# INTELLIGIBILITY OF OLDER TALKERS ON THE NU6 TEST

**Huiwen Goy [*1], M. Kathleen Pichora-Fuller [†1,2] and Pascal van Lieshout [‡1,2,3,4]**
[1]Department of Psychology, University of Toronto, 3359 Mississauga Rd, Mississauga ON L5L 1C6
[2]Toronto Rehabilitation Institute, 550 University Ave, Toronto ON M5G 2A2
[3]Department of Speech-Language Pathology, University of Toronto, 500 University Ave, Toronto ON M5G 1V7
[4]Rehabilitation Sciences Institute, University of Toronto, 500 University Ave, Toronto ON M5G 1V7

## 1 Introduction

Most speech intelligibility tests use recordings of younger adult talkers speaking in ideal conditions. However, talker variation is greater in the real world and includes factors such as age and speech changes due to the environment. Older talkers have rarely been recorded, even though it is well known that speech communication in noise is especially challenging for older adults. The first purpose of this study was to investigate the range of intelligibility scores in a sample of older talkers with different perceived speech quality. The second purpose was to determine whether older talkers could increase their intelligibility to a similar extent as younger talkers in past studies, when older talkers were asked to speak in noise or to speak clearly.

## 2 Method

### 2.1 Recording of speech materials

**Talkers**

Talkers were six female and two male adults (mean age = 79.6 years; range 73 to 88). Their hearing thresholds were similar to the 50th percentile for 70-year-olds [1]. Talkers had a variety of speech characteristics and were selected from a previous study in which listeners rated them on their suitability as an audiobook reader [2] (Table 1). Talkers were native Canadian English speakers and were in average to excellent health with no neurological or speech disorders.

**Table 1:** Perceived audiobook reader quality (Reader Scale of 1 to 5; 5 is best) and acoustic measures of target words in the Normal talking condition.

| Talker | Reader | $F_0$ (Hz) | Duration (ms) |
|---|---|---|---|
| Female 1 | 1.75 | 222 | 744 |
| Female 2 | 2.25 | 183 | 745 |
| Female 3 | 2.66 | 180 | 632 |
| Female 4 | 3.04 | 167 | 610 |
| Female 5 | 4.19 | 190 | 461 |
| Female 6 | 3.93 | 208 | 665 |
| Male 1 | 3.00 | 147 | 623 |
| Male 2 | 3.75 | 155 | 606 |

**Speech material and recording procedure**

Talkers recorded sentences from the Northwestern University Auditory Test No. 6 (NU6) [3], which consists of four lists of 50 monosyllabic keywords presented after a standard carrier phrase (e.g., *Say the word boat*). There were four recording conditions: speaking "normally, in your most comfortable voice" in quiet (Quiet) and in babble noise (Noise), and speaking "as if to someone with hearing loss" in quiet (HL) and in babble noise (HL+Noise).

Talkers were seated in an IAC sound-attenuating booth with a Sennheiser Linear E825S microphone placed 6 cm from their lips. Speech was recorded using the MS2, PA5 and RP2.1 components of the Tucker-Davis Technologies System III and the Avaaz Time-Frequency Representation program running on a Dell Precision 360 computer. In the noise conditions, multi-talker babble from the WIN test [4] was presented to talkers binaurally through Sennheiser 265 Linear headphones at 78 dB SPL. The same headphones were also worn by talkers in quiet.

Talkers recorded each NU6 list in four talking conditions over two sessions. During each session, talking conditions were always in a fixed order: Quiet, Noise, HL and HL+Noise. The recording order of the four NU6 lists was counterbalanced across talkers and words within each list were randomized. Sentences were re-recorded as needed at the end of the list if there were speech errors or hesitations. Sentences were then spliced out of the raw recordings and RMS-equated to 0.05 Pa using PRAAT [5].

### 2.2 Listener testing

**Listeners**

Listeners were 32 younger adults (mean age = 18.4 years, SD = 0.9) who were students enrolled in an introductory psychology course. Listeners had learned English in North America by the age of five years and they had average to excellent general health with no known neurological or speech disorders. They had normal hearing thresholds (≤ 20 dB HL from 250 to 8000 Hz), with no inter-aural differences greater than 15 dB, except for one participant who had an inter-aural difference of 20 dB at 4000 Hz.

**Experimental design and test procedure**

Each of the four NU6 lists was split in half to create eight half-lists of 25 words each. Every listener heard eight half-lists, each spoken by a different talker. Within those eight half-lists, every listener was exposed to two sets of all four talking conditions (though not every talker was paired with every talking condition for a given listener). The order of the four talking conditions and the order of the eight talkers were rotated through the set of 32 listeners.

Listeners were seated in an IAC sound-attenuating booth. Stimuli were presented to the right ear through Sennheiser 265 Linear headphones using a Dell Precision 360 desktop computer and the RP2.1, PA5, SM5 and HB7 components of the Tucker-Davis Technologies System III.

* huiwen.goy@utoronto.ca
† k.pichora.fuller@utoronto.ca
‡ p.vanlieshout@utoronto.ca

Babble noise from the WIN test [4] was presented continuously at 70 dB SPL, while speech was presented at 68 dB SPL with a 3-sec interval between sentences. Listeners reported the final word of every sentence and were asked to guess if they were unsure. The experimenter scored responses as they were made and also taped them to enable later confirmation of the scoring.

**Data analysis**

Percent correct scores were converted into RAU scores for statistical analyses [6]. A linear mixed-effects model was constructed for the effects of Talker and Talking Condition on word recognition accuracy scores, with listeners as the random effect. The model had an unstructured covariance matrix and the degrees of freedom were estimated using the between-within method. Various models were compared using ANOVA to determine whether individual fixed effects and their interaction explained additional variance in word recognition accuracy scores. Main effects were investigated using independent samples *t*-tests (Talker) and paired *t*-tests (Talking Condition) with Holm correction.

## 3 Results

Comparing a model that included both fixed effects with a model that included both fixed effects and an interaction term, the interaction term did not explain additional variance, $\chi^2(1) = 0.0011$, $p = 0.97$. Comparing a model that included both fixed effects with models that only included either fixed effect, the model with both fixed effects explained more variance than the model with either Talker or Talking Condition alone, $\chi^2(1) = 15.45$, $p < 0.001$ and $\chi^2(1) = 33.45$, $p < 0.001$, respectively. Thus, there was a significant main effect of Talker, $b = -0.62$, SE = 0.11, $t(7) = -5.98$, $p < 0.001$, and a significant main effect of Talking Condition, $b = 0.86$, SE = 0.22, $t(3) = 3.98$, $p = 0.01$, but no significant Talker × Talking Condition interaction.

Word recognition accuracy was best for talker F2 (Female 2), who was more intelligible than any other talker, $p$'s $< 0.05$. Female talkers with the lowest reader quality scores (F1, F2) were more intelligible than female talkers with the highest scores (F5, F6), $p$'s $< 0.005$. There was no significant difference between the male talkers (Figure 1).

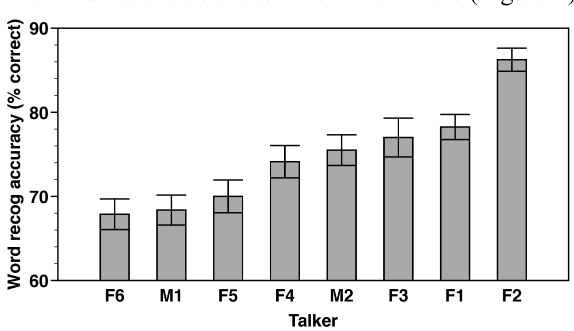


**Figure 1:** Mean word recognition accuracy scores of eight talkers, with standard error bars. F = Female; M = Male.

Word recognition performance was worse when talkers spoke in the Quiet condition than when they spoke in the Noise or HL+Noise conditions ($p$'s $< 0.05$). No other pairs of conditions were significantly different (Figure 2).

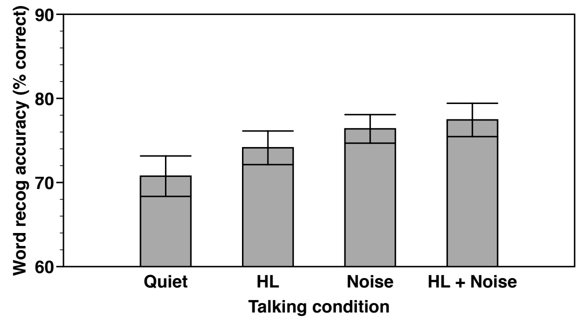


**Figure 2:** Mean word recognition accuracy scores for four talking conditions, with standard error bars.

## 4 Discussion

When speaking normally, older talkers showed a wide range of intelligibility scores (61 to 84%), similar to younger talkers in past research (65 to 85%) [7]. The "clear speech" of older talkers was not more intelligible than their normal speech, unlike younger talkers in most studies [8]. However, the benefit from speech produced in noise was similar to that of younger talkers under similar listening conditions [9]. Older talkers whose speech was perceived to be higher quality in quiet were actually less intelligible in noise. Future directions include testing older listeners to determine whether the characteristics of older talkers affect older listeners in a similar way as younger listeners.

## Acknowledgements

This study was funded by an NSERC grant (PI: KPF).

## References

[1] International Organization for Standardization. *Acoustics: Statistical distribution of hearing thresholds as a function of age (ISO 7029:2000)*. 2000. Geneva, Switzerland: Technical Committee ISO/TC 43 Acoustics.

[2] Goy H, Pichora-Fuller MK, Van Lieshout P. (2016). Effects of age on speech and voice quality ratings. *J Acoust Soc Am*. 2016;139(4):1648-1659. doi:10.1121/1.4945094

[3] Tillman TW, Carhart R. *An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6 (SAM-TR-66-55)*. 1966. Brooks Air Force Base, TX: USAF School of Aerospace Medicine.

[4] Wilson RH. Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *J Am Acad Audiol*. 2003;14(9):453-470.

[5] Boersma P, Weenink D. *Praat: doing phonetics by computer*. [Software] 2016. Available from http://www.praat.org

[6] Studebaker GA. A "rationalized" arcsine transform. *J Speech Hear Res*. 1985;28(3):455-462. doi:10.1044/jshr.2803.455

[7] Bond ZS, Moore TJ. A note on the acoustic-phonetic characteristics of inadvertently clear speech. *Speech Comm*. 1994;14:325-337. doi:10.1016/0167-6393(94)90026-4

[8] Smiljanić R, Bradlow A. Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Lang Linguist Compass*. 2009;3(1):236-264. doi:10.1111/j.1749-818x.2008.00112.x

[9] Pittman AL, Wiley TL. Recognition of speech produced in noise. *J Speech Lang Hear Res*. 2001;44(3):487-496. doi:10.1044/1092-4388(2001/038)