



Füllgrabe, Christian and Rosen, Stuart (2016) On the (un)importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers in Psychology*, 7 . 1268/1-1268/8. ISSN 1664-1078

Access from the University of Nottingham repository:

http://eprints.nottingham.ac.uk/38656/1/Fullgrabe_Rosen_FP16.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the Creative Commons Attribution licence and may be reused according to the conditions of the licence. For more details see: <http://creativecommons.org/licenses/by/2.5/>

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk



On The (Un)importance of Working Memory in Speech-in-Noise Processing for Listeners with Normal Hearing Thresholds

Christian Füllgrabe^{1*} and Stuart Rosen²

¹ Medical Research Council Institute of Hearing Research, The University of Nottingham, Nottingham, UK, ² Speech, Hearing and Phonetic Sciences, University College London, London, UK

OPEN ACCESS

Edited by:

Jerker Rönnerberg,
Linköping University, Sweden

Reviewed by:

Mary Rudner,
Linköping University, Sweden
Thomas Lunner,
The Swedish Institute for Disability
Research, Sweden
Kathryn Arehart,
University of Colorado Boulder, USA

*Correspondence:

Christian Füllgrabe
christian.fullgrabe@nottingham.ac.uk

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Psychology

Received: 29 March 2016

Accepted: 09 August 2016

Published: 30 August 2016

Citation:

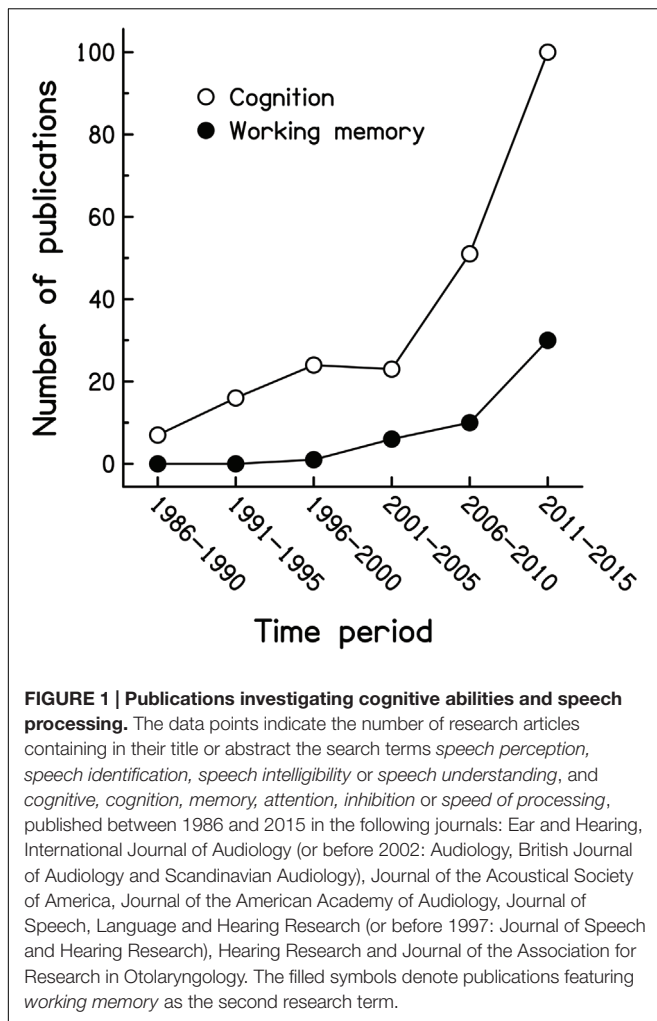
Füllgrabe C and Rosen S (2016)
On The (Un)importance of Working
Memory in Speech-in-Noise
Processing for Listeners with Normal
Hearing Thresholds.
Front. Psychol. 7:1268.
doi: 10.3389/fpsyg.2016.01268

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 the processing of speech in noise (SiN). The psychological construct that has received much interest in recent years is working memory. Empirical evidence indeed confirms the association between WM capacity (WMC) and SiN identification in older hearing-impaired listeners. However, some theoretical models propose that variations in WMC are an important predictor for variations in speech processing abilities in adverse perceptual conditions for *all* listeners, and this notion has become widely accepted within the field. To assess whether WMC also plays a role when listeners without hearing loss process speech in adverse listening 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, using sentence material routinely used in audiological and hearing research. A meta-analysis revealed that, for young listeners with audiometrically normal hearing, individual variations in WMC are estimated to account for, on average, less than 2% of the variance in SiN identification scores. This result cautions against the (intuitively appealing) assumption that individual variations in WMC are predictive of SiN identification independently of the age and hearing status of the listener.

Keywords: working memory, speech perception in noise, aging, normal hearing, hearing loss, supra-threshold auditory processing, sentence identification, reading-span test

INTRODUCTION

Over the past decades, there has been growing interest in the role of individual differences in cognitive functioning in speech processing, reflected by a noticeable increase in the number of scientific publications on this topic (see **Figure 1**). Such work reflects the emergence of the new interdisciplinary research field of Cognitive Hearing Science (e.g., Arlinger et al., 2009), focussing on understanding the interplay of auditory and cognitive processes in speech perception, primarily in adverse circumstances. Not only are key scientific issues at stake, there are also important clinical implications in trying to provide effective rehabilitation to people suffering from problems with spoken communication.



Working Memory and Its Role in Complex Cognition

Amongst the different cognitive abilities investigated, working memory (WM) has received considerable attention in recent years (see filled symbols in **Figure 1**). WM is considered by many psychologists as a “cognitive primitive,” due to its moderate-to-very-strong associations with different aspects of hot (i.e., emotion-laden; Klein and Boals, 2001) and cold cognition, such as reasoning (Barrouillet, 1996), attentional control (Das-Smaal et al., 1993), comprehension (Daneman and Merikle, 1996), and fact recall and pronoun referencing (Daneman and Carpenter, 1980). Over the years, different definitions have been given for this theoretical construct but it is generally agreed that the capacity of the WM system (WMC) can be reliably assessed by so-called complex span tasks. These require participants to perform a complex activity while concurrently trying to retain new information. For example, in one of the most widely used WM tasks, the Reading-Span (RSpan) test (Baddeley et al., 1985), visually presented sentences have to be read and their plausibility judged, while trying to remember parts of their content for recall after a variable number of sentences.

The Role of Working Memory in Speech Perception

Given the strong and systematic link between WM and higher-order complex behavior, it is hardly surprising that performance on complex span tasks has also been used to explain individual variability in understanding speech in noise (SiN).

For example, a series of audiological research studies investigated whether individual differences in WMC, measured by the version of the RSpan test developed by Rönnberg et al. (1989), can help predict unaided (Lunner, 2003; Rudner et al., 2011) and aided (Lunner, 2003; Foo et al., 2007; Rudner et al., 2008, 2009, 2011) speech perception in hearing-impaired (HI) listeners, and explain the user-dependent success of different types of signal-processing performed by the hearing aid (e.g., dynamic range or frequency compression; Souza et al., 2015). Mainly moderate, sometimes even strong correlations between SiN identification and RSpan scores were consistently reported. Surprisingly, when referring to these findings to corroborate the role of WM in SiN perception, it is generally not mentioned that the cited studies were conducted with HI listeners who, on average, were aged over 65 years.

Furthermore, on the basis of an extensive review of behavioral studies concerned with the effects of cognitive factors on SiN perception in HI and normal-hearing (NH) listeners, Akeroyd (2008) concluded, too, that cognitive functioning is associated with SiN identification, and that WMC, especially when measured by the RSpan test, is the best cognitive predictor. However, these conclusions were based solely on the results from HI listeners (namely the relevant citations in the paragraph above), a fact generally not acknowledged when citing this reference.

A similar assumption that the same crucial cognitive processes are at work in all listeners, independently of their age and hearing status, is made in recent models of speech/language processing (e.g., Rönnberg, 2003; Heald and Nusbaum, 2014). For example, according to the latest instantiation of the Ease of Language Understanding (ELU) 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. The greater the mismatch, the more effortful listening becomes. Both internal distortions (i.e., related to the integrity of the auditory, linguistic and cognitive systems) and external distortions (e.g., background noise) are supposed to contribute to the mismatch. Consequently, it is assumed within this framework that WMC also plays a role when NH listeners have to process spoken language in acoustically adverse conditions. While no experimental evidence supporting this claim has actually been provided, this notion has become widely accepted within the field.

STUDY SURVEY

To assess the claim that individual variability in WMC accounts for differences in SiN identification even in the absence of hearing

loss, we surveyed studies administering the RSpan test¹ and a measure of SiN identification to participants with audiometrically normal hearing sensitivity.

To ensure consistency with experimental conditions in investigations of HI listeners, only studies presenting sentence material routinely used in audiological and hearing research against spatially co-located background maskers were considered. In addition, we only examined studies in which the effect of age was controlled for, 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. The effect of age was controlled for either by restricting the analysis to a narrow age range, or by statistically partialling out the effect of age when using data from participants across a wider age range. Based on a request posted on the Auditory List² and a general literature search, we were able to compile data from 19 published and unpublished studies that complied with our inclusion criteria³. Since several studies measured SiN identification against different types of background maskers or for different performance levels, a total of 41 data sets was entered into the meta-analysis (see **Figure 2**). For each data set, the Pearson correlation coefficient (r ; diamonds) and associated 95 and 99% confidence intervals (CIs; black and red horizontal lines, respectively) are indicated, as well as the performance level at which the participants were tested, the type of masker, the sentence material⁴, the age range of the sample and the sample size. Within each of the three sections of **Figure 2**, data sets are organized by decreasing performance level (i.e., increasing difficulty). For identical performance levels, data sets are ordered by masker type, representing presumed increasing masker complexity, from “simple” notionally steady

noise⁵ through sinusoidally or speech-envelope-modulated noise to speech babble. Interestingly, some of the studies for which the data were reanalyzed on our request (indicated by an asterisk against them) did not even report the correlation between WMC and SiN identification in NH listeners in their original publication.

Across all data sets, the observed r values varied widely from -0.29 to 0.64 , with almost a quarter of the values being negative, indicating that sometimes low-WMC individuals showed better SiN identification than individuals with high WMC. CIs were rather large, suggesting that studies were underpowered (albeit not necessarily designed to assess this specific relationship), and, in most cases, the intervals included the value zero.

Seemingly in contradiction with the ELU-model prediction of higher WM involvement for speech identification in increasingly adverse listening conditions, there was no obvious trend for more consistent or stronger correlations in more difficult listening conditions (i.e., at lower performance levels). In fact, there is some (descriptive) evidence of stronger associations between WMC and SiN identification in easier listening conditions [see results in section I for the same listeners in high- and low-performance-level conditions in Koelewijn et al. (2012) and Carroll et al. (2016)]. However, this trend was based on results for two performance levels only, and it was not observed consistently across studies (Zekveld et al., 2011; Stenbäck et al., 2016) or even within the same study (Koelewijn et al., 2012).

Moreover, comparisons across different data sets obtained for similar performance levels did not show that inter-individual variability in WMC were more consistently or strongly associated with SiN identification for more complex maskers or target speech, as has previously been speculated (e.g., Rönnberg et al., 2010; Smith and Pichora-Fuller, 2015). For example, for young NH listeners, operating at a performance level of 50%-correct, the correlation for simple relatively predictable HINT sentences presented in a steady noise was 0.58 (Moradi et al., 2014) but only: (i) 0.14 in spectro-temporally and linguistically more complex babble noise (Ellis and Rönnberg, personal communication), and (ii) -0.01 for the linguistically more complex and unpredictable IIEEE sentences also presented in steady noise (Banks et al., 2015).

At the same time, the strength of the correlation varied even for studies using very similar test conditions and participant groups. For example, at a performance level of 50%-correct for IIEEE sentences presented in a steady noise masker, the correlation for young NH listeners was either -0.29 (Schoof and Rosen, 2014) or -0.01 (Banks et al., 2015). This illustrates the dependence of the results on the particular sample used (and its size) and cautions against basing conclusions as to the role of individual differences in WMC in SiN identification on observations from single small-scale studies.

As there was a sufficiently large number of data sets from studies restricting their sample to young listeners (aged 18–40 years), a random-effects meta-analysis model was used to estimate the average correlation among these studies. This kind

¹Most studies used the RSpan test originally developed by Rönnberg et al. (1989) but some administered a shorter version of the test. However, there seems to be no differences in mean performance between the two test versions (Classon, 2013).

²<http://www.auditory.org/>

³Data from a further two studies were not included in the meta-analysis due to the failure to obtain re-analysed data and the authors' explicit wish for us not to use their data.

⁴Description of the different sentence lists used in the studies entered into the meta-analysis:

ASL – Adaptive Sentence List (MacLeod and Summerfield, 1990): Predictable simple four- to six-word sentences (e.g., “The boiled egg was soft.”).

HINT – Hearing In Noise Test (Nilsson et al., 1994; Hällgren et al., 2006): Predictable simple three- to seven-word everyday sentences (e.g., “Strawberry jam is sweet.”).

GÖSA – Göttinger sentence test (Kollmeier and Wesselkamp, 1997): High-predictability three- to seven-word (mean = 5) everyday sentences (e.g., “The dispute has ended.”).

VU98 – (Versfeld et al., 2000): Eight- or nine-syllable everyday sentences (e.g., “The shop is within walking distance.”).

IEEE – Institute of Electrical and Electronics Engineers Harvard sentences (Rothauser et al., 1969; Killion et al., 2004): Low-predictability five-keyword sentences (e.g., “A white silk jacket goes with any shoes.”).

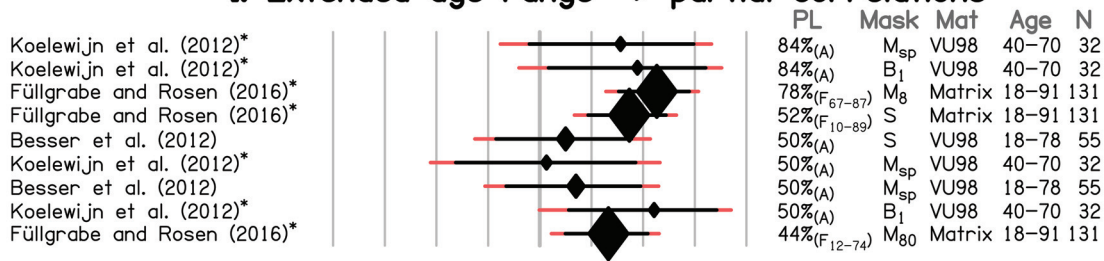
OLACS – Oldenburg Linguistically and Audiologically Controlled Sentences (Uslar et al., 2013): Seven-word sentences of varying linguistic complexity (e.g., “The little boy greets the nice father.” “The farmer, whom the teachers catch, smiles.”).

Matrix – Matrix sentences (Hagerman, 1982; Vlaming et al., 2011): Low-redundancy five-word sentences with the same syntactic structure (name-verb-number-adjective-object; e.g., “Nina wants some big beds.”).

In comparison, the cited investigations involving older HI listeners used HINT and Matrix sentences.

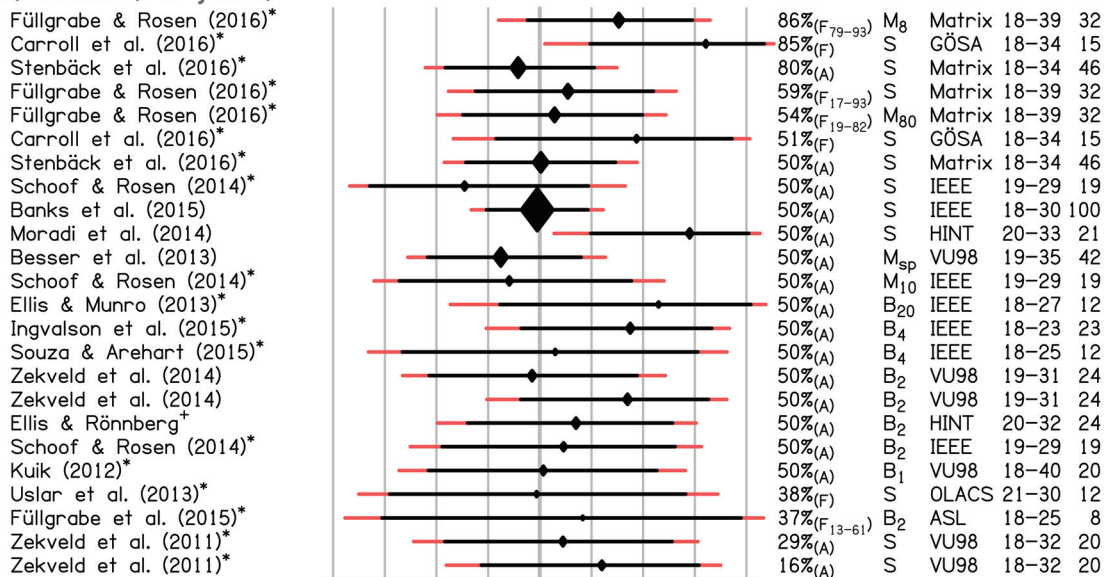
⁵Background noise on which no amplitude modulation is impressed is often referred to as a “steady” or “stationary” masker. However, even such notionally steady maskers contain intrinsic random amplitude fluctuations that impede speech perception (Stone et al., 2011, 2012).

I. Extended age range → partial correlations



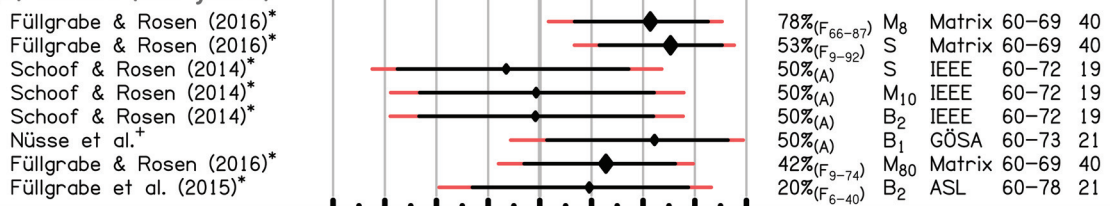
II. Limited age range → bivariate correlations

(A) YOUNG (≤40 years)



AVERAGE

(B) OLDER (≥60 years)



-0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8
Correlation coefficient

FIGURE 2 | A forest plot for a meta-analysis of studies investigating the association between WMC and speech-in-“noise” identification in NH listeners after controlling for the effect of age by (I) computing partial correlations or (II) using a limited age range [younger listeners aged ≤40 years (A) vs. older listeners aged ≥60 years (B)]. Shown in the plot are Pearson correlation coefficients (diamonds with their relative sizes indicating the study’s sample size) and associated 95% (black) and 99% (red) confidence intervals. Several studies contributed more than one correlation due to multiple listening conditions, varying in masker type or performance level, also indicated in the Figure (with the exception of the 2014 study by Zekveld et al. (2011) in which the target speech and masker babble were produced by speakers either of the same gender or of different genders). When necessary, the sign of the correlation was changed so that a positive correlation represents better performance on the two tasks. An average for correlations based only on young NH listeners is provided (circle). Also given in the figure are source references (* indicates re-analyzed published data; + indicates unpublished data, personal communication), experimental conditions (performance level, PL; type of masker, Mask; type of sentence material, Mat) and participant details (age range, Age; number of participants, N). Masker: S – notionally steady noise, M_s or M_{sp} – noise modulated by an X-Hz sinusoidal amplitude modulation or a speech envelope, B_x – X-talker babble. PL: X%_(A) – adaptive procedure tracking the speech reception threshold corresponding to X%-correct identification, X%_(F_{Z-Y}) – constant stimuli procedure using several fixed SNRs yielding an overall average performance level of X% with average performance for each of the different SNRs ranging from Z to Y%-correct identification, X%_(F) – constant stimuli procedure using a single fixed SNR, yielding an average performance level of X%. In some cases, the modulation depth of the amplitude-modulated noises was only 10%, which is hardly above detection threshold (e.g., Füllgrabe et al., 2005). Therefore, those maskers are labeled as steady rather than modulated.

of analysis has the advantage not only of assuming that the true treatment effect differs from study to study, but also accounts for the fact that multiple measures can arise from the same study (e.g., where different maskers have been used in the same listeners). The analysis was performed using the R package *metafor* (Viechtbauer, 2010) and a transformation of the r values to Fisher's z scale. Across all 24 data sets, the average r value was 0.12. In other words, individual variations in WMC in young people with audiometrically normal hearing are estimated to account for, on average, less than 2% of the variance in SiN identification scores.

Given the considerably smaller number of data sets in each of the two other categories, involving older listeners, we did not compute a summary statistic. However, it is noteworthy that in the largest study included in the survey, using listeners from a wide age range, significant correlations between WMC and SiN identification were found for unmodulated and modulated background noises (see section I of **Figure 2**), and when averaged across maskers, even after partialling out the effects of age and hearing sensitivity ($r = 0.39$; $p \leq 0.001$; as reported in Füllgrabe and Rosen, 2016). However, separate correlational analyses for each age group in this study revealed that the strength of the association differed across age groups, with the youngest listeners (18–39 years) showing the weakest and a non-significant correlation ($r = 0.18$; $p = 0.162$) while stronger and significant correlations were observed for the middle-aged (40–59 years) to old-old (70–91 years) age groups (all $r \geq 0.44$; all $p \leq 0.011$). A linear regression of SiN identification scores against age, RSpan scores and their interaction showed that the slope of the linear dependence of SiN identification performance on RSpan scores indeed increased significantly with age ($p \leq 0.001$). This illustrates the moderating effect of age on the relationship between WMC and SiN identification, cautioning that the statistical control of the effect of age by computing partial correlations is not necessarily appropriate.

DISCUSSION AND CONCLUSION

Contrary to common lore and model predictions, this meta-analysis failed to find consistent evidence that, in adverse listening conditions, WMC (as measured by the RSpan test) is a reliable and strong predictor of SiN identification in young listeners with normal hearing thresholds. Recent experimental work on the perception of interrupted speech, another form of signal degradation, is consistent with this finding (Benard et al., 2014; Nagaraj and Knapp, 2015).

It could be argued that the cognitive and speech tests used in the studies surveyed here are suboptimal or inappropriate measures of WMC and SiN processing, respectively (e.g., Besser et al., 2012; Sörqvist and Rönnberg, 2012; Keidser et al., 2015). However, both the conclusions of many empirical studies, showing a link between WMC and SiN processing, and the predictions of the ELU model are based on performance obtained on these very tests.

Another criticism could be made regarding the fact that SiN identification was predominantly assessed for performance levels

close to 50% correct, obscuring the possibility that WMC and SiN identification are linked to a greater extent than reported here at other performance levels. Indeed, according to the ELU model, a greater mismatch between sensory and mental representations, and hence a higher involvement of WM-based identification processes, is predicted as speech-to-noise ratios become less favorable. However, this does not seem to be borne out by the collected results. Alternatively, it has also been argued that WM-based restorative processes in older HI (Lunner and Sundewall-Thorén, 2007; Larsby et al., 2008, 2012) and young NH listeners (Stenbäck et al., 2015) might only be effective in conditions where the acoustic signal is not “too” degraded, suggesting a non-monotonic relationship between WMC and SiN identification. While this seems an interesting proposition, the collected results do not indicate the existence of such “sweet spots” for cognitive involvement.

Hence, all things considered, the results of this meta-analysis caution against the (intuitively appealing) assumption that individual variations in WM determine SiN processing in all its forms and independently of the age and hearing status of the listener.

Despite the inconsequential degree to which WMC can predict SiN identification performance in young NH listeners, the reported results should not to be interpreted as evidence against the involvement of cognition in speech and language processing in those listeners *per se*. First, individual differences in WMC have sometimes been shown to explain some of the variability in performance in more linguistically complex tasks, such as the comprehension of conversations (Keidser et al., 2015; but see Smith and Pichora-Fuller, 2015, for contrary results for the comprehension of narratives). Second, different cognitive measures, probing individually the hypothesized sub-processes of WM (e.g., inhibition, shifting, updating; Miyake et al., 2000) or other domain-general cognitive primitives (e.g., processing speed) might prove to be better predictors of SiN processing abilities than the RSpan test (e.g., Sörqvist et al., 2010; Rudner et al., 2011).

It is also important to emphasize that the here reported findings for young NH listeners are not incompatible with the body of evidence showing significant correlations between WMC and SiN identification in primarily older HI listeners. Our own data for NH listeners sampled from across the entire adult lifespan (Füllgrabe and Rosen, 2016) revealed that WMC becomes important for SiN identification from middle age onward, with the oldest listeners (≥ 70 years) showing the strongest correlation and differing significantly from the youngest age group. One possible explanation for an increasing cognitive involvement in terms of WMC with age, in addition to the loss of audibility, is the accumulation of age-related changes in supra-threshold auditory processing (e.g., sensitivity to temporal-fine-structure and temporal-envelope cues; Schneider and Pichora-Fuller, 2001; Füllgrabe et al., 2003, 2015), sometimes from as early as mid-life (Füllgrabe, 2013). Changes in the coding fidelity of single neurons or across a neural population (Henry and Heinz, 2013; Sergeyenkov et al., 2013; Bharadwaj et al., 2014; Lopez-Poveda, 2014), which are not detected by a conventional audiometric assessment, have

indeed been associated with degraded sensory representations of the acoustic speech signal. These internal distortions could then call for more WM-based compensatory mechanisms to enable activation of the appropriate representations in long-term memory. Why, however, such age-related internal changes in coding fidelity would result in a greater reliance on WMC for SiN identification than an increase in the amount of energetic and/or informational masking is unclear. Possibly, this discrepancy could be due to secondary changes in the precision of the phonological representations stored in long-term memory, following long-standing auditory processing deficits (e.g., Andersson, 2002; Classon et al., 2013), thus providing a top-down contribution to the mismatch between sensory and mental representations. Clearly, further reflections on the nature and source of listening adversity (see Mattys et al., 2012) are needed to generate oriented hypotheses that can be tested experimentally.

From a clinical perspective, a cognitive assessment (e.g., of WMC) may still prove helpful in improving the prediction of aided SiN identification performance for older audiological patients. Future evidence based on new large samples, independent of those repeatedly investigated in previous studies (Foo et al., 2007; Rudner et al., 2008, 2009, 2011), could further specify the role and importance of cognition in audiological practice.

In conclusion, even though the question of a general vs. specialized WM system in language comprehension is not new (Caplan and Waters, 1999) and it has been speculated that differences in tasks and their processing demands activate different sub-components of the WM system, the less-discerning general opinion is that variation in WMC (often assessed by a single measure) can explain differences in performance on a variety of speech tasks. Currently available data from independent research groups do not confirm this assumption for the frequently used task of sentence identification. However, this is not to say that the processing of SiN does not involve a range of cognitive abilities, including WM. For example, it is possible that, even when individual differences exist, the WMC of most individuals is sufficient for the

purpose of SiN identification. Systematic efforts are therefore required to establish under which acoustic and linguistic conditions the different cognitive abilities come into play (e.g., Fedorenko, 2014; Smith and Pichora-Fuller, 2015; Heinrich and Knight, 2016). Finally, the results of this meta-analysis clearly highlight the need for a consistent and explicit labeling of the participant characteristics (such as age and hearing status) when reporting results and caution against the untested generalization of research findings from one participant group to another.

AUTHOR CONTRIBUTIONS

CF collated, analyzed and plotted the data, and wrote the paper. SR analyzed the data, and revised and commented on the paper.

FUNDING

The Medical Research Council Institute of Hearing Research is supported by the Medical Research Council (grant number U135097130). This work was also supported by the Oticon Foundation (Denmark).

ACKNOWLEDGMENTS

Portions of this paper were presented at the 2015 International Symposium on Hearing in Groningen, NL. We are indebted to our colleagues Kathy Arehart, Briony Banks, Jana Besser, Rebecca Carroll, Rachel Ellis, Erin Ingvalson, Inga Holube, Lisa Kilman, Thomas Koelewijn, Theresa Nüsse, Tim Schoof, Pamela Souza, Victoria Stenbäck, Verena Uslar, Anna Warzybok, and Adriana Zekveld for sharing and reanalyzing their data. We also thank Tom Campbell, Alexander Francis, Gitte Keidser, Rebecca Millman, Daniel Oberfeld-Twistel and Valeriy Shafiro for stimulating discussions, and Oliver Zobay for statistical advice.

REFERENCES

- Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47(Suppl. 2), S53–S71. doi: 10.1080/14992020802301142
- Andersson, U. (2002). Deterioration of the phonological processing skills in adults with an acquired severe hearing loss. *Eur. J. Cogn. Psychol.* 14, 335–352. doi: 10.1080/09541440143000096
- Arlinger, S., Lunner, T., Lyxell, B., and Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scand. J. Psychol.* 50, 371–384. doi: 10.1111/j.1467-9450.2009.00753.x
- Baddeley, A., Logie, R., Nimmo-Smith, I., and Brereton, N. (1985). Components of fluent reading. *J. Mem. Lang.* 24, 119–131. doi: 10.1016/0749-596x(85)90019-1
- Banks, B., Gowen, E., Munro, K. J., and Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *J. Acoust. Soc. Am.* 137, 2015–2024. doi: 10.1121/1.4916265
- Barrouillet, P. (1996). Transitive inferences from set-inclusion relations and working memory. *J. Exp. Psychol. Learn.* 22, 1408–1422. doi: 10.1037/0278-7393.22.6.1408
- Benard, M. R., Mensink, J. S., and Baskent, D. (2014). Individual differences in top-down restoration of interrupted speech: links to linguistic and cognitive abilities. *J. Acoust. Soc. Am.* 135, EL88–EL94. doi: 10.1121/1.4862879
- Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—a review. *Trends Amplif.* 17, 75–93. doi: 10.1177/1084713813495459
- Besser, J., Zekveld, A. A., Kramer, S. E., Rönnberg, J., and Festen, J. M. (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, 194–209. doi: 10.1044/1092-4388(2011/11-0008)
- Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., and Shinn-Cunningham, B. G. (2014). Cochlear neuropathy and the coding of supra-threshold sound. *Front. Syst. Neurosci.* 8:26. doi: 10.3389/fnsys.2014.00026

- Caplan, D., and Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behav. Brain Sci.* 22, 77–94.
- Carroll, R., Warzybok, A., Kollmeier, B., and Ruigendijk, E. (2016). Age-related differences in lexical access relate to speech recognition in noise. *Front. Psychol.* 7:990. doi: 10.3389/fpsyg.2016.00990.
- Classon, E. (2013). *Representing Sounds and Spellings: Phonological Decline and Compensatory Working Memory in Acquired Hearing Impairment*. Ph.D. thesis, LiU Tryck, Linköping.
- Classon, E., Rudner, M., and Rönnberg, J. (2013). Working memory compensates for hearing related phonological processing deficit. *J. Commun. Disord.* 46, 17–29. doi: 10.1016/j.jcomdis.2012.10.001
- Daneman, M., and Carpenter, P. A. (1980). Individual-differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6
- Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychon. Bull. Rev.* 3, 422–433. doi: 10.3758/BF03214546
- Das-Smaal, E. A., De Jong, P. F., and Koopmans, J. R. (1993). Working memory, attentional regulation and the Star Counting Test. *Pers. Individ. Dif.* 14, 815–824. doi: 10.1016/0191-8869(93)90094-J
- Ellis, R. J., and Munro, K. J. (2013). Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition? *Int. J. Audiol.* 52, 14–22. doi: 10.3109/14992027.2012.721013
- Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Front. Psychol.* 5:335. doi: 10.3389/fpsyg.2014.00335
- Foo, C., Rudner, M., Rönnberg, J., and 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, 618–631. doi: 10.3766/jaaa.18.7.8
- Füllgrabe, C. (2013). Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *Am. J. Audiol.* 22, 313–315. doi: 10.1044/1059-0889(2013)12-0070
- Füllgrabe, C., Meyer, B., and Lorenzi, C. (2003). Effect of cochlear damage on the detection of complex temporal envelopes. *Hear. Res.* 178, 35–43. doi: 10.1016/S0378-5955(03)00027-3
- Füllgrabe, C., Moore, B. C. J., Demany, L., Ewert, S. D., Sheft, S., and Lorenzi, C. (2005). Modulation masking produced by second-order modulators. *J. Acoust. Soc. Am.* 117, 2158–2168. doi: 10.1121/1.1861892
- Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6:347. doi: 10.3389/fnagi.2014.00347
- Füllgrabe, C., and Rosen, S. (2016). “Investigating the role of working memory in speech-in-noise identification for listeners with normal hearing,” in *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, eds P. Van Dijk, D. Başkent, E. Gaudrain, E. De Kleine, A. Wagner, and C. Lanting (New York, NY: Springer International Publishing), 29–36.
- Hagerman, B. (1982). Sentences for testing speech intelligibility in noise. *Scand. Audiol.* 11, 79–87. doi: 10.3109/01050398209076203
- Hällgren, M., Larsby, B., and Arlinger, S. (2006). A Swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition. *Int. J. Audiol.* 45, 227–237. doi: 10.1080/14992020500429583
- Heald, S. L., and Nusbaum, H. C. (2014). Speech perception as an active cognitive process. *Front. Syst. Neurosci.* 8:35. doi: 10.3389/fnsys.2014.00035
- Heinrich, A., and Knight, S. (2016). “The contribution of auditory and cognitive factors to intelligibility of words and sentences in noise,” in *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, eds P. Van Dijk, D. Başkent, E. Gaudrain, E. De Kleine, A. Wagner, and C. Lanting (New York, NY: Springer International Publishing), 37–45.
- Henry, K. S., and Heinz, M. G. (2013). Effects of sensorineural hearing loss on temporal coding of narrowband and broadband signals in the auditory periphery. *Hear. Res.* 303, 39–47. doi: 10.1016/j.heares.2013.01.014
- Keidser, G., Best, V., Freeston, K., and Boyce, A. (2015). Cognitive spare capacity: evaluation data and its association with comprehension of dynamic conversations. *Front. Psychol.* 6:597. doi: 10.3389/fpsyg.2015.00597
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., and Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 116, 2395–2405. doi: 10.1121/1.1784440
- Klein, K., and Boals, A. (2001). The relationship of life event stress and working memory capacity. *Appl. Cogn. Psychol.* 15, 565–579. doi: 10.1002/Acp.727
- Koelwijn, T., Zekveld, A. A., Festen, J. M., Rönnberg, J., and Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *Int. J. Otolaryngol.* 2012:865731. doi: 10.1155/2012/865731
- Kollmeier, B., and Wesselkamp, M. (1997). Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *J. Acoust. Soc. Am.* 102, 2412–2421. doi: 10.1121/1.419624
- Kuik, A. M. (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.
- Larsby, B., Hällgren, M., and Lyxell, B. (2008). The interference of different background noises on speech processing in elderly hearing impaired subjects. *Int. J. Audiol.* 47(Suppl. 2), S83–S90. doi: 10.1080/14992020802301159
- Larsby, B., Hällgren, M., and Lyxell, B. (2012). “Working memory capacity and lexical access in speech recognition in noise,” in *Proceedings of the 3rd International Symposium on Auditory and Audiological Research (ISAAR 2011): Speech Perception and Auditory Disorders* (Ballerup: The Danavox Jubilee Foundation), 95–102.
- Lopez-Poveda, E. A. (2014). Why do I hear but not understand? Stochastic undersampling as a model of degraded neural encoding of speech. *Front. Neurosci.* 8:348. doi: 10.3389/fnins.2014.00348
- Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42(Suppl. 1), S49–S58. doi: 10.3109/14992020309074624
- Lunner, T., and Sundewall-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, 604–617. doi: 10.3766/jaaa.18.7.7
- MacLeod, A., and Summerfield, Q. (1990). A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use. *Br. J. Audiol.* 24, 29–43. doi: 10.3109/03005369009077840
- Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734
- Moradi, S., Lidestam, B., Saremi, A., and Rönnberg, J. (2014). Gated auditory speech perception: effects of listening conditions and cognitive capacity. *Front. Psychol.* 5:531. doi: 10.3389/fpsyg.2014.00531
- Nagaraj, N. K., and Knapp, A. N. (2015). No evidence of relation between working memory and perception of interrupted speech in young adults. *J. Acoust. Soc. Am.* 138, EL145–E150. doi: 10.1121/1.4927635
- Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J. Acoust. Soc. Am.* 95, 1085–1099. doi: 10.1121/1.408469
- Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42(Suppl. 1), S68–S76. doi: 10.3109/14992020309074626
- Rönnberg, J., Arlinger, S., Lyxell, B., and Kinnefors, C. (1989). Visual evoked potentials: relation to adult speechreading and cognitive function. *J. Speech Hear. Res.* 32, 725–735. doi: 10.1044/jshr.32.04.725
- Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031
- Rönnberg, J., Rudner, M., Lunner, T., and Zekveld, A. A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise Health* 12, 263–269. doi: 10.4103/1463-1741.70505

- Rothauser, E. H., Chapman, W. D., Guttman, N., Hecker, M. H. L., Nordby, K. S., Silbiger, H. R., et al. (1969). IEEE recommended practice for speech quality measurements. *IEEE Trans. Audio Electroacoust.* 17, 225–246.
- Rudner, M., Foo, C., Rönnerberg, J., and Lunner, T. (2009). Cognition and aided speech recognition in noise: specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scand. J. Psychol.* 50, 405–418. doi: 10.1111/j.1467-9450.2009.00745.x
- Rudner, M., Foo, C., Sundewall-Thorén, E., Lunner, T., and Rönnerberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users. *Int. J. Audiol.* 47(Suppl. 2), S91–S98. doi: 10.1080/14992020802304393
- Rudner, M., Rönnerberg, J., and Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *J. Am. Acad. Audiol.* 22, 156–167. doi: 10.3766/jaaa.22.3.4
- Schneider, B. A., and Pichora-Fuller, M. K. (2001). Age-related changes in temporal processing: implications for speech perception. *Sem. Hear.* 22, 227–240. doi: 10.1055/s-2001-15628
- Schoof, T., and Rosen, S. (2014). The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Front. Aging Neurosci.* 6:307. doi: 10.3389/fnagi.2014.00307
- Sergeyenko, Y., Lall, K., Liberman, M. C., and Kujawa, S. G. (2013). Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. *J. Neurosci.* 33, 13686–13694. doi: 10.1523/JNEUROSCI.1783-13.2013
- Smith, S. L., and Pichora-Fuller, M. K. (2015). Associations between speech understanding and auditory and visual tests of verbal working memory: effects of linguistic complexity, task, age, and hearing loss. *Front. Psychol.* 6:1394. doi: 10.3389/fpsyg.2015.01394
- Sörqvist, P., Ljungberg, J. K., and Ljung, R. (2010). A sub-process view of working memory capacity: evidence from effects of speech on prose memory. *Memory* 18, 310–326. doi: 10.1080/09658211003601530
- Sörqvist, P., and Rönnerberg, J. (2012). Episodic long-term memory of spoken discourse masked by speech: what is the role for working memory capacity? *J. Speech Lang. Hear. Res.* 55, 210–218. doi: 10.1044/1092-4388(2011/10-0353)
- Souza, P. E., Arehart, K. H., Shen, J., Anderson, M., and Kates, J. M. (2015). Working memory and intelligibility of hearing-aid processed speech. *Front. Psychol.* 6:526. doi: 10.3389/fpsyg.2015.00526
- Stenbäck, V., Hällgren, M., and Larsby, B. (2016). Executive functions and working memory capacity in speech communication under adverse conditions. *Speech Lang. Hear.* 1–9. doi: 10.1080/2050571X.2016.1196034
- Stenbäck, V., Hällgren, M., Lyxell, B., and Larsby, B. (2015). The Swedish Hayling task, and its relation to working memory, verbal ability, and speech-recognition-in-noise. *Scand. J. Psychol.* 56, 264–272. doi: 10.1111/sjop.12206
- Stone, M. A., Füllgrabe, C., Mackinnon, R. C., and Moore, B. C. J. (2011). The importance for speech intelligibility of random fluctuations in “steady” background noise. *J. Acoust. Soc. Am.* 130, 2874–2881. doi: 10.1121/1.3641371
- Stone, M. A., Füllgrabe, C., and Moore, B. C. J. (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *J. Acoust. Soc. Am.* 132, 317–326. doi: 10.1121/1.4725766
- Uslar, V. N., Carroll, R., Hanke, M., Hamann, C., Ruigendijk, E., Brand, T., et al. (2013). Development and evaluation of a linguistically and audiotically controlled sentence intelligibility test. *J. Acoust. Soc. Am.* 134, 3039–3056. doi: 10.1121/1.4818760
- Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *J. Acoust. Soc. Am.* 107, 1671–1684. doi: 10.1121/1.428451
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 36, 1–48. doi: 10.18637/jss.v036.i03
- Vlaming, M. S. M. G., Kollmeier, B., Dreschler, W. A., Martin, R., Wouters, J., Grover, B., et al. (2011). HearCom: hearing in the Communication Society. *Acta Acust. United Acust.* 97, 175–192. doi: 10.3813/AAA.918397
- Zekveld, A. A., Rudner, M., Johnsrude, I. S., Festen, J. M., Van Beek, J. H., and Rönnerberg, J. (2011). The influence of semantically related and unrelated text cues on the intelligibility of sentences in noise. *Ear Hear.* 32, e16–e25. doi: 10.1097/AUD.0b013e318228036a
- Zekveld, A. A., Rudner, M., Kramer, S. E., Lyzenga, J., and 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. doi: 10.3389/fnins.2014.00088

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer MR and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 Füllgrabe and Rosen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.