Estimates of decision weights and internal noise in the masked discrimination of vowels by young and elderly adults

Lynn Gilbertson
Department of Communication Sciences and Disorders, University of Wisconsin, Whitewater, Wisconsin 53901, USA
gilbertl@uww.edu

Robert A. Lutfi and Jungmee Lee
Auditory Behavioral Research Lab, Department of Communication Sciences and Disorders, University of Wisconsin, Madison, Wisconsin 53706, USA
robert.lutfi@wisc.edu, jungmee.lee@wisc.edu

Abstract: Gilbertson and Lutfi [(2014). Hear. Res. **317**, 9–14] report that older adults perform similarly to younger adults on a masked vowel discrimination task when the fundamental frequency (F0) of target and masker vowel differ but that the older adults perform more poorly when the F0 is the same. This paper presents an alternative analysis of those data to support the conclusion that the poorer performance of older adults is due to an increase in the decision weight on masker reflecting poorer selective attention in noise of older adults.

[QJF]
Date Received: February 3, 2015 **Date Accepted:** April 17, 2015

1. Introduction

Understanding speech amid background noise often becomes more difficult with increasing age (Helfer and Wilber, 1988). Researchers have compared the performance between young and old adults to understand how age affects perception and function during such difficult listening tasks (Helfer and Freyman, 2008; Huang *et al.*, 2010; Humes *et al.*, 1994; Van Rooij and Plomp 1990; Tun and Wingfield, 1999). Evidence from these studies suggests that older adults may have greater difficulties due to changes in two broad factors associated with aging. The first is degradation in the internal neural representation of the speech, generally described as *internal noise*. Common sources of internal noise include reduced hearing sensitivity, slowing of neural transmission, poor echoic memory, general lapses in attention, and/or the distracting effects of the noise (see reviews by Schneider *et al.*, 2002; Pichora-Fuller, 2003). The second factor is difficulty ignoring the masker and *selectively attending* to the target (Tun *et al.*, 2002). Here the internal representation of the speech is presumed intact, but the listener is judged effectively to "confuse" the masker for the target on a disproportionate number of trials.

Lutfi and Liu (2011), following Berg (1990), have described a general method for identifying the relative influence of selective attention and internal noise in masking studies. The method involves computing a relative decision weight on the masker from the relation of the participant's trial-by-trial response to perturbations in target and masker parameters. Errors resulting from the masker decision weight are taken to reflect limits in selective attention while all other errors are taken to reflect the influence of internal noise. The present study takes this approach in an attempt to evaluate the relative influence of selective attention and internal noise for young and old adults in the masked discrimination of vowels.

2. Method

2.1 Stimuli

The stimuli, procedure, and participants of Gilbertson and Lutfi (2014) were used for the present study. The stimuli were masker-target-masker triads of synthesized vowels presented in a two-interval, forced-choice task (Neary, 1989; Hawks and Miller 1995; Klatt and Klatt, 1990). Each vowel had a total duration of 250 ms, and the vowels in the triadic sequences were separated from each other by 20 ms silent intervals. Each triad was separated by a 0.5 s silent period. A total of 20 vowels were synthesized. The vowels were composed of a second formant frequency (F2) that ranged from 1 to 2 kHz in 50 Hz increments and a first formant frequency (F1) that was fixed at 250 Hz. Vowels with a high F2 were perceived as /i/ and vowels with a low F2 were perceived as an /u/. In one condition, the vowel triads (both masker and target) had a male fundamental frequency (F0) of 132 Hz. In a second condition, the target vowel in the triad had a male F0 of 132 Hz, and the two flanking maskers had a female F0 of 220 Hz. All stimuli were played at a 44 100 Hz sampling rate with 16 bit resolution using a MOTU 896 audio interface. The vowels were presented at approximately 70 dB sound pressure level (SPL) diotically over Beyerdynamic DT 990 headphones to participants seated in a double-walled, Industrial Acoustics (IAC), sound-attenuated chamber.

2.2 Procedure

Target and masker vowels were selected at random from the 20 samples to compose both triads in a trial. The masker vowels were the same within a triad but were selected independently of the target vowel. Neither target nor masker was allowed to be repeated across the two intervals of a single trial. The participant was asked to choose, via a mouse click, the triad that contained the target vowel that was perceptually most similar to /i/. Interval one or two was equally likely to contain the correct response. Participants were presented visual feedback on a computer monitor designating the response as correct or incorrect. Initially the participant completed a 50 trial practice session with the vowels randomly selected and played in isolation to familiarize the participant with the stimuli. A second practice session of 50 trials was administered with the vowel triads; however, the target vowel was 20 dB higher in level than the masker vowels. Responses were collected for both practice sessions but were not included in the present analysis. The experimental task consisted of 800 trials for both of the comparison conditions. The order of task completion was randomized across participants.

2.3 Participants

Participants included 17 adults ages 19–30 yr (mean = 22 yr, 10 females and 6 males) and 16 adults ages 56–79 yr (mean = 65 yr, 10 females and 6 males). Standard air conduction thresholds were evaluated bilaterally. All young adults were verified to have thresholds of 20 dB hearing level (HL) or better at all test frequencies. All older adults were verified to have thresholds of 30 dB HL or better at 250, 500, 1000, and 2000 Hz (cf. Russo and Pichora-Fuller, 2008; Stewart and Wingfield, 2009). The young adults self-reported that they were free from cognitive impairment. The older adults were verified to have normal cognitive function for their age by a physician, by neuropsychological testing, and by MRI as required for their membership in the normal aging control group for studies conducted through the Wisconsin Alzheimer's Disease Research Center (ADRC).

2.4 Analysis

Perturbation analysis was used to evaluate the relative influence of selective attention as given by the obtained decision weights and internal noise [for complete development of this analysis, the reader is referred to the papers by Lutfi and Liu (2011) and

Gilbertson and Lutfi (2014)]. A logistic regression was performed on the trial-by-trial data for the probability of an interval one response with the interval difference (first-second) in F2 of target and masker as predictor variables. Let c_T and c_M denote, respectively, the target and masker regression coefficients for the difference in F2,

$$\text{logit}[P(R=1)] = c_T\Delta T + c_M\Delta M + e, \tag{1}$$

where $P(R=1)$ is the probability of a first interval response, ΔT and ΔM are the differences in F2 of target and masker across the two intervals (first minus second), and e is the regression error taken to reflect internal noise. The relative weight on the masker, taken to reflect selective attention, was then given by

$$w = \frac{c_M}{c_T + |c_M|}. \tag{2}$$

Note here that a positive weight indicates some form of "confusion" of target and masker while a negative weight indicates a relative comparison between target and masker, both cases resulting in potential errors. The listener achieving optimal performance would give a weight of zero to the masker and for a weight of zero anything less than optimal performance would be attributed to internal noise (reflected in the error associated with the regression). Realistically, listener performance would be expected to be influenced by some combination of nonoptimal weight and internal noise. Our interest then is in their relative contribution as estimated from the regression.

3. Results and conclusion

Group differences in the relative effect on performance of internal noise and decision weights were evaluated by comparing for each participant their obtained d' to the d' that would have been obtained if the only limit on performance was that participant's decision weight. This was achieved by substituting each participant's decision weight into Eq. (3) to compute a percent correct score based on the weight in absence of internal noise

$$\text{PCweight} = 100(1 – w + 0.5w). \tag{3}$$

The percent correct score was then converted to a d' score and designated d'_{weight}. The advantage of comparing d'_{weight} values for the two groups (as opposed to comparing the decision weights as in Gilbertson and Lutfi, 2014) is that the effect of the decision weights on performance for the two groups can be directly evaluated. To this end, the d'_{weight} scores are plotted against the d'_{obtained} scores for each listener of each group in Fig. 1. The left panel gives the data for the condition in which target and masker shared the same F0, and the right panel gives the data for the condition in which target and masker F0 differed. The data for the younger and older adults are given as filled and open symbols, respectively, and the mean for each group by the star symbols.

The figure makes evident a large degree of variability in d'_{obtained} and d'_{weight} for both age groups and both listening conditions. Such large individual differences are not uncommon in masking experiments (cf. Lutfi and Liu, 2011). Notwithstanding, Gilbertson and Lutfi (2014) report for a randomly selected subset of these listeners that the decision weights were highly replicable over time. An two-way analysis of variance (ANOVA) (F0 × group) of the d'_{obtained} scores revealed a main effect for the F0 condition, scores being significantly lower in the F0-Same condition [$F_{(1,62)} = 8.263$, $p = 0.005\,54$]. The interaction of listener group and condition was not significant [$F_{(1,62)} = 0.279$, $p = 0.599$]. However, an ad-hoc comparison (using Tukey HSD method) of the group difference for the F0 same condition was significant with the elderly group showing poorer performance ($p = 0.032$).

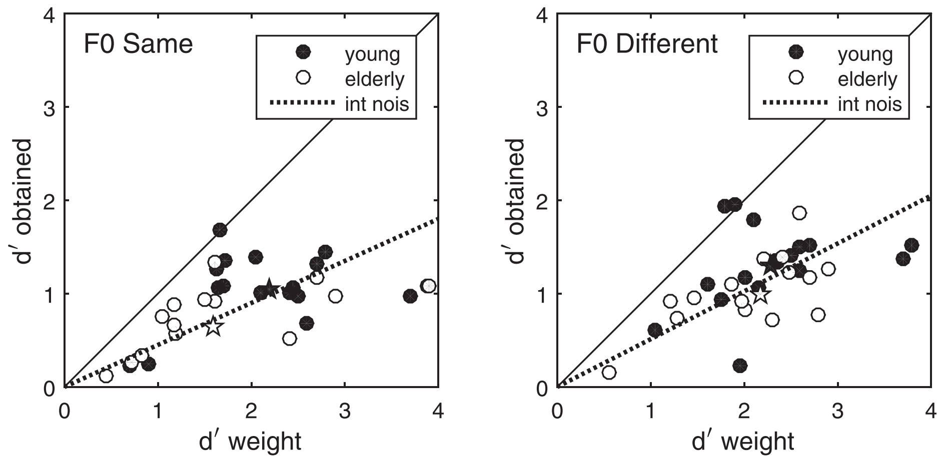

Fig. 1. Values of d' obtained are plotted against computed values of d'_{weight} for both young (solid symbols) and elderly (open symbols) adults for F0-same (left panel) and F0-different (right panel) conditions (see text for explanation of d'_{weight}). Dotted line gives expectation for the data based on a constant level of internal noise.

Plotting the data as in Fig. 1 makes immediately evident the relative extent to which decision weights and internal noise impact these performance differences. If listener decision weights are the only factor affecting performance, then all of the data will fall on a line with slope of 1 (the positive diagonal in the figure). The extent to which internal noise affects performance is reflected in the data falling on a line with the same intercept (0,0) but slope less than 1 (cf. Lutfi and Liu, 2011). Clearly, internal noise has an impact on performance for both groups inasmuch as almost all of the data fall below the diagonal. However, the effect of the internal noise is estimated to be the same for both groups inasmuch as the means for both groups fall on the same line with slope less than 1 (dotted line in the figure). Indeed for the F0-same condition, the difference in d'_{obtained} is seen to be entirely due to the difference in d'_{weight}. The results suggest that while there are large individual differences among listeners, the predominant factor affecting age differences in performance in the masked vowel discrimination task is the ability to selectively attend to the target when the masker and target share the same F0. That is, at least, to the extent that the decision weight can be taken as a measure of selective attention to the target. Additional studies are needed to identify the factors responsible for the large individual differences among listeners. Notwithstanding, the outcome demonstrates the utility of perturbation analysis as a tool for evaluating factors associated with age differences in listener performance in difficult listening tasks.

Acknowledgments

This research was supported by NIDCD Grant No. R01 DC001262-20 and the Wisconsin Alzheimer's Disease Research Center.

References and links

Berg, B. (**1990**). "Observer efficiency and weights in a multiple observation task," J. Acoust. Soc. Am. **88**, 149–158.

Gilbertson, L., and Lutfi, R. A. (**2014**). "Correlations of decision weights and cognitive function for the masked discrimination of vowels by young and old adults," Hear. Res. **317**, 9–14.

Hawks, J. W., and Miller, J. D. (**1995**). "A formant bandwidth estimation procedure for vowel synthesis," J. Acoust. Soc. Am. **97**, 1343–1344.

Helfer, K., and Freyman, R. (**2008**). "Aging and speech-on-speech masking," Ear Hear. **29**, 87–98.

Helfer, K., and Wilber, L. (**1988**). "Speech understanding and aging," J. Acoust. Soc. Am. **83**, 859–893.

Huang, Y., Xu, L., Wu, X., and Li, L. (**2010**). "The effect of voice cuing on releasing speech from informational masking disappears in older adults," Ear Hear. **31**, 579–583.

Humes, L., Watson, B., Christensen, L., Cokely, C., Halling, D., and Lee, L. (**1994**). "Factors associated with individual differences in clincal measures of speech recognition among the elderly," J. Speech Hear. Res. **37**, 465–474.

Klatt, D. H., and Klatt L. C. (**1990**). "Analysis, synthesis and perception of voice quality variations among male and female talkers," J. Acoust. Soc. Am. **87**, 820–856.

Lutfi, R. A., and Liu, C. (**2011**). "Target enhancement and noise cancellation in the identification of a rudimentary sound source in noise," J. Acoust. Soc. Am. **129**, EL52–EL56.

Nearey, T. M. (**1989**). "Static, dynamic, and relational properties in vowel perception," J. Acoust. Soc. Am. **85**, 2088–2113.

Pichora-Fuller, M. K. (**2003**). "Cognitive aging and auditory information processing," Int. J. Audiol. **42**, 26–32.

Russo, F., and Pichora-Fuller, M. (**2008**). "Tune in or tune out: Age-related differences in listening to speech in music," Ear Hear. **29**, 746–760.

Schneider, B. A., Daneman, M., and Pichora-Fuller, M. K. (**2002**). "Listening in aging adults: From discourse comprehension to psychoacoustics," Can. J. Exp. Psychol. **56**(3), 139–152.

Stewart, R., and Wingfield, A. (**2009**). "Hearing loss and cognitive effort in older adults' report accuracy for verbal materials," J. Am. Acad. Aud. **20**, 147–154.

Tun, P., and Wingfield, A. (**1999**). "One voice too many: Adult age differences in language processing with different types of distracting sounds," J. Gerontol. Psychol. Sci. **54B**, 317–327.

Tun, P. A., O'Kane, G., and Wingfield, A. (**2002**). "Distraction by competing speech in young and older adult listeners," Psychol. Aging **17**(3), 453–467.

Van Rooij, J. C. G. M., and Plomp, R. (**1990**). "Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses," J. Acoust. Soc. Am. **88**, 2611–2624.