Exceeding individual working memory capacity restrains aided speech recognition performance - effects in complex listening situations and effects of acclimatization

THOMAS LUNNER[1,2,3], ELISABET SUNDEWALL-THORÉN[1], MARY RUDNER[3], CATHARINA FOO[3] AND JERKER RÖNNBERG[3]

[1]*Oticon A/S, Research Centre Eriksholm, DK-3070 Snekkersten, Denmark*

[2]*Section of Technical Audiology, Department of Clinical and Experimental Medicine, Linköping University, S-58185, Sweden*

[3]*Swedish Institute of Disability Research and Department of Behavioural Sciences and Learning, Linköping University, S-58183, Sweden*

Two experiments were carried out that investigated individual working memory capacity and speech recognition performance in noise. Experiment 1 (Lunner and Sundewall-Thorén, 2007) investigated relations between individual working memory capacity and aided speech performance with fast and slow release times and in steady-state and modulated noise backgrounds. Experiment 2 (Rudner, Foo, Rönnberg and Lunner, 2007) investigated relations between individual working memory capacity and aided speech recognition performance under matched conditions (testing with acclimatized hearing-aid release times) and mismatched conditions (testing with new/unacclimatized hearing-aid release times). The results in both experiments indicate that if individual working-memory capacity is exceeded - either because of 'cognitive overload' due to acoustic variations or because of testing with 'cognitively mistuned' hearing-aid settings - speech recognition performance drops. Furthermore, the results suggest that laboratory testing under steady-state conditions may underestimate the role of cognition.

INTRODUCTION

Speech understanding and explicit cognitive processing

In an optimum listening situation, the speech signal is processed effortlessly and automatically. This means that the cognitive processing involved is largely unconscious and implicit. However, listening conditions are often suboptimal (e.g background sounds, reverberation, peripheral hearing loss), which means that implicit cognitive processes may be insufficient to uncover the meaning in the speech stream. Resolving ambiguities among previous speech elements and constructing expectations of prospective exchanges in the dialogue are examples of the complex processes that may arise. These processes are effortful and conscious and thus involve explicit cognitive processing.

Working memory and complex test conditions

A reasonable hypothesis is that working memory may be involved in speech recogni-

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.). ISBN: 87-990013-1-4. Print: Centertryk A/S.

tion in complex listening situations, since working memory includes both storage and processing aspects of incoming stimuli. Working memory is thought of as a capacity-limited system that stores recent visual-spatial, phonological and episodic information and at the same time provides a computational mental workspace in which the just stored information can be manipulated and integrated with knowledge stored in long-term memory.

EXPERIMENT 1: SIMPLE AND COMPLEX TEST CONDITIONS

The test conditions under which different signal-processing schemes are evaluated may be of importance for an individual hearing impaired person's apparent ability to utilize the signal processing. Under relatively steady-state conditions, performance is rather well predicted by the Speech Intelligibility Index (SII). However, the SII fails to predict speech intelligibility accurately in more complex situations. This suggests the importance of cognitive functions, such as attention and working memory, in more complex and natural listening situations.

Experiment 1 (Lunner and Sundewall-Thorén, 2007) investigates relations between individual working memory capacity and aided speech performance with fast and slow release times and in steady-state and modulated noise backgrounds.

Methods Experiment 1

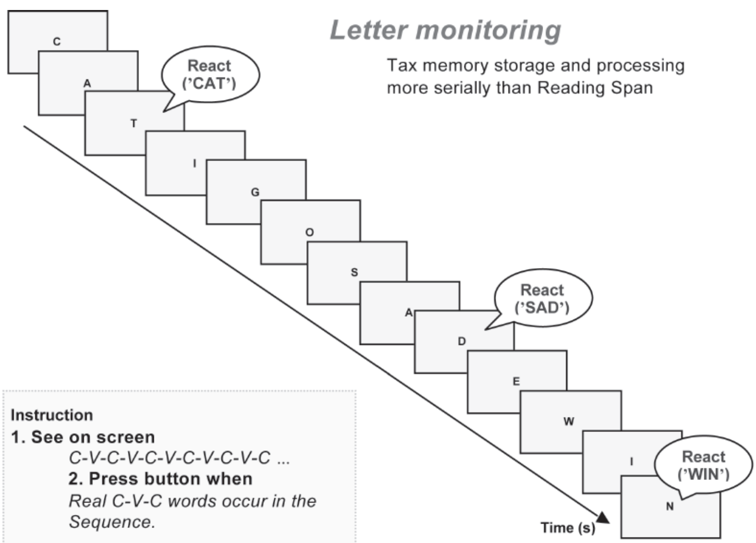

Fig. 1: Schematic of the Visual Letter Monitoring test (VLM).

Speech recognition in noise tests: Dantale 2 (Wagener *et al*, 2003). Background noises: Speech weighted unmodulated noise (Wagener *et al*, 2003) and 2-talker modulated noise (Dreschler *et al,* 2001). Hearing instruments and settings: Digifocus with fixed attack time settings of 10 ms and release time settings of either 40 ms (fast) or 640 ms (slow) in both channels. Cognitive tests/Working memory test: Danish Visual Letter Monitoring test (see, Fig. 1), VLM (Knutson *et al,* 1991; Gatehouse *et al* 2005, Lunner and Sundewall-Thorén, 2007). An outline of the test procedure is shown in Fig. 2. The number of participants was 23 with symmetric mild to moderate hearing loss.

Visit 1		Visit 2		Visit 3
Cognitive test	10 weeks	*Time Constant 1*	10 weeks	
Letter Monitoring		*Speech recognition test*		*Speech recognition test*
PTA6		Training		Training
Time Constant 1		Unmodulated noise		Unmodulated noise
		Modulated noise		Modulated noise
		Time Constant 2		

Fig. 2: Overview of experimental program Experiment 1.

Results Experiment 1

Fig. 3 shows the speech recognition in noise outcome for the four test conditions; (a) slow release time and unmodulated noise, (b) slow release time and modulated noise, (c) fast release time and unmodulated noise, (d) fast release time and modulated noise, for the test subjects grouped according to the visual cognitive test performance. Fig. 3 shows that under steady-state conditions (slow release time and unmodulated background noise) speech recognition in noise performance is similar across cognitive groups, while under fluctuating conditions (fast release times and modulated background noise) cognitively high performing improve their speech recognition in noise in contrast to the cognitively low performing who perform worse.

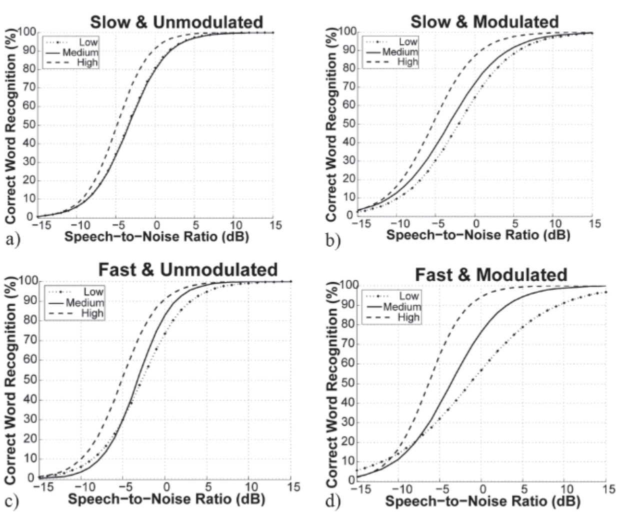

Fig. 3: Psychometric functions for cognitively *High, Medium* and *Low* performing subjects. a) *Slow* time-constant and unmodulated noise b) *Slow* time-constant and modulated noise c) *Fast* time-constant and unmodulated noise d) *Fast* time-constant and modulated noise. *(Figure from Lunner and Sundewall-Thorén, 2007)*.

Fig. 4 shows that pure tone hearing thresholds explains most variance under steady-state condition, while the cognitive measure explains most variance under fluctuation conditions. The results in Experiment 1 suggest that in a condition with slow-acting compression and unmodulated noise the test subjects' cognitive capacities are active, but do not exceed the capacity limit of most individual listeners. Thus, the individual peripheral hearing loss restrains performance and performance may be explained by audibility. Possession of greater cognitive capacity confers relatively little benefit. However, in the complex situation with fast-acting compression and varying background noise, much more cognitive capacity is required for successful listening. Thus, individual cognitive capacity restrains performance and speech-in-noise performance may, at least partly, be explained from individual working-memory capacity.

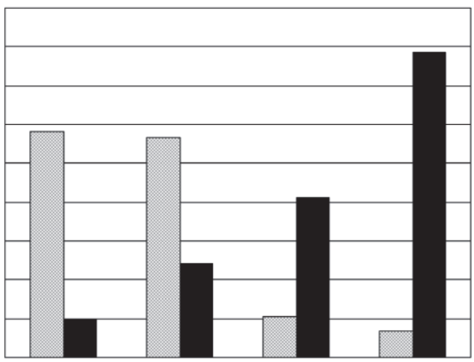

Fig. 4: Explained speech recognition in noise variance for different test conditions of background noise and hearing aid time constants. VLM=Cognitive test. PTA(6)=hearing loss. (*Figure from Lunner & Sundewall-Thorén, 2007*).

EXPERIMENT 2: AUDITORY ACCLIMATIZATION TO FAST AND SLOW RELEASE TIMES.

In view of the role of cognition in complex listening situations, it may be appropriate to consider auditory acclimatization (Gatehouse, 1992; Arlinger *et al,* 1996) in cognitive terms. According to this point of view, acclimatization can be considered a learning process, leading to increased automatization of the speech recognition process, and thus to a change in the balance of implicit and explicit cognitive processing, with a trend towards more implicit and less explicit processing. The interplay between implicit and explicit cognitive processing in language understanding is described by Rönnberg (2003) in terms of a working memory framework for Ease of Language Understanding (ELU).

Experiment 2 (Rudner *et al.*, 2007) investigates relations between individual working memory capacity and aided speech recognition performance under matched conditions (testing with acclimatized hearing-aid release times) and mismatched conditions (testing with new/unacclimatized hearing-aid release times).

Methods Experiment 2

The subjects were tested pre- and post 9 weeks of hearing aid use with fast and slow time constants, as in Experiment 1. Sixteen of the group of 32 hearing impaired test subjects used fast time constants, and 16 used slow time contstants, during the 9 week hearing aid use period. Speech recognition in noise tests: Hagerman sentences (Hagerman and Kinnefors, 1995) and Swedish HINT (Hällgren *et al,* 2006). Background noises: Speech weighted unmodulated noise (Hagerman, 1982) and 1-talker modulated noise (Hagerman, 2002). Hearing instruments and settings: Digifocus with fixed attack time settings of 10 ms and release time settings of either 40 ms (fast) or 640 ms (slow) in both channels. Cognitive tests/Working memory test: Swedish Reading Span (Daneman and Carpenter, 1980; Rönnberg, 2003; Lunner, 2003). Fig. 5 shows a schematic of the visual reading span test.

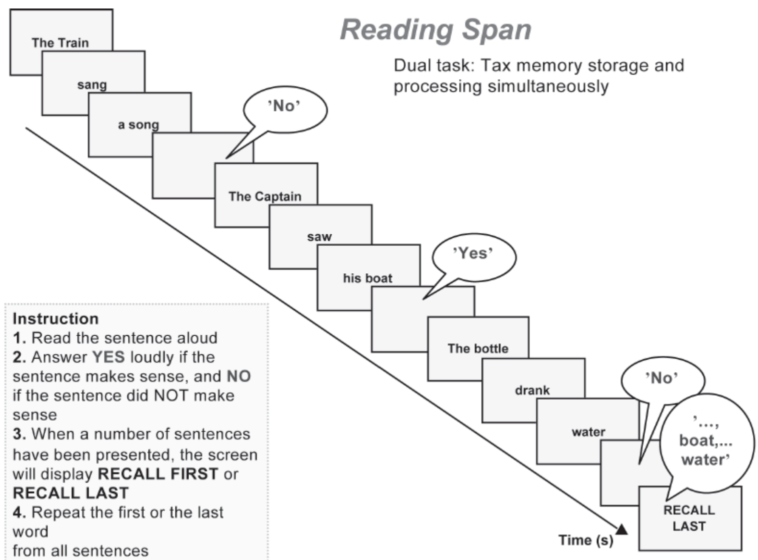

Fig. 5: Schematic of the Reading Span test.

Results Experiment 2

Table 1 shows the correlations between the visual Reading Span test and the auditory Hagerman sentences, for the different test conditions. The results indicate, regardless of type of noise and time constant, that under matched conditions (after 9 weeks) the correlations are *decreased* compared to before the acclimatization period, while in the correlations are *increased* under mismatch conditions. The same pattern was also indicated after correcting for pure tone average and age. The results from Experiment 2 indicate that switching to the new/unacclimatized setting, after a 9-week period with the acclimatized setting, will lead to more explicit cognitive processing and possibly poorer performance on aided speech recognition in noise if cognitive resources prove insufficient. For persons with lower working memory capacity, this may lead to poorer aided speech recognition in noise.

On the other hand, speech recognition in noise using the trained setting will become more implicit, and possibly more successful, due to acclimatization. This indicates that dual capacity (the Reading Span test, but not the Letter monitoring test) may contribute to the ability to reconcile differences between phonological information in the input signal and phonological representations in long-term memory, at least when the speech material is relatively predictable (i.e. Hagerman sentences, but not HINT). This finding supports the predictions of the working-memory framework for ELU (Rönnberg, 2003), regarding mismatch in circumstances when the message is relatively predictable. Moreover, we found that differences in correlation strength between match and mismatch conditions persisted in partial correlations with the effect of age removed. This suggests that the effect of mismatch is independent of age.

Furthermore, the correlations between Reading Span and Hagerman sentences replicates the findings by Lunner (2003).

Test setting	Fast						Slow					
Noise	Unmod			Mod			Unmod			Mod		
Experienced setting	Fast *n=15*	Slow *n=16*	None† *n=31*	Fast *n=15*	Slow *n=16*	None† *n=31*	Fast *n=15*	Slow *n=15*	None† *n=32*	Fast *n=15*	Slow *n=16*	None† *n=32*
Uncorrected	-.52*	**-.71****	-0.67**	-0.40	**-.73****	-0.65**	**-.69****	-0.21	-0.41*	**-.75****	-0.57*	-0.61**
Corrected for PTA_7	-0.36	**-0.59***	-0.58**	-0.17	**-.67****	-0.59**	**-0.61***	-0.15	-0.44*	**-.68****	-0.41	-0.50**
Corrected for PTA_7 and age	-0.21	**-0.44**	-0.47*	0.04	**-0.45**	-0.47*	**-0.44**	0.18	-0.43*	**-.56***	-0.15	-0.32**

Table 1: Correlations (Pearson's r) between performance on reading span and the Hagerman sentences in unmodulated and modulated noise, with fast and slow test settings, after nine weeks experience of either the fast or slow setting. Correlations a); partial correlations with the influence of PTA7 removed b); and with the influence of both PTA7 and age removed c). Bold underlined type indicates mismatch. * Correlation is significant at the 0.05 level (2-tailed). (Table form Rudner *et al*, 2007). † None indicates correlation coefficients before training at T1 reported in Foo *et al* (2007).

DISCUSSION AND CLINICAL IMPLICATIONS

The results in both experiments indicate that if individual working-memory capacity is exceeded - either because of 'cognitive overload' due to acoustic variations or because of testing with 'cognitively mistuned' hearing-aid settings - speech recognition performance drops. Furthermore, the results suggest that laboratory testing under steady-state conditions may underestimate the role of cognition.

Thus, we argue for the evaluation of hearing aids in more complex listening situations to better evaluate the influence of working-memory capacity limits. Speech recognition testing of hearing-impaired test subjects may then include natural variations in rate of speech (Tun and Wingfield, 1999), messages with different levels of context (Pichora-Fuller *et al*, 1995), different levels of comprehension (Hannon and Daneman, 2001), including 'natural' variations in reverberation and background noise, as well as 'natural' spatial listening conditions (Li *et al*, 2004; Freyman, 1999).

In addition, when switching to new/unacclimatized release time settings, nine weeks of experience seem to enable the cognitive system to adapt to the new release settings. Therefore it seems reasonable to suggest cognitive acclimatization periods before evaluating speech recognition performance with new release time settings, and possibly also with other signal processing schemes that may affect phonological representations.

REFERENCES

Arlinger, S., Gatehouse, S., Bentler, R. A., Byrne, D., Cox, R. M., Dirks, D. D. (**1996**). "Report of the Eriksholm Workshop on Auditory Deprivation and Acclimatization," Ear and Hearing, **17**(3), 87S-98S.

Daneman, M., Carpenter, P. A.. (**1980**). "Individual differences in Working Memory and reading," J Verb Learn Verb Behav **19**, 450-466.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., Westermann, S.. (**2001**) "ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment," International Collegium for Rehabilitative Audiology, Audiology **40**, 148-157.

Freyman, R. L., Helfer, K. S., McCall, D. D., Clifton, R. K.. (**1999**). "The role of perceived spatial separation in the unmasking of speech," J. Acoust. Soc. Am. **106**(6):3578-3588.

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," J. Am. Acad. Audiol. **18,** 618-631.

Gatehouse, S. (**1992**). "The time course and magnitude of perceptual acclimatization to frequency responses: Evidence from monaural fitting of hearing aids," J. Acoust. Soc. Am., **92**, 1258-1268.

Gatehouse, S., Naylor, G., Elberling, C. (**2006**). "Linear and non-linear hearing aid fittings – 2. Patterns of candidature," Int. J. Audiol. **45**, 153-171.

Hannon, B., Daneman, M. (**2001**). "A new tool for measuring and understanding individual differ-ences in the component process of reading comprehension," J. Educ. Psychol. **93**(1), 103-128.

Hagerman, B. (**1982**). "Sentences for testing speech intelligibility in noise," Scand. Audiol., **11**, 79–87.

Hagerman, B., and Kinnefors, C. (**1995**). "Efficient adaptive methods for measurements of speech reception thresholds in quiet and in noise," Scand. Audiol. **24**, 71–77.

Hagerman, B. (**2002**). "Speech recognition in slightly and fully modulated noise for hearing impaired subjects," Int. J. Audio., **41**(6), 321-329.

Hällgren, M., Larsby, B., and Arlinger, S. (**2006**). "A Swedish version of the Hearing in Noise Test (HINT) for measurement of speech recognition," Int. J. Audiol., **45**, 227-237.

Knutson, J. F., Schartz, H. A., Gantz, B. J., Tyler, R. S., Hinrichs, J. V., Woodworth, G. (**1991**). "Psychological change following 18 months of cochlear implant use," Ann Otol Rhinol Laryngol **100**:877-882.

Li, L., Daneman, M., Qi, J. G., Schneider, B.A. (**2004**) "Does the information content of an irrelevant source differentially affect spoken word recognition in younger and older adults?," J. Exp. Psychol. Hum. Percept. Perform. **30**(6), 1077-1091.

Lunner, T. (**2003**). "Cognitive function in relation to hearing aid use," Int. J. Audiol. **42**(Suppl 1), S49-S58.

Lunner, T., Sundewall-Thorén, E. (**2007**). "Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid," J. Am. Acad. Audiol., **18,** 604-617.

Pichora-Fuller, M. K., Schneider, B. A., Daneman, M. (**1995**). "How young and old adults listen to and remember speech in noise," J. Acoust. Soc. Am. **97,** 593-608.

Rudner, M., Foo, C., Rönnberg, J., Lunner, T. (**2007**). "Phonological mismatch makes aided speech recognition in noise cognitively taxing," Ear. Hear., **28**, 879-892.

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

Tun ,P. A., Wingfield, A. (**1999**). "One voice too many: adult age differences in language processing with different types of distracting sounds," J. Gerontol. B. Psychol. Sci. Soc. Sci. **54**(5), 317-327.

Wagener, K. C., Josvassen, J. L., Ardenkjaer, R. (**2003**). "Design, optimization and evaluation of a Danish sentence test in noise," Int. J. Audiol. **42,** 10-17.