How Linguistic Closure and Verbal Working Memory Relate to Speech Recognition in Noise—A Review

Jana Besser, MSc[1], Thomas Koelewijn, PhD[1], Adriana A. Zekveld, PhD[1,2,3], Sophia E. Kramer, PhD[1], and Joost M. Festen, PhD[1]

Abstract

The ability to recognize masked speech, commonly measured with a speech reception threshold (SRT) test, is associated with cognitive processing abilities. Two cognitive factors frequently assessed in speech recognition research are the capacity of working memory (WM), measured by means of a reading span (Rspan) or listening span (Lspan) test, and the ability to read masked text (linguistic closure), measured by the text reception threshold (TRT). The current article provides a review of recent hearing research that examined the relationship of TRT and WM span to SRTs in various maskers. Furthermore, modality differences in WM capacity assessed with the Rspan compared to the Lspan test were examined and related to speech recognition abilities in an experimental study with young adults with normal hearing (NH). Span scores were strongly associated with each other, but were higher in the auditory modality. The results of the reviewed studies suggest that TRT and WM span are related to each other, but differ in their relationships with SRT performance. In NH adults of middle age or older, both TRT and Rspan were associated with SRTs in speech maskers, whereas TRT better predicted speech recognition in fluctuating nonspeech maskers. The associations with SRTs in steady-state noise were inconclusive for both measures. WM span was positively related to benefit from contextual information in speech recognition, but better TRTs related to less interference from unrelated cues. Data for individuals with impaired hearing are limited, but larger WM span seems to give a general advantage in various listening situations.

Speech understanding is influenced by both auditory and cognitive capabilities (Akeroyd, 2008; Houtgast & Festen, 2008; Wingfield & Tun, 2007). Experimental research found that cognitive abilities, such as explicit processing of signal semantics and context or the suppression of irrelevant information, are crucial especially in adverse listening conditions (Schneider, Daneman, & Pichora-Fuller, 2002). People with a larger cognitive capacity are also better in adapting to new listening situations like unfamiliar hearing aid settings (Foo, Rudner, Rönnberg, & Lunner, 2007; Lunner, 2003). These empirical findings have been theoretically discussed and modeled in the Ease of Language Understanding (ELU) framework describing the involvement of active cognitive processing during speech recognition in situations where automatic processing is insufficient to decode the signal (Rönnberg, 2003; Rönnberg, Rudner, Foo, & Lunner, 2008; Stenfelt & Rönnberg, 2009). The growing body of empirical and theoretical insights feed the relatively new field of cognitive hearing science (Arlinger, Lunner, Lyxell, & Pichora-Fuller, 2009). In this field, researchers are striving to gain detailed insights into the complex interactions between cognitive and auditory mechanisms during speech recognition.

Additionally, the field aims to develop new clinical tests to complement existing audiologic diagnostics and rehabilitation strategies.

A review by Akeroyd (2008), including 20 experimental studies, investigated the relationship between various cognitive abilities and speech reception in noise. This review identified verbal working memory (WM) capacity as measured by the reading span (Rspan) test as one of the factors most powerful in predicting the recognition of masked speech. Furthermore, the survey cited two studies that had found strong associations between the ability to recognize partially masked text, measured by the then new text reception threshold (TRT) test, and speech recognition in noise for

[1]VU University Medical Center, Amsterdam, Netherlands
[2]The Swedish Institute for Disability Research, Sweden
[3]Linköping University, Linköping, Sweden

Corresponding Author:
Jana Besser, MSc, Department of ENT/Audiology, EMGO+ Institute, VU University Medical Center, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands.
Email: j.besser@vumc.nl

normal-hearing listeners (George et al., 2007; Zekveld, George, Kramer, Goverts, & Houtgast, 2007). Both the Rspan and the TRT test have been used increasingly in hearing research in recent years, and for both tests associations with speech recognition in adverse conditions have been confirmed repeatedly. There are indications that the TRT depends on cognitive aspects of speech recognition that differ from those measured by the Rspan (Besser, Zekveld, Kramer, Rönnberg, & Festen, 2012). To promote a better understanding of the specific abilities tapped into by these tests and their predictive values for speech recognition in noise, the current article gives an overview of the findings from experimental studies that have examined the relationship of speech recognition in noise with the TRT, the Rspan, or its auditory counterpart, the listening span (Lspan). For the studies including Rspan or Lspan tests, the selection was limited to articles published after the survey by Akeroyd (2008). Throughout the current article, we refer to speech perception abilities with "speech recognition" rather than terms suggesting a deeper level of semantic processing, such as "speech understanding" or "speech comprehension." In most of the reviewed studies speech recognition was measured with tasks of repeating the presented speech material, which does not by definition require understanding of the semantic meaning. Nonetheless, it should be noted that the many associations between speech recognition and cognitive measures observed in the discussed research suggest that listeners employed strategies of active listening and understanding.

The current article additionally presents a research study examining differences and similarities of verbal WM capacity when assessed with the Rspan compared to the Lspan test. The study was motivated by earlier findings of modality differences in verbal WM. Intramodality correlations between verbal tasks seem to be stronger than intermodality correlations (Humes, Burk, Coughlin, Busey, & Strausner, 2007). Accordingly, it is highly relevant to study modality-specific WM functions in hearing research and to further seek evidence for the findings reported so far. In the current study, we examined test-retest reliability of both Lspan and Rspan tests and modality differences in test performance in a population of young adults with normal hearing. Furthermore, their associations with the recognition of masked text and speech were examined, that is, associations within and across test modalities.

Verbal Working Memory Capacity

WM is a cognitive component managing the temporary storage and real-time processing of information (Baddeley, 1992; Oberauer, Suss, Wilhelm, & Wittman, 2003). WM is involved in tasks that require higher cognitive processing, such as reasoning, comprehension, and problem solving (Engle, 2002). WM capacity is commonly assessed with so called *span* tests (Bopp & Verhaeghen, 2005; Daneman &

Carpenter, 1980). An essential feature of WM-span tests is that they tap into *storage* and *manipulation* of incoming information simultaneously by means of a dual task (Shelton, Elliott, Hill, Calamia, & Gouvier, 2009). They are therefore also referred to as complex span tests (e.g., Engle, 2010; Vergauwe, Barrouillet, & Camos, 2010). A widely used span test is the Rspan test, originally developed by Daneman and Carpenter (1980). In this test, participants read short lists of sentences, judge the veracity of each sentence immediately after reading it, and recall each sentence's last word at the end of the sentence list. In many subsequent studies, adaptations to the original version of the Rspan test were applied (e.g., Baddeley, Logie, Nimmosmith, & Brereton, 1985; La Pointe & Engle, 1990; Rönnberg, Arlinger, Lyxell, & Kinnefors, 1989; Turner & Engle, 1989; Waters & Caplan, 1996).

In hearing research, the Rspan test version by Baddeley et al. (1985) is frequently used. It consists of short 5-word sentences with a subject-verb-object (SVO) syntax, half of which are semantically incorrect. The test has been transferred to other languages, that is, Swedish (Rönnberg, 1990) and Danish (Rudner, Foo, Sundewall-Thoren, Lunner, & Rönnberg, 2008). The Swedish and Danish test versions do not use SVO sentences, but sentences that consist of three words each. Both in the version by Baddeley et al. and in the Scandinavian versions, sentences are presented in a three-step fashion, displaying one sentence part at a time. The secondary task is judging the semantic correctness of each sentence directly after presentation. Furthermore, the words to-be-recalled can be either sentence-initial or sentence-final, which increases the memory load during test completion. Table 1 provides an overview of research examining the relationship of verbal WM capacity and masked speech recognition.

Verbal WM Span and Speech Recognition in Noise

Several of the studies listed in Table 1 investigated the association between verbal WM capacity and speech recognition with interfering talkers. Individuals with higher verbal WM capacity tend to perform better in such conditions than individuals with low WM capacity. This finding is consistent across different SRT test performance levels, with different numbers of interfering talkers (1, 2, or multiple), for people with and without hearing impairment, and whether complete sentences or single words from high- or low-context sentences should be repeated (Arehart et al., 2013; Desjardins & Doherty, 2013; Ellis & Munro, 2013; Koelewijn et al., 2012). An exception might be young individuals with normal hearing, for which Zekveld, Festen, et al. (2013) did not find an association of Rspan or Lspan with speech recognition with an interfering talker, neither at low nor high speech intelligibility levels. For listening conditions with a spatial separation of target and masking speech, there are inconsistent findings regarding the role of Rspan (Neher et al., 2009, 2011). For

Table 1. Hearing Research Including an Rspan and/or Lspan Test.

Study	WM span test & material	Speech recognition tests	Other tests and measures	Association span— speech recognition	Findings related to span test
Arehart, Souza, Baca, and Kates (2013) N = 26; 62-92 years; HI	Rspan Outcome: % correct targets; total number of targets not reported	SRT_{QUIET}; SRT_{TALK}-(8-talker babble) at 5 SNRs (−10 – 10 dB); low-context IEEE sentences; aided testing: 9 frequency-compression settings; 1 control condition; keyword recognition		Not reported	Group division into low- and high-span by median split; controlling for age and hearing loss, better baseline intelligibility for high-span group; high-span group less sensitive to signal distortion
Baldwin and Ash (2011)[a] N = 80; 18-31 (20.6) years; NH N = 26; 60-80 (68.6) years; NH	Lspan, Rspan random set size; outcome: mean proportion of correct words across sets; Lspan at 5 presentation levels (45-65 dB A)	SRT_{QUIET}; at 7 presentation levels; two-syllable spondee words	TMT-B (Trail Making Test; perceptual speed)	Old: Lspan − SRT_{QUIET} (β = .51** − .75**)[b]; young: Lspan − SRT_{QUIET} (β = −.51 − .43**); strength depending on presentation level; Rspan − SRT_{QUIET} not reported	Lspan old group more affected by presentation level; young and old groups same Rspan performance; SRT_{QUIET} predicted Lspan performance at all presentation levels for old
Besser et al. (current article) N = 55; 18-78 (44.0) years; NH	Rspan 5-word SVO sentences	SRT_{STEADY}, SRT_{FLUCT}; sentences: Versfeld, Daalder, Festen, and Houtgast (2000)	5 TRT versions; LDST (processing speed); self-report how pleasant or taxing TRT versions were	Rspan − SRT_{FLUCT} (r = −.42**), controlling for age (r = −.14)	Rspan M = 19.7; Rspan associated with processing speed and TRT versions, also when controlling for age
Besser et al. (2012) N = 42; 19-35 (24.4) years; NH	Rspan, Lspan 5-word SVO sentences (SVO)	SRT_{FLUCT}; sentences: Versfeld et al. (2000)	TRT_{500}	Rspan − SRT_{FLUCT} ns; Lspan − SRT_{FLUCT} ns	Rspan M = 22; Lspan M = 26.1; Rspan and Lspan associated; Lspan and TRT associated
Desjardins and Doherty (2013) N = 15; 18-25 (21.7) years; YNH N = 15; 55-77 (66.9) years; ONH N = 16; 59-76 (68.2) years; OHI	Rspan 5-word SVO sentences; outcome: % correct targets	SRT_{STEADY}, SRT_{TALK} (two- or six-talker); high- and low-context sentences; outcome: % correct sentence-final words; 75% correct level; aided testing for OHI	Visual motor tracking (DPRT); perceived ease of listening during R-SPIN; DSST (Digit-symbol substitution test; processing speed); Stroop test (selective attention)	Rspan − SRT_{STEADY} (r = −.45**); Rspan − SRT six-talker (r = −.52**); Rspan − SRT two-talker (r = −.55**); calculated for complete study sample	Rspan M = 54.8% (YNH), M = 49.4% (ONH), M = 41.8% (OHI); group differences Rspan YNH vs. OHI, but not others; Rspan associated with perceived effort in SRT_{STEADY} and two-talker
Ellis and Munro (2013) N = 15; 18-50 (26.3) years; NH (self-report)	Rspan outcome: % correct targets	SRT_{TALK} (multitalker babble); IEEE sentences (Rothauser et al., 1969); SNRs +12 − −15 dB (3-dB steps); 3 signal-processing conditions; outcome: % correct keywords across SNRs	Trail making test (TMT A and B; perceptual speed)	Rspan − SRT unprocessed (r = .45*); Rspan − SRT frequency compressed 1 (r = .2); Rspan − SRT frequency compressed 2 (r = .1); all one-tailed	Rspan M = 47.4; Rspan score and Rspan errors associated with SRT for unprocessed speech
Kjellberg and Ljung (2008)[c] N = 32; 18-34 years; NH (self-report)	Rspan 3-word sentences	SRT_{FLUCT}, words and sentences (Hagerman, 1982); SNR + 4 dB (noise condition) and +27 dB (control condition)	Recall of words presented in noise or control condition; recognition of sentences from speech in noise test; perceived effort during SRT	Not reported	Sentence recognition not associated with Rspan; Rspan associated with word recall and less reduction in recall by noise
Koelewijn, Zekveld, Festen, Rönnberg, and Kramer (2012) N = 32; 40-70 (51.3) years; NH	Rspan, Lspan 5-word SVO sentences	SRT_{FLUCT}, SRT_{TALK}; 50% and 84% correct level; sentences: Versfeld et al. (2000)	TRT_{500}; size-comparison span (SICspan); pupillometry during SRTs; subjective ratings of effort, performance, and motivation level for SRTs	Rspan − SRT_{TALK} at 50% correct level (r = −.50**); Rspan − SRT_{TALK} at 84% correct level (r = −.46**); Rspan − SRT_{FLUCT} at 84% correct level (r = −.36*)	Rspan M = 15.5, Lspan M = 21.4; Rspan and Lspan associated mutually and with TRT; Rspan associated with SICspan; Rspan predictor of SRT_{TALK}

Table 1. (continued)

Study	WM span test & material	Speech recognition tests	Other tests and measures	Association span—speech recognition	Findings related to span test
Ljung and Kjellberg (2009)[c] $N = 32$; 18-35 years; NH (self-report)	Rspan 3-word sentences	words and sentences in noise (Hagerman, 1982); SNR +15 dB; 2 virtual rooms with short/long reverberation	perceived effort during speech recognition tests; recall of words presented in noise or control condition; recognition of sentences from speech in noise test	No SRT assessed	Rspan not associated with recall or recall error measures
Neher et al. (2009) $N = 20$; 28-84 (60.0) years; HI	Rspan outcome: total correct; no further description available	SRT_{STEADY}, SRT_{TALK} (two-talker) at fixed SNR (50% correct level); aided testing; 3 spatial conditions for masker and speech (colocated, 180° front-back, 50° left-right); Dantale II sentences (Wagener, Josvassen, & Ardenkjaer, 2003)	Test of everyday attention (TEA; all subtests except 2 and 8)	Rspan − SRT_{STEADY} ($r = -.58^{**}$); Rspan − SRT front-back ($r = -.72^{**}$), persisted when controlling for age; Rspan − SRT left-right ($r = -.52^*$), ns when controlling for age	Rspan $M = 23.6$; Rspan predicted SRT front-back better than other included predictor (aided hearing acuity at high frequencies)
Neher, Laugesen, Jensen, and Kragelund (2011) $N = 23$; 60-78 (67.0) years; HI	Rspan 3-word sentences	SRT_{TALK} (two-talker) at fixed SNR (50% correct level); unaided testing with amplification; 3 spatial conditions for masker and speech (colocated, 180° front-back, 45° left-right); Dantale II sentences (Wagener et al., 2003)	spectral ripple discrimination; frequency range for interaural phase difference detection; visual elevator test (attention switching); sound localization	Rspan − SRT front-back ($r = -.08$); others also ns, no details reported	Rspan $M = 22.6$; Rspan not associated with any other measure (auditory, age, or elevator test)
Ng, Rudner, Lunner, Pedersen, and Rönnberg (2013)[d] $N = 26$; 32-65 (59.0) years; HI	Rspan 3-word sentences, 24 in total; outcome: total correct	Sentence-final word identification and recall; quiet, steady-state noise (SSN), four-talker (4T) masker; aided testing; Swedish HINT (Hällgren, Larsby, & Arlinger, 2006); fixed SNR (84% correct level); aided; 3 noise reduction (NR) conditions		not reported	Rspan $M = 10.4$; Rspan associated with recall of words in quiet, SSN with/without NR, 4T with NR ($r = .47 - .58^{**}$); Rspan associated with recall of words in primacy and asymptote, but not recency positions; high-span group more affected by noise, but benefited more from NR; no association with age or PTA
Rudner, Foo, Rönnberg, and Lunner (2009) $N = 32$; 51-80 (70.0) years; HI	Rspan 3-word sentences	SRT_{STEADY}, SRT_{FLUCT}; sentences: Hagerman and Kinnefors (1995) and Swedish HINT, Hällgren et al. (2006); 50% and 80% correct level, only outcomes at 80% used for analyses; aided testing; fast/slow compression, accustomed and unaccustomed	letter monitoring (LM)	Rspan − SRT_{STEADY} ($r -.31 - -.64^{**}$); Rspan − SRT_{FLUCT} ($r -.39 - -.80^*$); strength depending on speech material and compression setting	Rspan $M = 23.9$; Rspan associated with age, hearing acuity, and LM; associated with and best predictor of most Hagerman measures when test and experience setting were incongruent; association with HINT outcomes at slow-compression rates
Rudner, Rönnberg, and Lunner (2011) $N = 30$; (70.0) years; HI	Rspan 3-word sentences	SRT_{STEADY}, SRT_{FLUCT}; sentences: Hagerman and Kinnefors (1995) and Swedish HINT (Hällgren et al., 2006); 50% and 80% correct level; aided and unaided testing; fast/slow compression; before and after experience with compression setting; analyses on postexperience data only	letter monitoring (LM)	Rspan − SRT_{STEADY} at 50% correct ($r -.37^* - -.60^*$; Rspan − SRT_{STEADY} at 80% correct ($r -.51^{**} - -.53^{**}$); Rspan − SRT_{FLUCT} at 50% correct ($r -.38^* - -.53^*$); Rspan − SRT_{FLUCT} at 80% correct ($r -.31 - -.52^{**}$); strength depending on speech material, and compression setting	Rspan $M = 23.9$; Rspan associated with most measures of speech recognition; group division into low ($M = 18.7$) and high ($M = 28.4$) Rspan; high-span mostly better SRT results than low-span group; correlations similar for aided and unaided testing

Table 1. (continued)

Study	WM span test & material	Speech recognition tests	Other tests and measures	Association span—speech recognition	Findings related to span test
Rudner, Lunner, Behrens, Thoren, and Rönnberg (2012), Experiment 2[e] N = 30; 51-80 (70.0) years; HI	Rspan 3-word sentences	SRT$_{STEADY}$, SRT$_{FLUCT}$; aided testing; sentences: Hagerman and Kinnefors (1995); 50% and 80% correct level	perceived listening effort in both noises at SNRs −2, +4, +10 with speech fixed at 65 dB SPL	not reported	Rspan M = 23.9; controlling for WM capacity, lower SNRs give higher effort; Rspan associated with lower effort
Sörqvist and Rönnberg (2012)[f] N = 44; (25.9) years; NH	Rspan 3-word sentences	speech recognition with normal and rotated speech masker; SNR +5 dB; task: select stimulus from 4 visually presented options	episodic long-term memory (LTM) of spoken discourse; SICspan	no SRT assessed	Rspan M = 17.7; Rspan associated with SICspan; Rspan predictor of LTM in normal speech, controlling for LTM in rotated speech
Zekveld, Rudner, et al. (2011), Experiment 2 N = 20; 18-32 (22.0) years; NH	Rspan 5-word SVO sentences	SRT$_{STEADY}$; 29% and 16% correct level; with/without semantically (un)related word cues	TRT$_{ORIGINAL}$	Rspan – SRT 29% correct with unrelated cues (r = −.50*); other associations ns, no details reported	Rspan M = 26.1
Zekveld, Rudner, Johnsrude, Heslenfeld, and Rönnberg (2012) N = 18; 20-29 (23.6) years; NH	Rspan 5-word SVO sentences	SRT$_{STEADY}$; sentences: Versfeld et al. (2000); 29% correct level; SRTs with/without semantically (un)related word cues during fMRI scanning at estimated 29% correct level	TRT$_{ORIGINAL}$ fMRI scans during SRT tests	ns, no details reported	Rspan M = 25.3; Rspan associated with SRT benefit from related cues, but not with word recognition score; low-span group higher brain activation during SRT test; Rspan not associated with TRT
Zekveld, Festen, and Kramer (2013) N = 24; 18-27 (22.0) years; NH	Rspan, Lspan 5-word SVO sentences	SRT$_{TALK}$; sentences: Versfeld et al. (2000); 29% and 71% correct level	TRT$_{CENTER}$ (29% and 71% performance); pupillometry; surprise cued recall test	ns, no details reported	Rspan M = 22.0, Lspan M = 26.4; no significant associations between tests
Zekveld, Rudner, Johnsrude, and Rönnberg (2013) N = 18; 20-32 (23.0) years; NH	Rspan 5-word SVO sentences	SRT$_{STEADY}$, SRT$_{FLUCT}$, SRT$_{TALK}$; sentences: Versfeld et al. (2000); with semantically related or nonword cues; 29% and 71% correct level	TRT$_{ORIGINAL}$; SICspan; 2-alternative-forced-choice recognition of presented SRT sentences	not reported	Rspan Median = 21; Rspan associated with SRT$_{TALK}$ benefit from related cues; higher benefit of related cues in sentence recognition for high-span people
Zekveld and Kramer (2013), Experiment 1[g] N = 24; 18-26 (22.0) years; NH	Lspan 5-word SVO sentences	SRT$_{TALK}$, nonadaptive (1%, 50%, 99% correct level), and speech in quiet (no threshold); sentences: Versfeld et al. (2000)	TRT$_{ORIGINAL}$; pupillometry; cued recall test of TRT and SRT stimuli; self-rated listening effort and SRT performance	not reported	Lspan M = 25.3; Lspan associated with shorter peak latency for speech in quiet

Note. If not otherwise specified, the following test properties applied: Rspan and Lspan tests included 54 sentences, presented with increasing set-sizes, outcome was the total number of correctly recalled targets, targets were sentence-initial or sentence-final words, semantic judgments of sentences requested. SRTs and TRTs were assessed for complete sentences at a performance level of 50% correct responses. TRTs assessed with sentences by Versfeld et al. (2000), TRT versions described in Section "Linguistic processing ability." Abbreviations: SRT$_{STEADY}$ = SRT in a steady-state masker; SRT$_{FLUCT}$ = SRT in a fluctuating masker; SRT$_{TALK}$ = SRT with one or several interfering talkers; SRT$_{QUIET}$ = speech reception without a masker; M = mean score; SVO = subject-verb-object sentence structure, ns = nonsignificant. Associations marked * significant at the .05 level and marked ** significant at the .01 level.
[a]Rather than speech recognition in noise, this study examined the effect of sound level on speech recognition in quiet and influences of different factors on Lspan performance. [b]SRT results in this study were recoded, such that positive associations represent better performance on both measures. [c]The study examined abilities to remember speech perceived with a background masker rather than SRTs in noise. [d]The study assessed associations of Rspan with recall of words presented with a masker rather than with masked word recognition.[e]The study examined perceived effort during the recognition of masked speech rather than speech recognition performance as such. [f]The study focused on mutual associations of measures of WM capacity and their relationship to the ability to remember speech perceived in noise. [g]The study focused on associations between cognitive measures and pupil dilation during SRT tests, rather than on speech recognition performance as such.

people with a hearing impairment, it was found that high-span individuals have better speech recognition in multitalker maskers than low-span individuals when frequency compression is added. They are also less severely affected by increased levels of distortion due to larger signal modification (Arehart et al., 2013).

Sentence recognition in steady-state noise was associated with Rspan scores in all studies including participants with impaired hearing (Neher et al., 2009; Rudner et al., 2009, 2011). For NH participants, associations were only found at low intelligibility levels in situations including a clear mismatch between contextual cues and the stimulus content (Zekveld, Rudner, et al., 2011). In the one study looking at word recognition rather than sentence recognition in steady-state noise, both HI and NH participants with higher Rspan scores had better word recognition (Desjardins & Doherty, 2013).

The study results presented in Table 1 suggest that there are associations between speech recognition in fluctuating noise and Rspan performance in groups including adults with normal hearing of all ages (Besser et al., 2012; Koelewijn et al., 2012). However, in both studies mentioned, the association was governed by age, and Rspan performance was not a significant predictor of the SRT in regression models controlling for age. For HI participants, the hearing aid compression settings in combination with the sentence material used in the speech recognition task seem to be a determining factor for the role of verbal WM capacity in speech recognition with fluctuating maskers. In a study with older adults with impaired hearing, Rudner et al. (2009) observed that while the mismatch between accustomed compression setting and the setting during testing was the most influential parameter on the association with Rspan for one test corpus, the compression setting itself was decisive for a different corpus. In another study (Rudner et al., 2011), where there were no mismatch conditions because participants were accustomed to both fast and slow acting compression, speech recognition in fluctuating noise was related to Rspan performance independent of the compression setting used during testing and of the test corpus. The higher number of SRT conditions related to Rspan in Rudner et al. (2011), compared to Rudner et al. (2009), could also be related to the performance level in the speech recognition tasks for which the analyses were performed. Rudner et al. (2009) analyzed associations with Rspan for SRTs with 80% correct responses, whereas Rudner et al. (2011) examined relationships at 50% and 80% correct levels and found stronger associations for the former.

Verbal WM Span in Relation to Other Measures

Not only the ability to recognize masked speech but also the perceived effort during speech recognition appears to be related to Rspan performance. People with higher span scores rated word or sentence recognition in various

maskers as less effortful than people with a smaller Rspan (Desjardins & Doherty, 2013; Rudner et al., 2012). In contrast, Rspan was not related to perceived effort that changed with the SNR used during speech recognition testing (Rudner et al., 2012).

The benefit in speech recognition that people received from semantically related word cues, that is, context information, presented along with the masked speech signal, was also better for people with higher Rspan scores (Zekveld et al., 2012; Zekveld, Rudner, et al., 2013). Additionally, participants with higher Rspan scores had better SRTs in the presence of textual cues that were unrelated to the content of the target sentences (Zekveld, Rudner, et al., 2011). However, the disadvantage in SRTs introduced by semantically unrelated word cues compared to conditions without semantic interference was not related to Rspan performance (Zekveld et al., 2012). People with larger verbal WM capacity were also better in recalling words that had earlier been presented to them in silence, noise, or a four-talker masker (Kjellberg & Ljung, 2008; Ng et al., 2013). Remarkably, people with a larger span showed less decline in recall of stimuli from the noise compared to the silence condition in the Kjellberg and Ljung study, whereas high-span people's recall was disrupted more by masking in the study by Ng et al. (2013). It should be noted that Kjellberg and Ljung did not find these effects when testing with target sentences rather than words. When applying noise reduction mechanisms during aided testing, Ng et al. (2013) found that participants with a large Rspan were able to recall masked words equally well as words presented in quiet. People with a small Rspan did not benefit from noise reduction to the same degree.

A connection between WM capacity and age was found for people with impaired (Rudner et al., 2012) and normal hearing (Besser et al., 2012), where older ages were associated with lower span scores. However, other studies did not observe such an association (Arehart et al., 2013; Koelewijn et al., 2012; Ng et al., 2013). Rspan scores appear to be unaffected by impaired hearing. One of the listed studies (Desjardins & Doherty, 2013) and a study by Classon, Rudner, and Rönnberg (2013) included both NH and HI individuals. The Rspan scores of the two groups were comparable for people in the same age range. Also the Rspan scores of other HI groups (Rudner et al., 2009, 2011, 2012) were at the same level as those of NH participants of other studies at comparable ages (e.g., Neher et al., 2011).

Measuring Verbal WM Capacity With Spoken Versus Written Stimuli

As Table 1 demonstrates, the most widely used WM span test in hearing research is the Rspan test, that is, a test measuring verbal WM capacity with written stimuli. Only a few of the listed studies (Baldwin & Ash, 2011; Koelewijn et al.,

2012; Zekveld, Festen, et al., 2013) used an Lspan test next to an Rspan test, and only one study included the Lspan test as the only measure of verbal WM capacity (Zekveld & Kramer, 2013). Results of the studies including both Rspan and Lspan tests suggest that Lspan scores in NH individuals are generally somewhat higher than Rspan scores (Koelewijn et al., 2012; Zekveld, Festen, et al., 2013). Modality differences between visual and auditory verbal WM are also supported by neuroimaging research (Crottaz-Herbette, Anagnoson, & Menon, 2004). Also, in a study on speech and text recognition, Humes et al. (2007) found that correlations were stronger for different tasks in the same modality than for similar tasks in different modalities. Accordingly, it may make sense to use an auditory test for verbal WM capacity in studies examining hearing functions. On the other hand, it is commonly reasoned that WM capacity should be assessed by means of Rspan rather than Lspan tasks to avoid confounding effects of hearing acuity. However, insights into modality specific processing, also in combination with hearing impairment, could be gained by parallel testing in both modalities. We therefore conducted a study to further examine measures of WM capacity with textual or spoken stimuli in relation to each other and to the recognition of masked speech and text. The experiment is described in the following sections.

Experimental Method

Span tests. We used parallel Rspan and Lspan tests in Dutch, that were comparable to the earlier described existing versions of the Rspan test in English (Baddeley et al., 1985), Swedish (Andersson, Lyxell, Rönnberg, & Spen, 2001; Rönnberg, 1990), and Danish (Rudner et al., 2008).

Development of the test material. Two lists of test sentences were created, consisting of 54 sentences each, and one list of 10 practice sentences. All sentences were grammatically correct and had a subject-verb-object structure: article—noun—verb—article—noun. The nouns were singular or plural and animate or inanimate with definite or indefinite articles. None of the included verbs and nouns was used more than once.

Half of the test sentences were semantically incorrect. The sentences were assembled in such a way that semantic judgments could only be made after receiving the sentence's last word and could not be based on the animacy of the subject. This resulted in four different *types* of sentences: correct with an animate or inanimate subject and incorrect with an animate or inanimate subject. For each sentence type, 27 sentences were created. The sentence types were evenly spread over both test lists and over the 54 positions within each list. Furthermore, the order in which semantically correct and incorrect sentences were presented differed between lists.

The primary task in the span tests was recalling a number of target words from sets of 3 to 6 sentences. Example: the

river bordered the country; the comment surprised the eggs; the writer read the newspaper. Target words could be either all subject or all object nouns from the set (country, eggs, newspaper, or river, comment, writer). The sequence of target words to recall (subject vs. object nouns) was fixed and different for List 2 compared to List 1.

Both lists were carefully controlled for number of syllables and word frequencies. Information about word frequencies was gained from the Celex (Center for Lexical Information) database of the Max-Planck Institute for psycholinguistics (http://celex.mpi.nl/). Only nouns from the 5% most frequent nouns were used. This cutoff was chosen to exclude unfamiliar words that might be harder to comprehend and recall. We matched the frequency of the subject and the object nouns in each sentence. High-frequency verbs were used in all sentences.

Testing procedure. During testing, three sentence sets for each set size were presented in the order of increasing set size (3×3, 3×4, 3×5, and 3×6 sentences). Test sentences were interleaved by pauses of 1,750 ms, in which participants rated the semantic correctness by saying "right" or "wrong." After each set, participants received a written instruction on a computer screen to orally report either all first or all last nouns of the sentences in the set. Recall time was limited to 80 s. Via the user interface of the test software, the examiner recorded all oral responses on a second screen that was out of sight for the participant. The outcome score was the total number of correctly recalled target words. Prior to the test, participants were presented with 10 practice sentences divided into three sets (2×3 and 1×4 sentences).

During the Rspan test, the stimuli were displayed on a computer screen with a text size of 46 pt. Sentences were presented in a three-step fashion, displaying one part of the sentence (subject/verb/object) at a time. Sentence parts were displayed for 800 ms each with a blank interval of 75 ms between the parts, as in Andersson et al. (2001).

During the Lspan test, the stimuli were presented diotically through headphones at a level of 65 dB(A). The test sentences were recorded (16-bit, 44.1 kHz) such that they sounded naturally but still had a short pause between the sentence parts, similar to the Rspan test. The average sentence duration was 2.56 s ($SD = 0.15$ s). The sentences were equated with respect to their average RMS value.

Speech reception threshold (SRT). Participants' ability to recognize speech in noise was measured with the SRT test (Plomp & Mimpen, 1979). During the test, spoken sentences were presented diotically through headphones in a fluctuating noise. The spectral shape of the noise equaled the long-term average spectrum of the target speech (Festen & Plomp, 1990). After each sentence presentation, the listener was asked to literally repeat the sentence presented. The overall sound level was kept constant at 65 dB(A) during the test, but the signal-to-noise ratio (SNR) was varied adaptively to
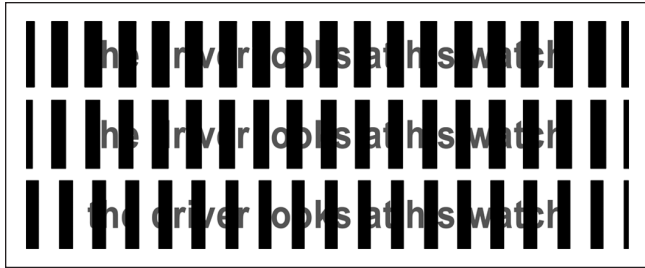
Figure 1. "The driver looks at his watch."
Note. English translation of a TRT sentence at 52% (top row), 58% (middle row), and 64% (bottom row) of unmasking. For printing purposes, the red color of the text as applied in the test was transferred to grey in this illustration.

Table 2. Descriptive Statistics Per Test Session for Reading Span (Rspan), Listening Span (Lspan), Speech Reception Thresholds (SRT), and Text Reception Thresholds (TRT): Means, Standard Deviations (*SD*), Minimum Scores (Min), Maximum Scores (Max), and Score Ranges.

		Mean	*SD*	Min	Max	Range
Rspan	Session 1	20.7	4.4	13.0	32.0	19.0
	Session 2	23.4	4.4	17.0	33.0	20.0
Lspan	Session 1	24.4	4.6	15.0	34.0	19.0
	Session 2	27.8	5.0	19.0	39.0	16.0
SRT	Session 1	−10.0	1.3	−12.9	−7.3	5.6
(dB SNR)	Session 2	−10.9	1.2	−14.7	−7.9	6.8
TRT (%)	Session 1	59.0	3.2	52.5	66.6	14.1
	Session 2	57.3	3.6	49.4	64.6	15.2

estimate the threshold at which the listener was able to repeat 50% of the sentences literally. Lower SRTs indicate better performance. SRTs were assessed with lists of 13 ordinary Dutch sentences consisting of 5 to 9 syllables (Versfeld et al., 2000).

Text reception threshold (TRT). Participants' ability to read fragmentary text was assessed with the TRT_{500} test (Besser et al., 2012). Sentences selected from the same corpus as the SRT sentences (Versfeld et al., 2000) were presented in red font with a white background on a PC screen and partially masked with a pattern of black vertical bars, see Figure 1. Participants were asked to read the sentences aloud. The percentage of masking was adaptively varied between trials to determine the percentage of unmasked text required to read 50% of the sentences correctly. Lower TRTs indicate better performance.

Participants

Forty-two young adults (19-35 years, *M* = 24.4, *SD* = 4.7) with normal hearing participated. Their pure-tone hearing thresholds did not exceed 20 dB HL at any of the octave frequencies 125 to 8000 Hz in both ears. The mean pure-tone average (PTA) over the tested frequencies was 3.5 dB HL for the right ear and 5.3 dB HL for the left ear. Participants were native speakers of Dutch with a higher professional or university level of education and no diagnosis of dyslexia. All of them had normal or corrected-to-normal vision. Furthermore, they were screened for near-vision acuity with Bailey and Lovie (1980) word charts.

Procedure

Participants completed a test and a retest session of 60 min each. The interval between the sessions was 2 to 4 weeks. In both sessions, the auditory speech and visual text reception thresholds were assessed first and thereafter the WM spans. The order of the threshold tests, the span tests, and the test lists was separately balanced across participants. Within

participants, test order was reversed in Session 2 compared to Session 1 within test types (threshold and span). The test material used in the two sessions was identical. However, the material used in one modality of the threshold and span tests in Session 1 was used in the other modality in Session 2. SRT and TRT tests were each conducted three times in a row per session with three different sentence lists. For the data analyses, SRTs and TRTs were averaged over the three test runs in each session.

Results

An analysis of the descriptive data (see Table 2) of all tests revealed no skewed variable distributions and no extreme univariate outliers, neither for the averaged data of both test sessions nor within test sessions, that is, all data points were within 3 *SD*s of the group mean. Intraclass Correlation Coefficients (ICCs) were calculated for the span tests to determine their test-retest reliability. The ICC of the Lspan was .66 and that of the Rspan was .69. These values are close to .70 and are therefore considered satisfactory (Aaronson et al., 2002; De Vet, Terwee, Knol, & Bouter, 2006).

We aimed to create span-test lists that would be interchangeably applicable in both the visual and the auditory modality without affecting the test results. To investigate whether the material fulfilled these requirements, a repeated measures ANOVA was conducted analyzing the main effects and interactions of test modality (reading vs. listening), test session (Session 1 vs. Session 2), and test material (List 1 vs. List 2) on span-test performance. The analysis revealed a significant main effect of test modality, $F(1, 41) = 53.76$, $p < .001$. Lspan scores were better than Rspan scores. There was also a significant main effect of test session, $F(1, 41) = 40.62$, $p < .001$. Span performance was better in Session 2. No main effect of test list was observed, $F(1, 41) = 1.34$, $p = .253$, regardless of test modality and test session. No significant interactions (all $F < 1$) between any of the included factors were found.

An ANOVA examining the effects of test modality and test session on the semantic-judgment task also revealed a significant main effect of test modality, $F(1, 41) = 6.47$, $p < .05$. The number of correct semantic judgments was very high for both tests, but still slightly better for the Lspan (98.0%) than for the Rspan (97.1%). There was also a main effect of test session, $F(1, 41) = 18.92$, $p < .001$, with better semantic judgments in Session 2 (98.1%) than in Session 1 (96.9%). There were no significant interactions between the factors.

In a methodological review of WM span tasks, Conway et al. (2005) argue that the most preferable scoring procedure for span tasks is the partial-credit unit (PCU) scoring because of its good psychometric properties. In PCU scoring, the proportion of correctly recalled targets is calculated for each test set and then averaged over all sets. Performing the ANOVA described above with PCU-scores yielded the same pattern of results as with the total span scores.

We calculated Pearson correlations between Lspan, Rspan, SRT, and TRT for the two test sessions separately. In Session 1, the only significant correlation was between the Rspan and the Lspan ($r = .49$, $p < .01$), which were also significantly correlated in Session 2 ($r = .60$, $p < .01$). In Session 2, we additionally observed a significant correlation between the TRT and the Lspan ($r = -.40$, $p < .05$). The negative correlation coefficient indicates that participants performing better on the TRT test also performed better on the Lspan test.

Discussion of Experimental Results

In the current study sample, the span tests had satisfactory test-retest reliabilities in both modalities. It should be noted that test-retest reliabilities might deviate in other populations, including older adults. Outcome scores of the span tests were independent of the applied test list. Accordingly, test scores achieved with the two lists can be directly compared to each other, in both modalities. This makes the tests well suited for use in the field of audiology and any other research investigating verbal processing within and across modalities in relation to WM capacity. The high percentage of correct semantic judgments confirms that the participants were able to comprehend the sentences well and test results are unlikely to represent perceptual effects. The mean Rspan scores observed in the present study compare well with those obtained with similar Rspan tests in earlier studies (see Table 1).

Rspan and Lspan scores were strongly associated, but the study also revealed a clear difference between participants' absolute performance on the Lspan and the Rspan test with an advantage for the auditory modality of about 4 target words. It is possible that this difference is related to the manner of stimulus presentation. During the Rspan test, sentences were split up into three parts (subject/verb/object) and only one of the parts was displayed at a time, creating a

rather uncommon reading situation. However, the present results reproduce earlier findings of superior recall performance with auditory compared to visual stimuli (e.g., Goolkasian & Foos, 2002). Thus they are more likely to reflect a general modality effect in verbal WM.

In the present study, performance on the span tests was better in Session 2. The training effect was the same for both sentence lists in both modalities, probably caused by greater familiarity with the test material in Session 2. The improvement of performance in Session 2 could also have been an effect of increased familiarity with the task rather than with the material. However, in this case an increase in performance from the first to the second span test in Session 1 would also have been expected, which was not the case.

None of the span tests was associated with speech recognition performance in fluctuating noise. In the current homogenous sample, intraindividual test-retest variance was relatively large compared to the interindividual variance in the test scores, which reduced the statistical power of the tests. The lack of correlations replicates results of earlier studies including only young people with a high educational level and normal hearing acuity (Zekveld et al., 2012; Zekveld, Festen, et al., 2013). There is evidence for the assumption that cognitive resources like WM capacity influence speech understanding performance mainly in adverse listening conditions, where automatic phonological signal decoding needs to be assisted by explicit cognitive processing (Rönnberg et al., 2008; Rönnberg, Rudner, Lunner, & Zekveld, 2010). The present study included only high-functioning young adults. For them the SRT test may not have been challenging enough to reveal effects of WM capacity on SRT performance. In line with this, Zekveld, Rudner, et al. (2011) found SRT performance in young normal-hearing individuals to be related to TRT and Rspan solely in conditions where word cues were presented that semantically interfered with the target SRT sentence. In other words, Rspan made a difference only in strongly adverse listening conditions in which the suppression of irrelevant information was required. These findings should not be interpreted such that speech comprehension in young people with good hearing depends less on cognitive capacity than in other populations. Rather, the listening conditions might need a higher level of disturbance for young normal-hearing individuals to show dependency on cognitive functioning. To tell whether the modality of stimulus presentation plays a role not only for the absolute test outcome but also in the WM span's relationship to the recognition of masked speech, participant groups with a wider variability in cognitive capability and hearing acuity need to be tested.

Linguistic Processing Ability

As stated in the introductory para of this article, next to span tests measuring verbal WM capacity, another test has been found to be a powerful predictor of speech recognition in

noise, that is, the text reception threshold (TRT) test. In this section, we therefore present and discuss hearing research including the TRT test. The first version of the TRT test was developed by Zekveld et al. (2007) as a visual analog of the auditory SRT test. It is assumed that the same modality aspecific linguistic abilities govern the ability to fill in the missing information in the visual and the auditory test, which is supported by a strong association between SRT and TRT (see Table 3 for details). Similar to the SRT, participants have to make use of linguistic closure to decode the fragmentary linguistic information presented to them and recognize the sentence.

Recently, Besser et al. (2012) developed four new versions of the initial test (TRT$_{ORIGINAL}$) by Zekveld et al. (2007) to increase the overlap in abilities tapped by the TRT and SRT tests. In the TRT$_{ORIGINAL}$, sentences are being built up word by word until the complete sentence is written on the screen, where it remains for 3500 ms. For the TRT$_{500}$ this duration was shortened to 500 ms. In the TRT$_{MEMORY}$ two sentences are presented consecutively in TRT$_{500}$ fashion before the participant is allowed to repeat them. During the TRT$_{CENTER}$ and the TRT$_{WORD}$ tests, sentences are not being built up, but only one word at a time is being displayed, either in the center of the screen (TRT$_{CENTER}$) or at the spatial position they would have if the whole sentence was being displayed (TRT$_{WORD}$).

Table 3 provides an overview of research including the TRT test. The TRT$_{ORGINAL}$ is the most widely used version of the test so far. TRT scores, that is, mean values of unmasking to obtain a performance level of 50% correct responses, range from 54.7% (Zekveld & Kramer, 2013, Experiment 2) to 58.2% (George et al., 2007) in participants with normal hearing and normal vision. Study populations with a lower average age (e.g., Zekveld et al., 2012; Zekveld & Kramer, 2013; Zekveld, Rudner, et al., 2011) seem to have better mean TRTs than groups with a higher average age (e.g., George et al., 2007), suggesting a connection between TRT outcome and age. However, analyses of associations between age and TRT outcomes within study groups with wide age ranges are inconclusive. While Besser et al. (2012) and Zekveld, Kramer, et al. (2011) found associations between age and different TRT versions in groups of NH individuals with age ranges of 18 to 78 and 46 to 73 years, respectively, Koelewijn et al. (2012) and Zekveld et al. (2007) did not find such associations in NH groups with comparable age ranges. Also, in a secondary analysis of combined data from the studies by Zekveld et al. (2007, 2008, 2009), no association between TRT performance and age was found ($r = .16$, $p > .1$) in a large group of NH individuals at ages between 18 and 78 years (Kramer, Zekveld, & Houtgast, 2009).

TRT and Speech Recognition in Noise

The association between TRT and the recognition of masked speech has been confirmed many times, also when age

effects were statistically controlled for. Different types of noise were used in the studies listed in Table 3. Associations were examined for SRTs in modulated noise, recognition of speech presented against a background of an interfering speaker, and SRTs in stationary noise.

For speech recognition in modulated maskers, all but the current study and the study by Koelewijn et al. (2012) found a significant association with the TRT, both in groups with normal hearing (Besser et al., 2012; George et al., 2007; Zekveld et al., 2007) and with impaired hearing (Zekveld, George, et al., in press). In the HI group of George et al. (2007), TRT was a significant predictor of SRT performance when auditory temporal resolution was controlled for, underlining the importance of auditory factors in speech recognition. In the NH group of the same study, the TRT was the best predictor of SRT performance in modulated noise, and in Besser et al. (2012) different TRT versions were the only significant predictors of speech recognition in modulated noise next to age, whereas other cognitive factors (Rspan and processing speed) did not predict SRT performance.

Two studies examined the TRT's relationship to speech recognition with an interfering talker. While Koelewijn et al. (2012) found an association between the TRT and SRTs with an interfering speaker in a NH group, Zekveld, Festen, et al. (2013) did not find such an association. These contrasting findings might be attributable to the fact that the sample size of Koelewijn et al. (2012) was larger and participants were considerably older (40-70 years) and more heterogenous in their educational background than those of Zekveld, Festen, et al. (2013), with young university students as the only participants.

The relationship between the TRT and speech recognition in steady-state noise is not clear-cut. Associations between these measures were found for NH people by Besser et al. (2012), George et al. (2007), Krull et al. (2013; young NH group), Zekveld et al. (2007), and Zekveld, Kramer, et al. (2011). Studies including participants with a hearing impairment did not find significant associations between the TRT and speech recognition in steady-state noise for this group (George et al., 2007; Krull et al., 2013; Zekveld, Kramer, et al., 2011). Possibly, the relationship between TRT and SRT was overruled by auditory factors in these groups. However, also in some of the NH groups associations were nonsignificant (Goverts et al., 2011; Krull et al., 2013 [old NH group]; Zekveld et al., 2012). In the study by Zekveld, Rudner, et al. (2011) a significant association was observed, but only for SRTs with 29% correct responses in which the sentences were presented together with words unrelated to the semantic content of the target sentence. Generally, the association between the TRT and the SRT in steady-state noise is the most stable in NH individuals of middle age or older, whereas associations vary in other groups. The study by Krull et al. (2013) is a special case. They found associations between the TRT and speech recognition in steady-state noise in a group of young NH participants, but not in

Table 3. Hearing Research Including a TRT Test.

Study	TRT version & test material	Speech recognition tests	Other tests and measures	Association TRT—speech recognition	Findings related to TRT
Besser et al. (2012) N = 55; 18-78 (44.0) years; NH	$TRT_{ORIGINAL}$, TRT_{500}, TRT_{CENTER}, TRT_{WORD}, TRT_{MEMORY}	SRT_{STEADY}, SRT_{FLUCT}	Rspan; LDST (general processing speed); self-report of how pleasant/taxing the TRT versions were	TRT – SRT_{FLUCT} (r = .46** – .59**); TRT – SRT_{STEADY} (r = .23 – .49**); strength depending on TRT version, most sign. when controlling for age	Other tests and measures ranging from 56.0 ($TRT_{ORIGINAL}$) to 64.3 (TRT_{MEMORY}); TRT predictor of SRT_{FLUCT}; TRT associated with Rspan, mostly also when controlling for age
Besser et al. (currrent article) N = 42; 19-35 (24.4) years; NH	TRT_{500}	SRT_{FLUCT}	Rspan; Lspan	*ns*, no details reported	TRT *M* = 58.2; TRT and Lspan associated (Session 2)
George et al. (2007) N = 13; 53-78 (63.5) years; NH N = 21; 46-81 (65.5) years; HI	$TRT_{ORIGINAL}$	SRT_{STEADY}, SRT_{FLUCT} (16-Hz block-modulated); unaided testing with amplification; sentences: Plomp and Mimpen (1979)	auditory spectral and temporal resolution	NH: TRT – SRT_{STEADY} (r = .61*), TRT – SRT_{FLUCT} (r = .80**); HI: TRT – SRT_{STEADY} (r = .34), TRT – SRT_{FLUCT} (r = .42)	TRT *M* = 58.2 (NH), *M* = 58.8 (HI); TRT best predictor of SRTs in NH; in HI, TRT predictor of SRT_{FLUCT} when accounting for temporal resolution
Goverts, Huysmans, Kramer, de Groot, and Houtgast (2011)[a] N = 13; 20-25 (22.0) years; NH, native N = 10; 19-45 (29.0) years; NH, nonnative	$TRT_{ORIGINAL}$ sentences with/without lexical, semantic, and syntactic distortions	SRT_{STEADY}		not reported	undistorted TRT Median = 56.2 (native), Median = 65.2 (nonnative); both groups sensitive to TRT distortions in all linguistic domains
Koelewijn et al. (2012) N = 32; 40-70 (51.3) years; NH	TRT_{500}	SRT_{FLUCT}, SRT_{TALK}; 50% and 84% correct level	Rspan; Lspan; SICspan; pupillometry during SRTs; ratings of effort, motivation, performance for SRTs	TRT – SRT_{TALK} 50% correct (r = .67**), 84% correct (r = .54**); TRT – SRT_{FLUCT} 50% correct (r = .19), 84% correct (r = .25)	TRT *M* = 59.8; TRT associated with peak pupil dilation during SRT_{TALK} (50% correct); TRT associated with Rspan, Lspan, SICspan
Krull, Humes, and Kidd. (2013) N = 25; 18-32 (24.0) years; NH N = 20; 65-84 (73.0) years; NH N = 21; 60-85 (75.0) years; HI	$TRT_{ORIGINAL}$ stimuli: 4-5-letter words from the R-SPIN test; Bilger, Nuetzel, Rabinowitz, and Rzeczkowski (1984)	steady-state noise (SSN) at +6 dB SNR; removal of temporal segments (INT); removal of spectral segments (FILT); unaided testing with amplification for HI; all with monosyllabic words from R-SPIN test	text recognition in visual Gaussian noise; cognitive tests for WM capacity (elderly only): memory updating (MU), sentence span, spatial short-term memory	TRT – SSN old (r = .08), young (r = .40*); TRT – INT old (r = .11), young (r = .55**); TRT – FILT old (r = .21), young (r = .49*)	TRT associated with all auditory measures in young NH group, but with none in older groups; TRT associated with text recognition in Gaussian noise

Table 3. (continued)

Study	TRT version & test material	Speech recognition tests	Other tests and measures	Association TRT—speech recognition	Findings related to TRT
Mishra, Lunner, Stenfelt, Rönnberg, and Rudner (2013)[b] N = 21; 22-54 (29.5) years; NH	$TRT_{ORIGINAL}$ stimuli: Swedish HINT sentences (Hällgren et al., 2006)		CSCT (cognitive spare capacity test, for executive functions); Rspan; letter memory test; Simon task	no SRT assessed	TRT M = 58.4; TRT associated with updating, audiovisual, and overall score of CSCT; TRT associated with semantic judgments in Rspan
Zekveld et al. (2007) N = 34; 19-78 (34.0) years; NH	$TRT_{ORIGINAL}$	SRT_{STEADY}, SRT_{FLUCT} (16-Hz block-modulated)		TRT − SRT_{STEADY} ($r = .58^{**}$); TRT − SRT_{FLUCT} ($r = .58^{**}$); also when controlling for age	TRT M = 55.4
Zekveld, Kramer, Vlaming, and Houtgast (2008)[c] N = 18; 19-31 (23) years; NH	$TRT_{ORIGINAL}$ and text in random dots with (TRT+S) and without (TRT) auditory sentence presentation at 3 SNRs; sentences: Plomp and Mimpen (1979)	SRT_{STEADY} with (SRT+T) and without (SRT) display of masked target sentence at 3 masking levels	Audiovisual reception threshold test (AVRT): simultaneous presentation of text and speech, adaptive masking in both modalities	not reported	TRT M = 53.6; simultaneous presentation of masked text enhanced speech perception and vice versa
Zekveld, Kramer, Kessens, Vlaming, and Houtgast (2009) N = 22; 18-28 (21) years; YNH N = 22; 45-65 (55) years; MA-NH N = 30; 46-69 (57) years; MA-HI	$TRT_{ORIGINAL}$ sentences: Plomp and Mimpen (1979)	SRT_{QUIET}, SRT_{STEADY} with and without textual support; aided testing for MA-HI	rating of effort during SRT tests; spatial WM span (SSP)	not reported	TRT M = 56.6 (YNH), M = 57.5 (MA-NH), M = 59.5 (MA-HI); TRT not associated with rated effort or SRT benefit from textual support
Zekveld, Kramer, and Festen (2011) N = 38; 46-73 (55.0) years; NH N = 36; 45-72 (61.0) years; HI Data from Zekveld, Kramer, and Festen (2010) added for analyses: N = 38; 19-31 (23.0) years; NH	$TRT_{ORIGINAL}$, TRT_{500}; sentences: Plomp and Mimpen (1979)	SRT_{STEADY}; 50%, 71%, 84% correct level; unaided testing	pupillometry during SRTs; LDST (general processing speed); vocabulary size; self-rated listening effort and SRT performance	NH: SRT_{STEADY} −$TRT_{ORIGINAL}$ ($r = .38^{**}$), TRT_{500} ($r = .35^*$); HI: SRT_{STEADY} − $TRT_{ORIGINAL}$ ($r = −.07$), TRT_{500} ($r = −.06$); for NH also sign. when controlling for age	$TRT_{ORIGINAL}$ M = 57.1 (NH), M = 58.2 (HI); TRT_{500} M = 58.4 (NH), M = 60.6 (HI); better TRT associated with higher cognitive load during SRT; in NH, lower TRT associated with higher age
Zekveld, Rudner, et al. (2011), Experiment 2 N = 20; 18-32 (22.0) years; NH	$TRT_{ORIGINAL}$	SRT_{STEADY} with semantically related, unrelated, or nonword text cues; 29% and 16% correct level	Rspan	TRT − SRT_{STEADY} at 29% correct with unrelated cues ($r = .60^{**}$); details for others not reported	TRT M = 54.9

(continued)

Table 3. (continued)

Study	TRT version & test material	Speech recognition tests	Other tests and measures	Association TRT—speech recognition	Findings related to TRT
Zekveld et al. (2012) N = 18; 20-29 (23.6) years; NH	TRT_{ORIGINAL}	SRT_{STEADY} (29% correct level); SRT_{STEADY} nonadaptive with semantically related, unrelated, or nonword text cues, at estimated 29% correct level	Rspan, fMRI during SRTs	TRT not associated (p > .05) with % correct words in SRT_{STEADY}	TRT M = 54.9; TRT associated with less interference from unrelated cues; TRT associated with brain activation in left angular gyrus
Zekveld, Festen, et al. (2013) N = 24; 18-27 (22.0) years; NH	TRT_{CENTER} with adapted colors; 29%, 71% correct level, for analyses average of 29% and 71% used	SRT_{TALK} 29% and 71% correct level	pupillometry during SRT and TRT; Rspan; Lspan; incidental cued recall test of TRT and SRT stimuli	ns, no details reported	TRT at 29% M = 63.2, at 71% M = 80.8; pupil peak and mean dilation during TRT smaller at 71% than at 29% correct
Zekveld, George, Houtgast, and Kramer (in press) N = 32; 46-83 (66.4) years; HI	TRT_{ORIGINAL}	SRT_{STEADY}; SRT_{FLUCT} (16-Hz block-modulated noise); unaided testing with amplification; sentences: Plomp and Mimpen (1979)	Amsterdam Inventory for Auditory Disability and Handicap (AIADH); sustained visual attention; Spatial working memory	TRT − SRT, z score transformed over noise types (r = .47**)	TRT M = 60.3; TRT associated with AIADH measures of sound detection, sound discrimination, and speech intelligibility
Zekveld, Rudner, et al. (2013) N = 18; 20-32 (23.0) years; NH	TRT_{ORIGINAL}	SRT_{STEADY}, SRT_{FLUCT}, SRT_{TALK} with semantically related or nonword cues; 29% and 71% correct level	Rspan; SICspan; 2-alternative-forced-choice recognition of presented SRT sentences	not reported	TRT not associated with SRT benefit from related cues
Zekveld and Kramer (2013), Experiment 1 N = 24; 18-26 (22.0) years; NH	TRT_{ORIGINAL}	SRT_{TALK}, nonadaptive (1%, 50%, 99% correct level), and speech in quiet (no threshold); sentences: Versfeld et al. (2000)	Lspan; pupillometry during SRT; cued recall test of TRT and SRT stimuli; self-rated listening effort and SRT performance	not reported	TRT M = 57.8; TRT associated with peak pupil dilation amplitude in 1% correct nonadaptive SRT; TRT associated with Lspan
Zekveld and Kramer (2013), Experiment 2 N = 13; (23.0) years; NH	TRT_{ORIGINAL}	SRT_{TALK} nonadaptive at 9 SNRs; adaptive SRT_{TALK} at 29% and 71% correct level	incidental cued recall test of previously presented SRT stimuli	not reported	TRT M = 54.7; no other TRT results reported

Note. If not otherwise specified, the following test properties applied: SRTs and TRTs were assessed for complete sentences at a performance level of 50% correct responses. TRTs assessed with sentences by Versfeld et al. (2000). SRT_{STEADY} = SRT in a steady-state masker; SRT_{FLUCT} = SRT in a fluctuating masker; SRT_{TALK} = SRT with one or several interfering talkers; SRT_{QUIET} = speech reception without a masker; M = mean score; ns = nonsignificant. Associations marked * significant at the .05 level and marked ** significant at the .01 level.
[a]Main purpose of the study was to examine the effects of language proficiency on TRT performance, rather than associations of TRT and SRT. [b]The study was set up to evaluate the cognitive spare capacity test, rather than examining associations with speech recognition in noise. [c]Main purpose of the study was to examine the benefit in speech comprehension received from simultaneously displayed textual information.

their older NH and HI groups. Different from all other reported studies, Krull et al. (2013) used word recognition rather than sentence recognition for both speech and text recognition tasks. The former relies more on vocabulary access than on linguistic abilities as processing of syntax and semantic coherence required in sentence recognition. In addition to its relationship with objective measures of speech recognition, the TRT was also found to be associated with subjective abilities to understand speech in quiet and in noise and to detect and discriminate sounds (Zekveld, George, et al., in press).

TRT in Relation to Cognitive Measures and Language Proficiency

In many of the described research studies, TRT tests were applied in combination with measures of cognitive ability and/or capacity, especially with tests of WM functions. All TRT versions developed by Besser et al. (2012) correlated with Rspan performance and most of them also when age was controlled for. Also Koelewijn et al. (2012) found associations between the TRT and both auditory and textual measures of verbal WM capacity, that is, Rspan, Lspan, and SICspan, that persisted when correcting for age as a confounder. However, in studies including young individuals only, TRT results were not associated with Rspan (Zekveld et al., 2012; Zekveld, Festen, et al., 2013), whereas it was associated with Lspan (Zekveld & Kramer, 2013). Mishra et al. (2013) did not calculate associations between TRT and Rspan performance, but found the TRT to be associated with the semantic judgment task of the Rspan test and with audio-visual tasks of memory and executive functioning of the cognitive spare capacity test. They interpreted these results as support for the TRT's significance in sentence recognition in terms of meaning processing and its value for estimating skills to integrate visual cues in audiovisual tasks of speech recognition. While associations were observed between the TRT and tests of verbal WM, no associations have been found with spatial WM (Zekveld, George, et al., in press), suggesting that the shared variance of TRT and verbal WM tests is caused by verbal processing abilities rather than a general processing capacity.

TRT performance has also been found to be associated with measures of pupillometry. Individuals with better TRTs tend to have larger pupil responses when trying to recognize sentences with an interfering talker (Koelewijn et al., 2012) or in stationary noise (Zekveld, Kramer, et al., 2011). However, in groups including only young NH individuals, an association with pupil responses during SRT testing was only observed at an extremely low intelligibility level (Zekveld & Kramer, 2013), but not at higher performance levels (Zekveld, Festen, et al., 2013). The pupil dilation response and its latency express processing load. It is reasonable to assume that the higher load observed during sentence recognition in people with better linguistic closure is a consequence of more coordinated brain activity in these people because larger cognitive capabilities as such are not likely to lead to increased processing load.

Similar to speech recognition in noise, TRT performance appears to be driven by language proficiency with clear advantages for native compared to nonnative speakers of the test language (Goverts et al., 2011). The authors obtained parallel results for SRT and TRT tests, where group differences in language ability lead to the same kind of differences in performance in both modalities, for example, regarding sensitivity to linguistic distortions in the test sentences.

Discussion

The current article provides an overview of recently published research examining the relationship of linguistic processing ability as measured by the TRT test and verbal WM capacity as measured by means of the Rspan or Lspan test with the recognition of masked speech. The studies included in the current review examined associations in a wide range of participant groups, including people with impaired or normal hearing at different ages and educational levels. Speech recognition was tested in steady-state or fluctuating maskers or with one or several interfering talkers, with complete sentences or single target words, with and without hearing aids, at different SNRs and presentation levels. Across these divergent study characteristics, a few general patterns of results could be distinguished.

The most robust association between Rspan and speech recognition was observed for speech in speech maskers with correlation coefficients around $r = .5$, independent of the number of interfering talkers in the masker (Desjardins & Doherty, 2013; Ellis & Munro, 2013; Koelewijn et al., 2012). In the studies by Desjardin and Doherty (2013) and Ellis and Munro (2013), participants did not need to repeat the complete test sentences, but only certain key words to complete the task. Furthermore, these two studies made use of two, six, or many talkers in the masker. The more talkers the masker contains, the less semantic interference with the target sentence the speech masker introduces because words or sentences of multitalker maskers are harder to identify. Accordingly, they become more alike fluctuating nonspeech maskers. The associations between Rspan and SRTs in fluctuating noise were less evident than between Rspan and SRTs with speech maskers. In the two studies that observed associations with SRTs in fluctuating noise, these associations disappeared when age as a confounder was controlled for (Besser et al., 2012; Koelewijn et al., 2012). It is not clear whether the associations between Rspan and the SRTs in multitalker babble (Ellis & Munro, 2013) and a six-talker masker (Desjardins & Doherty, 2013) would have persisted had age been controlled for in the analyses. In other words, it is hard to determine whether the semantic interference introduced by the masker drives the observed associations with the Rspan or whether the use of a speech rather than a nonspeech masker is sufficient to disclose such associations. It should be noted that for HI listeners two earlier studies (Foo et al., 2007; Lunner, 2003), discussed in the review by Akeroyd (2008), found associations of Rspan with speech recognition in fluctuating noise also when controlling for age. Furthermore, Arehart et al. (2013) found that Rspan was not associated with age. Age and Rspan explained variance in speech recognition by HI people in multitalker maskers with frequency compression independent of each other.

In contrast to the Rspan, the associations found for the TRT with speech recognition in fluctuating nonspeech maskers are more evident and were not confounded by age to the

same degree (Besser et al., 2012; George et al., 2007; Zekveld et al., 2007). Actually, in George et al. (2007), TRT was the only predictor of SRTs in fluctuating noise in the NH group, and in Besser et al. (2012) TRT predicted SRTs in fluctuating noise together with age, whereas Rspan was not a predictor. Likewise, TRT was related to speech recognition with interfering speech, independent of age, in a study that examined this relationship in a group of participants between 40 and 70 years of age (Koelewijn et al., 2012).

Regarding speech recognition in steady-state noise, associations with both the Rspan and the TRT vary for NH people, but are most apparent for the TRT in participants of middle or higher ages. Generally, speech recognition seems to be less prone to influences by WM span and TRT in groups of young individuals with normal hearing, which was also confirmed by the experimental results of the current study. Nonetheless, young NH people were a frequently chosen study sample in the experiments presented in Table 1 and Table 3, either in combination with other groups (Baldwin & Ash, 2011; Desjardins & Doherty, 2013; Goverts et al., 2011; Krull et al., 2013; Zekveld et al., 2009) or in isolation (Kjellberg & Ljung, 2008; Ljung & Kjellberg, 2009; Sörqvist & Rönnberg, 2012; Zekveld et al., 2008, 2012; Zekveld, Festen, et al., 2013; Zekveld & Kramer, 2013; Zekveld, Rudner, et al., 2011, 2013). However, of the studies in which data analyses were conducted and reported for the young NH separately, only a few discovered associations for TRT or WM span with speech recognition. Baldwin and Ash (2011) found Lspan to be associated with speech recognition in quiet at the lowest presentation level (45 dBA). Speech recognition in noise was not examined in that study. In Zekveld, Rudner, et al. (2011), Rspan and TRT were both associated with SRTs in steady-state noise at the 29% correct level, but only when the SRT sentences were presented together with textual cues that were semantically unrelated to the sentence content. The only instance of the listed studies, in which masked speech recognition as such was related to TRT performance in young NH people, is the study by Krull et al. (2013). Notably, the speech and text recognition tasks in this study included words rather than sentences. Accordingly, the tasks were more dependent on low-level (letter/phoneme perception) and midlevel (word perception) abilities of perceptual closure than on higher level processes like semantic and syntactic integration as required in sentence recognition. Possibly, contextual processing is generally not a dominating factor in TRT performance. For instance, TRT was not associated with speech recognition supported by textual cues that were semantically related to the SRT test sentences (Zekveld et al., 2012) or with SRT benefit resulting from the contextual cues (Zekveld, Rudner, et al., 2013).

Few of the presented studies included participants with a hearing impairment and if HI people were included, the investigation of speech recognition performance was often combined with questions of, for example, spatial listening abilities (Neher et al., 2009, 2011) or the influence of signal

manipulation and hearing aid settings (Arehart et al., 2013; Rudner et al., 2009, 2011). Consequently, it is hard to sketch an overall picture of the relationships between speech recognition, WM span, and TRT in these groups. However, the results by Neher et al. (2009) and the studies by Rudner et al. and Arehart et al. suggest that people with a higher Rspan receive more benefit from the spatial separation of target and interfering signal, perform better in various SRT tasks, and are better in handling listening situations with unfamiliar signal processing and a large amount of signal distortion. These findings are in line with results from earlier research with HI listeners where associations between Rspan and speech recognition in both fluctuating and steady-state noise were found independent of age (Foo et al., 2007; Lunner, 2003). They are also in line with the idea that highly degraded signals require allocation of processing resources to early, peripheral stages of signal decoding. As a consequence, large-capacity individuals retain more resources for later high-level processing tasks, which is reflected in their better overall performance (Rönnberg et al., 2008). It should be noted that participants with impaired hearing were generally tested either with hearing aids or with amplified signals to compensate for reduced hearing acuity (see Tables 1 and 3). According to earlier research—see Akeroyd (2008) for an overview—restored audibility is a prerequisite for associations between cognitive factors and speech reception in noise to emerge. Interestingly, one of the studies (Rudner et al., 2011) applied both aided and unaided testing for the HI participants and found that correlations between Rspan and SRTs in various masking conditions were similar for aided and unaided testing. However, post hoc analyses with the participants split into a high-span and a low-span group revealed that Rspan interacted with intelligibility level, noise type, and compression setting differently for aided versus unaided testing of speech recognition. On one hand, these results suggest that WM capacity generally plays a role in speech recognition in HI individuals. On the other hand, they suggest that a high Rspan gives an advantage in specific listening situations, and these situations are different for aided than for unaided listening.

Concerning the TRT, auditory suprathreshold factors should be regarded when examining associations with speech recognition in HI people. For example, TRT only predicted SRTs in fluctuating noise, when auditory temporal resolution was accounted for (George et al., 2007). For SRTs in steady-state noise, that are generally believed to depend more on bottom-up perceptual processes than SRTs in modulated maskers, no associations with TRT were found at all in HI groups. None of the studies included young individuals with a hearing impairment. It would be highly interesting to examine whether cognitive variables as TRT and WM span play a larger role in young HI than in young NH groups.

When looking at Rspan and TRT results in terms of absolute outcome scores, there are no indications that these differ between groups of people with normal and impaired hearing.

A study that included NH and HI individuals (Classon et al., 2013) found that the Rspan scores of the two groups were comparable, and also the Rspan scores of other HI groups (Rudner et al., 2009, 2011, 2012) were at the same level as those of NH participants of other studies in the same age range. This suggests that there are no general differences in verbal WM capacity between individuals with normal or impaired hearing. It is not clear if this also holds for span scores assessed with auditory stimuli since none of the listed studies used an auditory test for measuring WM capacity in HI people. This could be a topic for future research. Also the mean TRT outcomes of people with a hearing impairment (George et al., 2007; Zekveld, George, et al., in press) were comparable to those of NH people in the same or other studies.

Another question is the TRT's and WM spans' mutual relationship and the influence of age thereon. As mentioned earlier, TRT was associated with Rspan and Lspan independent of age in one study (Koelewijn et al., 2012), whereas in another study, the two TRT versions (TRT_{500} and $TRT_{ORIGINAL}$) most frequently used in hearing research were only associated with Rspan, when age was not controlled for (Besser et al., 2012). In young NH individuals, TRT and Rspan were not significantly correlated with each other, but with a correlation coefficient of $r = -.33$, the association might have reached significance in a bigger sample (Zekveld et al., 2012). TRT and Lspan were found to be associated in young NH participants, like the results of the current study and those of Zekveld and Kramer (2013) showed. Generally, it can be stated that TRT and WM spans share some variance and do not measure fully independently operating cognitive functions. The relationship appears to be influenced by age in some cases, but does not differ in general for different age groups. This was confirmed by additional analyses of the data published in Besser et al. (2012). We performed these to examine the possibility of age being an effect modifier in the relationship between TRT and Rspan, that is, the question whether the interdependence of the measures differs for different age groups. For none of the five TRT versions an effect modification of age was observed in the association with the Rspan.

The data presented in Table 1 and Table 3 reveal some differences between TRT and Rspan with regard to the integration of semantic context in speech recognition. For example, higher Rspan scores were connected with more benefit in speech recognition from word cues semantically related to the SRT sentences, but not to the interference from semantically unrelated word cues. For the TRT, the opposite was true. People with a better TRT had an SRT advantage in conditions with semantically unrelated sentence context (Zekveld et al., 2012). These results are in line with the observation that TRT is associated with performance on the SICspan test, which measures the ability to ignore irrelevant information during processes of information storage in WM (Koelewijn et al., 2012). TRT and WM span have also been tested in relation to the processing load experienced during speech recognition tasks. Better TRT performance was found to be associated with higher mental processing load, that is, cognitive activity, during SRT testing, whereas no such patterns were observed for Rspan or Lspan (Koelewijn et al., 2012; Zekveld & Kramer, 2013; Zekveld, Kramer, et al., 2011).

In summary, the results suggest that the Rspan and Lspan measures of verbal WM capacity are indicators of abilities to handle listening situations that require complex verbal processing. This is manifested in the span tests' associations with speech recognition in speech maskers, which introduce a higher level of interference than nonspeech maskers, in associations with the ability to integrate contextual information during speech recognition, and—for HI people—in associations with the ability to handle unfamiliar hearing aid settings and make use of spatial cues in speech recognition. The WM spans' representation of complex processing abilities is probably also the cause of its sensitivity to age and the influence of age on its relationship with other measures. However, the presented data also suggest that at sufficiently high levels of distortion, WM capacity is called upon in speech recognition in all age groups.

The TRT appears to represent abilities of linguistic processing activated in speech and text recognition. As a consequence, the TRT's relationship to SRTs is less dependent on the task complexity. It is therefore also less affected by age. In contrast to WM capacity, linguistic abilities have been shown to remain stable or increase (vocabulary) over the adult life span (Verhaeghen, 2003; Wingfield & Grossman, 2006). The age effects observed in some of the studies in relation to the TRT and its association with the SRT may be a result of decreased perceptual acuity in one or the other modality. TRT and SRT overlap in their dependency on abilities of linguistic closure, but differ in the domain of perceptual processes. This also explains why it is necessary to control for auditory processing abilities when examining the relationship between TRT and SRT in HI individuals.

Conclusions

Lspan and Rspan are capacity measures indicating the potential to perform well in situations requiring complex verbal processing, such as the integration of simultaneous streams of verbal information. The TRT on the other hand is an ability measure indicating individual competences in linguistic closure. In NH participants of middle or higher ages, TRT and Rspan were both associated with the recognition of speech in speech, but of the two only TRT predicted recognition performance in fluctuating maskers. This was primarily a consequence of the confounding effect of age on the associations between Rspan scores and SRTs in fluctuating noise. Also for steady-state maskers, the association between TRT and SRT is more evident than that between Rspan and SRT. Participant groups of young ages with normal hearing appear

to be less suitable than older groups to examine associations between speech recognition, WM span, and TRT. For hearing impaired individuals, a higher WM span appears to provide for better abilities to perform well in various listening situations. None of the studies examined relationships of verbal WM span, TRT, and speech recognition in young individuals with a hearing impairment. Also differences in associations with Lspan compared to Rspan have received little attention so far. These are topics of interest for future research.

Acknowledgments

Declaration of Conflicting Interests

Funding

References

Aaronson, N., Alonso, J., Burnam, A., Lohr, K. N., Patrick, D. L., Perrin, E., & Stein, R. E. K. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research*, *11*, 193-205.

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. *International Journal of Audiology*, *47*(Suppl. 2), S53-S71.

Andersson, U., Lyxell, B., Rönnberg, J., & Spen, K.-E. (2001). Cognitive correlates of visual speech understanding in hearing-impaired individuals. *Journal of Deaf Studies and Deaf Education*, *6*, 103-116.

Arehart, K. H., Souza, P., Baca, R., & Kates, J. M. (2013). Working memory, age, and hearing loss: Susceptibility to hearing aid distortion. *Ear and Hearing*, *34*, 251-260. doi:10.1097/AUD.0b013e318271aa5e

Arlinger, S., Lunner, T., Lyxell, B., & Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scandinavian Journal of Psychology*, *50*, 371-384.

Baddeley, A. (1992). Working memory. *Science*, *255*, 556-559.

Baddeley, A., Logie, R., Nimmosmith, I., & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language*, *24*, 119-131.

Bailey, I. L., & Lovie, J. E. (1980). The design and use of a new near-vision chart. *American Journal of Optometry and Physiological Optics*, *57*, 378-387.

Baldwin, C. L., & Ash, I. K. (2011). Impact of sensory acuity on auditory working memory span in young and older adults. *Psychology and Aging*, *26*, 85-91.

Besser, J., Zekveld, A. A., Kramer, S. E., Rönnberg, J., & 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. *Journal of Speech, Language, and Hearing Research*, *55*, 194-209.

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research*, *27*, 32-48.

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *Journals of Gerontology Series B-Psychological Sciences and Social Sciences*, *60*, 223-233.

Classon, E., Rudner, M., & Rönnberg, J. (2013). Working memory compensates for hearing related phonological processing deficit. *Journal of Communication Disorders*, *46*, 17-29.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769-786.

Crottaz-Herbette, S., Anagnoson, R. T., & Menon, V. (2004). Modality effects in verbal working memory: Differential prefrontal and parietal responses to auditory and visual stimuli. *Neuroimage*, *21*, 340-351.

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

De Vet, H. C. W., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, *59*, 1033-1039.

Desjardins, J. L., & Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. *Ear and Hearing*, *34*, 261-272. doi:10.1097/AUD.0b013e31826d0ba4.

Ellis, R. J., & Munro, K. J. (2013). Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition? *International Journal of Audiology*, *52*(1), 14-22. doi:10.3109/14992027.2012.721013

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19-23.

Engle, R. W. (2010). Role of working-memory capacity in cognitive control. *Current Anthropology*, *51*, S17-S26.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, *88*, 1725-1736.

Foo, C., Rudner, M., Rönnberg, J., & Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *Journal of the American Academy of Audiology*, *18*, 618-631.

George, E. L. J., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M., & Houtgast, T. (2007). Auditory and non-auditory factors affecting speech reception in noise by older listeners. *Journal of the Acoustical Society of America*, *121*, 2362-2375.

Goolkasian, P., & Foos, P. W. (2002). Presentation format and its effect on working memory. *Memory & Cognition*, *30*, 1096-1105.

Goverts, S. T., Huysmans, E., Kramer, S. E., de Groot, A. M., & Houtgast, T. (2011). On the use of the distortion-sensitivity approach in examining the role of linguistic abilities in speech understanding in noise. *Journal of Speech, Language, and Hearing Research*, *54*, 1702-1708.

Hagerman, B. (1982). Sentences for testing speech intelligibility in noise. *Scandinavian Audiology*, *11*, 79-87.

Hagerman, B., & Kinnefors, C. (1995). Efficient adaptive methods for measuring speech reception threshold in quiet and in noise. *Scandinavian Audiology*, *24*, 71-77.

Hällgren, M., Larsby, B., & Arlinger, S. (2006). A Swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition. *International Journal of Audiology*, *45*, 227-237.

Houtgast, T., & Festen, J. M. (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *International Journal of Audiology*, *47*, 287-295.

Humes, L. E., Burk, M. H., Coughlin, M. P., Busey, T. A., & Strausner, L. E. (2007). Auditory speech recognition and visual text recognition in younger and older adults: Similarities and differences between modalities and the effect of presentation rate. *Journal of Speech, Language, and Hearing Research*, *50*, 283-303.

Kjellberg, A., & Ljung, R. (2008). Recall of words heard in noise. *Applied Cognitive Psychology*, *22*, 1088-1098.

Koelewijn, T., Zekveld, A. A., Festen, J. M., Rönnberg, J., & Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *International Journal of Otolaryngology*, *2012*, 1-11. doi:10.1155/2012/865731.

Kramer, S. E., Zekveld, A. A., & Houtgast, T. (2009). Measuring cognitive factors in speech comprehension: The value of the Text Reception Threshold test as a visual equivalent of the SRT test. *Scandinavian Journal of Psychology*, *50*, 507-515.

Krull, V., Humes, L. E., & Kidd, G. R. (2013). Reconstructing wholes from parts: Effects of modality, age, and hearing loss on word recognition. *Ear and Hearing*, *34*(2), e14-23. doi:10.1097/AUD.0b013e31826d0c27

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology-Learning Memory and Cognition*, *16*, 1118-1133.

Ljung, R., & Kjellberg, A. (2009). Long reverberation time decreases recall of spoken information. *Journal of Building Acoustics*, *16*, 301-312.

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *International Journal of Audiology*, *42*(Suppl. 1), S49-S58.

Mishra, S., Lunner, T., Stenfelt, S., Rönnberg, J., & Rudner, M. (2013). Visual information can hinder working memory processing of speech. *Journal of Speech, Language, and Hearing Research*. Advance online publication. doi:10.1044/1092-4388(2012/12-0033). Retrieved from http://jslhr.asha.org/cgi/rapidpdf/1092-4388_2012_12-0033v1.pdf

Neher, T., Behrens, T., Carlile, S., Jin, C., Kragelund, L., Specht Petersen, A. S., & van Schaik, A. (2009). Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, and cognition. *International Journal of Audiology*, *48*, 758-774.

Neher, T., Laugesen, S., Jensen, N. S., & Kragelund, L. (2011). Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *Journal of the Acoustical Society of America*, *130*, 1542-1558.

Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., & Rönnberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *International Journal of Audiology*, *52*, 433-441.

Oberauer, K., Suss, H. M., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, *31*, 167-193.

Plomp, R., & Mimpen, A. M. (1979). Improving the reliability of testing the speech reception threshold for sentences. *Audiology*, *18*, 43-52.

Rönnberg, J. (1990). Cognitive and communicative function: The effect of chronological age. *European Journal of Cognitive Psychology*, *2*, 253-275.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology*, *42*(Suppl. 1), S68-S76.

Rönnberg, J., Arlinger, S., Lyxell, B., & Kinnefors, C. (1989). Visual evoked potentials: Relation to adult speechreading and cognitive function. *Journal of Speech and Hearing Research*, *32*, 725-735.

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology*, *47*(Suppl. 2), S99-S105.

Rönnberg, J., Rudner, M., Lunner, T., & Zekveld, A. A. (2010). When cognition kicks in: Working memory and speech understanding in noise. *Noise and Health*, *12*, 263-269.

Rothauser, E. H., Chapman, W. D., Guttman, N., Nordby, K. S., Silbiger, H. R., Urbanek, G. E., & Weinstock, M. (1969). IEEE recommended practice for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics*, *17*, 227-246.

Rudner, M., Foo, C., Rönnberg, J., & 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. *Scandinavian Journal of Psychology*, *50*, 405-418.

Rudner, M., Foo, C., Sundewall-Thoren, E., Lunner, T., & Rönnberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users. *International Journal of Audiology*, *47*, S91-S98.

Rudner, M., Lunner, T., Behrens, T., Thoren, E. S., & Rönnberg, J. (2012). Working memory capacity may influence perceived effort during aided speech recognition in noise. *Journal of the American Academy of Audiology*, *23*, 577-589.

Rudner, M., Rönnberg, J., & Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *Journal of the American Academy of Audiology*, *22*, 156-167.

Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K. (2002). Listening in aging adults: From discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology*, *56*, 139-152.

Shelton, J. T., Elliott, E. M., Hill, B. D., Calamia, M. R., & Gouvier, W. D. (2009). A Comparison of Laboratory and Clinical Working Memory Tests and Their Prediction of Fluid Intelligence. *Intelligence*, *37*, 283.

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

Stenfelt, S., & Rönnberg, J. (2009). The Signal-Cognition interface: Interactions between degraded auditory signals and cognitive processes. *Scandinavian Journal of Psychology*, *50*, 385-393.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127-154.

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological Science*, *21*, 384-390.

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging*, *18*, 332-339.

Versfeld, N. J., Daalder, L., Festen, J. M., & Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *Journal of the Acoustical Society of America*, *107*, 1671-1684.

Wagener, K., Josvassen, J. L., & Ardenkjaer, R. (2003). Design, optimization and evaluation of a Danish sentence test in noise. *International Journal of Audiology*, *42*, 10-17.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology*, *49*, 51-79.

Wingfield, A., & Grossman, M. (2006). Language and the aging brain: Patterns of neural compensation revealed by functional brain imaging. *Journal of Neurophysiology*, *96*, 2830-2839.

Wingfield, A., & Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *Journal of the American Academy of Audiology*, *18*, 548-558.

Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2013). Task difficulty differentially affects two measures of processing load: The pupil response during sentence processing and delayed cued recall of the sentences. *Journal of Speech, Language, and Hearing Research*. Advance online publication. doi:10.1044/1092-4388(2012/12-0058). Retrieved from http://jslhr.asha.org/cgi/rapidpdf/1092-4388_2012_12-0058v1.pdf

Zekveld, A. A., George, E. L. J., Houtgast, T., & Kramer, S. E. (in press). Cognitive abilities relate to self-reported hearing disability. *Journal of Speech, Language, and Hearing Research*.

Zekveld, A. A., George, E. L., Kramer, S. E., Goverts, S. T., & Houtgast, T. (2007). The development of the text reception threshold test: A visual analogue of the speech reception threshold test. *Journal of Speech, Language, and Hearing Research*, *50*, 576-584.

Zekveld, A. A., & Kramer, S. E. (2013). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*. Manuscript submitted for publication.

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear and Hearing*, *31*, 480-490.

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear and Hearing*, *32*, 498-510.

Zekveld, A. A., Kramer, S. E., Kessens, J. M., Vlaming, M. S., & Houtgast, T. (2009). The influence of age, hearing, and working memory on the speech comprehension benefit derived from an automatic speech recognition system. *Ear and Hearing*, *30*, 262-272.

Zekveld, A. A., Kramer, S. E., Vlaming, M. S., & Houtgast, T. (2008). Audiovisual perception of speech in noise and masked written text. *Ear and Hearing*, *29*, 99-111.

Zekveld, A. A., Rudner, M., Johnsrude, I. S., Festen, J. M., van Beek, J. H., & Rönnberg, J. (2011). The influence of semantically related and unrelated text cues on the intelligibility of sentences in noise. *Ear and Hearing*, *32*, e16-e25.

Zekveld, A. A., Rudner, M., Johnsrude, I. S., Heslenfeld, D. J., & Rönnberg, J. (2012). Behavioral and fMRI evidence that cognitive ability modulates the effect of semantic context on speech intelligibility. *Brain and Language*, *122*, 103-113.

Zekveld, A. A., Rudner, M., Johnsrude, I. S., & Rönnberg, J. (2013). Working memory mediates the facilitative effect of semantically related cues on intelligibility. Manuscript submitted for publication.