sounds during a learning task induced spectrotemporal plasticity across the tonotopic map in primary auditory cortex. This may be due to the background sounds emphasizing the contrast between target and non-target stimuli. Although these issues need to be explored in human listeners, they suggest that learning outcomes may be mediated by the listening conditions, and that background noise does not necessarily have a wholly negative effect on auditory learning.

1.3. *Current study*

To date, auditory-neurophysiological studies of speech processing in noise (and disorders thereof) have been conducted predominantly in children ages 8–15 years old. Children this age have often received prolonged instruction in language and literacy, and most have either been diagnosed with a learning problem (LP) or "cleared" as typical learners. An ideal approach to investigate auditory processing and its disorders would also measure neural activity in preschoolers. The preschool years are a time of rapid auditory learning and development, and a crucial age for identification of children who may begin to lag behind their peers with respect to language milestones. This approach could also facilitate developmental research aimed at discovering how auditory neurophysiology matures interactively with auditory perception and cognition.

---

[2] We note that this response has often been referred to by our group as the cABR (auditory brainstem response to complex sounds).