Talker identification: Effects of masking, hearing loss, and age

Virginia Best,[1,a)] Jayne B. Ahlstrom,[2] Christine R. Mason,[1] Elin Roverud,[1]
Tyler K. Perrachione,[1] Gerald Kidd, Jr.,[1] and Judy R. Dubno[2]

[1]*Department of Speech, Language and Hearing Sciences, Boston University, Boston, Massachusetts 02215, USA*
[2]*Department of Otolaryngology-Head and Neck Surgery, Medical University of South Carolina, Charleston,
South Carolina 29425, USA*

The ability to identify who is talking is an important aspect of communication in social situations and, while empirical data are limited, it is possible that a disruption to this ability contributes to the difficulties experienced by listeners with hearing loss. In this study, talker identification was examined under both quiet and masked conditions. Subjects were grouped by hearing status (normal hearing/sensorineural hearing loss) and age (younger/older adults). Listeners first learned to identify the voices of four same-sex talkers in quiet, and then talker identification was assessed (1) in quiet, (2) in speech-shaped, steady-state noise, and (3) in the presence of a single, unfamiliar same-sex talker. Both younger and older adults with hearing loss, as well as older adults with normal hearing, generally performed more poorly than younger adults with normal hearing, although large individual differences were observed in all conditions. Regression analyses indicated that both age and hearing loss were predictors of performance in quiet, and there was some evidence for an additional contribution of hearing loss in the presence of masking. These findings suggest that both hearing loss and age may affect the ability to identify talkers in "cocktail party" situations.
© 2018 Acoustical Society of America. https://doi.org/10.1121/1.5024333

I. INTRODUCTION

Sensitivity to non-linguistic features of speech is important for successful communication in many social situations. Such features might include the acoustic characteristics of the talker's voice (which would indicate who is talking at any given moment), variations in dialect, and suprasegmental cues (e.g., prosody) that might indicate the mood of a talker and the general intention of their speech (whether they are telling a joke, asking a question, etc.). In this study, we examined the ability of listeners to identify talkers by their voice (hereafter referred to as "talker identification") in quiet and in the presence of competing sounds like those one might encounter in social situations. In addition to the inherent value of this ability, there is also some evidence that talker identification might indirectly influence speech understanding in noisy settings. For example, several studies have shown that the speech of familiar talkers is better understood in noise than the speech of unfamiliar talkers (Nygaard and Pisoni, 1998; Souza *et al.*, 2013). Moreover, listeners are better able to segregate competing utterances if one of them is in a familiar voice (Johnsrude *et al.*, 2013; Souza *et al.*, 2013).

Subjective reports indicate that talker identification might play a role in the difficulties experienced by listeners with hearing loss in social communication settings. For example, a pertinent question in the well-known Speech, Spatial and Qualities of Hearing Scale asks: "Do you find it easy to recognize different people you know by the sound of each one's voice?" (Gatehouse and Noble, 2004). While this question was not rated particularly poorly among listeners

with hearing loss (average score of 7.8 on a scale from 0 to 10), the question was shown to have one of the strongest correlations with the experience of handicap ($r = 0.53$, or $r = 0.59$ when controlling for audiometric thresholds). As noted by Gatehouse and Noble (2004, p. 91): "It is plausible to invoke the idea of social competence… in seeing how these qualities help to influence the handicap experience. Failures of identification, such as of persons or their mood, may add to a sense of embarrassment and reinforce a desire to avoid social situations, as will the effort needed to engage in conversation."

From a research standpoint, very little evidence is available about this issue. There are multiple ways in which talker identification in typical listening situations could be affected by hearing loss. First, the encoding of features required to reliably identify talkers and discriminate between talkers may be disrupted. This could affect talker identification even under quiet conditions, but may also be exacerbated by the presence of competing "noise" (i.e., unwanted sources of sound). Second, talker identification may be difficult primarily in the presence of other talkers because of the need to segregate the competing voices before identification can take place, or because of the need to identify more than one talker. This effect might be particularly relevant in older listeners who are thought to have particular difficulty segregating and inhibiting competing sound sources (e.g., Janse, 2012).

There is an extensive literature characterizing the acoustic features relevant to normally hearing listeners' perception of talker identity in quiet conditions (e.g., Remez *et al.*, 2007; Latinus and Belin, 2011; Schweinberger *et al.*, 2014) and the neural mechanisms involved (e.g., Formisano *et al.*, 2008; Perrachione *et al.*, 2009; Chandrasekaran *et al.*, 2011;

a)Electronic mail: ginbest@bu.edu

Latinus *et al.*, 2013). However, the literature is quite limited with respect to the study of talker identification in the presence of noise (see Razak *et al.*, 2017). Moreover, very few studies on this topic have included younger or older listeners with hearing loss. The limited findings bearing on this issue that are available suggest that older listeners are poorer than younger listeners at categorizing the sex of the talker (Schvartz and Chaterjee, 2012) and recognizing voices in quiet (Yonan and Sommers, 2000). In listeners fitted with cochlear implants, there are some studies that suggest that the severely reduced spectral resolution available through an implant disrupts the ability to categorize the sex of the talker (e.g., Fuller *et al.*, 2014) and to identify the talker (Vongphoe and Zeng, 2005).

The ability to segregate competing voices in order to understand a target message is adversely affected by both hearing loss and advanced age (Summers and Leek, 1998; Mackersie *et al.*, 2001; Mackersie, 2003). Interestingly, the task of segregating one voice from another often inherently involves talker identification (see Helfer and Freyman, 2009), but talker identification is rarely measured in the presence of competing voices (see Razak *et al.*, 2017). To our knowledge, only one study has explicitly assessed talker identification in the presence of a competing voice in listeners with hearing loss (Rossi-Katz and Arehart, 2009). In that study, it was reported that older listeners with hearing loss were poorer than younger listeners at recognizing a previously learned talker in the presence of a competing talker. However, the identification task was a rather easy one (identification from a set of two male and two female voices chosen to be highly distinctive) and the performance of listeners was close to ceiling in quiet. Thus, it is not possible from those data to determine whether the limitation in the older group was due to the presence of the competing talker, or a more general difficulty with talker identification that could only be observed when performance was lower than ceiling. Moreover, the contribution of hearing loss separately from age to talker identification was not reported.

The aim of the current study was to measure the ability of listeners to identify talkers on the basis of their voice in quiet, in noise, and in the presence of a competing talker. These three conditions in combination allowed us to assess *talker identification*, as well as *susceptibility to noise*, and issues related to the *segregation* of competing voices. A 2 × 2 group design that included younger and older listeners, with and without hearing loss, provided the means to untangle the effects of age and hearing loss on talker identification, noise susceptibility, and segregation. We hypothesized that hearing loss would adversely affect talker identification in quiet, by disrupting the features needed to discriminate voices from each other, and that this effect would be exacerbated in the presence of masking. In older listeners, we expected that cognitive limitations might lead to specific challenges in the competing talker condition.

II. METHODS

A. Participants

Thirty-two listeners were recruited across two locations (Boston University, BU, and the Medical University of South Carolina, MUSC). These listeners were recruited as four groups of eight on the basis of their age and hearing status. The younger normal-hearing (YNH) group included six females and two males, and ranged in age from 21 to 35 yr (mean age 26 yr). The younger hearing-impaired (YHI) group included four females and four males, and ranged in age from 19 to 41 yr (mean age 24 yr). The older normal-hearing (ONH) group included seven females and one male, and ranged in age from 60 to 84 yr (mean age 70 yr). The older hearing-impaired (OHI) group included four females and four males, and ranged in age from 69 to 84 yr (mean age 75 yr).

Hearing losses in both hearing-impaired (HI) groups were bilateral, symmetric, and sensorineural. Mean audiograms for each group (collapsed across left and right ears) are shown in Fig. 1. For the purposes of statistical analysis, the four-frequency average hearing loss (4FAHL; mean threshold at 0.5, 1, 2, and 4 kHz averaged across the ears) was calculated for each listener. Average 4FAHLs were 52 dB and 37 dB for the YHI and OHI groups, respectively (cf. 4 dB and 9 dB for YNH and ONH, respectively). As a result of our recruiting procedure, which covered these four listener groups, age and hearing loss were not significantly correlated in the final sample of 32 participants ($r = -0.08$, $p = 0.671$).

Seven of the eight YHI and six of the eight OHI were regular hearing aid wearers. For this experiment, however, hearing losses in all HI participants were compensated for by the application of linear gain to the stimuli prior to presentation. This gain was set on an individual basis according to the NAL-RP prescription rule (Byrne *et al.*, 1991; Dillon, 2012), which provides frequency-dependent amplification using a modified half-gain rule based on audiometric thresholds.

B. Stimuli

Target stimuli were simple questions drawn from a corpus of questions and answers recorded in the laboratory (see Best *et al.*, 2016). The corpus contains 227 questions that
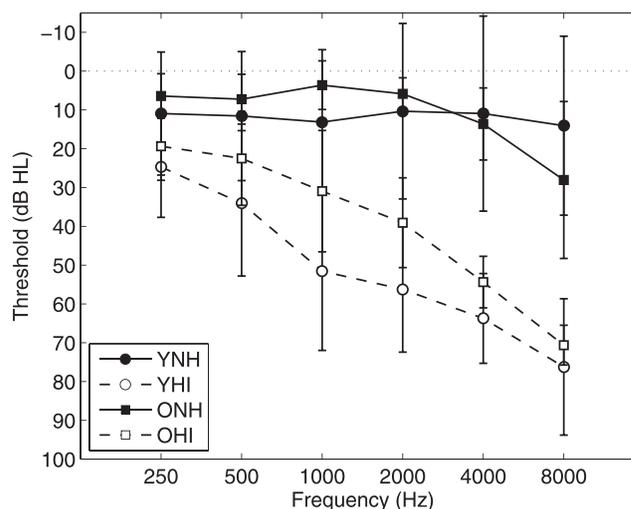

FIG. 1. Mean audiograms for each group (collapsed across left and right ears). Error bars show ±1 standard deviation across the eight participants in each group.

fall into different categories (e.g., days of the week, opposites, simple arithmetic). Recordings of each question were available from 12 male and 12 female young-adult talkers with American accented English. From each set of 12, 4 talkers were chosen arbitrarily to be targets and the remaining 8 were used in the competing talker condition. Figure 2 shows the mean fundamental frequency (F0) for each of the target and competing talkers to illustrate the range across talkers and show where the target talkers fell within the set of all talkers. The mean F0 was calculated across all 227 questions, using the software package Praat (Boersma and Weenink, 2009).

During testing, talker identification was measured (1) in quiet, (2) in the presence of speech-shaped, steady-state noise, and (3) in the presence of a single same-sex competing talker. The speech-shaped noise was created separately for female and male talkers such that the spectrum of the noise matched the average long-term spectrum of the entire set of questions spoken by all 12 talkers. The noise and competing talker conditions were tested at four target-to-masker ratios (TMRs), which were achieved by varying the level of the target talker relative to the fixed noise and competing talker level of 65 dB sound pressure level (SPL). The range of TMRs, chosen on the basis of pilot testing to cover the sloping portion of the psychometric function, were slightly different for the noise (-10, -5, 0, $+5$ dB) and competing talker (-7.5, -2.5, $+2.5$, $+7.5$ dB) conditions.

Stimuli were presented diotically over headphones (Sennheiser HD280 Pro, Wedemark, Germany) via an RME HDSP 9632 soundcard (Haimhausen, Germany) or via a Lynx TWO-B soundcard (Costa Mesa, CA) and a Tucker-Davis-Technologies HB5 headphone buffer (Alachua, FL). The listener was seated in a sound-treated booth. Responses were given by clicking with a mouse on a grid showing the numbers 1–4 (corresponding to the four target talkers). Stimulus delivery and recording of responses were controlled using MATLAB (MathWorks Inc., Natick, MA).

C. Procedures

The same experimental protocol was followed at the two locations (BU and MUSC). Each listener completed two sessions of two to three hours, one with female talkers and one with male talkers (order counterbalanced within groups). Listeners were encouraged to take regular breaks throughout a session to avoid fatigue. Within each session, training and testing were completed using a protocol adapted from Perrachione et al. (2014).

Training was conducted in two phases. In the first phase, in which listeners heard the different talkers for the first time, active and passive trials were interleaved to facilitate learning (Wright et al., 2010). A single question was chosen, and listeners were presented with 16 trials of this question. The first eight of these were passive trials in which a single utterance of the question spoken by each talker was presented in a consecutive order (and then repeated), with simultaneous visual cues identifying the number associated with each talker. The second eight were active trials in which the four utterances were presented twice in a random order, and the listener was required to identify the talker. Correct-answer feedback was given on each of these trials. This passive/active procedure was completed for 15 different questions (240 trials total). The second training phase consisted only of active trials, and was designed to reinforce the learning from the first phase. In this phase, the same 15 questions used for the first phase were intermixed, such that each question spoken by each of the 4 talkers was presented twice in a random order, for a total of 120 trials. The listener was required to identify the talker after each presentation, and correct-answer feedback was given. Together, the two training phases took about 30–45 min to complete.

The testing phase was divided into two identical halves. Each half consisted of a block of 40 trials in quiet, followed by a block of 160 trials in the presence of noise (40 trials per TMR), followed by a block of 160 trials in the presence of a competing talker (40 trials per TMR). No feedback was provided. Questions in this phase were drawn randomly on each trial, using different categories of questions from those used for the training phases. Note that in the competing talker condition, since the two talkers both uttered questions that were novel (i.e., not used during training), the only way the participant could identify the target was on the basis of it being one of the learned voices. Two of the YHI subjects took substantially longer for training and testing than the
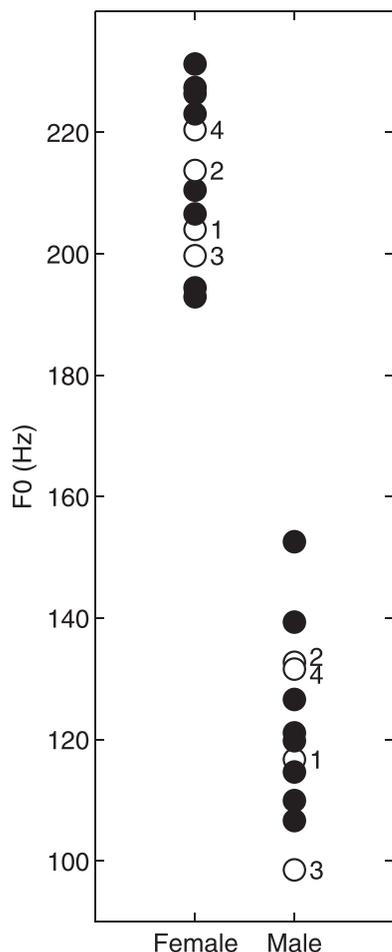

FIG. 2. Mean F0 for each of the 24 talkers used in the experiment (left column: females; right column: males). Open symbols correspond to target talkers, which are labeled with the numbers 1–4 within each group.

other subjects, and were only able to complete the first half of the testing phase in each session. For these subjects, scores were calculated based on trials from that first half only, whereas for the rest of the subjects, scores were calculated based on trials pooled across the first and second halves.[1]

For statistical analysis, scores in the different conditions were converted to rationalized arcsine units (Studebaker, 1985). Before this conversion, scores were adjusted for chance performance on this task (by scaling scores of 25%–100% into the range 0–1).

III. RESULTS

A. Training

Average scores for the first training phase (active trials only) and the second training phase are shown in Fig. 3 for the four listener groups and the female and male talker sessions. There was a clear difference in scores between the female and male talker sessions, with poorer scores for the female talkers, and thus the data for the two sessions are considered separately here and in the analyses that follow. At this stage, however, we have no reason to believe that the effect of talker sex reflects an inherent difference in difficulty rather than the similarity among voices within the two pools of talkers selected arbitrarily for this study. Two-way mixed analyses of variance (ANOVAs) were conducted on the training scores with a between-subjects factor of group and a within-subjects factor
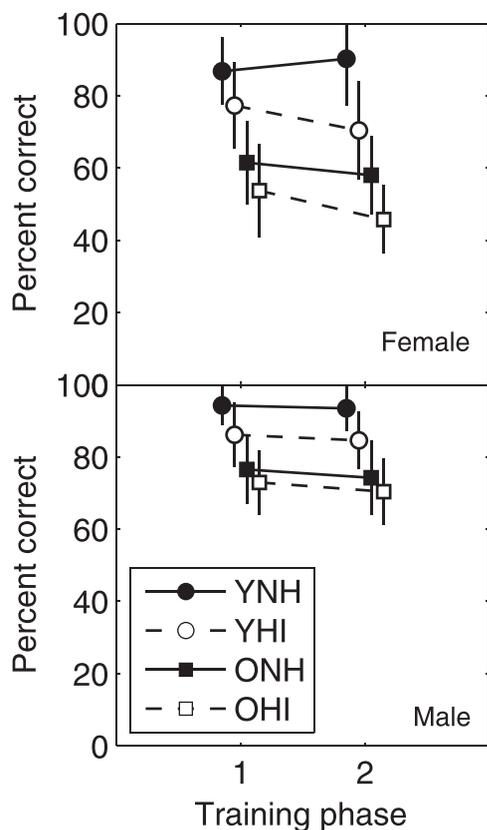

FIG. 3. Mean scores for each group in the first (left) and second (right) training phases. Error bars show ± 1 standard deviation across the eight participants in each group.

TABLE I. Results of the ANOVAs conducted on the training scores for both female and male talker sessions. Shown are the degrees of freedom (DF) along with F and p values.

	Female			Male		
	DF	F	p	DF	F	p
Group	3,28	20.357	<0.001	3,28	15.189	<0.001
Training phase	1,28	3.813	0.061	1,28	2.017	0.167
Group × training phase	3,28	6.191	0.002	3,28	0.193	0.901

of training phase (Table I). These analyses revealed that, for both female and male talker sessions, the scores for the four groups were significantly different, but scores did not differ significantly for the two training phases. *Post hoc* comparisons (Bonferroni, $p < 0.05$) indicated that scores for the YNH were better than scores for all of the other groups, and that scores for the YHI were better than for the OHI. A significant two-way interaction between group and training phase was also observed for the female talker session only, reflecting the fact that all groups performed slightly worse in the second phase of training except for the YNH group who showed a small improvement. The reason for this pattern is not clear.

B. Group mean performance

Figure 4 shows mean scores for the four groups in the testing phase. Shown are scores in quiet (left panel) and in the noise and competing talker conditions (middle and right panels) as a function of TMR. As expected, scores were poorer in both kinds of competition than in quiet. Moreover, although the TMR ranges were different between the noise and competing talker conditions, interpolation suggests that scores were poorer for the competing talker (e.g., average of 43% vs 61% at 0 dB TMR).

One-way ANOVAs conducted on the quiet scores found a significant main effect of group for both the female and male talker sessions (Table II). *Post hoc* comparisons (Bonferroni, $p < 0.05$) indicated that scores for the YNH were significantly better than scores for all of the other groups, which were not significantly different from each other. Two-way mixed ANOVAs conducted on scores in the noise and competing talker conditions (with a within-subjects factor of TMR and a between-subjects factor of group) found significant main effects of group and TMR for both conditions for both the female and male talker sessions (Table II). As was found in quiet, *post hoc* comparisons (Bonferroni, $p < 0.05$) indicated that scores for the YNH were significantly better than scores for all of the other groups, which were not significantly different from each other. A significant two-way interaction was observed in all conditions except for the female competing talker condition, suggesting that the differences between groups varied somewhat with TMR (although this might be in part due to ceiling and floor effects at high and low TMRs).

C. Effects of hearing loss and age on performance

Because of the large individual variability that was evident within our four listener groups, even in quiet, multiple
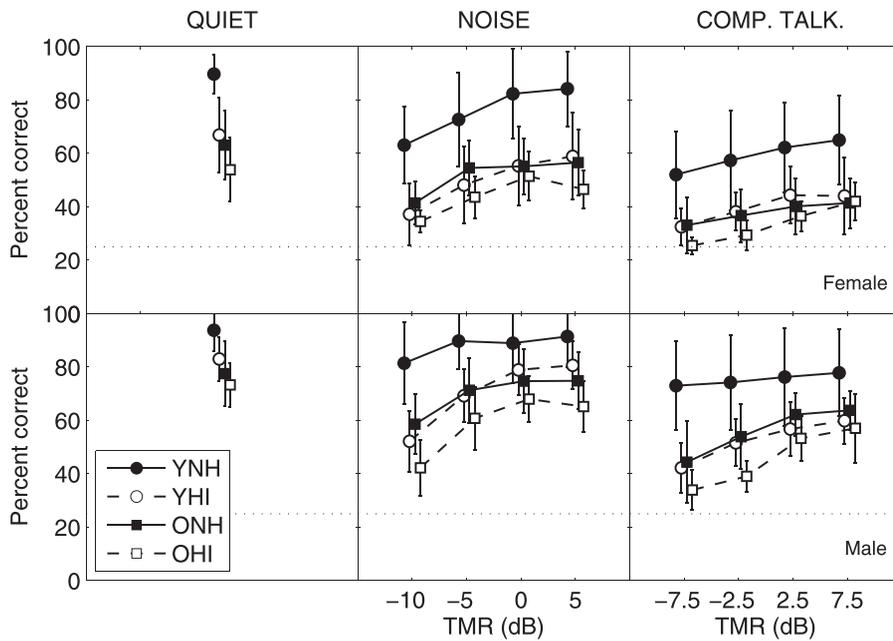
FIG. 4. Mean scores for each group in quiet (left), and mean scores in the presence of noise (middle) and competing talker (right) as a function of TMR. Error bars show ±1 standard deviation across the eight participants in each group.

regression analyses were conducted on scores in the testing phase in order to investigate in more detail the contributions of age and hearing loss. Figure 5 shows scores in quiet for the female (top row) and male (bottom row) talker sessions plotted as a function of age (left panels) and hearing loss (right panels). Results of the regression analyses on these scores (Table III) indicated that both age and hearing loss were significant predictors, with scores declining as age or hearing loss increased, for both female and male talker sessions.[2]

Inspection of Fig. 4 and Table II suggests that a potentially close relationship exists between quiet and noise/competing talker performance, in that the group differences found in the quiet condition, generally, were also apparent in the noise and competing talker conditions. To examine this relationship more closely, Fig. 6 shows noise (left) and competing talker (right) scores (averaged across TMR for each participant) plotted against quiet scores for the female (top row) and male (bottom row) talker sessions. These plots reveal that noise and competing talker scores were strongly associated with quiet scores. Thus, quiet scores were

included along with age and hearing loss as predictors in multiple regression analyses on the noise and competing talker scores (Table III). These analyses confirmed that scores in quiet were strong predictors of scores in both noise and competing talker conditions for both the female and male talker sessions. For the female talker session, the analyses indicated that the contributions of age and hearing loss were no longer significant in the noise/competing talker conditions when quiet scores were in the model. For the male talker session, however, the contribution of hearing loss persisted in both the presence of noise and the competing talker. It is somewhat puzzling that this additional contribution of hearing loss was only evident in the easier male talker session, but we note that many scores for the more difficult

TABLE II. Results of the ANOVAs conducted on the scores for each condition (quiet, noise, and competing talker) for both female and male talker sessions. Shown are the degrees of freedom (DF) along with F and p values.

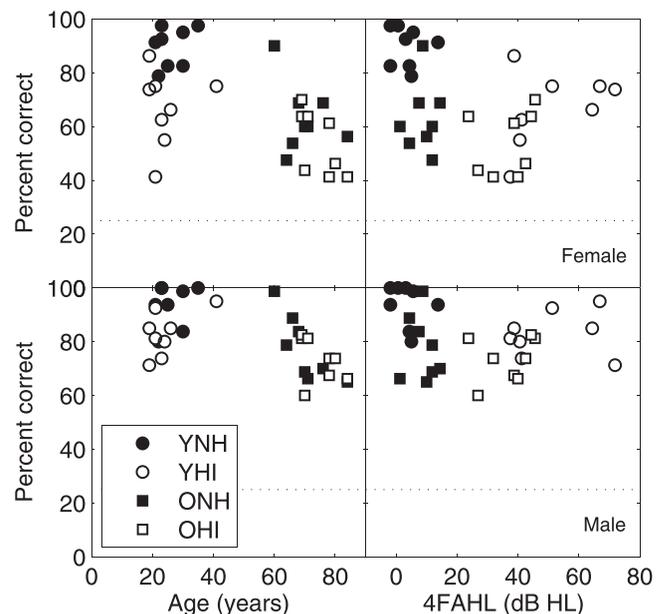|	Female			Male		
	DF	F	p	DF	F	p
Quiet						
Group	3,31	14.161	<0.001	3,31	8.299	<0.001
Noise						
Group	3,28	10.749	<0.001	3,28	11.673	<0.001
TMR	3,84	88.001	<0.001	3,84	88.434	<0.001
Group × TMR	9,84	2.384	0.019	9,84	2.764	0.007
Comp. Talk.						
Group	3,28	9.237	<0.001	3,28	12.233	<0.001
TMR	3,84	22.589	<0.001	3,84	29.805	<0.001
Group × TMR	9,84	1.271	0.265	9,84	2.285	0.024

FIG. 5. Individual scores in quiet for the female (top row) and male (bottom row) talker sessions plotted against age (left panels) and 4FAHL (right panels).

J. Acoust. Soc. Am. **143** (2), February 2018

Best *et al.*   1089

TABLE III. Results of the multiple regression analysis conducted on the scores for each condition (quiet, noise, and competing talker) for both female and male talker sessions. Shown are the R^2 values of the overall model fit, along with standardized coefficients (β) and p values for each predictor.

	Female		Male	
	β	p	β	p
Quiet	Model $R^2 = 0.50$		Model $R^2 = 0.42$	
Age	−0.61	<0.001	−0.57	<0.001
4FAHL	−0.41	0.004	−0.36	0.016
Noise	Model $R^2 = 0.79$		Model $R^2 = 0.92$	
Quiet	0.92	<0.001	0.84	<0.001
Age	0.13	0.279	−0.09	0.203
4FAHL	−0.08	0.433	−0.18	0.007
Comp. Talk	Model $R^2 = 0.72$		Model $R^2 = 0.86$	
Quiet	0.87	<0.001	0.82	<0.001
Age	0.04	0.797	−0.02	0.829
4FAHL	0.02	0.892	−0.22	0.008

female talker session were near chance for the ONH group, leaving little room to see an additional effect of hearing loss in older listeners.

D. Error patterns

To examine the pattern of errors made by participants in each condition, confusion matrices were generated and are shown in Fig. 7. These matrices show, for each talker in the set, the distribution of responses given on trials in which that talker was the target. Figure 7 shows that certain talkers (e.g., female 4, and male 3) were more robustly identified than others, which is seen both in the high proportion correct
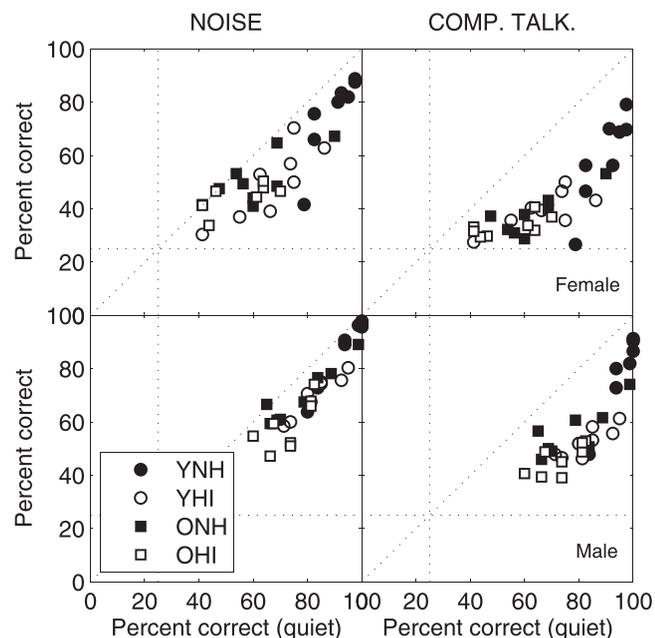

FIG. 6. Individual scores (averaged across TMR) plotted against quiet scores for the female (top row) and male (bottom row) talker sessions for both noise (left column) and competing talker (right column) conditions.

for those targets, and the low proportion of responses to those targets when they were not presented. An examination of Fig. 2 indicates that these two talkers had relatively extreme values of F0 (male 3 having the lowest of any of the talkers, and female 4 having one of the highest). This suggests that F0 may have been one of the vocal parameters, presumably from a complex set, that participants used to make decisions about talker identity in this experiment.

IV. DISCUSSION

In this study, we examined talker identification in quiet and in two kinds of interference (steady-state noise and a single same-sex competing talker). Both kinds of interference had a detrimental effect on talker identification scores, and there was evidence that this effect was greater when the interference was a competing talker. This suggests that the necessity to segregate similar talkers in the competing talker condition caused additional interference over that caused by a steady-state noise masker, as it does for the task of speech intelligibility (e.g., see discussion of the distinction between energetic and informational masking for speech in Kidd and Colburn, 2017). An additional aim was to understand the extent to which hearing loss and advanced age affect talker identification in quiet and with these two types of competing sounds. Broadly, our findings demonstrate that the ability to identify talkers is affected by both factors. For our participants, both increased hearing loss and increased age were associated with poorer talker identification. However, there were indications that these two factors exerted subtly different adverse effects.

The effect of hearing loss was apparent in quiet, but there was some evidence for an additional effect of hearing loss in the presence of competing sounds. In other words, hearing loss appeared to affect talker identification and susceptibility to noise. Given that the additional effect of hearing loss was significant only for the male talkers (and not the female talkers) it is a result that warrants further investigation. If it holds, this increased susceptibility to noise could be explained on the basis of reduced spectral and temporal resolution, typical in this population, which would further disrupt the features needed to discriminate between and identify the different talkers. While we have no direct evidence for this, there are certainly numerous studies in the literature showing that hearing loss disrupts the ability of listeners to use F0 and to discriminate vowel sounds, especially in the presence of masking (e.g., Summers and Leek, 1998; Arehart et al., 2005).

The effect of age was prominent under both quiet and masked conditions, but there was no evidence that it was stronger in the presence of competing sounds. This result is consistent with a broad disruption to the ability to identify voices, rather than a disruption to the representation of particular voice features. Moreover, there was no evidence for a specific effect of age in the competing talker condition as per our original hypothesis. Thus, it appears that the difficulties older listeners have inhibiting competing talkers for the task of speech recognition do not translate to the task of talker identification.
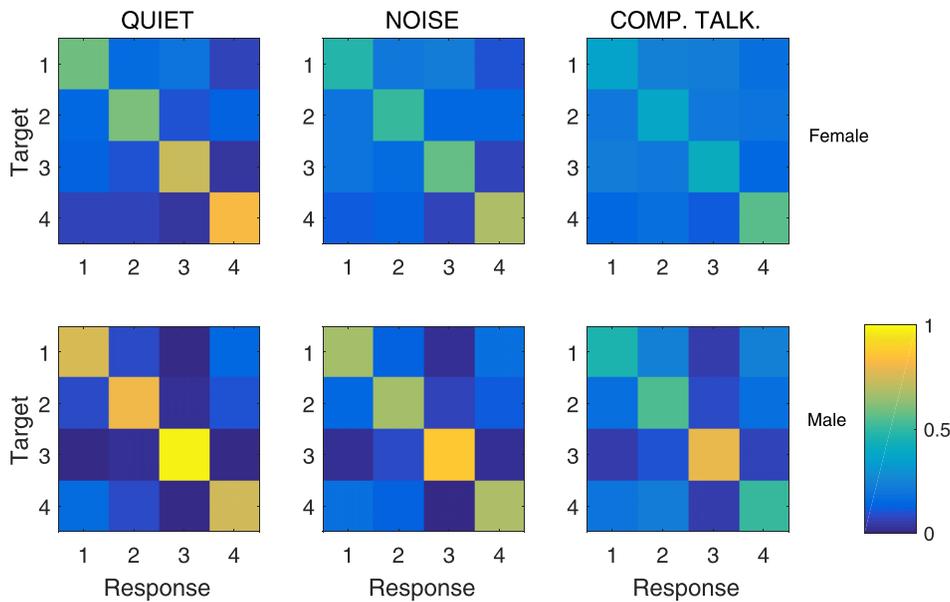
FIG. 7. (Color online) Confusion matrices showing, for each target talker, the distribution of responses across the four talkers (collapsed across subjects and TMRs). The color map indicates a proportion between zero and one, with each row summing to one. Female and male talker sessions are shown in the top and bottom rows, while the quiet, noise, and competing talker conditions are shown in the left, middle, and right columns, respectively.

One possible explanation for these effects is that memory limitations in older listeners affect the learning of new voices. In other words, while the immediate identification of talkers may be intact, the ability to store the voice information associated with a particular talker and to use stored representations to discriminate between similar voices may be disrupted. Certainly, there are numerous studies showing that older listeners have trouble remembering the *content* of heard speech. To give one recent example, Schurman *et al.* (2014) used a 1-back task to show that older listeners were poorer at identifying a sentence than younger listeners when asked to hold the sentence in memory, despite equivalent performance in the case of immediate recall. There are also studies showing that older listeners are poorer at recalling the voice sex for identified sentences (Kausler and Puckett, 1981; Spencer and Raz, 1995).

To a first approximation, these effects of hearing loss and age may have different physiological origins (peripheral and central, respectively). However, while the findings here are consistent with this speculation, further experiments that examine discrimination and memory independently are needed to confirm this conclusion. Moreover, some of the observed effects may be specific to our experimental design. For example, the decision was made to use unfamiliar voices in this experiment, to enable a consistent set of stimuli to be used for all listeners. However, this added the requirement that listeners learn the voices, which may have placed more emphasis on cognitive factors than if familiar voices had been used (e.g., Wenndt, 2016). In addition, we provided all listeners with a fixed amount of training (in terms of number of trials). Within this relatively brief training phase, listeners had to learn the identifying features of each talker, the distinguishing features between talkers in the set, and the labels associated with each talker. After this fixed amount of training, we found a strong effect of age in the testing phase, but it is not known whether the older listeners would have "caught up" to the younger listeners (with equivalent hearing losses) had they been allowed more time to train.

Finally, we observed a remarkable range of abilities across our 32 participants. While some of these individual differences were shown to relate to age and hearing loss, there was also substantial variability that was not explained by these factors. Of interest in future work will be to determine whether an individual's ability to identify talkers is reflected in other aspects of "cocktail party listening" such as understanding speech in noise, following multiperson conversations, or remembering who said what.

ACKNOWLEDGMENTS

[1]A statistical comparison of scores for the first and second halves found no consistent changes that would lead us to believe that either ongoing learning effects, or fatigue effects, affected the results.

[2]It is clear from Fig. 5 that the distribution of ages in our sample is not uniform. Specifically, as a result of recruiting participants into younger and older groups, there is a gap in the distribution between ages 41 and 60. Thus, our regression analysis cannot be considered to be predictive for middle-aged listeners.

Arehart, K. H., Rossi-Katz, J., and Swensson-Prutsman, J. (**2005**). "Double-vowel perception in listeners with cochlear hearing loss: Differences in fundamental frequency, ear of presentation, and relative amplitude," J. Speech Lang. Hear. Res. **48**, 236–252.

Best, V., Streeter, T., Roverud, E., Mason, C. R., and Kidd, G. (**2016**). "A flexible question-and-answer task for measuring speech understanding," Trends Hear. **20**, 1–8.

Boersma, P., and Weenink, D. (**2009**). "Praat: Doing phonetics by computer (version 6.0.20) [computer program]," http://www.praat.org (Last viewed September 4, 2016).

Byrne, D. J., Parkinson, A., and Newall, P. (**1991**). "Modified hearing aid selection procedures for severe-profound hearing losses," in *The*

Vanderbilt Hearing Aid Report II, edited by G. A. Studebaker, F. H. Bess, and L. B. Beck (York, Parkton, MD), pp. 295–300.

Chandrasekaran, B., Chan, A. H. D., and Wong, P. C. M. (**2011**). "Neural processing of what and who information in speech," J. Cognit. Neurosci. **23**, 2690–2700.

Dillon, H. (**2012**). *Hearing Aids* (Boomerang, Turramurra, Australia).

Formisano, E., De Martino, F., Bonte, M., and Goebel, R. (**2008**). " 'Who' is saying 'what'? Brain-based decoding of human voice and speech," Science **322**, 970–973.

Fuller, C. D., Gaudrain, E., Clarke, J. N., Galvin, J. J., Fu, Q. J., Free, R. H., and Başkent, D. (**2014**). "Gender categorization is abnormal in cochlear implant users," J. Assoc. Res. Otolaryng. **15**, 1037–1048.

Gatehouse, S., and Noble, W. (**2004**). "The Speech, Spatial and Qualities of Hearing Scale (SSQ)," Int. J. Audiol. **43**, 85–99.

Helfer, K. S., and Freyman, R. L. (**2009**). "Lexical and indexical cues in masking by competing speech," J. Acoust. Soc. Am. **125**, 447–456.

Janse, E. (**2012**). "A non-auditory measure of interference predicts distraction by competing speech in older adults," Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn. **19**, 741–758.

Johnsrude, I. S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H. P., and Carlyon, R. P. (**2013**). "Swinging at a cocktail party: Voice familiarity aids speech perception in the presence of a competing voice," Psychol. Sci. **24**, 1995–2004.

Kausler, D. H., and Puckett, J. M. (**1981**). "Adult age differences in memory for sex of voice," J. Gerontol. **36**, 44–50.

Kidd, G., and Colburn, H. S. (**2017**). "Informational masking in speech recognition," in *The Auditory System at the Cocktail Party*, edited by J. C. Middlebrooks and J. Z. Simon (Springer Nature, New York), pp. 75–109.

Latinus, M., and Belin, P. (**2011**). "Human voice perception," Curr. Biol. **21**, R143–R145.

Latinus, M., McAleer, P., Bestelmeyer, P. E. G., and Belin, P. (**2013**). "Norm-based coding of voice identity in human auditory cortex," Curr. Biol. **23**, 1075–1080.

Mackersie, C. L. (**2003**). "Talker separation and sequential stream segregation in listeners with hearing loss: Patterns associated with talker gender," J. Speech Lang. Hear. Res **46**, 912–918.

Mackersie, C. L., Prida, T. L., and Stiles, D. (**2001**). "The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss," J. Speech Lang. Hear. Res **44**, 19–28.

Nygaard, L. C., and Pisoni, D. B. (**1998**). "Talker-specific learning in speech perception," Percept. Psychophys. **60**, 355–376.

Perrachione, T. K., Pierrehumbert, J. B., and Wong, P. (**2009**). "Differential neural contributions to native- and foreign-language talker identification," J. Exp. Psychol. Hum. Percept. Perform. **35**, 1950–1960.

Perrachione, T. K., Stepp, C. E., Hillman, R. E., and Wong, P. C. (**2014**). "Talker identification across source mechanisms: Experiments with laryngeal and electrolarynx speech," J. Speech Lang. Hear. Res **57**, 1651–1665.

Razak, A., Thurston, E. J., Gustainis, L. E., Kidd, G. J., Swaminathan, J., and Perrachione, T. K. (**2017**). "Talker identification in three types of background noise," in *173rd Meeting of the Acoustical Society of America*, Boston, MA.

Remez, R. E., Fellowes, J. M., and Nagel, D. S. (**2007**). "On the perception of similarity among talkers," J. Acoust. Soc. Am. **122**, 3688–3696.

Rossi-Katz, J., and Arehart, K. H. (**2009**). "Message and talker identification in older adults: Effects of task, distinctiveness of the talkers' voices, and meaningfulness of the competing message," J. Speech Lang. Hear. Res. **52**, 435–453.

Schurman, J., Brungart, D., and Gordon-Salant, S. (**2014**). "Effects of masker type, sentence context, and listener age on speech recognition performance in 1-back listening tasks," J. Acoust. Soc. Am. **136**, 3337–3349.

Schvartz, K. C., and Chaterjee, M. (**2012**). "Gender identification in younger and older adults: Use of spectral and temporal cues in noise-vocoded speech," Ear Hear. **33**, 411–420.

Schweinberger, S. R., Kawahara, H., Simpson, A. P., Skuk, V. G., and Zäske, R. (**2014**). "Speaker perception," WIREs Cogn. Sci. **5**, 15–25.

Souza, P., Gehani, N., Wright, R., and McCloy, D. (**2013**). "The advantage of knowing the talker," J. Am. Acad. Audiol. **24**, 689–700.

Spencer, W. D., and Raz, N. (**1995**). "Differential effects of aging on memory for content and context: A meta-analysis," Psychol. Aging **10**, 527–539.

Studebaker, G. A. (**1985**). "A 'rationalized' arcsine transform," J. Speech. Lang. Hear. Res. **28**, 455–462.

Summers, V., and Leek, M. R. (**1998**). "F0 processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss," J. Speech. Lang. Hear. Res. **41**, 1294–1306.

Vongphoe, M., and Zeng, F. G. (**2005**). "Speaker recognition with temporal cues in acoustic and electric hearing," J. Acoust. Soc. Am. **118**, 1055–1061.

Wenndt, S. (**2016**). "Human recognition of familiar voices," J. Acoust. Soc. Am. **140**, 1172–1183.

Wright, B. A., Sabin, A. T., Zhang, Y., Marrone, N., and Fitzgerald, M. B. (**2010**). "Enhancing perceptual learning by combining practice with periods of additional sensory stimulation," J. Neurosci. **30**, 12868–12877.

Yonan, C. A., and Sommers, M. S. (**2000**). "The effects of talker familiarity on spoken word identification in younger and older listeners," Psychol. Aging **15**, 88–99.