

Investigating the Role of Working Memory in Speech-in-noise Identification for Listeners with Normal Hearing

Christian Füllgrabe and Stuart Rosen

Abstract With the advent of cognitive hearing science, increased attention has been given to individual differences in cognitive functioning and their explanatory power in accounting for inter-listener variability in understanding speech in noise (SiN). The psychological construct that has received most interest is working memory (WM), representing the ability to simultaneously store and process information. Common lore and theoretical models assume that WM-based processes subtend speech processing in adverse perceptual conditions, such as those associated with hearing loss or background noise. Empirical evidence confirms the association between WM capacity (WMC) and SiN identification in older hearing-impaired listeners. To assess whether WMC also plays a role when listeners without hearing loss process speech in acoustically adverse conditions, we surveyed published and unpublished studies in which the Reading-Span test (a widely used measure of WMC) was administered in conjunction with a measure of SiN identification. The survey revealed little or no evidence for an association between WMC and SiN performance. We also analysed new data from 132 normal-hearing participants sampled from across the adult lifespan (18–91 years), for a relationship between Reading-Span scores and identification of matrix sentences in noise. Performance on both tasks declined with age, and correlated weakly even after controlling for the effects of age and audibility ($r=0.39$, $p\leq 0.001$, one-tailed). However, separate analyses for different age groups revealed that the correlation was only significant for middle-aged and older groups but not for the young (<40 years) participants.

Keywords Aging · Audiometrically normal · Correlations · Cognition · Noise · Older listeners · Reading-span test · Speech intelligibility · Working-memory capacity · Young listeners · Matrix sentences

C. Füllgrabe (✉)

MRC Institute of Hearing Research, Science Road, Nottingham NG7 2RD, UK
e-mail: christian.fullgrabe@ihr.mrc.ac.uk

S. Rosen

UCL Speech, Hearing & Phonetic Sciences, 2 Wakefield Street, London WC1N 2PF, UK
e-mail: stuart@phon.ucl.ac.uk

© The Author(s) 2016

P. van Dijk et al. (eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, Advances in Experimental Medicine and Biology 894, DOI 10.1007/978-3-319-25474-6_4

1 Introduction

Recent years have seen an increased interest in the role of individual differences in cognitive functioning in speech and language processing and their interaction with different types of listening tasks and conditions. The psychological construct that has received the most attention in the emerging field of cognitive hearing science is working memory (WM), possibly because it has been shown to be involved in a wide range of complex cognitive behaviours (e.g. reading comprehension, reasoning, complex learning). WM can be conceptualised as the cognitive system that is responsible for active maintenance of information in the face of ongoing processing and/or distraction. Its capacity (WMC) is generally assessed by so-called complex span tasks, requiring the temporary storage and simultaneous processing of information. For example, in one of the most widely used WM tasks, the Reading-Span test (Baddeley et al. 1985), visually presented sentences have to be read and their semantic correctness judged (processing component), while trying to remember parts of their content for recall after a variable number of sentences (storage component).

A growing body of evidence from studies using mainly older hearing-impaired (HI) listeners indeed confirms that higher WMC is related to better unaided and aided speech-in-noise (SiN) identification, with correlation coefficients frequently exceeding 0.50 (Lunner 2003; Foo et al. 2007; Lunner and Sundewall-Thorén 2007; Arehart et al. 2013). In addition, high-WMC listeners were less affected by signal distortion introduced by hearing-aid processing (e.g. frequency or dynamic-range compression).

Consistent with these results, models of speech/language processing have started incorporating active cognitive processes (Rönnberg et al. 2013; Heald and Nusbaum 2014). For example, according to the Ease of Language Understanding model (Rönnberg et al. 2013), any mismatch between the perceptual speech input and the phonological representations stored in long-term memory disrupts automatic lexical retrieval, resulting in the use of explicit, effortful processing mechanisms based on WM. Both internal distortions (i.e., related to the integrity of the auditory, linguistic, and cognitive systems) and external distortions (e.g. background noise) are purportedly susceptible to contribute to the mismatch. Consequently, it is assumed that WMC also plays a role when individuals with normal hearing (NH) have to process spoken language in acoustically adverse conditions.

However, Füllgrabe et al. (2015) recently failed to observe a link between Reading-Span scores and SiN identification in older listeners (≥ 60 years) with audiometrically NH (≤ 20 dB HL between 0.125 and 6 kHz), using a range of target speech (consonants and sentences), maskers (unmodulated and modulated noise, interfering babble), and signal-to-noise ratios (SNRs).

2 Study Survey

To assess the claim that individual variability in WMC accounts for differences in SiN identification even in the absence of peripheral hearing loss, we surveyed published and unpublished studies administering the Reading-Span test and a measure

of SiN identification to participants with audiometrically NH. To ensure consistency with experimental conditions in investigations of HI participants, only studies presenting sentence material “traditionally” used in hearing research (i.e., ASL, Hagerman, HINT, IEEE, QuickSIN, or Versfeld sentences) against co-located background maskers were considered. In addition, we only examined studies in which the effect of age was controlled for (either by statistically partialling it out or by restricting the analysis to a “narrow” age range), in order to avoid inflated estimates of the correlation between WMC and SiN tasks caused by the tendency for performance in both kinds of tasks to worsen with age. Figure 1 summarizes the results of this survey.

Correlation coefficients in the surveyed studies are broadly distributed, spanning almost half of the possible range of r values (i.e., from -0.29 to 0.58). Confidence intervals (CIs) are generally large and include the null hypothesis in 21/25 and 24/25 cases for CIs of 95 and 99%, respectively, suggesting that these studies are not appropriately powered. For the relatively small number of studies included in this survey, there is no consistent trend for stronger correlations in more complex and/or informationally masking backgrounds or at lower SNRs, presumably corresponding to more adverse listening conditions.

Across studies restricting their sample to young (18–40 years) participants, the weighted average r value is 0.12, less than 2% of the variance in SiN identification. According to a power calculation, it would require 543 participants to have an 80% chance of detecting such a small effect with $p=0.05$ (one-tailed)!

3 Analysis of Cohort Data for Audiometrically Normal-Hearing Participants

Given the mixed results from previous studies based on relatively small sample sizes, we re-analysed data from a subset of a large cohort of NH listeners taking part in another study.

3.1 Method

Participants were 132 native-English-speaking adults, sampled continuously from across the adult lifespan (range = 18–91 years). Older (≥ 60 years) participants were screened using the Mini Mental State Examination to confirm the absence of cognitive impairment. All participants had individual audiometric hearing thresholds of ≤ 20 dB HL at octave frequencies between 0.125 and 4 kHz, as well as at 3 kHz, in the test ear. Despite clinically “normal” audibility, the pure-tone average (PTA) for the tested frequency range declined as a function of age ($r=0.65$, $p\leq 0.001$, one-tailed). Since changes in sensitivity even in the normal audiometric range can affect SiN identification (Dubno and Ahlstrom 1997), PTA is treated as a possible confounding variable in analyses involving the entire age group.

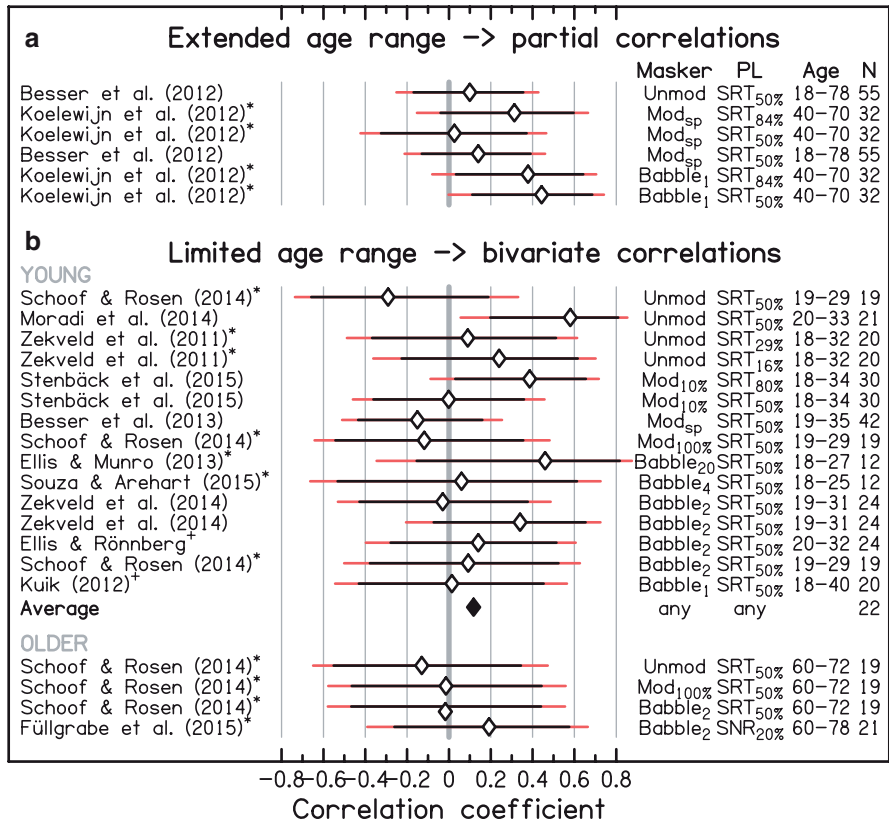


Fig. 1 Comparison of Pearson correlation coefficients (diamonds) and associated 95% (black) and 99% (red) confidence intervals for studies investigating the association between WMC and speech-in-“noise” identification in NH participants after controlling for the effect of age by **a** computing partial correlations, or **b** using a limited age range. When necessary, the sign of the correlation was changed so that a positive correlation represents good performance on the two tasks. A weighted average for correlations based only on young NH listeners is provided (multiple r values for the same study sample are entered as their average). Source references (* indicates re-analysed published data; + indicates unpublished data, personal communication) and experimental (type of masker (*Masker*); performance level (*PL*) and participant (age range (*Age*); number of participants (*N*)) details are given in the figure. Masker: *Unmod* unmodulated noise, *Mod_{x%}* or *sp* noise modulated by an X% sinusoidal amplitude modulation or a speech envelope, *Babble_x* X-talker babble. PL: *SRT_{x%}* adaptive procedure tracking the speech reception threshold corresponding to X%-correct identification, *SNR_{x%}* fixed SNR levels yielding, on average, X%-correct identification

WMC was assessed by means of the computerized version of the Reading-Span test (Rönnberg et al. 1989). Individual sentences were presented in three parts on a computer screen to be read aloud and judged as plausible or implausible. After three to six sentences, either the first or last word of each of the sentences had to be recalled. WMC corresponded to the number of correctly recalled words in any order.

SiN identification was assessed using the English version of the Matrix sentence test (Vlaming et al. 2011). Each target sentence, presented monaurally at 70 dB SPL, followed a fixed syntactic structure (proper noun—verb—numeral—adjective—noun) but had low semantic redundancy. The noise maskers had the same long-term spectrum as the target sentences and were either unmodulated or 100% sinusoidally amplitude modulated at 8 or 80 Hz. Target and masker were mixed together at SNRs ranging from -3 to -15 dB, and the mixture was lowpass-filtered at 4 kHz.

3.2 Results and Discussion

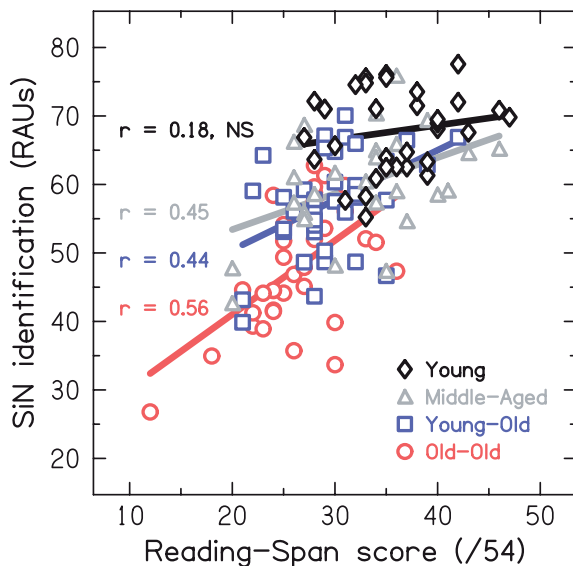
Identification scores were transformed into rationalized arcsine units (RAUs) and averaged across masker types and SNRs to reduce the effect of errors of measurement and to yield a composite intelligibility score representative of a range of test conditions.

Confirming previous results for audiometrically NH listeners (Füllgrabe et al. 2015), Reading-Span and SiN identification scores showed a significant decline with age, with Pearson's $r = -0.59$ and -0.68 (both $p \leq 0.001$, one-tailed), respectively. The scatterplot in Fig. 2 shows that, considering all ages, performances on the tasks were significantly related to each other ($r = 0.64$, $p \leq 0.001$, one-tailed). This association remained significant after partialling out the effects of age and PTA ($r = 0.39$, $p \leq 0.001$, one-tailed), contrasting with the results of Besser et al. (2012), using a cohort including only a few ($N = 8$) older (≥ 60 years) participants, but being roughly consistent with those reported by Koelewijn et al. (2012) for a cohort comprised of middle-aged and older (≥ 40 years) participants (see Fig. 1a).

To further investigate the age dependency of the association between WMC and SiN identification, participants were divided into four age groups: “Young” (range = 18–39 years, mean = 28 years; $N = 32$), “Middle-Aged” (range = 40–59 years, mean = 49 years; $N = 26$), “Young-Old” (range = 60–69 years, mean = 65 years; $N = 40$), and “Old-Old” (range = 70–91 years, mean = 77 years; $N = 34$). Separate correlational analyses for each age group revealed that the strength of the association differed across groups (see Fig. 2). Consistent with the overall trend seen in Fig. 1, the correlation was weak and non-significant in the group of young participants ($r = 0.18$, $p = 0.162$, one-tailed). In contrast, the correlations were moderately strong and significant in the three older groups (all $r \geq 0.44$, all $p \leq 0.011$, one-tailed). Comparing the different correlation coefficients, after applying Fisher's r -to- z transformation, revealed a significant difference between the Young and Old-Old group ($z = -1.75$, $p = 0.040$, one-tailed). There was no evidence for a difference in variance between these groups (Levene's test, $F_{(1,64)} < 1$, $p = 0.365$).

The age-related modulation of the strength of the correlation between WMC and SiN perception could be due to the different performance levels at which the age groups operated in this study (mean identification was 68, 60, 57, and 48 RAUs for the Young, Middle-Aged, Young-Old, and Old-Old group, respectively). However,

Fig. 2 Scatterplot relating SiN identification averaged across background noises and SNRs to Reading-Span scores for the four age groups. The best linear fit to the data (*thick lines*) and associated bivariate Pearson correlation coefficients for each age group are given in the figure



when performance only for the two lowest SNRs (corresponding to 46 RAUs) was considered, WMC was still not associated with SiN identification in the young participants ($r=0.04$, $p=0.405$, one-tailed).

4 Conclusions

Taken together, the reported results fail to provide evidence that, in *acoustically* adverse listening situations, WMC (as measured by the Reading-Span test) is a reliable and strong predictor of SiN *intelligibility* in young listeners with normal hearing. The new data presented here suggest that WMC becomes more important with age, especially in the oldest participants. One possible explanation for this increasing cognitive involvement with age could be the accumulation of age-related deficits in liminary but also supraliminary auditory processing (e.g. sensitivity to temporal-fine-structure and temporal-envelope cues; Füllgrabe 2013; Füllgrabe et al. 2015), resulting in under-defined and degraded internal representations of the speech signal, calling for WM-based compensatory mechanisms to aid identification and comprehension.

Our findings do not detract from the practical importance of cognitive assessments in the prediction of SiN identification performance in older HI listeners and the possible interaction between cognitive abilities and hearing-aid processing. Nor do they argue against the involvement of cognition in speech and language processing in young NH listeners *per se*. First, individual differences in WMC have been shown to explain some of the variability in performance in more linguistically complex task (such as in the comprehension of dynamic conversations; Keidser et al.

2015), presumably requiring memory or attentional/inhibitory processes associated with WMC (Conway et al. 2001; Kjellberg et al. 2008). Second, different cognitive measures, probing the hypothesized sub-processes of WM (e.g. inhibition, shifting, updating) or other domain-general cognitive primitives (e.g. processing speed) might prove to be better predictors of SiN processing abilities than the Reading-Span test.

In conclusion, and consistent with recent efforts to establish if and under which conditions cognitive abilities influence the processing of spoken language (e.g. Fedorenko 2014; Heinrich and Knight, this volume), the current results caution against the assumption that WM necessarily supports SiN identification independently of the age and hearing status of the listener.

Acknowledgments We would like to thank our colleagues who shared and reanalysed their data, and Dr. Oliver Zobay for his statistical advice. The MRC Institute of Hearing Research is supported by the Medical Research Council (grant number U135097130). This work was also supported by the Oticon Foundation (Denmark). CF is indebted to Prof. Brian Moore for granting access to the test equipment of his laboratory.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution-Noncommercial 2.5 License (<http://creativecommons.org/licenses/by-nc/2.5/>) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

The images or other third party material in this chapter are included in the work's Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

References

- Arehart KH, Souza P, Baca R, Kates JM (2013) Working memory, age, and hearing loss: susceptibility to hearing aid distortion. *Ear Hear* 34(3):251–260
- Baddeley A, Logie R, Nimmo-Smith I, Brereton N (1985) Components of fluent reading. *J Mem Lang* 24(1):119–131
- Besser J, Zekveld AA, Kramer SE, Rönnerberg J, Festen JM (2012) New measures of masked text recognition in relation to speech-in-noise perception and their associations with age and cognitive abilities. *J Speech Lang Hear Res* 55(1):194–209
- Besser J, Koelewijn T, Zekveld AA, Kramer SE, Festen JM (2013) How linguistic closure and verbal working memory relate to speech recognition in noise—a review. *Trends Amplif* 17(2):75–93
- Conway ARA, Cowan N, Bunting MF (2001) The cocktail party phenomenon revisited: the importance of working memory capacity. *Psychon Bull Rev* 8(2):331–335
- Dubno JR, Ahlstrom JB (1997) Additivity of multiple maskers of speech. In: Jesteadt W (ed) *Modeling sensorineural hearing loss*. Lawrence Erlbaum Associates, Hillsdale, pp 253–272
- Ellis RJ, Munro KJ (2013) Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition? *Int J Audiol* 52(1):14–22
- Fedorenko E (2014) The role of domain-general cognitive control in language comprehension. *Front Psychol* 5:335
- Foo C, Rudner M, Rönnerberg J, Lunner T (2007) Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *J Am Acad Audiol* 18(7):618–631

- Füllgrabe C (2013) Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *Am J Audiol* 22(2):313–315
- Füllgrabe C, Moore BC, Stone MA (2015) Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front Aging Neurosci* 6:347
- Heald SL, Nusbaum HC (2014) Speech perception as an active cognitive process. *Front Syst Neurosci* 8:35
- Keidser G, Best V, Freeston K, Boyce A (2015) Cognitive spare capacity: evaluation data and its association with comprehension of dynamic conversations. *Front Psychol* 6:597
- Kjellberg A, Ljung R, Hallman D (2008) Recall of words heard in noise. *Appl Cognit Psychol* 22(8):1088–1098
- Koelewijn T, Zekveld AA, Festen JM, Rönnerberg J, Kramer SE (2012) Processing load induced by informational masking is related to linguistic abilities. *Int J Otolaryngol* 2012:865731. 65731.
- Kuik AM (2012) Speech reception in noise: on auditory and cognitive aspects, gender differences and normative data for the normal-hearing population under the age of 40. Bachelor's thesis. Vrije Universiteit Amsterdam, Amsterdam
- Lunner T (2003) Cognitive function in relation to hearing aid use. *Int J Audiol* 42(Suppl 1):49–58
- Lunner T, Sundevall-Thorén E (2007) Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J Am Acad Audiol* 18(7):604–617
- Moradi S, Lidestam B, Saremi A, Rönnerberg J (2014) Gated auditory speech perception: effects of listening conditions and cognitive capacity. *Front Psychol* 5:531
- Rönnerberg J, Arlinger S, Lyxell B, Kinnefors C (1989) Visual evoked potentials: relation to adult speechreading and cognitive function. *J Speech Hear Res* 32(4):725–735
- Rönnerberg J, Lunner T, Zekveld A, Sörqvist P, Danielsson H, Lyxell B, Dahlstrom O, Signoret C, Stenfelt S, Pichora-Fuller MK, Rudner M (2013) The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front Syst Neurosci* 7:31
- Souza P, Arehart K (2015) Robust relationship between reading span and speech recognition in noise. *Int J Audiol* 54: 705–713
- Stenbäck V, Hällgren M, Lyxell B, Larsby B (2015) The Swedish Hayling task, and its relation to working memory, verbal ability, and speech-recognition-in-noise. *Scand J Psychol* 56(3):264–272
- Vlaming MSMG, Kollmeier B, Dreschler WA, Martin R, Wouters J, Grover B, Mohammadh Y, Houtgast T (2011) HearCom: hearing in the communication society. *Acta Acust United Acust* 97(2):175–192
- Zekveld AA, Rudner M, Johnsrude IS, Festen JM, van Beek JH, Rönnerberg J (2011) The influence of semantically related and unrelated text cues on the intelligibility of sentences in noise. *Ear Hear* 32(6):e16–e25
- Zekveld AA, Rudner M, Kramer SE, Lyzenga J, Rönnerberg J (2014) Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. *Front Neurosci* 8:88