



# Auditory Verbal Working Memory as a Predictor of Speech Perception in Modulated Maskers in Listeners With Normal Hearing

DOI:  
[10.1044/2017\\_JSLHR-S-16-0105](https://doi.org/10.1044/2017_JSLHR-S-16-0105)

**Document Version**  
Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

## **Citation for published version (APA):**

Millman, R. E., & Mattys, S. L. (2017). Auditory Verbal Working Memory as a Predictor of Speech Perception in Modulated Maskers in Listeners With Normal Hearing. *Journal of Speech, Language, and Hearing Research*, 60(5), 1236-1245. [https://doi.org/10.1044/2017\\_JSLHR-S-16-0105](https://doi.org/10.1044/2017_JSLHR-S-16-0105)

**Published in:**  
Journal of Speech, Language, and Hearing Research

## **Citing this paper**

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

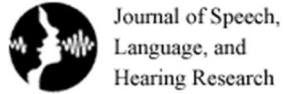
## **General rights**

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## **Takedown policy**

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [uml.scholarlycommunications@manchester.ac.uk](mailto:uml.scholarlycommunications@manchester.ac.uk) providing relevant details, so we can investigate your claim.





**Auditory verbal working memory as a predictor of speech perception in modulated maskers in normal-hearing listeners**

Journal:	<i>Journal of Speech, Language, and Hearing Research</i>
Manuscript ID	JSLHR-S-16-0105.R2
Manuscript Type:	Research Article
Date Submitted by the Author:	11-Oct-2016
Complete List of Authors:	Millman, Rebecca; University of Manchester, Manchester Centre for Audiology and Deafness Mattys, Sven; University of York, Department of Psychology
Keywords:	Memory, Speech perception, Aging

SCHOLARONE™  
Manuscripts

view

1  
2  
3 **Auditory verbal working memory as a predictor of speech perception in modulated**  
4 **maskers in normal-hearing listeners<sup>a)</sup>**  
5  
6  
7

8  
9  
10 Rebecca E. Millman<sup>b)</sup> and Sven L. Mattys<sup>c)</sup>  
11

12  
13  
14  
15 <sup>a)</sup> Portions of this work were presented at the Third International Conference on Cognitive  
16 Hearing Science for Communication, Linköping, Sweden, 2015.  
17

18  
19  
20  
21  
22 <sup>b)</sup> Manchester Centre for Audiology and Deafness, School of Health Sciences, Faculty of  
23 Biology, Medicine, and Health, University of Manchester, Manchester, M13 9PL, UK.  
24

25  
26  
27  
28  
29 <sup>c)</sup> Department of Psychology, University of York, York, YO10 5DD, UK.  
30  
31

32  
33  
34 Submitted to: Journal of Speech, Language, and Hearing Research  
35

36  
37  
38 Running title: Auditory memory for speech in noise  
39

40  
41  
42  
43 Author to whom correspondence should be addressed: Rebecca E. Millman  
44

45  
46  
47  
48 Telephone: +44 (0) 161 275 3387 Fax: +44 (0) 161 275 3373  
49

50  
51  
52  
53 Electronic-mail: rebecca.millman@manchester.ac.uk  
54  
55  
56  
57  
58  
59  
60

## Abstract

**Purpose:** Background noise can interfere with our ability to understand speech. Working memory capacity (WMC) has been shown to contribute to the perception of speech in modulated noise maskers. WMC has been assessed with a variety of auditory and visual tests, often pertaining to different components of working memory. This study assessed the relationship between speech perception in modulated maskers and components of auditory verbal working memory (AVWM) over a range of signal-to-noise ratios.

**Method:** Speech perception in noise and AVWM were measured in 30 listeners (age range 31-67 years) with normal hearing. AVWM was estimated using forward digit recall, backward digit recall, and non-word repetition.

**Results:** After controlling for the effects of age and average pure-tone hearing threshold, speech perception in modulated maskers was related to individual differences in the phonological component of working memory (as assessed by non-word repetition) but only in the least favorable SNR. The executive component of working memory (as assessed by backward digit) was not predictive of speech perception in any conditions.

**Conclusions:** AVWM is predictive of the ability to benefit from temporal dips in modulated maskers: Listeners with greater phonological WMC are better able to correctly identify sentences in modulated noise backgrounds.

## Introduction

Older adult listeners frequently experience difficulties in understanding speech presented in a noisy background. Difficulties may result from reduced audibility (e.g., Humes, 2002), age (e.g., Dubno, Horwitz, & Ahlstrom, 2002), perceptual deficits in temporal resolution, spectral resolution, or cochlear compression (e.g., Gregan, Nelson, & Oxenham, 2013), and deficits in the cognitive systems underpinning speech perception (CHABA, 1988; Wingfield, Tun, & McCoy, 2005). Among the cognitive abilities related to the perception of speech in noisy backgrounds, working memory has attracted substantive attention (e.g., Gatehouse, Naylor, & Eberling, 2003; Lunner, 2003; Akeroyd, 2008; Rönnberg, Rudner, Lunner, & Zekveld, 2010; Mattys, Davis, & Bradlow, 2012; Füllgrabe & Rosen, 2016). Working memory may be conceptualized as the system that underpins the temporary storage and manipulation of information required for cognitive tasks (e.g., Baddeley, 1992).

One view argues that age-related differences in cognition can be accounted for by deficits in auditory processing (e.g., Schneider & Pichora-Fuller, 2000; McCoy, Tun, Cox, Colangelo, Stewart, & Wingfield, 2005; Humes, Busey, Craig, & Kewley-Port, 2015). On this account, age-related decline in auditory processing depletes working memory resources so that fewer resources are available for storage and manipulation of information. Consistent with this view, age-related differences in cognitive tasks may be largely eliminated when the perceptual difficulty of the tasks are taken into account (e.g., Pichora-Fuller, Schneider, & Daneman, 1995; Murphy, Craik, Li, & Schneider, 2000).

An alternative view, as proposed by the Ease of Language Understanding (ELU) model (e.g., Rönnberg, 2003; Rönnberg et al., 2010; Rönnberg, Danielsson, Rudner,

*Millman and Mattys (2016)*

1  
2  
3 Arlinger, Sternäng, Wahlin, & Nilsson, 2011; Rönnberg, Lunner, Zekveld, Sörqvist,  
4  
5 Danielsson, Lyxell, Dahlström, Signoret, Stenfelt, Pichora-Fuller, & Rudner, 2013),  
6  
7 argues that working memory is always deployed when speech is presented against a noisy  
8  
9 background, even at favorable signal-to-noise ratios (SNRs). The ELU model predicts  
10  
11 reliance on working memory in perceptually adverse conditions because working  
12  
13 memory is thought to: 1) Compensate for the ambiguity caused by degraded phonological  
14  
15 representations, 2) Contribute to predictive coding, either before or during the early  
16  
17 portions of a spoken stimulus, and 3) Facilitate explicit semantic expectations of a  
18  
19 sentence as it unfolds (e.g., Rönnberg et al., 2013). The role of working memory may  
20  
21 become even more important when speech is presented in a fluctuating masker because  
22  
23 listeners with greater cognitive capacity are more able to take advantage of the temporal  
24  
25 dips in the masker (e.g., Gatehouse et al., 2003). Working memory may play a greater  
26  
27 role in fluctuating maskers in older listeners because age-related deficits in speech  
28  
29 perception are exacerbated when the background noise is modulated (e.g., Takahasi &  
30  
31 Bacon, 1992; Dubno et al., 2002). The dependence on working memory is increased  
32  
33 further when the masker consists of speech rather than meaningless fluctuating noise  
34  
35 (Koelewijn, Zekveld, Festen, Rönnberg, & Kramer, 2012; Zekveld, Rudner, Johnsrude, &  
36  
37 Rönnberg, 2013a). In such challenging listening situations, working memory is thought to  
38  
39 support maintenance of a mental representation of a spoken sentence, whilst knowledge  
40  
41 of language and context are recruited to fill the gaps in the degraded auditory signal.  
42  
43  
44  
45  
46  
47  
48  
49

50  
51 The exact contribution of working memory to speech perception under adverse  
52  
53 conditions remains controversial, however. Part of the debate centers on handling  
54  
55 confounds such as age (e.g., Wingfield, Stine, Lahar, & Aberdeen, 1988; Salthouse,  
56  
57  
58  
59  
60

1  
2  
3 1994), pure-tone average thresholds (PTAs), and stimulus audibility (Baldwin & Ash,  
4 2011; McCoy et al., 2005; Wingfield et al., 2005). Other issues are related to the validity  
5  
6 of the tests used for measuring working memory. For instance, although the reading-span  
7  
8 test (R-SPAN, Daneman & Carpenter, 1980; Baddeley, Logie, Nimmo-Smith, &  
9  
10 Brereton, 1985) is a widely used measure of (visual) verbal working memory in hearing  
11  
12 research (e.g., Pichora-Fuller et al., 1995; Lunner, 2003; Akeroyd, 2008; Füllgrabe,  
13  
14 Moore, & Stone, 2015; Smith & Pichora-Fuller, 2015), recent work (e.g., Sörqvist,  
15  
16 Ljungberg, & Ljung, 2010; Sörqvist & Rönnerberg, 2012) has questioned the use of R-  
17  
18 SPAN as the sole suitable measure of working memory.  
19  
20  
21  
22  
23

24  
25 Because both reading and listening span tasks are highly correlated with reading  
26  
27 comprehension in young normal-hearing adults under quiet listening conditions (e.g.,  
28  
29 Daneman & Carpenter, 1980), it is often assumed that verbal working memory for  
30  
31 language comprehension is modality-independent. However, evidence of modality-  
32  
33 specific working memory (Humes, Burk, Couchlin, Busey, & Strauser, 2007) has  
34  
35 motivated comparisons of the R-SPAN with its auditory equivalent, the listening span (L-  
36  
37 SPAN; e.g., Koelewijn et al., 2012; Zekveld, Festen, & Kramer, 2013b; Besser,  
38  
39 Koelewijn, Zekveld, Kramer, & Festen, 2013). Specifically, data have shown age-related  
40  
41 deficits in performance on the L-SPAN but not on the R-SPAN (e.g. Pichora-Fuller et al.,  
42  
43 2005; Baldwin & Ash, 2011). Moreover, L-SPAN performance is not related to speech  
44  
45 perception in noise in older listeners with normal hearing (NH listeners), whereas R-  
46  
47 SPAN performance is (Koelewijn et al., 2012). Besser et al. (2013) also noted that L-  
48  
49 SPAN scores are generally higher than R-SPAN scores in NH listeners (Koelewijn et al.,  
50  
51 2012; Zekveld et al., 2013b). These discrepancies between R-SPAN and L-SPAN  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 outcomes further question the reliance on SPAN tests in assessing the contribution of  
4  
5  
6 working memory to speech perception in noise (e.g., Sörquist & Ronnberg, 2012).  
7

8 The primary objective of the present study was to determine whether age-related  
9  
10 changes in AVWM are associated with deficits in understanding speech in noise in  
11  
12 listeners with normal hearing. A speech in noise task was chosen as the dependent  
13  
14 measure in order to provide a listening situation that would promote reliance on working  
15  
16 memory, i.e., the ability to take advantage of the amplitude minima in modulated maskers  
17  
18 to detect speech cues. When temporal fluctuations are imposed on an unmodulated noise  
19  
20 masker, listeners sometimes benefit from the temporal dips in the masker, i.e., they  
21  
22 exhibit speech masking release (MR). Listeners may need components of AVWM to  
23  
24 retain glimpses of speech revealed during the temporal dips. Additional cognitive factors,  
25  
26 including linguistic abilities (e.g., Benard, Mensink, & Başkent, 2014; Avivi-Reich,  
27  
28 Daneman, & Schneider, 2014), which also involve memory processes, may also be used  
29  
30 to make inferences about the target speech based on partial acoustic and linguistic cues  
31  
32 obtained during the temporal dips.  
33  
34  
35  
36  
37  
38

39 Given the complex nature of the R-SPAN and L-SPAN tests, which tax both the  
40  
41 central executive (for processing) and phonological loop (for storage) components of  
42  
43 working memory (e.g., Baddeley et al., 1985; Gathercole, Pickering, Ambridge, &  
44  
45 Wearing, 2004), any relationship between verbal working memory capacity measured  
46  
47 using these tests and the ability to understand speech in noise remains ill-defined.  
48  
49

50 In this study, we aimed to isolate the sub-component(s) of AVWM that  
51  
52 contribute(s) to the ability to benefit from speech MR and speech perception in  
53  
54 modulated noise. Accordingly, standardized tests of AVWM, namely, forward digit  
55  
56  
57  
58  
59  
60



1  
2  
3 recall, backward digit recall, and non-word repetition, were used. These tests may be  
4  
5 considered in terms of a gradient of processing demands (e.g., Daneman & Merikle,  
6  
7 1996), with forward digit recall requiring little processing demand, backward digit span  
8  
9 requiring more processing demand, and non-word repetition requiring the greatest  
10  
11 processing demand. However, an alternative view based on Baddeley's (1992) model of  
12  
13 working memory proposes that these tests probe different sub-processes in AVWM:  
14  
15 Forward digit recall and backward digit recall both tap into the phonological component  
16  
17 of working memory, whereas backward digit recall additionally relies on an attention-  
18  
19 demanding manipulation of the digit sequence by the executive control (e.g., Alloway,  
20  
21 Gathercole, Willis, & Adams, 2004; Gathercole et al., 2004). In that conceptualization,  
22  
23 backward digit recall loads onto the same factor as more complex working memory such  
24  
25 as the SPAN tests (e.g., Alloway et al., 2004; Gathercole et al., 2004). The phonological  
26  
27 component of working memory is a common underlying factor in both digit recall and  
28  
29 non-word repetition, but these two tasks are also supported by long-term lexical  
30  
31 knowledge (e.g. Gathercole, 1995). Long-term lexical knowledge may play a lesser role  
32  
33 in non-word repetition than digit recall, depending on the degree of "wordlikeness" of the  
34  
35 non-words (e.g., Gathercole, Willis, Emslie, & Baddeley, 1991; Gathercole, 1995). The  
36  
37 aim of using these different tests of AVWM was to clarify the role of sub-processes in  
38  
39 AVWM in speech perception in modulated maskers.  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

## 50 **Methods**

### 51 ***Listeners***

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Thirty listeners, all native English speakers, participated in the study. Listeners (11 male) ranged in age from 31 to 67 years (mean age = 53.5 years, s.d. = 9.4 years). All listeners provided written informed consent and were paid for their participation in the study.

Pure-tone air-conduction audiometric thresholds were measured in accordance with the procedures recommended by the British Society of Audiology (British Society of Audiology, 2004). The listeners had audiometrically normal hearing in both ears, defined as audiometric thresholds of no more than 20 dB HL for octave frequencies between 0.5 and 4 kHz. Individual and mean audiograms of NH listeners are shown in Fig 1. Audiometric thresholds were measured for frequencies of 0.5, 1, 2, and 4 kHz only, because the stimuli were bandpass-filtered with cut-off frequencies of 0.5 and 4 kHz using a 6<sup>th</sup>-order Butterworth filter and a linear-phase FIR digital filter (Vickers, Moore, & Baer, 2001).

*Insert Figure 1 about here*

### **Procedure**

Audiometric thresholds, AVWM capacity, and speech perception were measured in one session. Listeners were seated in a double-walled sound-attenuating booth. For the measures of AVWM and speech perception, stimuli were delivered diotically through Sennheiser HD 650 headphones. Stimuli were played through an E-MU soundcard using custom MATLAB (The MathWorks Inc., Natick, MA) routines. The overall stimulus level was 65 dB SPL. Depending on the task, responses were made either orally (AVWM) or typed manually by the participants using a keyboard (speech perception).

### **Speech masking release (MR)**

The speech stimuli were IEEE sentences (Rothausser, Chapman, Guttman, Silbiger, Hecker, Urbanek, Nordby, & Weinstick, 1969) spoken by five female speakers. Some aspects of the experimental design for the measurement of speech perception were driven by considerations for an accompanying neuroimaging experiment, which is not included in the present study. The entire list of IEEE sentences was sorted based on duration: One hundred and twenty sentences of the IEEE shortest sentences (mean duration = 2.22 s, s.d. = 0.06 s) were used for the behavioral testing and a further thirty sentences were used for training prior to behavioral testing. Sentences contained four or five key words (mean number of key words = 4.92, s.d. = 0.27).

Sentences were normalised for long-term root mean square (RMS) amplitude. Key word identification was measured in a masking noise that was spectrally shaped to match the long-term power spectrum of the sentences. The noise was either unmodulated or 100 % modulated with a 2-Hz square-wave (50 % duty cycle). The “on” and “off” periods of the 2-Hz square-wave masker were both 250 ms. The on/off slopes of the square-wave were shaped with 5-ms cosine-squared ramps. The noise was shaped after scaling and the level of the modulated noise during the “on” periods was the same as that of the unmodulated noise (Gregan et al., 2013). The modulation frequency (4 Hz) of the modulated noise was similar to the speech rate (~4 Hz, e.g., Hughes, 2013) of the IEEE sentences.

Speech MR was defined as the difference in the correct response rate between the unmodulated and modulated masker conditions. In the case of speech MR, the optimal

1  
2  
3 masker modulation rate is dependent on the speech material. A modulation frequency of  
4  
5 8-10 Hz usually provides the greatest amount of speech MR (e.g., Miller & Licklider,  
6  
7 1950; Gustafsson & Arlinger, 1994). We chose 2 Hz because results from previous  
8  
9 studies show that a 2-Hz square-wave masker provides substantial benefit from sentence-  
10  
11 level speech MR (e.g., Nelson, Jin, Carney, & Nelson, 2003). Different noise samples  
12  
13 were selected for each presentation. When the noise was modulated, the first period of the  
14  
15 noise was always “on” and therefore the first 250-ms portion of each sentence was  
16  
17 masked. The speech-in-noise stimuli were ramped on and off using a 25-ms raised cosine  
18  
19 function.  
20  
21  
22  
23

24  
25 One method for measuring speech perception in noise is to compare the signal-to-  
26  
27 noise ratio (SNR) required for 50 % correct performance in the unmodulated masker with  
28  
29 the SNR required for 50 % correct performance in the modulated masker, i.e., the speech  
30  
31 reception threshold ( $SRT_{50}$ ). In the case of speech MR,  $SRT_{50}$  is often difficult to  
32  
33 calculate accurately because listeners with NH usually perform well even at low SNRs  
34  
35 when the masker is modulated: For some listeners, performance in modulated noise may  
36  
37 not fall below 50 % correct, even for very unfavorable SNRs (e.g., Gregan et al., 2013).  
38  
39 Therefore, an accepted alternative method (e.g., Nelson et al., 2003; Gregan et al., 2013;  
40  
41 Füllgrabe et al., 2015) is to calculate the difference in performance between the  
42  
43 unmodulated and modulated masker at fixed levels of SNR, which is the method we used  
44  
45 here. SNRs were fixed at -4, -8, and -12 dB. Negative SNRs were chosen because they  
46  
47 result in more speech MR (e.g., Bernstein & Grant, 2009; Bernstein & Brungart, 2011;  
48  
49 Gregan et al., 2013). Twenty sentences were randomly assigned to each of three SNR  
50  
51 conditions (-4, -8, and -12 dB) and two masker conditions (unmodulated, modulated).  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 The order of the SNR and masker conditions was randomized prior to the start of each  
4 participant's testing. Participants were instructed to type the words they could understand  
5 using a computer keyboard.  
6  
7  
8  
9

10 Speech perception was measured as the percentage of words correctly transcribed.  
11 Individual scores for correct identification of keywords were transformed into  
12 rationalized arcsine units (RAU) (Studebaker, 1985) for statistical analyses. Unless  
13 specified, the results depicted in figures and in the text are the untransformed percentages  
14 of keywords correctly identified.  
15  
16  
17  
18  
19  
20  
21  
22

### 23 **Auditory Verbal Working Memory (AVWM)**

24  
25 The primary aim of the present study was to isolate the components of AVWM  
26 that contribute to speech MR and speech perception in modulated noise in NH listeners.  
27 To this end, auditory stimuli from a subset of the Automated Working Memory  
28 Assessment (AWMA) tests (Pearson Education Limited, London, United Kingdom,  
29 2007) were employed. The AWMA tests used were non-word repetition (NWR), forward  
30 digit recall (Digit), and backward digit recall (BackDigit). As mentioned previously,  
31 these AWMA tests may be considered in terms of a gradient of processing demands (e.g.,  
32 Daneman & Merikle, 1996), with Digit entailing minimal processing demands and NWR  
33 requiring the greatest processing demands. They can also be described as probing  
34 different sub-components of AVWM. NWR is assumed to assess the phonological  
35 component of working memory (e.g., Gathercole et al., 1991; Gathercole et al., 2004).  
36 Both Digit and BackDigit assess the phonological component of working memory as  
37 well, but, importantly, BackDigit is also thought to require executive control (e.g.,  
38 Alloway et al., 2004; Gathercole et al., 2004) because the digits must undergo active  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 processing in order to be recalled in reverse order. The order of the AWMA tests was  
4  
5 counterbalanced across participants.  
6  
7

8 For each test, number sequences or non-word sequences of increasing length were  
9  
10 presented to listeners. The starting number of items was two. There were six trials per  
11  
12 length. The test was discontinued after 3 incorrect responses for a given length. For the  
13  
14 Digit and BackDigit tests, responses were scored as incorrect if verbal responses did not  
15  
16 correspond exactly to the target sequence of numbers. For the NWR test, responses were  
17  
18 scored as incorrect if the verbal response differed phonemically from the target non-word  
19  
20 by one phoneme or more. The maximum possible score was 36 for the NWR test, 48 for  
21  
22 the Digit test, and 36 for the BackDigit test.  
23  
24  
25  
26  
27  
28

## 29 Results

### 30 *Speech perception in noise maskers*

31  
32 Mean keyword identification scores are shown in Fig. 2. Consistent with previous  
33  
34 work (e.g., Nelson et al., 2003; Gregan et al., 2013), speech perception in the presence of  
35  
36 an unmodulated masker at low SNRs was poor. The lower SNRs (-8 and -12 dB) resulted  
37  
38 in floor performance in the unmodulated masker condition. Given the floor performance  
39  
40 in those conditions, speech MR was only measured for -4 dB SNR. For that condition, a  
41  
42 paired *t*-test (2-tailed) confirmed that listeners benefitted from speech masking release,  
43  
44 that is, speech perception in the modulated masker condition was significantly better than  
45  
46 in the unmodulated masker condition [35.6 %, s.d. = 9.3 %,  $t(29) = 20.34$ ,  $p < 0.005$ ].  
47  
48  
49  
50  
51  
52

53 The results also showed that speech perception in the modulated masker condition  
54  
55 improved as the SNR increased: From 23.8 % (s.d. = 6.5 %) at -12 dB SNR to 46.3 %  
56  
57  
58  
59  
60

1  
2  
3 (s.d. = 9.2 %) at -4 dB SNR. An Analysis of Variance (ANOVA) confirmed the increase  
4  
5  
6 in performance as a function of SNR [ $F(2,58) = 124.11, p < 0.005$ ].  
7  
8  
9

10 *Insert Fig. 2 about here*  
11  
12  
13  
14

15 The MR effect at -4 dB SNR is consistent with previous studies that used  
16  
17 unmodulated and square-wave modulated noise to measure speech MR (e.g., Nelson et  
18  
19 al., 2003; Gregan et al., 2013). Notable methodological differences in the present study  
20  
21 compared to those of Gregan et al. (2013) include: 1) A lower masker modulation  
22  
23 compared to those of Gregan et al. (2013) include: 1) A lower masker modulation  
24  
25 frequency for the modulated masker (2-Hz in our study vs. 10-Hz in theirs) and 2) The  
26  
27 temporal relationship of the speech and the masker (speech and masker were gated on  
28  
29 and off together in our study vs. the speech signal was temporally centered within a  
30  
31 longer duration masker in theirs). A 2-Hz square-wave modulated masker results in less  
32  
33 speech MR than a 10-Hz square-wave modulated masker (e.g., Nelson et al., 2003).  
34  
35 However, in the present study, the modulated noise was always in phase with the  
36  
37 sentence onset, which would make the amplitude minima in the modulated noise  
38  
39 predictable, given that only one modulation rate was tested. The predictability of the  
40  
41 temporal dips in the modulated masker would improve absolute performance in the  
42  
43 modulated masker condition and increase the amount of speech MR, relative to a  
44  
45 modulated masker with a random starting phase.  
46  
47  
48  
49  
50  
51  
52

53 ***Link between AVWM, hearing thresholds, and age***  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Fig. 3 shows the mean scores on the AVWM tests of NWR, Digit, and BackDigit.  
4  
5  
6  
7

8 *Insert Fig. 3 about here*  
9

10  
11  
12 Hierarchical multiple regression analyses were employed to further explore the  
13 predictive value of Age and PTA on each one of the sub-components of AVWM  
14 considered separately. Age was entered first, followed by PTA. Table I shows that Age  
15 and PTA showed different patterns of association with the sub-components of AVWM.  
16 NWR scores were reliably predicted by Age ( $R^2 = 0.25, p = 0.005$ ), but PTA did not  
17 further increase the model fit ( $\Delta R^2 = 0.05, p = 0.16$ ). Forward digit scores were also  
18 reliably predicted by Age in the first step of the model ( $R^2 = 0.15, p = 0.03$ ), but adding  
19 PTA did not improve the fit ( $\Delta R^2 = 0, p = 0.99$ ). As for BackDigit, Age did not make a  
20 contribution to performance ( $R^2 = 0.05, p = 0.25$ ) in the first step of the model. However,  
21 in the second step of the model, PTA did add to the model ( $\Delta R^2 = 0.14, p = 0.04$ ).  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35

36 In sum, consistent with the finding that working memory capacity declines with  
37 age (e.g., Wingfield *et al.*, 1988; Salthouse, 1994), the present results show that the sub-  
38 components of AVWM that relate to the phonological component of working memory  
39 (forward digit recall, non-word repetition) were affected by age. There was no significant  
40 effect of age on backward digit recall, which is in contrast with some previous studies  
41 (e.g., Tun, McCoy, & Wingfield, 2009). It is possible that the association between age  
42 and auditory forward and backward digit recall identified in those studies was  
43 confounded with age-related increases in hearing thresholds. In the present study, all  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 listeners had clinically-defined normal hearing (i.e.,  $\leq 20$  dB HL), and there was no  
4  
5 significant correlation between age and PTA ( $r = 0.024$ ,  $p = 0.90$ , 2-tailed significance).  
6  
7

8 We also found that the executive component of working memory, as measured by  
9  
10 backward digit recall, was influenced by hearing thresholds. This was in sharp contrast  
11  
12 with the lack of effect of hearing thresholds on forward digit recall and non-word  
13  
14 repetition, which are measures of the phonological component of AVWM that requires  
15  
16 little executive control (e.g., Gathercole et al., 2004). These differential predictive values  
17  
18 of age and PTA support the idea that the AVWM tests used here probed different sub-  
19  
20 processes in AVWM (e.g., Alloway et al., 2004; Gathercole et al., 2004), rather than  
21  
22 representing a gradient of processing demands (Daneman & Merikle, 1996). We return to  
23  
24 this distinction in the next section.  
25  
26  
27  
28  
29  
30  
31

### 32 ***Link between AVWM and speech perception in modulated maskers***

33  
34 Hierarchical multiple regression models were used to assess AVWM predictors of  
35  
36 speech MR and speech perception in modulated maskers. Variables thought to affect  
37  
38 speech perception in noise, namely Age (Dubno et al., 2002; George, Festen, & Houtgast,  
39  
40 2006) and PTA (e.g., Bacon, Opie, & Montoya, 1998; Desloge, Reed, Braida, Perez, &  
41  
42 Delhome, 2010), were entered in the first step of the model. The AVWM tests were  
43  
44 included in a stepwise-fashion in a second step because the relative contributions of sub-  
45  
46 components of AVWM to speech perception in modulated maskers are unknown.  
47  
48  
49

50 Analyses were carried out for MR at -4 dB SNR only because 1) This SNR is  
51  
52 comparable with that used by Gregan et al. (2013, -5 dB SNR) to assess perceptual  
53  
54 predictors of speech MR, and 2) The results for the -4 dB SNR condition were not  
55  
56  
57  
58  
59  
60

1  
2  
3 affected by floor performance. Results are shown in Table 2 (speech MR, -4 dB SNR).  
4  
5 The first step, which included Age ( $p = 0.50$ ) and PTA ( $p = 0.10$ ), did not significantly  
6  
7 predict speech MR ( $R^2 = 0.12$ ,  $p = 0.19$ ). The AVWM scores did not add to the model fit  
8  
9 (NWR,  $p = 0.46$ ; Digit,  $p = 0.33$ ; BackDigit,  $p = 0.64$ ).  
10  
11

12  
13 Regression analyses were also carried out for speech perception in the modulated  
14  
15 noise masker at each SNR (-4, -8, and -12 dB). Age and PTA were included in the first  
16  
17 step of the model and the scores on the AVWM tests (NWR, Digit, and BackDigit) were  
18  
19 added in the second step. These analyses test a central prediction of the ELU model (e.g.,  
20  
21 Rönnerberg, 2013), namely that working memory supports speech perception in noise, even  
22  
23 at more favorable SNRs. At an SNR of -4 dB (Table 3), the first step of the model,  
24  
25 including Age ( $p = 0.04$ ) and PTA ( $p = 0.17$ ), was predictive of performance ( $R^2 = 0.20$ ,  
26  
27  $p = 0.046$ ). The AVWM scores did not increase the model fit (NWR,  $p = 0.13$ ; Digit,  $p =$   
28  
29  $0.4$ ; BackDigit,  $p = 0.82$ ). At -8 dB SNR (Table 4), neither Age ( $p = 0.29$ ), nor PTA ( $p =$   
30  
31  $0.43$ ) predicted performance ( $R^2 = 0.06$ ,  $p = 0.41$ ). The AVWM scores did not improve  
32  
33 the model fit (NWR,  $p = 0.1$ ; Digit,  $p = 0.47$ ; BackDigit,  $p = 0.36$ ). At -12 dB SNR  
34  
35 (Table 5), Age ( $p = 0.12$ ) and PTA ( $p = 0.03$ ) were predictive of performance ( $R^2 = 0.23$ ,  
36  
37  $p = 0.03$ ). In the second step, only the phonological component of working memory, as  
38  
39 measured by NWR ( $p = 0.048$ ), added to the model fit ( $\Delta R^2 = 0.11$ ,  $p = 0.048$ ) for an  
40  
41 SNR of -12 dB (Digit,  $p = 0.33$ , BackDigit,  $p = 0.48$ ). Fig. 4 shows that better speech  
42  
43 perception in a modulated masker at -12 dB SNR was associated with higher NWR  
44  
45 scores ( $p = 0.04$ ).  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Insert Fig. 4 about here

1  
2  
3  
4  
5  
6 In sum, these analyses show that the phonological component of working memory  
7  
8 (NWR) is the sub-component of AVWM that contributes most to speech perception in  
9  
10 modulated noise. The lack of association between BackDigit and speech-in-noise  
11  
12 perception suggests that individual differences in central executive processing are not  
13  
14 significantly related to speech perception in modulated noise maskers.  
15  
16  
17  
18  
19

## 20 **General Discussion**

21  
22  
23 Speech perception in noise is influenced by both perceptual factors (e.g. Pichora-  
24  
25 Fuller et al., 1995; Gregan et al., 2013) and cognitive factors (CHABA, 1988; Wingfield,  
26  
27 et al., 2005; Rönnberg et al., 2013). There is some evidence of a link between (visual)  
28  
29 verbal working memory capacity and the ability of both normal-hearing (Koelewijn, et  
30  
31 al., 2012; Zekveld et al., 2013a; Füllgrabe & Rosen, 2016) and hearing-impaired listeners  
32  
33 (Lunner, 2003; Foo, Rudner, Rönnberg, & Lunner, 2007; Rönnberg et al., 2010, 2011,  
34  
35 2013; Rudner, Rönnberg, and Lunner, 2011) to understand speech in noisy backgrounds.  
36  
37 To date, however, there has been limited evidence that auditory verbal working memory  
38  
39 (AVWM) contributes to understanding speech in noise (for reviews, see Akeroyd, 2008;  
40  
41 Besser et al., 2013). In the present study, we examined the relationship between  
42  
43 individual sub-components of AVWM and the ability of NH listeners to take advantage  
44  
45 of the temporal dips in modulated maskers to aid speech perception. We hypothesized  
46  
47 that individual differences in AVWM would be related to listeners' ability to understand  
48  
49 speech played in modulated maskers. Indeed, understanding sentences played against  
50  
51 modulated background noise requires successive elements of a sentence, especially those  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 revealed during the amplitude minima in the masker, to be identified and remembered  
4  
5 until the syntactic structure of the sentence can be parsed and its meaning inferred (Lyxell  
6  
7 & Rönnberg, 1989; Rönnberg et al., 2013).  
8  
9

10  
11 Within the Ease of Language Understanding framework (ELU, Rönnberg, 2003;  
12  
13 Rönnberg et al., 2011, 2013), phonological input is matched to stored phonological  
14  
15 representations in semantic long-term memory. Because ambiguity between input and  
16  
17 stored representations is higher under perceptually adverse conditions, working memory  
18  
19 is recruited to resolve the mismatch. The ELU model predicts that working memory  
20  
21 contributes to speech perception in noise even at favorable SNRs (e.g., Rönnberg et al.,  
22  
23 2013). Therefore, we assessed the relationship between speech perception in modulated  
24  
25 maskers and components of AVWM over a range of signal-to-noise ratios. We found that  
26  
27 individual differences in the phonological component of working memory were linked  
28  
29 with the ability to perceive speech in modulated noise only when the SNR was very  
30  
31 unfavorable (i.e., -12 dB). Thus, our results support the link between working memory  
32  
33 and speech perception put forward in the ELU model, but they do not support the claim  
34  
35 that this relationship is manifest at more favorable SNRs (i.e., -4 or -8 dB), at least in  
36  
37 listeners with normal hearing and with the present stimuli.  
38  
39  
40  
41  
42

43  
44 Individual differences on the non-word repetition (NWR) test of AVWM were  
45  
46 associated with speech perception in modulated noise. NWR is a measure of the  
47  
48 phonological component of working memory and, to some extent, long-term lexical  
49  
50 knowledge (Gathercole & Adams, 1994; Gathercole, 1995). Listeners may rely on the  
51  
52 interplay between the phonological component of working memory and lexical  
53  
54 knowledge to support the perception of speech in modulated noise. In order for listeners  
55  
56  
57  
58  
59  
60

1  
2  
3 to benefit from temporal dips in modulated maskers, the phonological component of  
4  
5 working memory may be needed to retain glimpses of locally and temporarily  
6  
7 meaningless information in working memory. Together with a contribution from long-  
8  
9 term lexical knowledge, the phonological component of working memory may contribute  
10  
11 to improving speech perception in modulated noise through retention of speech cues  
12  
13 acquired during the temporal dips.  
14  
15

16  
17 This view of an interplay between the phonological component of working  
18  
19 memory and long-term lexical knowledge to support speech perception in modulated  
20  
21 noise is consistent with ELU predictions (e.g., Rönnberg et al., 2011, 2013). Indeed, the  
22  
23 network supporting speech perception in noise relies heavily on working memory and is  
24  
25 also critically dependent on additional cognitive resources, including linguistic abilities  
26  
27 (Benard et al., 2014), lexical and linguistic compensation (e.g., Avivi-Reich et al., 2015),  
28  
29 and phonological and lexical representations from semantic long-term memory  
30  
31 (Rönnberg et al., 2011; 2013). For example, linguistic context, retrieved from long-term  
32  
33 memory, rather than working memory, contributes to the restoration of interrupted speech  
34  
35 (Benard et al., 2014; Nagaraj & Knapp, 2015). Long-term semantic memory resources  
36  
37 may be deployed to interact with the information that has already been decoded in the  
38  
39 phonological component of working memory, i.e., the glimpses acquired during temporal  
40  
41 dips in modulated masker, to facilitate improved speech perception in modulated noise.  
42  
43  
44  
45  
46  
47

48 The results presented here suggest that uncovering a relationship between speech  
49  
50 perception in noise and AVWM depends on selecting the appropriate test of AVWM. In  
51  
52 this study, only the phonological component of working memory, as measured using non-  
53  
54 word repetition, was associated with speech perception in modulated noise. Considered in  
55  
56  
57  
58  
59  
60

1  
2  
3 terms of a gradient of processing demands (e.g., Daneman & Merikle, 1996), NWR may  
4  
5 be associated with speech perception in -12 dB SNR because NWR may entail the  
6  
7 greatest processing demands of the three AVWM tests we used. However, the variation  
8  
9 in the predictive values of Age and PTA on the individual AVWM tests suggests that  
10  
11 these tests might instead probe sub-components of AVWM with different requirements  
12  
13 for executive control, rather than simply different gradients of processing demands.  
14  
15

16  
17 We did not find an association between speech perception in modulated noise and  
18  
19 individual differences in the central executive component of working memory, as  
20  
21 measured by backward digit recall. If individual differences in central executive functions  
22  
23 are not predictive of speech perception in modulated noise, this may explain why some  
24  
25 previous studies have failed to identify a relationship between complex (visual) verbal  
26  
27 working memory measures involving the central executive, i.e., the reading span test, and  
28  
29 the perception of interrupted speech (e.g., Nagaraj & Knapp, 2015; Shafiro, Sheft, Risley,  
30  
31 & Gygi, 2015).  
32  
33  
34  
35  
36  
37  
38

## 39 **Summary**

40  
41 The present study provides support for differential roles of sub-components of  
42  
43 auditory verbal working memory in the abilities of listeners with normal hearing to use  
44  
45 the speech cues available during the temporal dips in modulated background noise.  
46  
47 Specifically, individual differences in the phonological component of working memory,  
48  
49 rather than the central executive component, were associated with speech perception in  
50  
51 modulated noise under adverse listening conditions.  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**ACKNOWLEDGMENTS**

This study was funded in part by the Wellcome Trust [ref: 105624] through the Centre for Chronic Diseases and Disorders (C2D2) at the University of York, UK.

For Peer Review

**REFERENCES**

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- Akeroyd, M. A.** (2008). Are individual differences in speech perception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing impaired adults. *International Journal of Audiology, 47*, 53-71.
- Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A-M.** (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology, 87*, 85-406.
- Avivi-Reich, M., Daneman, M, & Schneider, B. A.** (2014). How age and linguistic competence alter the interplay of perceptual and cognitive factors when listening to conversations in a noisy environment. *Frontiers in Systems Neuroscience, 8*, 21.
- Bacon, S. P., Opie, J. M., & Montoya, D. Y.** (1998). The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *Journal of Speech Language, and Hearing Research, 41*, 549-563.
- Baddeley, A., Logie, R., Nimmo-Smith, I., & Brereton, N.** (1985). Components of fluent reading. *Journal of Memory & Language, 24*, 119-131.
- Baddeley, A.** (1992). Working memory. *Science, 255*, 556-559.



- 1  
2  
3  
4  
5  
6 **Baldwin, C. L., & Ash, I. K.** (2011). Impact of sensory acuity on auditory working  
7  
8 memory span in young and older adults. *Psychology of Aging, 26*, 85-91.  
9  
10  
11  
12  
13 **Benard, M., Mensink, J., & Başkent, D.** (2014). Individual differences in top-down  
14  
15 restoration of interrupted speech: Links to linguistic and cognitive abilities.  
16  
17 *Journal of the Acoustical Society of America, 135*, EL88-EL94.  
18  
19  
20  
21  
22 **Bernstein, J. G. W., & Grant, K. W.** (2009). Auditory and auditory-visual  
23  
24 intelligibility of speech in fluctuating maskers for normal-hearing and hearing-  
25  
26 impaired listeners. *Journal of the Acoustical Society of America, 125*, 3358-3372.  
27  
28  
29  
30  
31  
32 **Bernstein, J. G. W., & Brungart, D. S.** (2011). Effects of spectral smearing and  
33  
34 temporal fine-structure distortion on the fluctuating-masker benefit for speech at a  
35  
36 fixed signal-to-noise ratio. *Journal of the Acoustical Society of America, 130*,  
37  
38 473-488.  
39  
40  
41  
42  
43  
44 **Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., & Festen, J. M.** (2013).  
45  
46 How linguistic closure and verbal working memory relate to speech recognition in  
47  
48 noise – A review. *Trends in Amplification, 17*, 75-93.  
49  
50  
51  
52  
53 **British Society of Audiology** (2004). *Recommended procedure: Pure tone air and bone*  
54  
55  
56  
57  
58  
59  
60

Millman and Mattys (2016)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

conduction thresholds audiometry with and without masking and determination of  
uncomfortable loudness levels. Reading, UK: British Society of Audiology.

**CHABA (Committee on Hearing, Bioacoustics, and Biomechanics).** (1988). Speech  
understanding and aging. *Journal of the Acoustical Society of America*, 83, 859-  
895.

**Daneman, M., & Carpenter, P. A.** (1980). Individual differences in working memory  
and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.

**Daneman, M., & Merikle, P. M.** (1996). Working memory and language  
comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3, 422-433.

**Desloge, J. G., Reed, C. M., Braida, L. D., Perez, Z. D., & Delhorne, L. A.** (2010).  
Speech reception by listeners with real and simulated hearing impairment: effects  
of continuous and interrupted noise. *Journal of the Acoustical Society of America*,  
128, 342-359.

**Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B.** (2002). Benefit of modulated maskers  
for speech recognition by younger and older adults with normal hearing. *Journal  
of the Acoustical Society of America*, 111, 2897-2907.

**Foo, C., Rudner, M., Rönnerberg, J., & Lunner, T.** (2007). Recognition of speech in

1  
2  
3 noise with new hearing instrument compression release settings requires explicit  
4  
5 cognitive storage and processing capacity. *Journal of the American Academy of*  
6  
7  
8 *Audiology*, 18, 553-566.  
9

10  
11  
12 **Füllgrabe, C., Moore, B. C. J., & Stone, M. A.** (2015). Age-group differences in  
13  
14 speech identification despite matched audiometrically normal hearing:  
15  
16 contributions from auditory temporal processing and cognition. *Frontiers in*  
17  
18 *Aging Neuroscience*, 6, 347.  
19

20  
21  
22 **Füllgrabe, C., & Rosen, S.** (2016). On the (un)importance of working memory in  
23  
24 speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers*  
25  
26 *in Psychology*, 7, 1268.  
27  
28  
29

30  
31  
32  
33 **Gatehouse, S., Naylor, G., & Eberling, C.** (2003). Benefits from hearing aids in  
34  
35 relation to the interaction between the user and the environment. *International*  
36  
37 *Journal of Audiology*, 42, S77-85.  
38  
39

40  
41  
42  
43 **Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D.** (1991). The influences of  
44  
45 number of syllables and wordlikeness on children's repetition of nonwords.  
46  
47  
48 *Applied Psycholinguistics*, 12, 349-367.  
49

50  
51  
52  
53 **Gathercole, S. E., & Adams, A.-M.** (1994). Children's phonological working memory:  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Contributions of long-term knowledge and rehearsal. *Journal of Memory &*  
4  
5  
6 *Language*, 33, 672-388.  
7  
8  
9

10 **Gathercole, S. E.** (1995). Is nonword repetition a test of phonological memory or lexical  
11  
12 knowledge? It all depends on the nonwords. *Memory & Cognition*, 23, 83-94.  
13  
14  
15

16  
17 **Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H.** (2004). The structure  
18  
19 of working memory from 4 to 15 years of age. *Developmental Psychology*. 40,  
20  
21 177-190.  
22  
23  
24

25  
26  
27 **George, E. L. J., Festen, J. M., & Houtgast, T.** (2006). Factors affecting masking  
28  
29 release for speech in modulated noise for normal-hearing and hearing impaired  
30  
31 listeners. *Journal of the Acoustical Society of America*, 120, 2295–2311.  
32  
33  
34

35  
36  
37 **Gregan, M. J., Nelson, P. B., & Oxenham, A. J.** (2013). Behavioral measures of  
38  
39 cochlear compression and temporal resolution as predictors of speech masking  
40  
41 release in hearing-impaired listeners. *Journal of the Acoustical Society of*  
42  
43 *America*, 134, 2895-2912.  
44  
45  
46

47  
48 **Gustafsson, H. Å., & Arlinger, S. D.** (1994). Masking of speech by amplitude-  
49  
50 modulated noise. *Journal of the Acoustical Society of America*, 95, 518-529.  
51  
52  
53

54  
55 **Hughes, M. L.** (2013). *Objective measures in cochlear implants*, 1<sup>st</sup> ed. (Plural  
56  
57  
58  
59  
60

1  
2  
3 Publishing, San Diego).  
4  
5  
6  
7

8 **Humes, L. E.** (2002). Factors underlying the speech-recognition performance of elderly  
9 hearing-aid wearers. *Journal of the Acoustical Society of America*, *112*, 1112-  
10 1132.  
11  
12  
13

14  
15  
16  
17 **Humes, L. E., Burk, M. H., Coughlin, M. P., Busey, T. A., & Strauser, L. E.** (2007).  
18 Auditory speech recognition and visual text recognition in younger and older  
19 adults: similarities and differences between modalities and the effects of  
20 presentation rate. *Journal of Speech, Language, and Hearing Research*, *50*, 283-  
21 303.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31

32 **Humes, L. E., Busey, T. A., Craig, J., & Kewley-Port, D.** (2015). Are age-related  
33 changes in cognitive function driven by age-related changes in sensory  
34 processing? *Attention Perception and Psychophysics*, *75*, 508-524.  
35  
36  
37  
38  
39  
40

41 **Koelewijn, T., Zekveld, A. A., Festen, J. M., Rönnberg, J., & Kramer, S. E.** (2012).  
42 Processing load induced by informational masking is related to linguistic abilities.  
43  
44  
45  
46 *International Journal of Otolaryngology*, *2012*, 1-11.  
47  
48  
49

50 **Lunner, T.** (2003). Cognitive function in relation to hearing aid use. *International*  
51 *Journal of Audiology*, *42*, S49-58.  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 **Lyxell, B., & Rönnberg, J.** (1989). Information-processing skills and speechreading.  
4  
5 *British Journal of Audiology*, 23, 339-347.  
6  
7

8  
9  
10 **Mattys, S. L., Davis, M. H., & Bradlow, A. R.** (2012). Speech recognition in adverse  
11  
12 conditions: A review. *Language and Cognitive Processes*, 27, 953-978.  
13  
14

15  
16  
17 **McCoy, S. L., Tun, P. A., Cox, C. L., Colangelo, M., Stewart, R. A., & Wingfield, A.**  
18  
19 (2005). Hearing loss and perceptual effort: Downstream effects on older adults'  
20  
21 memory for speech. *The Quarterly Journal of Experimental Psychology*, 58, 22–  
22  
23 33.  
24  
25

26  
27  
28  
29 **Miller, G. A., & Licklider, J. C. R.** (1950). The intelligibility of interrupted speech.  
30  
31 *Journal of the Acoustical Society of America*, 22, 167-173.  
32  
33

34  
35  
36 **Murphy, D. R., Craik, F. I. M., Li, K. & Schneider, B. A.** (2000). Comparing the  
37  
38 effects of aging and background noise on short-term memory performance.  
39  
40 *Psychology and Aging*, 14, 44-59.  
41  
42

43  
44  
45 **Nagaraj, N. K., & Knapp, A. N.** (2015). No evidence of relation between working  
46  
47 memory and perception of interrupted speech in young adults. *Journal of the*  
48  
49 *Acoustical Society of America*, 138, EL145-EL150.  
50  
51

52  
53  
54  
55 **Nelson, P. B., Jin, S. H., Carney, A. E., & Nelson, D. A.** (2003). Understanding speech  
56  
57  
58  
59  
60

1  
2  
3 in modulated interference: Cochlear implant users and normal-hearing listeners.

4  
5  
6 *Journal of the Acoustical Society of America*, 113, 961-968.

7  
8  
9  
10 **Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M.** (1995). How young and old  
11  
12 adults listen to and remember speech in noise. *Journal of the Acoustical Society of*  
13  
14  
15 *America*, 97, 593-608.

16  
17  
18  
19  
20 **Rönnberg, J.** (2003). Cognition in the hearing impaired and deaf as a bridge between  
21  
22 signal and dialogue: a framework and a model. *International Journal of*  
23  
24  
25 *Audiology*, 42, S68-S76.

26  
27  
28  
29  
30 **Rönnberg, J., Rudner, M., Lunner, T., & Zekveld, A. A.** (2010). When cognition  
31  
32 kicks in: Working memory and speech understanding in noise. *Noise & Health*,  
33  
34  
35 12, 263.

36  
37  
38  
39 **Rönnberg, J., Danielsson, H., Rudner, M., Arlinger, S., Sternäng, O., Wahlin, Å., &**  
40  
41 **Nilsson, L-G.** (2011). Hearing loss is negatively related to episodic and semantic  
42  
43 long-term memory but not to short-term memory. *Journal of Speech, Language,*  
44  
45  
46 *and Hearing Research*, 54, 705–726.

47  
48  
49  
50  
51 **Rönnberg, J., Lunner, T., Zekveld, A. A., Sörqvist, P., Danielsson, H., Lyxell, B.,**

1  
2  
3 **Dahlström, Ö, Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & Rudner, M.**  
4  
5 (2013). The Ease of Language Understanding (ELU) model: theoretical,  
6  
7 empirical, and clinical advances. *Frontiers in Systems Neuroscience*, 7, 31.  
8  
9

10  
11  
12 **Rothauser, E., Chapman, W., Guttman, N., Silbiger, H., Hecker, M., Urbanek, G.,**  
13  
14 **Nordby, K., & Weinstock, M.** (1969). IEEE recommended practice for speech  
15  
16 quality measurements. *IEEE Transactions in Acoustics*, 17, 225–246.  
17  
18

19  
20  
21  
22 **Rudner, M., Rönnberg, J., & Lunner, T.** (2011). Working memory supports listening  
23  
24 in noise for persons with hearing impairment. *Journal of the American Academy*  
25  
26 *of Audiology*, 22, 156-167.  
27  
28

29  
30  
31 **Salthouse, T. A.** (1994). The aging of working memory. *Neuropsychology*, 8, 535–543.  
32  
33

34  
35  
36 **Schneider, B. A., & Pichora-Fuller, M. K.** (2000). Implications of perceptual  
37  
38 deterioration for cognitive aging research,” *The Handbook of Aging and*  
39  
40 *Cognition*, 2<sup>nd</sup> ed. (Erlbaum, New York).  
41  
42

43  
44  
45 **Shafiro, V., Sheft, S., Risley, R., & Gygi, B.** (2015). Effects of age and hearing loss on  
46  
47 the intelligibility of interrupted speech. *Journal of the Acoustical Society of*  
48  
49 *America*, 137, 745-756.  
50  
51

52  
53  
54  
55 **Sörqvist, P., Ljungberg, J. K., & Ljung, R.** (2010). A sub-process view of working  
56  
57  
58  
59  
60



1  
2  
3 memory capacity: Evidence from effects of speech on prose memory. *Memory*,  
4  
5  
6 18, 310–326.  
7  
8  
9

10 **Sörqvist, P., & Rönnberg, J.** (2012). Episodic long-term memory of spoken discourse  
11  
12 masked by speech: What is the role for working memory capacity? *Journal of*  
13  
14 *Speech, Language, and Hearing Research*, 55, 210-218.  
15  
16  
17

18  
19  
20 **Smith, S. L., & Pichora-Fuller, M. K.** (2015). Associations between speech  
21  
22 understanding and auditory and visual tests of verbal working memory: Effects of  
23  
24 linguistic complexity, task, age, and hearing loss. *Frontiers in Psychology*, 6,  
25  
26 1394.  
27  
28  
29

30  
31  
32 **Studebaker, G. A.** (1985). A “rationalized” arcsine transform. *Journal of Speech and*  
33  
34 *Hearing Research*, 28, 455-462.  
35  
36  
37

38  
39 **Takahashi, G. A., & Bacon, S. P.** (1992). Modulation detection, modulation masking,  
40  
41 and speech understanding in noise in the elderly. *Journal of Speech and Hearing*  
42  
43 *Research*, 35, 1410-1421.  
44  
45  
46

47  
48 **Tun, P. A., McCoy, S., & Wingfield, A.** (2009). Aging, hearing acuity, and the  
49  
50 attentional costs of effortful listening. *Psychology of Aging*, 24, 761-766.  
51  
52  
53

54  
55 **Vickers, D. A., Moore, B. C. J., & Baer, T.** (2001). Effects of low-pass filtering on the  
56  
57  
58  
59  
60

1  
2  
3 intelligibility of speech in quiet for people with and without dead regions at high  
4  
5  
6 frequencies. *Journal of the Acoustical Society of America*, 110, 1164-1175.  
7  
8  
9

10 **Wingfield, A., Stine, E. L., Lahar, C. J., & Aberdeen, J. S. (1988).** Does the capacity  
11  
12 of working memory change with age? *Experimental Aging Research*, 14, 103–  
13  
14 107.  
15  
16

17  
18  
19 **Wingfield, A., Tun, P. A., & McCoy, S. L. (2005).** Hearing loss in older adulthood:  
20  
21 what it is and how it interacts with cognitive performance. *Current Directions in*  
22  
23 *Psychological Science*, 14, 144–148.  
24  
25  
26

27  
28  
29 **Zekveld, A. A., Rudner, M., Johnsrude, I. S., & Ronnberg, J. (2013a).** The effects of  
30  
31 working memory capacity and semantic cues on the intelligibility of speech in  
32  
33 noise. *Journal of the Acoustical Society of America*, 134, 2225-2234.  
34  
35  
36

37  
38  
39 **Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2013b).** Task difficulty differentially  
40  
41 affects two measures of processing load: The pupil response during sentence  
42  
43 processing and delayed cued recall of the sentences. *Journal of Speech,*  
44  
45 *Language, and Hearing Research*, 56, 1156-1165.  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**FIGURE CAPTIONS**

FIG. 1. Pure-tone air-conduction audiometry for the left ear (left panels) and right ear (right panels) of the listeners. The solid grey lines represent individual audiometric thresholds. The solid black lines represent mean audiometric thresholds for the listeners and error bars represent one standard error of the mean (SEM). The dashed dark grey line indicates the threshold of 20 dB HL that was considered the limit of normal hearing.

FIG. 2. Mean accuracy (%) for keyword identification of IEEE sentences as a function of SNR. Sentences were presented in an unmodulated noise masker (grey line) and a square-wave-gated noise masker (black line). Error bars show +/- one SEM.

FIG. 3. Mean AWMA raw scores. Memory tests were non-word repetition (NWR), forward digit recall (Digit) and backward digit recall (BackDigit). Error bars show +/- one SEM.

FIG. 4. Partial regression plots of the relationship between non-word repetition (NWR) and speech perception in modulated noise at -12 dB SNR (Speech Identification (RAU), -12 dB SNR), after controlling for the effects of Age and PTA.

TABLE I.

Linear model of predictors (Age and PTA) of sub-components of AVWM (NWR, Digit, and BackDigit). Model parameters shown include standardized beta coefficients ( $\beta$ ) and the significance value ( $p$ ). The  $R^2$  for the initial model step along with the change in  $R^2$  ( $\Delta R^2$ ) for the second step of the model are also reported. Values in bold indicate results at  $p \leq 0.05$ .

Model		NWR		Digit		Backdigit	
		$\beta$	$p$	$\beta$	$p$	$\beta$	$p$
Step 1	Age	<b>-0.5</b>	<b>0.005</b>	<b>-0.39</b>	<b>0.03</b>	-0.22	0.25
		$R^2 = 0.25, p = 0.005$		$R^2 = 0.15, p = 0.03$		$R^2 = 0.05, p = 0.25$	
Step 2		$\beta$	$p$	$\beta$	$p$	$\beta$	$p$
	Age	<b>-0.5</b>	<b>0.004</b>	<b>-0.39</b>	<b>0.03</b>	-0.21	0.24
	PTA	-0.23	0.16	0.001	0.99	<b>-0.37</b>	<b>0.04</b>
		$\Delta R^2 = 0.05, p = 0.16$		$\Delta R^2 = 0, p = 0.99$		$\Delta R^2 = 0.14, p = 0.04$	

TABLE 2.

Linear model of predictors of speech MR for an SNR of -4 dB for NH listeners. See Table 1 legend for details. Values in bold indicate results at  $p \leq 0.05$ .

speech MR (-4 dB SNR)			
Model		$\beta$	$p$
Step 1	Age	-0.12	0.50
	PTA	-0.31	0.10
$R^2 = 0.12$ for Step 1 ( $p = 0.19$ ).			
Excluded	NWR	-0.16	0.46
	Digit	0.2	0.33
	BackDigit	-0.1	0.64

**TABLE 3.**

Linear model of predictors of speech perception in a modulated masker for an SNR of -4 dB.

See Table 1 legend for details. Values in bold indicate results at  $p \leq 0.05$ .

---

---

speech perception in modulated masker (-4 dB SNR)			
Model		$\beta$	$p$
Step 1	<b>Age</b>	<b>-0.38</b>	<b>0.04</b>
	PTA	-0.24	0.17
Excluded	NWR	0.32	0.13
	Digit	0.16	0.40
	BackDigit	0.05	0.82

**$R^2 = 0.20$  for Step 1 ( $p = 0.046$ ).**

---

---

For Peer Review

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**TABLE 4.**

Linear model of predictors of speech perception in a modulated masker for an SNR of -8 dB.

See Table 1 legend for details. Values in bold indicate results at  $p \leq 0.05$ .

---



---

speech perception in modulated masker  
(-8 dB SNR)

Model		$\beta$	$p$
Step 1	Age	-0.2	0.29
	PTA	-0.15	0.43
Excluded	NWR	0.37	0.10
	Digit	0.15	0.47
	BackDigit	0.19	0.36

$R^2 = 0.06$  for Step 1 ( $p = 0.41$ ).

---



---

For Peer Review

TABLE 5.

Linear model of predictors of speech perception in a modulated masker for an SNR of -12 dB. See Table 1 legend for details. Values in bold indicate results at  $p \leq 0.05$ .

speech perception in modulated masker (-12 dB SNR)			
Model		$\beta$	$p$
Step 1	Age	-0.27	0.12
	<b>PTA</b>	<b>-0.39</b>	<b>0.03</b>
Excluded	NWR	0.4	0.048
	Digit	0.06	0.75
	BackDigit	-0.14	0.94
<b><math>R^2 = 0.23</math> for Step 1 (<math>p = 0.03</math>).</b>			
Step 2	Age	-0.07	0.71
	PTA	-0.3	0.08
	<b>NWR</b>	<b>0.4</b>	<b>0.048</b>
Excluded	Digit	-0.21	0.33
	BackDigit	-0.13	0.48
<b><math>\Delta R^2 = 0.11</math> for Step 1 (<math>p = 0.048</math>).</b>			



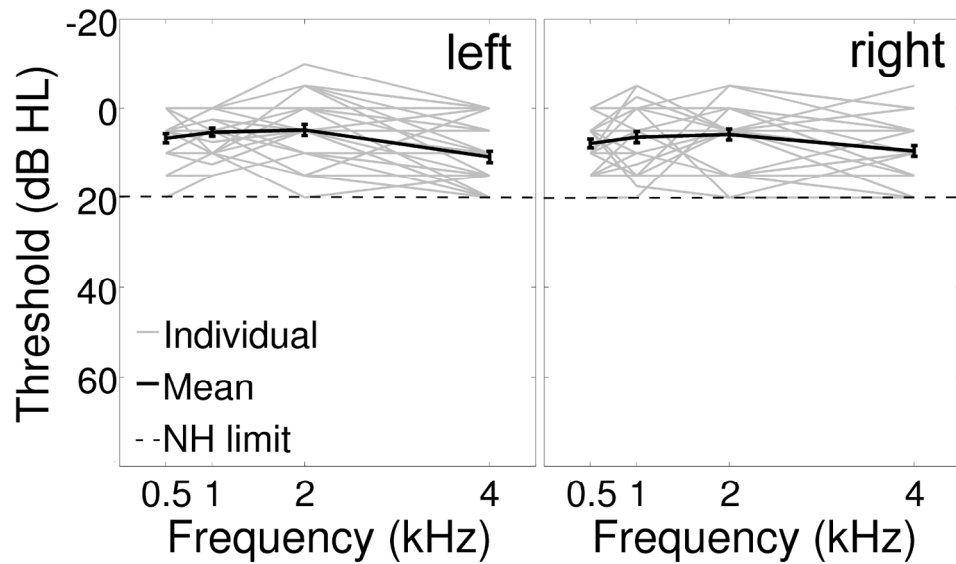


FIG. 1. Pure-tone air-conduction audiometry for the left ear (left panels) and right ear (right panels) of the listeners. The solid grey lines represent individual audiometric thresholds. The solid black lines represent mean audiometric thresholds for the listeners and error bars represent one standard error of the mean (SEM). The dashed dark grey line indicates the threshold of 20 dB HL that was considered the limit of normal hearing.

152x92mm (300 x 300 DPI)

Review

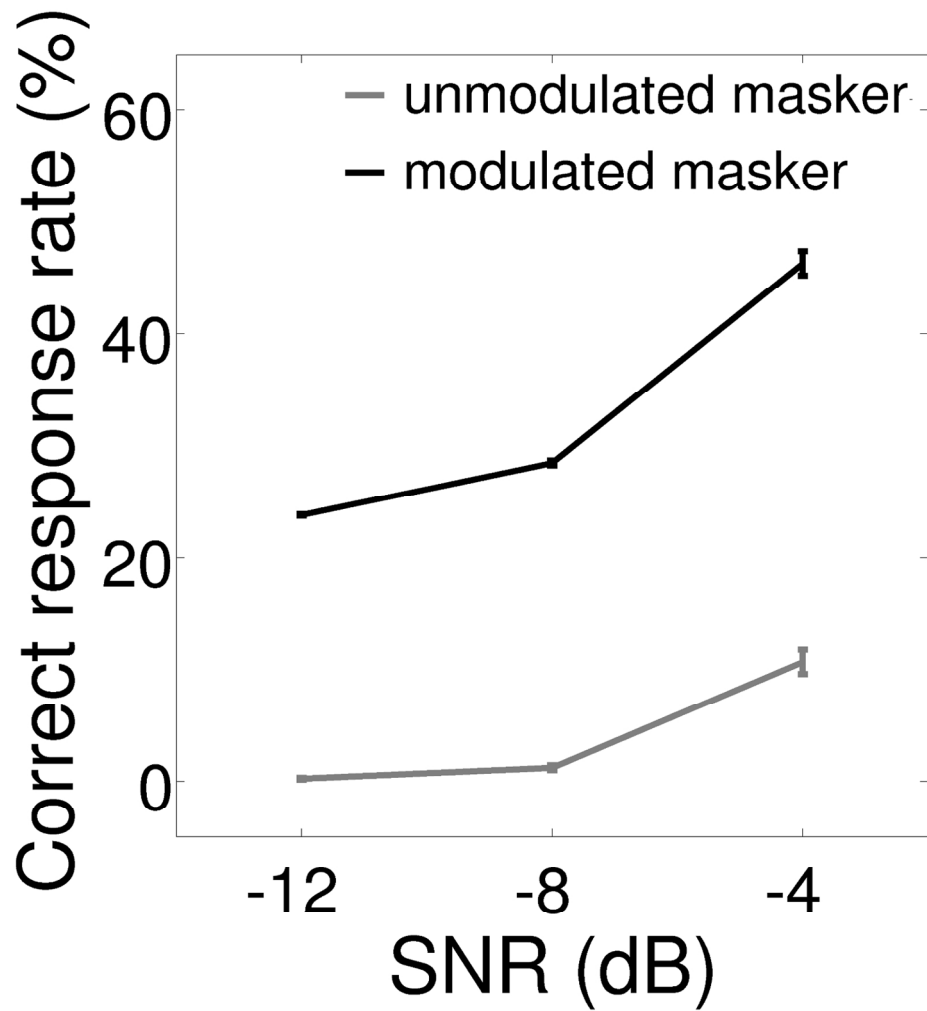


FIG. 2. Mean accuracy (%) for keyword identification of IEEE sentences as a function of SNR. Sentences were presented in an unmodulated noise masker (grey line) and a square-wave-gated noise masker (black line). Error bars show +/- one SEM.

146x152mm (300 x 300 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

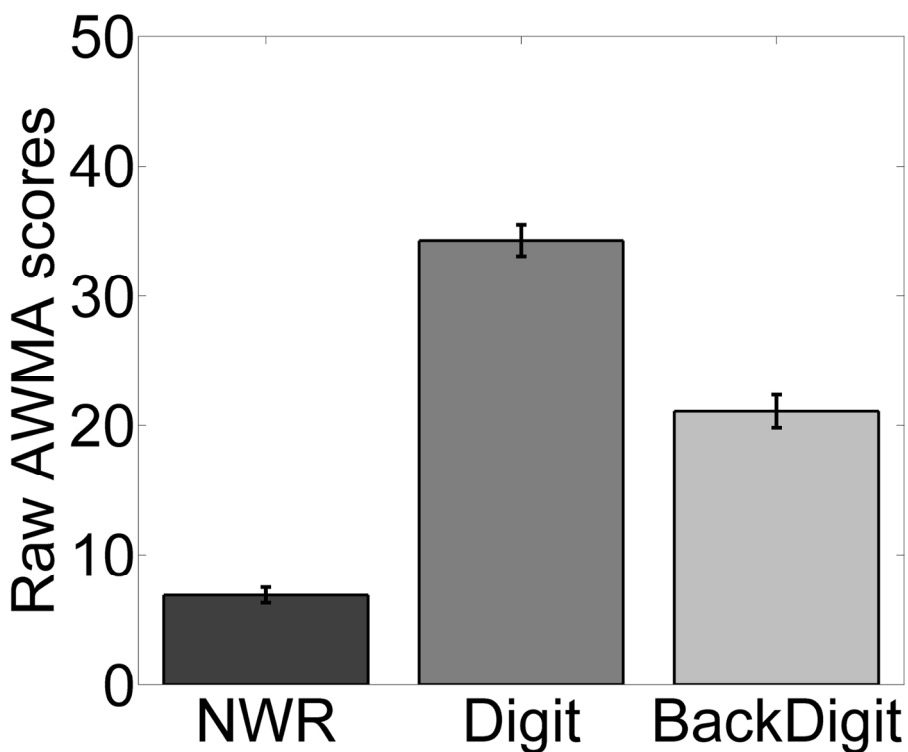


FIG. 3. Mean AWMA raw scores. Memory tests were non-word repetition (NWR), forward digit recall (Digit) and backward digit recall (BackDigit). Error bars show +/- one SEM.

152x121mm (300 x 300 DPI)

view

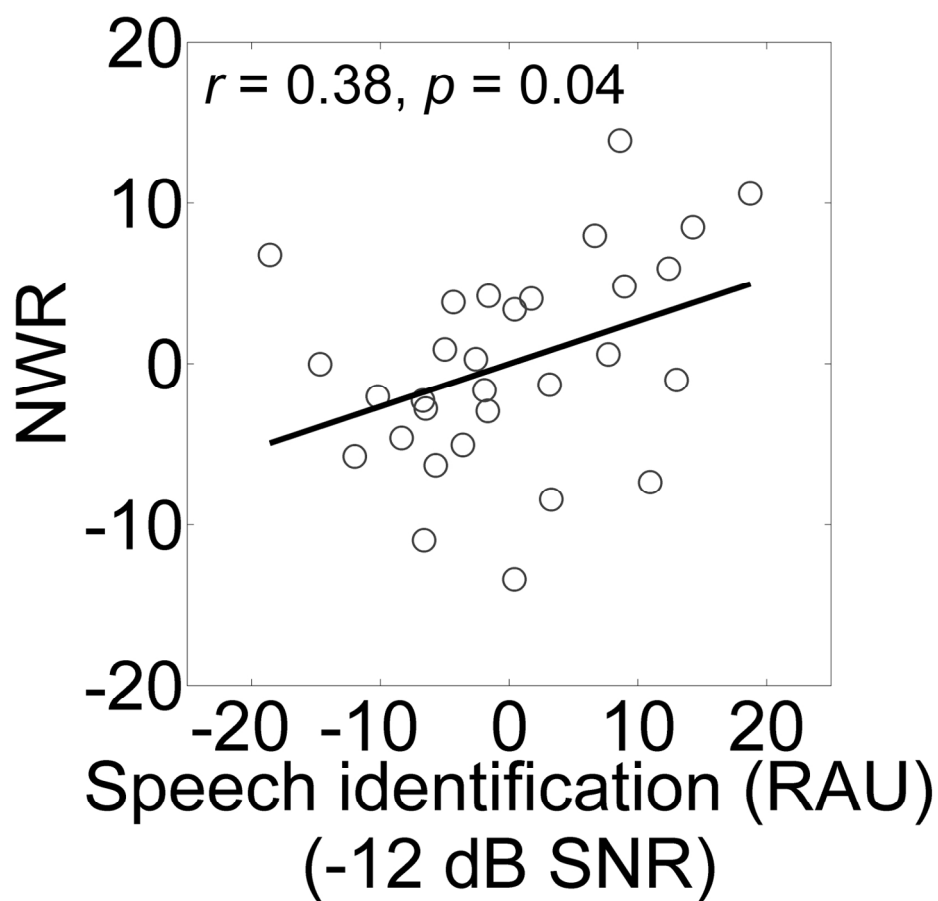


FIG. 4. Partial regression plots of the relationship between non-word repetition (NWR) and speech perception in modulated noise at -12 dB SNR (Speech Identification (RAU), -12 dB SNR), after controlling for the effects of Age and PTA.

152x141mm (300 x 300 DPI)