Hearing Research xxx (2018) 1–11

Contents lists available at ScienceDirect

Hearing Research

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

Research Paper

Effect of age and hearing loss on auditory stream segregation of speech sounds

Marion David [a, *], Alexis N. Tausend [a], Olaf Strelcyk [b], Andrew J. Oxenham [a]

[a] Department of Psychology, University of Minnesota, Minnesota, 55455, USA
[b] Sonova U.S. Corporate Services, 4520 Weaver Parkway, Warrenville, IL, 60555, USA

ARTICLE INFO

Article history:
Received 21 November 2017
Received in revised form
9 February 2018
Accepted 15 March 2018
Available online xxx

Keywords:
Auditory stream segregation
Hearing loss
Fundamental frequency
Vocal tract length

ABSTRACT

Segregating and understanding speech in complex environments is a major challenge for hearing-impaired (HI) listeners. It remains unclear to what extent these difficulties are dominated by direct interference, such as simultaneous masking, or by a failure of the mechanisms of stream segregation. This study compared older HI listeners' performance with that of young and older normal-hearing (NH) listeners in stream segregation tasks involving speech sounds. Listeners were presented with sequences of speech tokens, each consisting of a fricative consonant and a voiced vowel (CV). The CV tokens were concatenated into interleaved sequences that alternated in fundamental frequency (F0) and/or simulated vocal tract length (VTL). Each pair of interleaved sequences was preceded by a "word" consisting of two random tokens. The listeners were asked to indicate whether the word was present in the following interleaved sequences. The word, if present, occurred within one of the interleaved sequences, so that performance improved if the listeners were able to perceptually segregate the two sequences. Although HI listeners' identification of the speech tokens in isolation was poorer than that of the NH listeners, HI listeners were generally able to use both F0 and VTL cues to segregate the interleaved sequences. The results suggest that the difficulties experienced by HI listeners in complex acoustic environments cannot be explained by a loss of basic stream segregation abilities.

1. Introduction

Hearing loss is associated not just with a loss of audibility, but also with a reduced ability to "hear out" important sounds (such as speech) in complex acoustic backgrounds (Moore, 1998; Oxenham, 2018). These deficits lead to communication difficulties in everyday environments, and have been associated with social isolation (Mick et al., 2014) and accelerated cognitive decline with age (Gonzales et al., 2017).

Understanding speech in complex acoustic backgrounds relies on our ability to organize the incoming sounds into coherent perceptual streams. To be intelligible, the target speech must be integrated within a single auditory stream and segregated from the competing voices and other extraneous sounds (Bregman, 1990). Auditory stream segregation has been extensively investigated in normal-hearing (NH) listeners (for reviews, see Moore and Gockel, 2002, 2012) but less is known about auditory stream segregation in hearing-impaired (HI) listeners (Rose and Moore, 1997; Oxenham, 2008; Valentine and Lentz, 2008).

Many acoustic cues can be used by NH listeners for stream segregation. For instance, speech intelligibility in the presence of interfering speech can be improved by introducing a difference in fundamental frequency (F0) as small as 2 semitones between the competing speech streams (Bird and Darwin, 1998; Brokx and Nooteboom, 1982; Darwin et al., 2003). Another acoustic cue is related to the vocal tract length (VTL) of the talker, which reflects the talker size and influences judgements relating to the talker's gender and age (Smith and Patterson, 2005). For NH listeners, a difference in VTL greater than 8% between competing talkers has been found to induce an improvement in speech recognition (Darwin et al., 2003). Both these cues have been examined to some extent in HI listeners. Some studies have found a reduced ability of HI listeners to use F0 differences to segregate competing vowels (Arehart et al., 1997) or sentences (Summers and Leek, 1998). Mackersie et al. (2011) investigated the effect of F0 and VTL cues on the perceptual segregation of sentences by HI listeners. The listeners' task was to repeat keywords associated with a call sign

* Corresponding author.
E-mail address: david602@umn.edu (M. David).

https://doi.org/10.1016/j.heares.2018.03.017
0378-5955/

within sentences of the coordinate response measure (CRM) corpus, while ignoring a sentence spoken by a competing talker. The F0 difference between the target and competing talker was set to between 0 and 9 semitones and the VTL ratio was set to between 1 and 1.38. Mackersie et al. (2011) found that HI listeners were less able to use F0 cues than NH listeners and that they did not benefit from VTL cues.

Because these studies were carried out with stimuli that overlapped in time and frequency, it is not known to what extent the performance of HI listeners was limited by direct, or energetic, masking and to what extent it was mediated by effects related to streaming and/or informational masking (Durlach et al., 2003; Watson, 2005) in ways that cannot be accounted for by the overlap and interaction of the stimuli in the auditory periphery.

Studies using simple tonal stimuli to investigate streaming have come to differing conclusions. Rose and Moore (1997) found that the frequency separation needed to induce segregation between two alternating pure tones was generally greater for HI listeners than for NH listeners; however for HI listeners with unilateral hearing loss, the frequency separation needed for segregation was similar in both ears, suggesting that the relationship between streaming and peripheral hearing loss was not straightforward. Using broadband inharmonic tones, Valentine and Lentz (2008) found no significant differences between NH and HI groups in their ability to integrate or segregate alternating tone sequences based on their frequency separation.

To summarize, although HI listeners perform more poorly when listening to speech in competing speech backgrounds (Mackersie et al., 2011), they do not always perform differently from normal in non-speech tasks involving sequential stream segregation (Rose and Moore, 1997; Valentine and Lentz, 2008). It is possible that the difficulties experienced in speech tasks are related to direct (or energetic) masking when the target and masker overlap in time. Alternatively, it may be that the mechanisms involved in segregating highly spectro-temporally variable stimuli, such as speech, are not directly related to the mechanisms involved in segregating the simple repeating sequences typically used in auditory streaming studies. To address this question, we used a paradigm similar to that introduced by David et al. (2017a, 2017b). Speech sounds consisting of an unvoiced fricative consonant and a voiced vowel are arranged in interleaved sequences, and listeners are encouraged to listen either to the sequence as a single stream, or to attend to one of the two interleaved sequences. Our earlier studies showed that NH listeners were able to segregate such sequences based on differences in F0 between alternating speech sounds (David et al., 2017a) and on differences in spatial location and high-frequency spectral cues (David et al., 2017b). The present study investigated the performance of listeners with differing degrees of hearing loss in tasks involving the segregation of interleaved sequences of speech sounds based on differences in gender (male or female talker) or differences in either F0 or VTL, and compared their performance with age-matched listeners with normal hearing, as well as with younger listeners with normal hearing.

2. Listeners

Three groups of participants were tested: young NH listeners (YNH), older hearing-impaired listeners (AHI), and older NH listeners (ANH). The age range of the YNH group was 18–30 years (median age 20 years). The age range of the AHI group was 41.6–78 years (median age 67 years), and the age range of the ANH group was 46–73 years (median age 60 years). Normal hearing was defined as having audiometric thresholds ≤20 dB HL at all octave frequencies between 250 and 8000 Hz. Participants in the HI group had at least one threshold ≥25 dB HL across this frequency range. Not all participants took part in all experiments, so each experiment includes a description of its participants. Fig. 1 displays the audiograms for all the participants in this study. Only one ear of each participant was tested, and only the audiogram for that ear is shown in Fig. 1. For participants in the YNH and ANH groups, the better of the two ears was selected based on the average audiometric pure-tone thresholds estimated between 250 and 8000 Hz. For the AHI group, the poorer of the two ears was selected, based on the same criterion. All the listeners were native speakers of American English. They all provided written informed consent to participate in the study, and were paid for their participation. All protocols were approved by the Institutional Review Board of the University of Minnesota.

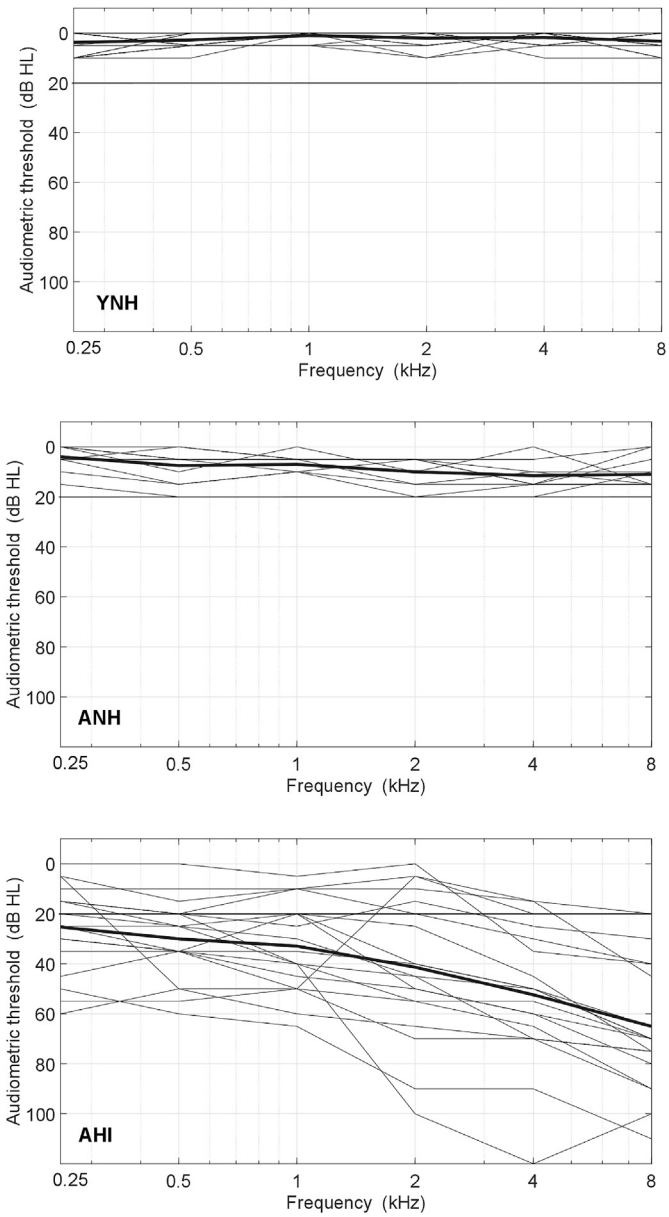

Fig. 1. Audiograms of the three groups of listeners: YNH (young normal-hearing listeners), ANH (older normal-hearing listeners), and AHI (older hearing-impaired listeners). The black lines correspond to the mean audiogram of all the participants in each group and the grey lines correspond to the individual audiograms. The pure tone averages (PTA), based on the thresholds at 0.5, 1, 2 and 4 kHz, were 1.9, 9.5, and 39.3 dB HL for the YNH, ANH, and AHI groups, respectively.

Please cite this article in press as: David, M., et al., Effect of age and hearing loss on auditory stream segregation of speech sounds, Hearing Research (2018), https://doi.org/10.1016/j.heares.2018.03.017