Memory of spoken discourse masked by speech

P. Sörqvist[1,2,3], J. Rönnberg[1,2]

[1] Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Sweden

[2] Department of Behavioural Sciences and Learning, Linköping University, Sweden, jerker.ronnberg@liu.se

[3] Faculty of Engineering and Sustainable Development, Department of Building, Energy and Environmental Engineering, University of Gävle, Sweden, patrik.sorqvist@hig.se

INTRODUCTION

When we listen to a talker in optimal conditions we can follow what is said on the basis of bottom-up processes mapping the speech automatically onto our mental lexicon. However, when the speech signal is distorted, as when the target speech is masked by another voice, the speech segments no longer match the phonological representations in long-term memory and top-down working memory processes must be engaged to interpret what is said (Rönnberg 2003; 2008; 2010). Since the capacity of working memory varies inter-individually, people with high working memory capacity (WMC) should be better able to process and subsequently remember the message in the target signal. However, thus far no research has addressed how WMC modulates the effect of a to-be-ignored speech signal on auditory-verbal memory processes. In contrast, the role of WMC in susceptibility to the effects of irrelevant speech on visual-verbal memory processes is well explored (Sörqvist 2010). Furthermore, in the context of dichotic listening, high-WMC individuals are better able to ignore a speech stream presented to the to-be-ignored ear (Conway et al. 2001) and better able to divide attention across both ears (Colflesh & Conway 2007). Based on this literature, WMC should modulate the disruption of episodic long-term memory of materials conveyed by a speech signal that is masked by another speech signal. The primary purpose of the present study was to investigate the role of WMC to episodic long-term memory of spoken materials in a cocktail party context. Moreover, we intended to test which of two complex-span tasks (i.e., reading span and size-comparison span, typical tasks that measure WMC) is the better predictor of the effect.

METHOD

Participants

A total of 44 students (*M* = 25.91 years, *SD* = 6.51) at the University of Gävle participated in this experiment. They all reported normal hearing and normal or corrected-to-normal vision, and they received a small honorarium as compensation for their participation.

Materials

The reading span (RSPAN) task was adopted from Daneman and Carpenter (1980) and the size-comparison span (SICSPAN) task was adopted from Sörqvist et al. (2010). In RSPAN, the participants read sentences and determine whether they are normal (e.g., the boy ate a sandwich) or absurd (e.g., the banana ate an apple). After reading a set of sentences, the participants were probed to recall the first (e.g., boy banana) or the last (e.g., sandwich apple) words presented in each sentence. The number of words recalled is used as a measure of WMC. In SICSPAN, the participants switched between comparing the size of objects (e.g., "Is ELEPHANT smaller than MOUSE?") and encoding words (e.g., LION) for later recall. After a set of words, the participants were asked to recall the to-be-recalled words and to avoid recalling words from the distractor activity. All words within a set (i.e., to-be-recalled words and comparison words) belong to the same semantic category (e.g., Animals). This enhances item confusion and makes selective retrieval of the to-be-recalled items (and selective encoding of the to-be-recalled items into the memory-set) difficult. Two stories were used to measure episodic long-term memory. One story was about a fictitious culture called the "Timads" and the other was about a fictitious culture called the "Lobiks". Both stories consisted of 10 short paragraphs about different topics and each story was approximately 7.5 minutes long. The participants listened to the stories and afterwards they answered 20 questions about each story respectively. To create a masking speech sound, a third story about a fictitious culture called the "Ansarians" was recorded, but spoken by a different male. The masking speech was spectrally inverted around 2 kHz to create rotated speech. All sound stimuli were presented binaurally through headphones. The target speech was presented at approximately 60 dB(A) L_{eq} and the masking speech (normal or rotated) was presented at approximately 55 dB(A) L_{eq}. To obtain a measure of how well the participants could hear what was said in the two listening conditions (normal and rotated masking speech), the participants also performed a hearing test. Twelve sentences from the two stories used in the episodic long-term memory test (6 from the Timads and 6 from the Lobiks) were used as target materials for the hearing test.

Design and Procedure

A within-participants design was used. The participants completed the tasks in three consecutive blocks; first, the reading span task and the size-comparison span task; second, the hearing test; and third, the two episodic long-term memory tasks. All tasks and sound conditions were counterbalanced between participants.

Table 1: Mean scores (in percentage) on the hearing test and the episodic long-term memory (LTM) test in two listening conditions (normal vs. rotated masking speech)

	Masking sound						
	Normal speech		Rotated speech		Difference		
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*	*t*
Hearing	.85	.15	.99	.04	.14	.14	6.69**
Episodic LTM	.18	.13	.28	.17	.10	.13	4.93**

Note: the *t*-tests are based on a within-participants design with 44 participants

** $p<0.01$

Table 2: Hierarchal regression analyses of the relation of size-comparison span and reading span to the effects of masking speech on episodic long-term memory

	Relation to episodic long-term memory disruption		
Variable	β[a]	*t*	ΔR^2
Predictors entered alone			
Size-comparison span	.37	3.68**	.14
Reading span	.25	2.21*	.06
Predictors entered simultaneously			
Size-comparison span	.34	2.83**	.08
Reading span	.08	0.63	< .01

* $p<0.05$; ** $p<0.01$

[a] Standardized regression coefficient for predictors of episodic long-term memory in the normal speech condition after control for episodic long-term memory in the rotated speech condition and hearing in the normal speech condition

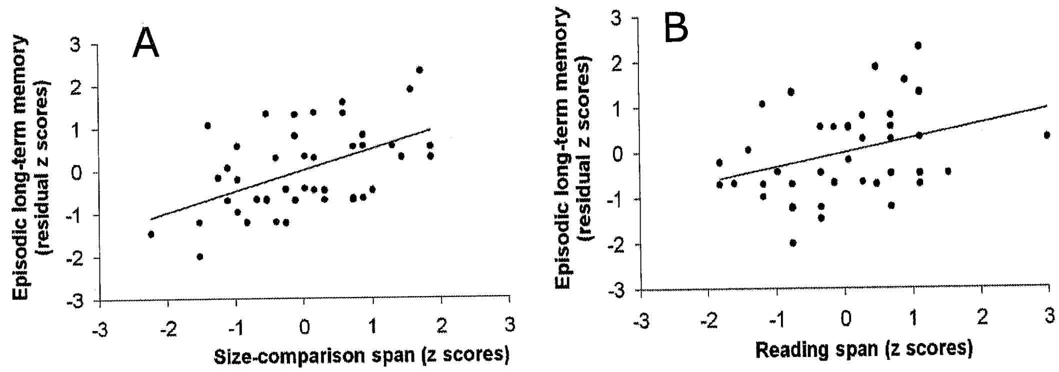

Figure 1: The figure illustrates the relationship between working memory capacity (Panel A: size-comparison span; Panel B: reading span) and episodic long-term memory for spoken discourse masked by speech (when variance explained by episodic long-term memory in the rotated-speech condition and hearing in the normal speech condition is statistically removed).

RESULTS

As can be seen in Table 1, performance on both the hearing test and the episodic long-term memory test was significantly impaired when a normal speech masked the target speech. This study intended to test if working memory capacity (WMC) modulates the effects of masking speech on episodic long-term memory of materials conveyed by a target speech signal. Moreover, we intended to test if SICSPAN is a stronger predictor of this relationship than RSPAN is. To this end, we used a series of residual analyses in the context of hierarchical regression analysis (cf. Sörqvist et al. 2010). The results are reported in Table 2 and Figure 1. Both WMC tasks predicted the effect significantly, but size-comparison span is the stronger predictor of the two.

DISCUSSION

The present study demonstrates an important role for cognitive-control processes in listening. Specifically, individual differences in WMC appear to underlie individual differences in susceptibility to the effects of to-be-ignored speech on auditory-verbal memory processes, consistent with previous investigations on cross-modal auditory distraction (Sörqvist 2010). This is particularly noteworthy because it suggests that individual differences in susceptibility to within-modal and cross-modal auditory distraction are mediated by the same top-down cognitive processes. Of particular interest here, size-comparison span (SICSPAN) appears to be a stronger predictor of the magnitude of this effect than reading span (RSPAN). When the participants perform SICSPAN they must recruit inhibitory processes to resolve semantic confusion (or competition taking place between category exemplars at encoding and at retrieval of target items) in contrast to RSPAN that does not require resolution of semantic confusion. This may be the reason why SICSPAN is the stronger predictor.

REFERENCES

Colflesh GJH, Conway ARA (2007). Individual differences in working memory capacity and divided attention in dichotic listening. Psychon Bull Rev 14: 699-703.

Conway ARA, Cowan N, Bunting MF (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. Psychon Bull Rev 8: 331-335.

Daneman M, Carpenter P (1980). Individual differences in working memory and reading. J Verbal Learn Verbal Behav 19: 450-466.

Rönnberg J (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. Int J Audiol 42: S68-S76.

Rönnberg J, Rudner M, Foo C, Lunner T (2008). Cognition counts: A working memory system for ease of language understanding (ELU). Int J Audiol 47: S171-S177.

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

Sörqvist P (2010). The role of working memory capacity in auditory distraction: A review. Noise & Health 12: 217-224.

Sörqvist P, Ljungberg JK, Ljung R (2010). A sub-process view of working memory capacity: Evidence from effects of speech on prose memory. Memory 18: 310-326.