Hearing Research 317 (2014) 9–14

Contents lists available at ScienceDirect

Hearing Research

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

Research paper

Correlations of decision weights and cognitive function for the masked discrimination of vowels by young and old adults

Lynn Gilbertson*,[1], Robert A. Lutfi [1]

Auditory Behavioral Research Lab, Department of Communication Sciences and Disorders, University of Wisconsin Madison, 482 Scott H Goodnight Hall, 1975 Willow Dr., Madison, WI 53706, USA

A R T I C L E I N F O

Article history:
Received 13 June 2014
Received in revised form
31 August 2014
Accepted 3 September 2014
Available online 27 September 2014

A B S T R A C T

Older adults are often reported in the literature to have greater difficulty than younger adults understanding speech in noise [Helfer and Wilber (1988). J. Acoust. Soc. Am, 859–893]. The poorer performance of older adults has been attributed to a general deterioration of cognitive processing, deterioration of cochlear anatomy, and/or greater difficulty segregating speech from noise. The current work used perturbation analysis [Berg (1990). J. Acoust. Soc. Am., 149–158] to provide a more specific assessment of the effect of cognitive factors on speech perception in noise. Sixteen older (age 56–79 years) and seventeen younger (age 19–30 years) adults discriminated a target vowel masked by randomly selected masker vowels immediately preceding and following the target. Relative decision weights on target and maskers resulting from the analysis revealed large individual differences across participants despite similar performance scores in many cases. On the most difficult vowel discriminations, the older adult decision weights were significantly correlated with inhibitory control (Color Word Interference test) and pure-tone threshold averages (PTA). Young adult decision weights were not correlated with any measures of peripheral (PTA) or central function (inhibition or working memory).

1. Introduction

With increasing age understanding speech in noise can become challenging, as evidenced by multiple studies comparing performance of younger and older adults in speech-in-noise recognition tasks (Bouma and Gootjes, 2011; Helfer and Wilber, 1988; Pichora-Fuller et al., 1995; Stewart and Wingfield, 2009; Tun and Wingfield, 1999). The often observed poorer performance of older adults in these studies has been attributed to a variety of factors, ranging from reduced hearing sensitivity (Humes et al., 1994; van Rooij and Plomp, 1990) to declining cognitive function with increasing age (Craik, 1965; Guerreiro et al., 2010; Inglis and Caird, 1963). Tun and Wingfield (1999), for example, asked younger and older adults to recall a target talker sentence presented simultaneously with different types of distracters (single talker, two talkers, babble, or white noise) at different levels of intensity. Unlike many of the younger adults, the older adults' word recall performance was negatively affected by the intensity and type of the distracter. The amount of variance in their data explained by measures of cognitive function and hearing sensitivity suggested that the difference in performance could be attributed both to a decrease in speed of processing with increasing age and to generally poorer hearing. Huang et al. (2010) used a priming paradigm to evaluate if familiarity with a target speaker's voice would reduce the amount of informational masking of the target speech in a background sound. The participants were asked to repeat a target sentence that was played simultaneously with two-talker babble (speech-in-speech condition) or steady speech-spectrum noise (speech-in-noise condition). The target sentences were syntactically correct, but were not semantically meaningful. On the priming trials, a sentence spoken by the target speaker was played in isolation prior to each target plus masker trial. The priming trials were compared to non-primed trials. Younger adults showed a significant release from masking when the primer was present in the speech-in-speech condition, but older adults did not show this effect, suggesting a failure to use an efficient decision strategy (however, also see Agus et al., 2009; Helfer and Freyman, 2008). Several other studies have used canonical correlation analysis, regression, and principal component analysis in an attempt to identify the relative

Abbreviations: F0, fundamental frequency; F1, first formant frequency; F2, second formant frequency; PTA, pure tone averages; SD, standard deviation; ADRC, Wisconsin Alzheimer's Disease Research Center; 2IFC, Two-interval, forced-choice

* Corresponding author. Present address: Department of Communication Sciences and Disorders, University of Wisconsin Whitewater, 800 W Main St., Roseman Hall 1016, Whitewater, WI 53190-1790, USA. Tel.: +1 262 472 5202.

E-mail addresses: gilbertl@uww.edu (L. Gilbertson), ralutfi@wisc.edu (R.A. Lutfi).

[1] Tel.: +1 608 262 6485, +1 608 262 3951.

http://dx.doi.org/10.1016/j.heares.2014.09.001
0378-5955/

importance of factors affecting differences in speech perception among young and older adults. The amount of variance with increasing age that can be explained by audiological test results (pure tone thresholds, speech recognition thresholds, etc.) varies in these studies from 48 to 75%. Cognitive measures by comparison (memory, speed of processing, IQ, etc) account for 24–33% of the variance (Humes et al., 1994; van Rooij and Plomp, 1990).

To date, research has focused on the relationship between different measures of cognitive function and tests of performance accuracy, such as percent correct and masked threshold. Most of the available research underscores group differences between young and old adults. However, there is a great deal of variation in the physical aging process of older adults and there is no known biomarker of age, making it difficult to group older adult's physical functions by chronological age (Dollemore, 2009). It is important to not only understand differences in hearing function between young and old adults, but also what factors might account for individual differences within these groups. The present study aimed to gain a better understanding of the individual differences in speech recognition performance among older listeners by measuring listener decision weights that reflect the relative reliance listeners place on target and masker (cf. Berg, 1990; Lutfi and Liu, 2011). Decision weights potentially offer greater insight into the reason for performance differences among individuals by providing an estimate of how listeners make use of information in the target and masker; that is, measuring *how* listeners perform the task in addition to *how well* they perform the task. This approach has not previously been taken to investigate the differences in masked discrimination of vowels among older adults, nor have decision weights been linked with cognitive measures in previous studies.

2. Materials and methods

2.1. Stimuli

The stimuli were sequences of synthesized steady vowels patterned after adult speakers of American English presented as masker-target-masker triads (Hawks and Miller, 1995; Klatt and Klatt, 1990; Nearey, 1989). This set of stimuli provided the experimental control necessary to calculate decision weights and in turn, to evaluate individual differences beyond percent correct. The decision weight calculation is explained in the calculation section. The synthesized vowels had frequencies below 3000 Hz which reduced the likelihood that presbycusis would confound any differences between young and old participants. Vowels were chosen over pure-tone complexes to allow for a slightly greater degree of generalization to “real world” speech-in-speech listening conditions. The masker-target-masker triads were modeled after the interleaved-word paradigm of Kidd et al. (2008) to minimize the amount of energetic masking, or failure of frequency analysis at the level of the cochlea, and focus on informational masking. For the purposes of this paper, informational masking is defined as masking that cannot be explained by known processes occurring at the auditory periphery, but rather is an effect of such factors as uncertainty regarding the acoustic properties of the masker, perceived similarity of target and masker, and attention and memory. Both informational and energetic masking contribute to the difficulty older adults report when listening for a target sound during noisy listening conditions. However, informational masking has been shown to have a profound detrimental influence on performance, specifically in speech-in-speech listening conditions (Brungart et al., 2006); therefore it was the focus of the current study. The interleaved-word paradigm greatly reduces the amount of energetic masking because the words no longer overlap simultaneously in frequency and time causing a breakdown at the level of the peripheral auditory system. Because the interleaved paradigm has been shown to elicit little or no energetic masking, the decrease in percent correct from listening to a single target sentence in isolation to listening for a target sentence interleaved temporally with a distracter sentence has been explained by informational masking.

All vowels had a duration of 250 ms and were separated from one another by 20 ms silent intervals. For such conditions, little forward or backward energetic masking is expected (Dorman et al., 1977). All vowels had a constant first-formant frequency (F1) = 250 Hz and a variable second-formant frequency (F2). F2 varied in 50-Hz increments from 1000 to 2000 Hz yielding 20 sounds. The perceived vowel varied from/i/as in “beet” (with a high F2) to/u/as in “boot” (with a low F2). In the first condition (F0-same), both the middle target vowel and the flanking masker vowels had a male fundamental frequency (F0) of 132 Hz. In the second condition (F0-different), the masker vowels had a female F0 of 220 Hz while the target vowel maintained the male F0 of 132 Hz. The F2 value was randomly selected with equal probability on each presentation from the range 1000–2000 Hz. The F2 of the target vowels were chosen independently from the F2 of the masker vowels. Within each triad the formant frequencies of the masker vowels were constrained to be the same. The vowels were presented diotically over Beyerdynamic DT 990 headphones to participants seated in a double-walled, Industrial Acoustics (IAC), sound-attenuated chamber. They were played at a 44,100 Hz sampling rate with 16 bit resolution using a MOTU 896 audio interface. The level of the vowels was calibrated so that the overall sound level at the eardrum was approximately 70 dB SPL (see Lutfi et al., 2008).

2.2. Procedure

Two randomly-selected, masker-target-masker triads made up each trial of a two-interval, forced-choice (2IFC) design. A silent period of 0.5 s separated the two intervals. Because the participants had no known background in acoustics it was impractical to ask them to discriminate F2. Therefore, in all conditions the participant was instructed to choose the interval containing the target vowel closest to an/i/. The participant response was counted as correct if the interval selected contained the middle vowel with the higher F2. The correct response was equally likely to be interval one or interval two. Participants made responses by clicking a mouse button while seated at a computer. Visual feedback was presented on the computer monitor after each response indicating whether the response was correct or incorrect. Before completing the test trials, each participant completed 50 practice trials in which they heard a single vowel in each interval of the 2IFC task and were asked to identify in which interval the vowel sounded more like an/i/. Participants were then asked to complete an additional 50 practice trials of the 2IFC task with the vowel triads. In this second practice session the target vowels were 20 dB higher in level than the masker vowels. Finally, the participant completed 16 blocks of 50 trials (800 test trials for each session). Each condition session took about 1 h. Participants were allowed to take breaks as needed between blocks of trials. The participants completed all 800 trials of condition one before moving on to condition two. The order of task completion was randomized across participants.

2.3. Participants

A total of 33 participants completed the study; 16 older adults (10 females and 6 males, ages 56–79 years, mean = 65) and 17 young adults (14 females and 3 males, ages 19–30 years, mean = 22). Pure-tone air conduction thresholds were measured