Speech recognition in reverberation and noise: the roles of the lexicon and age

Rebecca Carroll [1,3], *Anna Warzybok* [2,3]

[1] Institute of Dutch Studies, University of Oldenburg, Germany
[2] Medizinische Physik, University of Oldenburg, Germany
[3] Cluster of Excellence 'Hearing4all', University of Oldenburg, Germany

rebecca.carroll@uni-oldenburg.de, anna.warzybok@uni-oldenburg.de

Index Terms: speech recognition, lexical access time, age, vocabulary size, reverberation, noise

1. Introduction

Older adults tend to score worse in speech recognition (SR) tests compared to younger adults, even when their hearing is within the normal range [1,2]. Acoustically adverse conditions may exacerbate this difference [2, 3]. SR in acoustically adverse conditions is traditionally measured by SR tests that present speech signals masked by noise [2, 4, 5]. Reverberation has also been shown to constitute a (somewhat different) form of acoustic challenge [5, 6], although it is not quite clear how performance of older listeners with normal hearing (ONH) may be affected by this listening condition compared to younger listeners with normal hearing (YNH). Carroll et al. [7] found lexical access time and vocabulary size to predict speech-in-noise recognition scores. Differences were shown for age group and for signal-to-noise ratios (SNR), indicating that the role of lexical access may depend on the acoustic condition as well as population. Here, we asked,

1. Whether reverberated speech triggers similar SR patterns as speech in noise, and whether the combination of noise and reverberation exacerbates SR difficulties,
2. Whether ONH are affected differently from YNH by the different acoustically adverse listening conditions,
3. How individual differences may account for SR scores in the different acoustic conditions.

2. Methods

We tested SR in five acoustic conditions and two populations, and correlated them with individual differences measures.

2.1. Participants

A group of 22 younger (25.3 ± 4.1 yrs., 15 women) and a group of 22 older (67.7 ± 4.8 yrs., 15 women) native listeners of German, matched in education, participated in this study. Normal hearing was defined as having pure tone averages (PTA-4, 500, 1k, 2k, 4k Hz) of 20 dB HL or lower. The PTA-4 was 3.2 ± 3.0 dB HL for YNH, and 8.4 ±4.5 dB HL for ONH.

2.2. Speech material & acoustic conditions

Participants completed six lists of the Göttingen Sentence Test [4], a SR test with everyday sentences, twenty sentences per test list. SR scores were measured in % correct in several acoustic conditions. Sentences were presented (i) in a stationary speech-shaped noise at -4 and -6 dB SNR, (ii) in reverberation with reverberation times (τ) of 2.03 and 3.24 sec, and (iii) in a combination of noise (+7 dB SNR) and reverberation (τ = 3.25 sec). We used STI [8] predictions to determine SNR and τ values yielding about 50% or 80% correct SR per list.

2.3. Individual differences measures

We tested a small battery of visually presented individual differences measures, with a focus on the mental lexicon. The WST [9] and a German version of the revised Peabody Picture Vocabulary Test [10] were combined into a new construct variable VOCABULARY. Lexical access time (LA) was measured by a lexical decision test [6], and verbal working memory (WM) was measured using the German reading span test [11].

3. Results

YNH's SR scores followed those predicted by the STI [8] in reverberation (50%, 82%) and in noise (51%, 85%). ONHs' SR scores were significantly worse compared to YNHs', most obviously in reverberation with τ=3.24 sec (37% instead of 50%). SR for both listener groups was worst in the condition combining noise and reverberation (YNH=38%, ONH=29% instead of 50% correct). Both groups had comparable VOCABULARY and WM scores, but differed in LA time (ONH > YNH, F=14.08, p < .001, corrected R^2=0.71).

Hierarchical regression models per listener group and acoustic condition showed a complex picture, in which VOCABULARY and/or LA significantly predicted speech recognition in some of the noise and reverberation conditions, but never in the combination of noise and reverberation. There was a trend of PTA-4 to correlate with ONHs' speech recognition in reverberated speech, but this tendency did not hold when accounting for other individual measures. WM never correlated with SR scores.

4. Conclusions

We conclude that the way and degree with which SR relates to vocabulary knowledge and LA time depends on the acoustic condition and age. SR performance in a combination of noise and reverberation can be neither predicted by STI [8] with standard audiometric input nor by psycholinguistic measures.

5. Acknowledgements

This work was supported by the Cluster of Excellence EXC 1077/1 'Hearing4all' funded by the German Research Council.

6. References

[1] C. Füllgrabe, B.C. Moore, and M.A. Stone, "Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing", *Frontiers in Aging Neuroscience,* vol. 6, 347, 2015.

[2] T. Schoof, and S. Rosen, "The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners", *Frontiers in Aging Neuroscience,* vol. 6, article 307, 2014.

[3] L.E. Humes, T.A. Busey, J. Craig, and D. Kewley-Port, "Are age-related changes in cognitive function driven by age-related changes in sensory processing?", *Attention, Perception, and Psychophysics,* vol. 75, pp. 508-524, 2013.

[4] B. Kollmeier, and T. Wesselkamp, "Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment", *Journal of the Acoustical Society of America,* vol, 102, no.4, pp. 2412-2421, 1997.

[5] L.L. Wong, E.H.N. Ng, and S.D. Soli, "Characterization of speech understanding in various types of noise", *Journal of the Acoustical Society of America,* vol. 132, pp. 2642-2651, 2014.

[6] A. Warzybok, J. Rennies, T. Brand et al., "Effects of spatial and temporal integration of a single early reflection on speech intelligibility", *Journal of the Acoustical Society of America*, vol. 133, no. 1, pp. 269-282, 2013.

[7] R. Carroll, A. Warzybok, B. Kollmeier, and E. Ruigendijk, *"How and to what degree is speech intelligibility in noise modulated by vocabulary knowledge, lexical access time, working memory, and age?",* submitted manuscript, 2015.

[8] IEC 60268-16. Sound system equipment, part 16: Objective rating of speech intelligibility by Speech Transmission Index. International Electrotechnical Commission, Geneva, Switzerland, 2003.

[9] K.-H. Schmidt, and P. Metzler, *Wortschatztest*. Weinheim: Beltz, 1992.

[10] S. Buhlheller, and H.O. Häcker, "Deutschsprachige Fassung des PPVT-III für Jugendliche und Erwachsene". [German version of the PPVT-III for teenagers and adults], Frankfurt: Swets.

[11] R. Carroll, M. Meis, M. Schulte, et al., "Development of a German reading span test with dual task design for application in cognitive hearing research", *International Journal of Audiology,* vol. 54, no., 2, pp. 136-141, 2015.