

Effect of speaker age on speech recognition and perceived listening effort in older adults with hearing loss

Megan J. McAuliffe PhD^{1,2}
Phillipa J. Wilding MAud¹
Natalie A. Rickard PhD^{1,2}
Greg A. O'Beirne PhD^{1,2}

¹Department of Communication Disorders, University of Canterbury, Christchurch, New Zealand

²New Zealand Institute of Language, Brain & Behaviour, University of Canterbury, Christchurch, New Zealand

Keywords: aging, hearing loss, effort, speech recognition, speech, listener

Correspondence Address:

Dr Megan McAuliffe
Department of Communication Disorders
University of Canterbury
Private Bag 4800
Christchurch, New Zealand

Ph: +64 3 364 2987 ext. 7075

Fax: +64 3 364 2760

Email: megan.mcauliffe@canterbury.ac.nz

*Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.*

Abstract

Purpose: Older adults exhibit difficulty understanding speech that has been experimentally degraded. Age-related changes to the speech mechanism lead to natural degradations in signal quality. We tested the hypothesis that older adults with hearing loss would exhibit declines in speech recognition when listening to the speech of older adults, compared with the speech of younger adults, and would report greater amounts of listening effort in this task.

Methods: Nineteen individuals with age-related hearing loss completed speech recognition and listening effort scaling tasks. Both were conducted in quiet, when listening to high and low predictability phrases produced by younger and older speakers respectively.

Results: No significant difference in speech recognition existed when stimuli were derived from younger or older speakers. However, perceived effort was significantly higher when listening to speech from older adults, as compared to younger adults.

Conclusions: For older individuals with hearing loss, natural degradations in signal quality may require greater listening effort. However, they do not interfere with speech recognition – at least in quiet. Follow-up investigation of the effect of speaker age on speech recognition and listening effort under more challenging noise conditions appears warranted.

Introduction

For older adults, age-related reductions in the ability to understand speech have significant negative effects on communication. A variety of factors are known to influence speech understanding in this population including the degree of hearing loss, integrity of the central auditory pathway, and age-related declines in cognition (CHABA, 1988). However, signal-dependent factors also influence the speech understanding abilities of older adults. For example, older individuals consistently exhibit difficulty recognising rapid or time-compressed speech (e.g., Gordon-Salant & Fitzgibbons, 1993; Gordon-Salant & Fitzgibbons, 1999, 2001; Konkle, Beasley, & Bess, 1977; Vaughan & Letowski, 1997), as well as reverberant speech (e.g., Gordon-Salant & Fitzgibbons, 1993; Halling & Humes, 2000; Humes & Christopherson, 1991). Older individuals with hearing loss have also been shown to exhibit poorer recognition of synthesised speech compared to younger adults (Humes, Nelson, & Pisoni, 1991). Furthermore, older adults exhibit poorer recognition of accented speech compared to younger adults; though it appears that this group may not be disproportionately affected (Ferguson, Jongman, Sereno, & Keum, 2010; Gordon-Salant, Yeni-Komshian, & Fitzgibbons, 2010). Moreover, older adults' recall abilities are detrimentally affected by faster than normal speech rates (Stine, Wingfield, & Poon, 1986). However, one ubiquitous signal-dependent factor with the potential to negatively influence speech recognition is yet to be investigated – the naturally-occurring signal degradation associated with the aging voice. This preliminary investigation explores whether, and how, age of the speaker influences both speech recognition and perceived listening effort in older adults with hearing loss.

Both the temporal and spectral features of the speech signal undergo change as people age (Torre & Barlow, 2009; Zraick, Gregg, & Whitehouse, 2006). If one considers these age-related

*Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.*

changes as a natural degradation in speech signal quality, then it follows that such age-related changes may also influence the speech recognition abilities of older adults – particularly those with hearing loss. Any decline in speech recognition that occurs as a result of a naturally degraded speech signal may result in noteworthy effects to communication – effects particularly salient considering that the primary communication partners of older adults are likely to be older adults themselves. Anecdotally, older adults commonly report that the speech of young people is difficult to understand. While this observation clearly results from contextual as well as linguistic factors, it is possible that the faster speech rate commonly observed in younger speakers (e.g., Harnsberger, Shrivastav, Brown, Rothman, & Hollien, 2008; Jacewicz, Fox, O'Neill, & Salmons, 2009; Smith, Wasowicz, & Preston, 1987) may play a part in this observation. To our knowledge, research is yet to consider whether older listeners with hearing loss exhibit differential performance on speech recognition tasks when speech stimuli are presented from both younger and older speakers.

The quality of the speech signal is affected by aging. A number of studies have observed the ability of listeners to distinguish between young and older speakers based on auditory perception alone (e.g., Ptacek & Sander, 1966; Ryan & Burk, 1974; Shipp & Hollien, 1969). The characteristic “older voice” is thought to result from a complex interaction of sensorimotor changes that accompany aging – for example, speech motor changes (Liss, Weismer, & Rosenbek, 1990; Ryan & Burk, 1974) and declines in auditory feedback (Liss, et al., 1990) – and cognitive-linguistic factors such as slowed cognitive processing. As with age-related changes to other systems, there is much individual variation with regards to the rate and extent of the effect (Mueller, 1997). Significant differences exist between men and women (Gorham-Rowan &

Laures-Gore, 2006; Torre & Barlow, 2009), and the overall physical condition of the individual also plays a role in the speech changes observed (Ramig & Ringel, 1983).

Perceptual studies have indicated that older voices may be perceived as breathy, hoarse, unstable and different in pitch compared to younger voices (Gorham-Rowan & Laures-Gore, 2006; Mueller, 1997; Ptacek, Sander, Maloney, & Jackson, 1966). Acoustic studies confirmed these voicing changes noted perceptually. In general, the average F0 of females tends to drop with age, whereas the average F0 of older males tends to rise (Torre & Barlow, 2009). Studies have also demonstrated that fundamental frequency varies more for older than younger adults (Gorham-Rowan & Laures-Gore, 2006; Torre & Barlow, 2009; Xue & Deliyski, 2001) and that voicing instability (e.g., as measured through jitter, shimmer, and noise-to-harmonics ratios) is present in the speech signals of older adults (Xue & Deliyski, 2001).

Articulation also changes as people age. Increased segment durations (Benjamin, 1982; Smith, et al., 1987) have been a relatively consistent finding, along with reduced speed of performance on oral diadochokinetic tasks (Padovani, Gielow, & Behlau, 2009; Parnell & Amerman, 1987), presence of spirantization in the stop gap (Liss, et al., 1990) and reduced tongue-to-palate contact during fricative production (McAuliffe, Ward, & Murdoch, 2006). Age-related vowel centralization has been observed in a number of cases, particularly for very old individuals (Benjamin, 1982; Liss, et al., 1990). This generally matches with the perceptual impression of reduced consonant imprecision observed in older adults. Indeed, some authors have suggested that the speech of older healthy adults may exhibit mild characteristics of dysarthria, a speech motor disorder associated with neurologic impairment (Amerman & Parnell, 1990; Liss, et al., 1990; Ryan & Burk, 1974).

Conceivably, the changes to speech production that occur with age may affect listeners' perception – particularly that of older adults with hearing loss. For example, if we consider the finding that older adults with hearing loss demonstrate reduced ability to recognise speech that is experimentally degraded, then it follows that the common perceptual characteristics of phonatory instability associated with the aging voice could negatively influence speech recognition. As suggested by Torre and Barlow (2009), age-related changes to vowel formant frequencies may blur phoneme boundaries, affecting consonant recognition and hence perception. Furthermore, vowel centralisation – a phenomenon observed with aging – is also seen in individuals with dysarthria and, in this population, is related to reduced speech intelligibility (Weismer, 1997). On this basis, it is not unreasonable to suggest that the aging voice, while its degradations to signal quality would be considered mild, has the potential to negatively influence speech recognition in older individuals with hearing loss.

One caveat to this hypothesis relates to changes in speech rate with age. Speaking and reading rates have been shown to slow with age (Harnsberger, et al., 2008; Jacewicz, et al., 2009; Smith, et al., 1987). Therefore, it is possible that this reduced speech rate could improve speech recognition in older individuals with hearing loss due to the availability of increased processing time¹. When combined, the characteristics of a slower speech rate and increased variation in F0 may indeed highlight pertinent aspects of the speech signal and facilitate speech recognition. It is known that listeners tune in to the rhythmic components of speech to facilitate lexical segmentation (Cutler & Butterfield, 1992). If these were highlighted by older adult speakers, speech recognition of stimuli from older adults could be recognised with greater accuracy than that of younger adult speakers.

¹ Noting, however, that increasing inter-word intervals has been found to improve sentence recall in only a subset of older individuals with hearing loss (Gordon-Salant & Fitzgibbons, 1997).

The use of speech stimuli from older and younger speakers provides a unique opportunity to examine how naturally-induced changes to the speech signal affect speech recognition in older adults with hearing loss. However, we propose that the aging voice may affect not only speech recognition, but also the degree of effort required to recognize speech. Particularly for those with hearing loss, the additional effort required to concentrate on listening and understanding can result in considerable fatigue (see Anderson Gosselin & Gagné, 2010). In some cases listeners may score highly on speech recognition tasks, but report that substantial ‘mental effort’ was required to complete the task (Anderson Gosselin & Gagné, 2010). McCoy et al. (2005) provided quantitative evidence of this effect, finding that for participants with mild-to-moderate hearing loss the increased perceptual effort required to decipher words produced notable effects on recall performance. Hearing impairment (McCoy, et al., 2005), age (e.g., Larsby, Hallgren, Lyxell, & Arlinger, 2005), and listening environment all play a role in influencing the degree of perceived listening effort. We posit that the integrity of the speech signal is an additional component that may influence perceived effort. Evidence in support of this position is found in the study of Whitehill and Wong (2006) who observed that impaired voice quality was strongly positively correlated with ratings of listener effort. In addition, the presence of articulation errors and ‘slurred’ speech was also moderate-to-strongly correlated with higher degrees of listener effort. On this basis, it seems reasonable to suggest that speech produced by older adults may also require increased effort on the part of the listener to correctly perceive the message.

In summary, the range of speech characteristics affected by the process of aging combines to result in a mild degradation in acoustic signal quality. Conceivably, this natural degradation may increase the difficulties with speech recognition that older adults with hearing loss tend to experience, and/or result in an increased amount of required effort to accurately complete the

*Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.*

task of speech recognition. To our knowledge, research is yet to consider how the aging speech signal influences speech recognition or listening effort. This preliminary study aimed to determine whether the speech recognition abilities of older adults with hearing loss, and their degree of perceived listening effort, varied as a factor of speaker age. It was hypothesised that listeners would exhibit reduced speech recognition, and increased listening effort, when listening to the speech of older adults.

Method

Listener Participants

Nineteen individuals with age-related hearing loss participated in the study (10 males and nine females). All were native speakers of New Zealand English and aged between 60 and 87 years ($M = 71.4$ years, $SD = 8.48$ years). Degree of hearing loss was determined using behavioural pure tone audiometry. Calculations of the pure tone average (PTA) of thresholds at 500, 1000, 2000 and 4000, and 8000 Hz were conducted. A PTA of 20 dB HL or worse in the better hearing ear was required for participation. Hearing losses were also required to be symmetrical, with interaural differences in PTA of no greater than 19 dB at any frequency (Jerger, Jerger, & Pirozzolo, 1991). Participants with a hearing loss from childhood, a previous history of neurological disorder, dementia or other significant medical history were excluded from the study. Figure 1 shows the hearing thresholds (average of the left and right ears) of the participants. As can be seen, the participants exhibited a range of hearing loss severities; however, the pattern of sloping high frequency hearing loss appeared consistent and characteristic of age-related hearing loss (Gates & Mills, 2005). Sixteen of the 19 participants had Type A tympanograms in at least one ear, with none showing a Type B or Type C

tympanogram. These tympanometric findings, taken together with the configuration and symmetry of the audiometric results, and the lack of past history of ear disease, are consistent with primarily sensorineural hearing losses in all participants.

[insert Figure 1 near here]

Ten of the 19 participants were hearing aid owners; nine owned binaural hearing aids and one owned a unilateral hearing aid. Of these, only five reported wearing their hearing aids on a regular basis. These participants did not wear their hearing aids during the listening experiments. Participants were recruited from the client database at the University of Canterbury Speech and Hearing Clinic and each received a \$10 petrol voucher as compensation for their involvement in the project. Approval for this study was obtained from the University of Canterbury Human Ethics Committee and all individuals provided written consent prior to undertaking the study.

Speech Stimuli

Experimental speech stimuli were recorded from a total of eight speakers, four of whom were classified as ‘young’ (two males and two females, $M = 27$ years, $SD = 2.06$ years) and four of whom were classified as ‘older’ (two males and two females, $M = 80$ years, $SD = 5.80$ years). All had been selected as their speech was considered, by the authors, as representative of ‘young’ and ‘older’ New Zealand English speech respectively. All participants were free of colds or other respiratory issues that may have affected their speech at the time of the recording and reported that they were in good physical condition. In addition, participants reported no prior history of neurological disorder, speech or language disorder, or uncorrected hearing loss.

Speech samples were collected during a single one-hour session with each speaker. Two sets of 72 experimental phrases, all six syllables in length, were recorded. Set one comprised low

*Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.*

inter-word predictability phrases, which were chosen in order to lessen semantic and linguistic cues that might assist in speech understanding (Liss, Spitzer, Caviness, Adler, & Edwards, 1998). Set two included high inter-word predictability phrases, adapted from the Speech Perception in Noise (SPIN) Test (Kalikow, Stevens, & Elliott, 1977). Appendix A contains a list of the experimental phrases. Digital audio recordings of the speech samples were made in a sound treated room. During recordings, participants wore an Audix HT2 Headset Condenser Microphone placed approximately six centimetres from the mouth. Stimuli were recorded directly to a laptop computer using Sony Sound Forge Version 9.0 at a sampling rate of 48 kHz with 16 bits of quantisation.

Following recording, a phrase stimuli selection process was undertaken with the aim of selecting those phrases considered representative of ‘young’ and ‘older’ speech characteristics for use in the listening experiments. Similar procedures have been used in studies examining the perception of dysarthric speech (e.g., Borrie, et al., in press; Liss, et al., 1998). Firstly, the data set was screened and any phrases that exhibited hesitations or reading errors were omitted from the pool of phrases for possible inclusion in the perception experiment. Then the first and second authors, both speech pathologists with experience in motor speech disorders, judged the perceptual characteristics and acceptability of the phrases for use in the experiment. For the younger voices, the process was simply to select those phrases that exhibited normal speech and voice production. For the older voices, phrases that exhibited perceptual evidence of age-related speech characteristics including slower rate and voicing changes such as hoarseness and instability were selected as representative of ‘older’ voices for the experiment. The final experimental stimuli set consisted of 144 phrases – 72 from ‘young’ speakers and 72 from ‘older’ speakers, with an equivalent number of phrases from each of the eight speakers (i.e., 18

phrases per speaker). Each stimulus set of 72 phrases consisted of 36 high predictability and 36 low predictability phrases, nine of each predictability type from each speaker. The amplitudes of the selected speech samples were normalised to an RMS level of -18.5 dB (re: 0 dB full scale). An equal-loudness contour was used for the level calculation, and parts of the file with an amplitude less than -50 dB (calculated with an attack/release time of 200 ms) were ignored.

To confirm that speech production differences existed between the younger and older stimuli set, the selected phrases were subject to acoustic analysis. All acoustic measures were completed using TF32 (Milenkovic, 2002). The following analyses were conducted over the duration of the phrase – fundamental frequency (F0), variation in fundamental frequency (F0 SD), variation in amplitude (dB SD), and speech rate (in syllables per second). To perform this analysis, the beginning and end point of each phrase was selected by placing cursors on the first and last evidence of phonemes on the spectrographic display. In addition, pitch traces were visually inspected to identify apparent anomalies which were removed before analysis. Following the phrase level analyses, examination of the first, second and third formants (F1, F2 and F3) of the START, FLEECE and THOUGHT vowels for modern New Zealand English (Maclagan, 2009) was conducted. An equal number of vowels from each group, balanced between male and female, were analysed. Measures were taken from the temporal midpoints of each vowel using both spectrograms and waveform displays. While the inclusion of measures of jitter, shimmer, and noise-to-harmonics ratio would have augmented the perceptual voicing findings, the intent of the acoustic analysis was to confirm differences based on the experimental phrases employed in the study. The voice acoustic measures should be conducted on sustained phonation and such measures were not included in the study. The resultant data were combined

and mean and standard deviation values were calculated for young females, young males, older females, and older male speakers respectively (see Tables 1 and 2).

[insert Tables 1 & 2 near here]

Given that differences in acoustic parameters exist across males and females (Torre & Barlow, 2009), both groups were analysed independently. Therefore, the results of the young male and older male speakers and the results of the young female and older female groups were compared using independent groups *t*-tests. Given the number of comparisons completed, a more conservative alpha level of $p \leq 0.01$ was employed (Shearer, 1982). Statistical analysis revealed that the speech rates of the older males and older females were significantly slower than their younger counterparts (females, $t(70) = 9.51, p < 0.001$; males, $t(70) = 7.76, p < 0.001$), confirming our perceptual impressions, and consistent with prior research that has found slower speech rates in older individuals. The older females exhibited significantly greater pitch variation than the younger females, $t(70) = -3.68, p < 0.001$, which was not seen for males, $t(70) = -1.10, p = .28$. However, older men exhibited significantly increased F0 compared to younger males, $t(70) = -2.90, p < .01$, consistent with prior aging and voice research (Harnsberger, et al., 2008) though there were no reported differences for females, $t(70) = -1.25, p = .22$. There were no significant differences between groups for the parameter of amplitude variation (females, $t(70) = .78, p = .44$; males, $t(70) = -.46, p = .65$). For the vowel analysis, there were no significant differences in F1, F2 or F3 values across vowels when the young female and older female groups were compared ($p > 0.01$). When the male speakers were compared, only one significant difference was observed, with a significant raising of F1 for the FLEECE vowel in older males, $t(14) = -5.42, p < 0.001$. From this analysis, it appeared that there was minimal to no evidence of age-related vowel changes in the experimental stimuli.

***Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.***

Reliability of Acoustic Measures

Twenty percent of the acoustic data set (10% high predictability phrases and 10% low predictability phrases) was re-measured for reliability purposes. To determine intra-rater reliability, the investigator who conducted the initial measurements also completed the second set of reliability measures. For inter-rater reliability, an investigator not involved in conducting the original measurements completed the same analysis. Pearson product moment correlations examined the strength of relationships between these measures and absolute between-measure differences were also calculated. Analysis indicated that reliability was found to be acceptable. These data are presented in Table 3.

[insert Table 3 near here]

Procedures

All participants passed a pre-experiment cognitive screen – the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975) – and underwent several standard audiological assessments to confirm their suitability for participation in the study prior to completing the experimental tasks. The listening task took approximately 40 minutes to complete. The experimental tasks were completed using the University of Canterbury Perceptual Speech Ratings (UC-PSR) computer programme (O'Beirne, 2009).

For the completion of all tasks, participants were seated in front of a laptop. The experimental phrases were presented through Sennheiser HD280 Pro circumaural headphones. Prior to commencement of the experiment, a speech sample from a speaker (who was not included in the final stimuli set) was presented, and the participants were instructed to use the

*Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.*

on-screen sliding scale to adjust the volume until it was at a comfortable listening level. No further volume adjustments were allowed after this point. In addition, those participants who had not had previous experience with computers were given brief instruction on how to operate the mouse. The researcher operated the mouse on the behalf of those participants who were not comfortable doing so.

The participants were advised that they would hear some short phrases, which were to be spoken by both males and females of different ages. The phrases were presented one at a time and the participant controlled the rate of presentation, with the next phrase presented only when the subject clicked the “next” button. In addition they were told, in lay terms, that some of the phrases would make sense but others may not; however, that all phrases contained real English words. Participants were instructed to listen to each phrase and repeat it exactly as they heard it. The order of phrase presentation was randomly generated for each participant and each phrase was presented only once². Listeners were encouraged to give their best attempt if they were unsure of the complete phrase. Following each attempt the researcher typed their response into the computer, giving participants the chance to confirm that what was recorded was accurate. This procedure was used for all participants as the majority had indicated that they would not be comfortable typing their own responses.

After providing their response, participants were further instructed to rate how much effort was required to recognise each phrase using a computer-based listener effort scale. A 10 cm visual analogue scale was presented on the screen and participants were required to point the mouse to a location on a continuum, from “minimal effort” to “maximum effort” (as per

² With the exception of those used for reliability purposes which were presented after the experimental phrases were completed.

Whitehill & Wong, 2006). Effort ratings were recorded on the basis of the distance (in centimetres) from the left end of the scale to the marked point.

Each phrase transcription was then scored with regards to the percentage of words correct. Words correct were defined as those words that matched the intended target exactly. Exceptions to this were words that differed only by the tense “ed” or the plural “s”, substitutions between “a” and “the”, and homophones. For the effort rating task, the computer programme converted the visual analogue scale into a score of between 0 and 10.

Measurement Reliability

To assess the reliability of responses during the listening experiment, approximately 10% of all phrases (equal numbers of the low predictability and high predictability phrases) were repeated. For intelligibility, the average absolute difference between the first and second presentations of the stimuli was two percentage points, and the correlation was $r = .73, p < .001$. For effort, the mean absolute difference was .35 and the correlation was $r = .77, p < .001$. The reliability of both measures was considered acceptable.

Results

The mean speech recognition and perceived listening effort scores of the group of 19 older listeners are presented in Table 4. Mean and standard deviation scores are presented by age group of the speaker (i.e., younger versus older) and stimulus predictability (high predictability versus low predictability phrases). To model the effects of speaker age group and stimulus predictability, taking into account presentation volume, listener age, and listener PTA, we fit linear mixed effects models to the data in *R*, with listener and phrase as random effects (Baayen,

*Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.*

Davidson, & Bates, 2008). Mixed effects models are advantageous for repeated measures data as they enable simultaneous consideration of all factors that may influence the structure of a data set; while limiting the power of trends carried by specific individuals or phrases (Baayen, et al., 2008). In the current study, separate models were constructed for speech recognition and perceived listener effort.

[insert Table 4 near here]

For speech recognition, the dependent variable was first converted using an arcsine transformation, to account for the positive skew of the data set. The best-fit model for speech recognition is presented in Table 5. Factors tested included age of the speaker group, stimulus predictability, stimulus presentation volume, listener age, and listener PTA. As expected, the model shows that speech recognition was significantly reduced for low predictability phrases compared to high predictability phrases. However, there was no effect of speaker group, signalling that the participants exhibited a similar level of difficulty understanding the speech of both the young and older speaker groups. Given that the primary variable of interest, the effect of speaker age, was non-significant the determination of the effect of listener age or PTA upon the result was not of theoretical interest. However PTA did, in general, have a significant negative effect on speech recognition scores in the current data set. It should also be noted that a separate model was constructed that included listener age, but not PTA (as these co-varied to a moderate degree, $r = .44$, $p = .057$). Similarly, this model demonstrated that increased listener age also had a significant negative effect on speech recognition score. The inclusion of either variable in the model did not affect the primary finding – that speech recognition was not influenced by the age group of the speaker.

[insert Table 5 near here]

***Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.***

The best-fit model for perceived listening effort is presented in Table 6. As expected, stimulus predictability had a significant effect on listening effort scores, with higher perceived effort reported when listeners were presented with low predictability stimuli. Of greatest interest, the age group of the speaker had a significant effect on perceived listening effort. Results indicated that significantly higher levels of perceived listening effort were reported by the group when listening to the speech of older adults. Both age group of the speaker, and stimulus predictability, remained highly significant in models that include listener age and PTA as controls³. Only PTA approached significance, at approximately $p = .08$, across all models tested. There were no significant interactions between any of the factors. In sum, these results indicate that listeners reported significantly higher levels of perceived effort when listening to speech from the older adult speaker group.

[insert Table 6 near here]

Discussion

The study found that a group of 19 older listeners with hearing loss obtained similar speech recognition scores when transcribing the speech of older and younger adult speakers. While the aging speech signal did not result in demonstrable changes to speech recognition, perceived listening effort was significantly higher when listening to older, compared with younger, speakers. It appeared that some component of the aging speech signal caused listeners with hearing loss to perceive that additional effort was required to decipher it. This effect was robust, and independent of listeners' age and PTA.

³ It should be noted that the covariance of listener age and PTA is not problematic in relation to these findings, as these variables do not reach independence separately, together, or in interaction. Therefore, there is no case for including them together in the same model.

In contrast to our hypothesis, the older listeners studied did not demonstrate greater difficulty understanding the speech of older as opposed to younger speakers. This pattern of results held for both the high predictability and low predictability phrases. Therefore, it would seem that the lack of difference between the two speech conditions (i.e., young speech and older speech) should likely be attributed to either the speech stimuli employed in the investigation or the degree of difficulty of the perception task, or a combination of both. Similar to previous research (e.g., Halling & Humes, 2000), performance on the speech recognition task was influenced by both the listeners' age and their PTA.

Regarding the speech signal, the older speakers exhibited a slower speaking rate than the younger speakers. Perhaps this slower rate of speech allowed listeners with hearing loss additional processing time and, hence, compensated for the perceptible differences in voice and speech quality exhibited by the older speaker group. While the two speaker groups were perceptibly different, and exhibited significant acoustic differences on a range of parameters, it is possible that these signal differences were simply not large enough to result in differences to speech recognition scores. In designing this preliminary study, it was thought that the low predictability speech condition would tax the perceptual system of those with hearing loss and, potentially, be more likely to reveal any differences. However, this was not the case. While prior studies have reported reduced intelligibility with experimental degradations in signal quality (Gordon-Salant & Fitzgibbons, 1993; Halling & Humes, 2000; Humes, et al., 1991), it appears that the naturally-occurring signal changes were less marked and contributed to the non-significant result. In addition, the design of the current experiment, with the speech signal presented only in quiet, may have contributed to the lack of difference in speech recognition scores. Older listeners with hearing loss commonly report a higher degree of difficulty

*Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.*

understanding speech in noise, an observation supported by experimental findings (e.g., Gordon-Salant & Fitzgibbons, 2004; Humes, 2002; Larsby, et al., 2005). The inclusion of noise conditions in future studies would likely provide a more challenging speech recognition task that may reveal differences in speech recognition when young and older voices are compared.

While no significant differences in speech recognition were found in relation to speaker age, the individuals with hearing loss did report significantly higher levels of perceived effort when listening to the speech of the older adults. Though the degree of perceived effort required for the task was relatively low, it seems that the perceptual and acoustic differences between the young and older speech stimuli were enough to result in increased perceptions of effort when attempting to decipher the aging speech signal. Based on the findings of Whitehill and Wong (2006), who reported correlations between voice quality and perceived effort and imprecise articulation and perceived effort for dysarthric speakers, it seems plausible to suggest that the voicing differences inherent in the speech signal of our older speaker cohort may have contributed to the perception of increased effort. Potentially, greater processing or attentional resources were required when listening to the older group in order to maintain equivalent levels of speech recognition. This interpretation bears some relationship to the findings of McCoy and colleagues (2005) who reported that while individuals with hearing loss successfully recognised the final words in three-word recall sets, this appeared to come at a processing cost to recall performance. This finding was attributed to the degree of perceptual effort required to complete the task of speech recognition successfully.

The concept that listeners may have been challenged when they listened to the aging voice, but not enough to reach a critical threshold to affect intelligibility is an interesting finding. While the current results should be viewed as preliminary, it seems that a potential interaction

*Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.*

between the voicing signal and listening effort is a worthwhile line of enquiry. Perception of effort was likely to be affected by numerous variables including the individual's perception of task difficulty, as well as their attention levels and cognitive resources. Recently, the use of a dual-task paradigm to objectively examine listening effort has been proposed (Anderson Gosselin & Gagné, 2010) and initial research using this modality has proved promising (Anderson Gosselin & Gagne, 2011; Sarampalis, Kalluri, Edwards, & Hafter, 2009). Based on these findings, and the results of the current study, research using a dual task paradigm that includes both a challenging noise condition and variation in the speech signal would be a potentially fruitful way with which to investigate the effect of the aging speech signal upon listening effort.

In conclusion, the current study found that older listeners with hearing loss exhibited similar levels of speech understanding when listening to the speech of younger and older adults. While speech recognition was not differentially affected, the listeners with hearing loss reported higher levels of perceived effort when listening to the speech of their older adult counterparts. To our knowledge, this study is the first to examine how a naturally degraded speech signal, associated with age, affects perception for older adults with hearing loss. While its findings are preliminary, it appears that the aging speech signal does have some influence on perception – even if it is simply in the degree of effort required to decipher the signal. This study highlights the importance of considering both listener and speaker in a communication exchange. A greater degree of listening effort required during dialogue between older adults with hearing loss and aged speech has implications for the design of aural rehabilitation plans that go beyond simply improving audibility. The greater degree of listener effort required when conversing with an aged speaker may result in poorer comprehension or greater fatigue on the part of the listener, for

*Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.*

example, and therefore may impact greatly on that listener's inclination to engage in such conversations. Any clinical measures that reduce the degree of listener effort required, or otherwise compensate for it, may well promote communication exchanges between older adults. A better understanding of the intrinsic and extrinsic factors that contribute to listener effort is therefore warranted. Future studies would benefit from the inclusion of more challenging noise conditions and speech stimuli from very old speakers – to increase the level of task difficulty and hence further tax the perceptual system. More importantly perhaps, further investigation of listening effort is warranted. To confirm the current findings, or otherwise, the use of a dual-task paradigm seems an obvious next step.

References

- Amerman, J. D., & Parnell, M. M. (1990). Auditory impressions of the speech of normal elderly adults. *International Journal of Language & Communication Disorders*, 25(1), 35-43.
- Anderson Gosselin, P., & Gagne, J.-P. (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of Speech Language and Hearing Research*, 54, 944-958
- Anderson Gosselin, P., & Gagné, J.-P. (2010). Use of a dual-task paradigm to measure listening effort. *Canadian Journal of Speech-Language Pathology & Audiology*, 34(1), 43-51.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.
- Benjamin, B. J. (1982). Phonological performance in gerontological speech. *Journal of Psycholinguistic Research*, 11(2), 159-167.
- Borrie, S. A., McAuliffe, M. J., Liss, J. M., Kirk, C., O'Beirne, G., & Anderson, T. (in press). Familiarization conditions and the mechanisms that underlie improved recognition of dysarthric speech. *Language and Cognitive Processes*.
- Committee on Hearing and Bioacoustics and Biomechanics (CHABA). (1988). Speech understanding and aging. *Journal of the Acoustical Society of America*, 83, 859-895.
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, 31, 218-236.
- Ferguson, S. H., Jongman, A., Sereno, J. A., & Keum, K. A. (2010). Intelligibility of foreign-accented speech for older adults with and without hearing loss. *Journal of the American Academy of Audiology*, 21(3), 153-162.

***Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.***

- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198.
- Gates, G. A., & Mills, J. H. (2005). Presbycusis. *The Lancet, 366*(9491), 1111 - 1120.
- Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech and Hearing Research, 36*(6), 1276-1285.
- Gordon-Salant, S., & Fitzgibbons, P. J. (1997). Selected cognitive factors and speech recognition performance among young and elderly listeners. *Journal of Speech Language and Hearing Research, 40*, 423-431.
- Gordon-Salant, S., & Fitzgibbons, P. J. (1999). Profile of auditory temporal processing in older listeners. *Journal of Speech, Language and Hearing Research, 42*(2), 300-311.
- Gordon-Salant, S., & Fitzgibbons, P. J. (2001). Sources of age-related recognition difficulty for time-compressed speech. *Journal of Speech, Language and Hearing Research, 44*(4), 709-719.
- Gordon-Salant, S., & Fitzgibbons, P. J. (2004). Effects of stimulus and noise rate variability on speech perception by younger and older adults. *Journal of the Acoustical Society of America, 115*(4), 1808-1817.
- Gordon-Salant, S., Yeni-Komshian, G. H., & Fitzgibbons, P. J. (2010). Recognition of accented English in quiet by younger normal-hearing listeners and older listeners with normal-hearing and hearing loss. *Journal of the Acoustical Society of America, 128*(1), 444-455.
- Gorham-Rowan, M. M., & Laures-Gore, J. (2006). Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders, 39*(3), 171-184.

***Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.***

- Halling, D. C., & Humes, L. E. (2000). Factors affecting the recognition of reverberant speech by elderly listeners. *Journal of Speech Language and Hearing Research*, 43(2), 414-431.
- Harnsberger, J. D., Shrivastav, R., Brown, W. S., Rothman, H., & Hollien, H. (2008). Speaking rate and fundamental frequency as speech cues to perceived age. *Journal of Voice*, 22(1), 58-69.
- Humes, L. E. (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *Journal of the Acoustical Society of America*, 112(3), 1112-1132.
- Humes, L. E., & Christopherson, L. (1991). Speech identification difficulties of hearing-impaired elderly persons: The contributions of auditory processing deficits. *Journal of Speech and Hearing Research*, 34(3), 686-693.
- Humes, L. E., Nelson, K. J., & Pisoni, D. B. (1991). Recognition of synthetic speech by hearing-impaired elderly listeners. *Journal of Speech and Hearing Research*, 34(5), 1180-1184.
- Jacewicz, E., Fox, R. A., O'Neill, C., & Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language Variation and Change*, 21(2), 233-256.
- Jerger, J., Jerger, S., & Pirozzolo, F. (1991). Correlational analysis of speech audiometric scores, hearing loss, age, and cognitive abilities in the elderly. *Ear and Hearing*, 12(2), 103-109.
- Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61(5), 1337-1351.
- Konkle, D. F., Beasley, D. S., & Bess, F. H. (1977). Intelligibility of time-altered speech in relation to chronological aging. *Journal of Speech and Hearing Research*, 20(1), 108-115.

- Larsby, B., Hallgren, M., Lyxell, B., & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology, 44*, 131-143.
- Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *Journal of the Acoustical Society of America, 104*, 2457-2466.
- Liss, J. M., Weismer, G., & Rosenbek, J. C. (1990). Selected acoustic characteristics of speech production in very old males. *Journal of Gerontology: Psychological Sciences, 45*(2), P35-P45.
- Maclagan, M. (2009). Reflecting connections with the local language: New Zealand English. *International Journal of Speech-Language Pathology, 11*(2), 113-121.
- McAuliffe, M. J., Ward, E. C., & Murdoch, B. E. (2006). Speech production in Parkinson's disease: I. An electropalatographic investigation of tongue-palate contact patterns. *Clinical Linguistics and Phonetics, 20*(1), 1-18.
- McCoy, S. L., Tun, P. A., Clarke Cox, L., Colangelo, M., Stewart, R. A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *The Quarterly Journal of Experimental Psychology, 58A*(1), 22-33.
- Milenkovic, P. (2002). TF32. Madison, WI: University of Wisconsin-Madison, Department of Electrical and Computer Engineering.
- Mueller, P. B. (1997). The aging voice. *Seminars in Speech and Language, 18*(2), 159-169.
- O'Beirne, G. A. (2009). UC PSR - University of Canterbury Perceptual Speech Ratings. Christchurch: University of Canterbury.

***Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.***

- Padovani, M., Gielow, I., & Behlau, M. (2009). Phonarticulatory diadochokinesis in young and elderly individuals. *Arquivos de Neuro-psiquiatria*, *67*(1), 58-61.
- Parnell, M. M., & Amerman, J. D. (1987). Perception of oral diadochokinetic performances in elderly adults. *Journal of Communication Disorders*, *20*, 339-352.
- Ptacek, P. H., & Sander, E. K. (1966). Age recognition from voice. *Journal of Speech and Hearing Research*, *9*, 353-360.
- Ptacek, P. H., Sander, E. K., Maloney, W. H., & Jackson, C. C. R. (1966). Phonatory and related changes with advanced age. *Journal of Speech and Hearing Research*, *9*(3), 353-360.
- Ramig, L. A., & Ringel, R. L. (1983). Effects of physiological aging on selected acoustic characteristics of voice. *Journal of Speech and Hearing Research*, *26*(1), 22-30.
- Ryan, W. J., & Burk, K. W. (1974). Perceptual and acoustic correlates of aging in the speech of males. *Journal of Communication Disorders*, *7*(2), 181-192.
- Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language, and Hearing Research*, *52*(5), 1230-1240.
- Shearer, W. M. (1982). *Research procedures in speech, language, and hearing*. Baltimore, MD: Williams & Wilkins.
- Shipp, T., & Hollien, H. (1969). Perception of the aging male voice. *Journal of Speech and Hearing Research*, *12*(4), 703-710.
- Smith, B. L., Wasowicz, J., & Preston, J. (1987). Temporal characteristics of the speech of normal elderly adults. *Journal of Speech and Hearing Research*, *30*, 522-529.
- Stine, E. L., Wingfield, A., & Poon, L. W. (1986). How much and how fast: Rapid processing of spoken language in later adulthood. *Psychology and Aging*, *1*(4), 303-311.

***Accepted to Journal of Speech, Language & Hearing Research (10th October 2011).
Do not cite without permission.***

- Torre, P., III, & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders*, 42(5), 324-333.
- Vaughan, N. E., & Letowski, T. (1997). Effects of Age, Speech Rate, and Type of Test on Temporal Auditory Processing. *J Speech Lang Hear Res*, 40(5), 1192-1200.
- Weismer, G. (1997). Motor speech disorders. In W. J. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 191-219). Cambridge, MA: Blackwell.
- Whitehill, T. L., & Wong, C.-Y. (2006). Contributing factors to listener effort for dysarthric speech. *Journal of Medical Speech-Language Pathology*, 14(4), 335-340.
- Xue, S. A., & Deliyski, D. (2001). Effects of aging on selected acoustic voice parameters: Preliminary normative data and educational implications. *Educational Gerontology*, 27(2), 159-168.
- Zraick, R. I., Gregg, B. A., & Whitehouse, E. L. (2006). Speech and voice characteristics of geriatric speakers: A review of the literature and a call for research and training. *Journal of Medical Speech - Language Pathology*, 14(3), 133-142.

Table 1: Mean and standard deviation values (in parenthesis) of selected acoustic measures completed on the experimental phrases and divided into male and female, young versus older speaker groups respectively.

	F0	F0 SD	dB SD	SR
YW	181.00 (21.79)	35.22 (12.61)	12.79 (2.01)	4.15 (0.52)
YM	106.68 (15.15)	19.79 (6.25)	12.19 (2.54)	4.08 (0.57)
OW	186.46 (14.40)	46.04 (12.32)	12.38 (2.38)	3.11 (0.40)
OM	117.01 (15.06)	21.34 (5.69)	12.50 (3.22)	2.94 (0.67)

Note: YW = young women ($n = 2$), YM = young men ($n = 2$), OW = older women ($n = 2$), OM = older men ($n = 2$). Standard deviation is in parenthesis. F0 = fundamental frequency, FO SD = fundamental frequency standard deviation, db SD = amplitude standard deviation, SR = speech rate in syllables per second.

Table 2: Mean and standard deviation values (in parenthesis) for F1, F2 and F3 taken from the temporal mid-point of selected vowels from the experimental phrases and divided into male and female, young versus older speaker groups respectively.

	FLEECE			START			THOUGHT		
	F1	F2	F3	F1	F2	F3	F1	F2	F3
YW	371 (35)	2589 (155)	3337 (319)	853 (55)	1604 (58)	2950 (127)	483 (42)	913 (90)	2765 (349)
YM	264 (19)	2079 (44)	2771 (158)	727 (52)	1428 (118)	2446 (97)	399 (42)	750 (88)	2175 (119)
OW	370 (56)	2597 (202)	3398 (257)	916 (50)	1649 (86)	2962 (328)	506 (55)	928 (125)	2958 (230)
OM	340 (35)	2168 (123)	2654 (85)	682 (24)	1393 (73)	2442 (116)	394 (46)	827 (149)	2509 (387)

Note: YW = young women ($n = 2$), YM = young men ($n = 2$), OW = older women ($n = 2$), OM = older men ($n = 2$).

Table 3: Intra and inter-rater reliability measures for the acoustic parameters.

	Intra-rater		Inter-rater	
	Average AD	Pearson's r	Average AD	Pearson's r
F0 SD	2.75	0.989	2.37	0.991
dB SD	0.42	0.985	0.35	0.984
SR	0.14	0.942	0.09	0.986
F1	11.08	0.997	11.33	0.995
F2	34.43	0.998	37.33	0.998

Note: AD = absolute difference, FO SD = fundamental frequency standard deviation, db SD = amplitude standard deviation, SR = speech rate in syllables per second, F1 = F1 in Hz combined across vowels, F2 = F2 in Hz combined across vowels.

Table 4: Mean and standard deviation (in parenthesis) scores from the 19 older listeners with hearing loss for speech recognition and perceived listening effort.

	Younger Speakers	Older Speakers
Speech recognition (%)		
High predictability phrases	98.03 (08.76)	97.45 (10.16)
Low predictability phrases	84.60 (24.78)	84.63 (23.59)
Listening Effort (score)		
High predictability phrases	0.81 (1.79)	1.05 (1.96)
Low predictability phrases	1.85 (2.50)	2.17 (2.68)

Table 5: Coefficients of a linear mixed effects model for speech recognition, with listener and phrase as random effects.

	Estimate	Standard Error	t-value	Pr(> t)
(intercept)	1.919	0.097	19.725	0.0000
Group = young	0.010	0.0288	0.354	0.7237
Predict = low	-0.285	0.0288	-9.895	0.0000
PTA	-0.011	0.003	-4.501	0.0000

Table 6: Coefficients of a linear mixed effects model for listener effort, with listener and phrase as random effects.

	Estimate	Standard Error	t-value	Pr(> t)
(intercept)	-1.469	1.529	-0.961	0.3367
Group = young	-0.282	0.085	-3.325	0.0009
Predict = low	1.081	0.085	12.764	0.0000
PTA	0.070	0.041	1.714	0.0867

Figure 1: Pure-tone air conduction thresholds (averaged left and right ear) for the older listeners with hearing loss ($n = 19$).

