Linköping University Post Print

Cognition and hearing aids.

Thomas Lunner, Mary Rudner and Jerker Rönnberg

N.B.: When citing this work, cite the original article.

The definitive version is available at www.blackwell-synergy.com:

Thomas Lunner, Mary Rudner and Jerker Rönnberg, Cognition and hearing aids., 2009, Scandinavian journal of psychology, (50), 5, 395-403. http://dx.doi.org/10.1111/j.1467-9450.2009.00742.x

Copyright: Blackwell Publishing

Postprint available at: Linköping University Electronic Press http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-51867

Lunner et al.	Cognition	and	hearing	aids.
Edillier of all	COSIMULO	uii u	110011115	urus.

Cognition and hearing aids

Thomas Lunner^{1,2,3,} Mary Rudner^{3,4}, and Jerker Rönnberg^{3,4}

1 Oticon A/S, Research Centre Eriksholm, Snekkersten, Denmark

- 2. Department of Clinical and Experimental Medicine, Linköping University, Sweden
- 3 Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Sweden
 - 4 Department of Behavioural Sciences and Learning, Linköping University, Sweden

Address for correspondence:

Thomas Lunner, PhD

Oticon A/S Research Centre Eriksholm

Kongevejen 243, DK-3070 Snekkersten

Denmark

Telephone +45 48 29 89 18

Fax +45 49 22 36 29

tlu@oticon.dk

Abstract

The perceptual information transmitted from a damaged cochlea to the brain is more poorly specified than information from an intact cochlea and requires more processing in working memory before language content can be decoded. In addition to making sounds audible, current hearing aids include several technologies that are intended to facilitate language understanding for persons with hearing-impairment in challenging listening situations. These include directional microphones, noise reduction, and fast-acting amplitude compression systems. However, the processed signal itself may challenge listening to the extent that with specific types of technology, and in certain listening situations, individual differences in cognitive processing resources may determine listening success. Here, current and developing digital hearing aid signal processing schemes are reviewed in the light of individual working memory (WM) differences. It is argued that signal processing designed to improve speech understanding may have both positive and negative consequences, and that these may depend on individual WM capacity.

Introduction

Advances in hearing aid technology are of great potential benefit to persons with hearing impairment. It is estimated that approximately 15% of the western population have a hearing impairment of such an extent that they would benefit from amplified hearing by way of hearing aids. Modern hearing aids incorporate technologies such as multiple-band wide dynamic range compression, directional microphones, and noise reduction. Individual settings for most of these functions are primarily based on pure-tone thresholds. Therefore, persons with hearing impairment with the same audiogram will receive similar hearing aid fitting even though they may have different supra-threshold auditory abilities relating to different pathologies or individual cognitive abilities. The research community has acknowledged that successful (re)habilitation of persons with hearing impairment must be individualized and based on understanding of underlying mechanisms, especially the mechanisms of cochlear damage and language understanding. This paper is based on recent data suggesting that ease of language understanding is highly dependent on the individual's working memory (WM) capacity in challenging speech understanding conditions, and focuses especially on a discussion of how different types of signal processing concepts in hearing aids may support or challenge the storage and processing functions of WM.

The main issues delineated in this paper concern the trade-off between individual WM capacity - seen from an intra-individual as well as an inter-individual perspective - and the benefits and costs involved in more advanced signal processing, as well as factors that modulate the signal-cognition interaction. The paper ends with a rather radical suggestion of a concept that takes the individual storage and processing function of WM into account to steer the function of the signal processing in the hearing aid.

Working memory and individual differences

This section is largely inspired by Pichora-Fuller (2007), and sets up the framework of WM differences under listening conditions that challenge cognitive capacity in different ways. When listening becomes difficult, e.g. because of irrelevant sound sources interfering with the target signal or because of a poorly specified input signal due to hearing impairment, listening must rely more on prior knowledge and context than would be the case when the incoming signal is clear and undistorted. This shift from mostly bottom-up (signal-based) to mostly top-down (knowledge-based) processing is accompanied by a sense of listening being more effortful.

In a review of different models of WM, Miyake and Shah (1999) concluded that many fitted the following generic description: Working memory is those mechanisms or processes that are involved in the control, regulation, and active maintenance of task-relevant information in the service of complex cognition, including novel as well as familiar, skilled tasks.

The WM model for Ease of Language Understanding (ELU, Rönnberg 2003; Rönnberg, Rudner, Foo & Lunner, 2008) proposes that under favourable listening conditions, language input can be rapidly and implicitly matched to stored phonological representations in long-term memory, whereas under suboptimum conditions, it is more likely that this matching process may fail. In such a mismatch situation, the model predicts that explicit, or conscious, cognitive processes must be engaged to decode the speech signal. Thus, under taxing conditions, language understanding may be a function of explicit cognitive capacity; whereas under less taxing conditions it may not.

WM has been proposed to consist of a number of different components including processing buffers (Baddeley, 1986, 2000), and individual differences in WM function (e.g. Engle, Cantor & Carullo,

1992) could relate to any of them. Indeed, researchers have investigated a variety of properties that contribute to individual differences in WM (e.g., resource allocation, Just & Carpenter, 1992; buffer size, Cowan, 2001; Wilken & Ma, 2004; processing capacity, Halford, Wilson & Phillips, 1998; Feldman Barrett, Tugade & Engle, 2004). In the following discussion it is assumed that, within the capacity constraint, resources can be allocated to either processing or storage, or both. A simple additive model is assumed;

$$C=P+S \tag{1}$$

where C is the available individual WM capacity, P is the processing component of WM, and S is the storage component of WM. This is schematically illustrated in figure 1b, where the black bars illustrate the processing (P) component, and the grey bars illustrate the storage (S) component. For a given C, the additive relationship defines how much of either P or S that will be left if the other is used. If the processing and storage demands of a particular task exceed available capacity this may result in task errors, loss of information from temporary storage (temporal decay of memories, forgetting) or slower processing.

For any given individual, the greater the demands made on the processing function of WM, the fewer resources can be allocated to its storage function. For example, distorting the signal or reducing the signal-to-noise ratio (SNR) or the availability of supportive contextual cues (e.g., Pichora-Fuller, Schneider & Daneman, 1995) would all increase processing demands with possible consequent reduction of available storage capacity. Thus, recall of words or sentences is better when target speech can be clearly heard (Rabbitt, 1968; Tun & Wingfield 1999; Wingfield & Tun 2001; Pichora-Fuller *et al.*, 1995).

Complex WM tasks require simultaneous storage (maintaining information in an active state for later recall) and processing (manipulating information for a current computation; Daneman &

Carpenter, 1980). In the reading span task, a WM task based on sentence processing, the participant reads a sentence and completes a task that requires trying to understand the whole sentence (by reading it aloud, repeating it, or judging it for some property such as whether the sentence make sense or not). Following the presentation of a set of sentences, the respondent is asked to recall the target word (such as the first or last word in the sentence) of each sentence in the set. The number of sentences in the recall set is increased and recall errors noted as a function of number of sentences in the set (WM span, WMS). The span score typically reflects the maximum number of target words that are correctly recalled. Span size is significantly correlated with language comprehension (Daneman & Carpenter, 1980; Daneman & Merikle, 1996). WM span measured in this way can also vary within individuals as a function of task (figure 1b).

a. Inter-individual differences in working memory capacity					
Fred					
Mary					
b. Intra-individual differences in working memory span (WMS): Allocation of resources to processing vs storage varies with task					
	Processing	Storage			
Fred in quiet			WMS = 6		
Fred in noise			WMS = 4		
Fred in more noise			WMS = 2		

Figure 1. Schematic representations of inter-individual differences in working memory capacity (a) suggesting that two individuals may differ in their working memory capacity, and (b) intra-individual differences suggesting that for a given individual the allocation of the person's limited capacity to the processing and storage functions of working memory varies with task demands. (Adopted from Pichora-Fuller, 2007.)

Working Memory and Hearing Loss

Speech recognition performance is affected for people with hearing impairment even under relatively favorable external SNR conditions (e.g., Plomp, 1988, McCoy et al., 2005; van Boxtel et al., 2000; Larsby, Hällgren, Lyxell & Arlinger, 2005). For persons with hearing loss, perceived listening effort, (as assessed by ratings of subjective effort in different situations), may indicate the degree to which limited WM resources are allocated to perceptual processing (Rudner, Lunner, Behrens, Sundewall Thorén & Rönnberg, 2009). Higher levels of perceived effort may indicate fewer resources for information storage, suggesting that listeners who are hard of hearing would be poorer than listeners with normal hearing on complex auditory tasks involving storage. Indeed, results by Rabbitt (1990) suggest that listeners who are hard of hearing allocate more information processing resources to the task of initially perceiving the speech input, leaving fewer resources for subsequent recall.

Figure 2 shows results from an experiment by Lunner (2003) with 72 patients who had similar levels of hearing loss as indicated by pure-tone audiograms. The participants' hearing aids were adjusted to assure audibility of the target signal and their speech reception thresholds (SRT) in noise were determined. SRT was defined as the level at which 50% of words presented were correctly recalled. Individual WM capacity, as measured by the reading span test (Andersson, Lyxell, Rönnberg & Spens, 2001; Daneman & Carpenter, 1980; Rönnberg, 1990), accounted for 40% of the inter-individual variance. That is, WMS was a good predictor of SRT. These findings have been confirmed in subsequent studies (Foo, Rudner, Rönnberg & Lunner, 2007; Rudner, Foo, Sundewall-Thorén, Lunner & Rönnberg, 2008; Akeroyd, 2008).

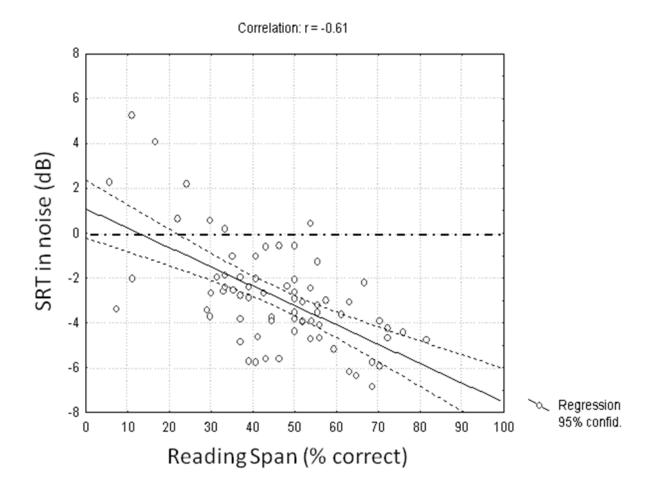


Figure 2. Scatterplot and regression line showing correlation between reading span and speech recognition in noise (n = 72). Shown are Pearsson correlations with 95% confidence limits for the correlation coefficient. Low (negative) SRT means high performance in noise. (Replotted from Lunner, 2003.)

Hearing aid signal processing and individual WM differences

Below, we review evidence indicating that some types of hearing aid signal processing may release WM resources, resulting in better storage capacity and faster information processing in challenging listening situations. However, hearing aid signal processing may also challenge listening by generating unwanted processing artifacts by distorting the auditory scene (i.e. the distinct auditory objects builds up the listening environment, see e.g. Shinn-Cunningham, 2008), or generating

audible artifacts and other unintended side-effects such as distortions of the target signal waveform (Stone & Moore, 2004; 2008), thereby taxing WM resources. The trade-off between WM benefits and signal processing artifacts may depend on the individually available cognitive resources, and therefore individual differences in cognitive processing resources may determine listening success with specific types of technology.

Inter-individual differences in capacity limitations constraining WM processing and storage may explain why one listening situation may be too challenging for one individual but not for another. Increases in WM span post hearing-aid intervention (i.e. intra-individual improvements in WM storage) would suggest that the intervention has resulted in listening becoming easier with fewer WM processing resources needing to be allocated.

Signal processing in hearing aids is designed to help users specifically in challenging listening situations. Usually the objective is, by some means, to remove signals that are less important in a particular situation and/or to emphasize or enhance signals that are more important. However, the consequences for the individual in terms of communicative benefit may depend on individual WM capacity. Several studies indicate that pure tone hearing threshold elevation is the primary determinant of speech recognition performance in quiet background conditions, e.g. in a conversation with one person or listening to the television under otherwise undisturbed conditions (see e.g., Dubno, Dirks & Morgan, 1984; Schum, Matthews & Lee, 1991; Magnusson, Karlsson & Leijon, 2001). Thus, in less challenging situations, individual differences in WM are possibly of secondary importance for successful listening. The individual peripheral hearing loss is the main constraint on performance, and the most important objective for the hearing aid signal processing is to make sounds audible. This can be by means of slow-acting compression (e.g. Dillon, 1996. Lunner, Hellgren, Arlinger & Elberling, 1997). A slow-acting compression system maintains near

constant gain-frequency response in a given speech/noise listening situation, and thus preserves the differences between short-term spectra in the speech signal. In less challenging listening situations, greater WM capacity confers relatively little benefit and the same is true of advanced signal processing designed to enhance target speech and/or to reduce interfering noise. In more challenging situations, however, signal processing designed to enhance speech and/or to reduce noise may – or may not - benefit the hearing aid user, depending on the implementation.

Even though speech recognition performance may not always be improved by the hearing aid signal processing, reductions in subjectively rated listening effort may result (e.g. Schulte *et al.*, 2009). SRT in noise is typically negative (see e.g. figure 2). Speech-to-noise ratios of 5dB or higher are realistic values for real-life conversation situations, such as conversing inside or outside urban homes (Pearsons, Bennett & Fidell, 1977). In such listening situations, conventional SRT tests are insensitive to signal processing improvements, and other measures such as subjective rating of listening effort (Schulte et al., 2009; Rudner *et al.*, 2009) or increases in WM span-scores post hearing-aid intervention may be a better predictor of hearing aid signal processing effects.

Directional microphones

Modern hearing aids can usually be switched between omni-directional and directional microphones. Directional microphone systems are designed to take advantage of the spatial differences between the relevant signal and noise. Directional microphones are more sensitive to sounds coming from the front than sounds coming from the back and the sides. The assumption is that because people usually turn their heads to face a conversational partner, frontal signals are most important, while sounds from other directions are of less importance. Several algorithms have been developed to provide maximum attenuation of moving or fixed noise source(s) behind the listener (see e.g. van den Bogaert *et al.*, 2008). Usually, switching between directional microphone and

omni-directional microphone takes place automatically in situations that are determined by the SNR-estimation algorithm to be beneficial for the particular type of microphone. The directional microphone usually comes into play when estimated SNR is below a given threshold value, and the target signal is estimated to be coming from the frontal position.

A review by Ricketts (2005) addressed the benefit of directional microphones compared to omnidirectional, showing that with the directional microphone, SNR improvement could be as high as 6-7 dB, and was typically 3-4 dB, in certain noisy environments. The noisy environments where directional benefit was seen were characterized by (a) no more than moderate reverberation, (b) the listener facing the sound source of interest, and (c) the distance to this source being rather short. The SRT in noise shows improvements in accordance with the SNR improvements (Ricketts, 2005). Thus, at least in particular situations, directional microphones give a clear and documented benefit.

However, if the target is not in front or if there are multiple targets, the attenuation of sources from directions other than frontal by directional microphones may interfere with the auditory scene (Shinn-Cunningham, 2008; Shinn-Cunningham & Best, 2008). In natural communication, the listener often switches attention to different locations. Therefore, omni-directional microphones may be preferred in situations requiring frequent shifts of attention or monitoring of sounds at multiple locations. Unexpected or unmotivated automatic switches between directional and omni-directional microphones may be cognitively disturbing if the switching interferes with the listening situation (Shinn-Cunningham & Best, 2008). Van den Bogaert *et al.* (2008) have shown that directional microphone algorithms substantially interfere with localization of target and noise sources, suggesting that directional microphones may, in addition to attenuating lateral sources, distort natural monitoring of sounds at multiple locations.

Sarampalis, Kalluri, Edwards, and Hafter (in press) investigated WM performance under different SNRs, ranging from -2 dB to +2 dB, to simulate the improvement in SNR by directional microphones compared to omni-directional microphones. The WM test was a dual-task paradigm with (a) a primary perceptual task involving repeating the last word of sentences presented over headphones, and (b) a secondary memory task involving recalling these words after each set of eight sentences (Pichora-Fuller *et al.*, 1995). The sentences were high- and low-context sentences from the Revised Speech Perception in Noise Test (Bilger et al., 1984). Performance on the secondary (memory) task improved significantly in the +2dB SNR condition which simulated directional microphones. The directional microphone intervention may have freed some WM resources, increasing storage capacity in the (tested) noisy situations.

Inter-individual and intra-individual differences in WM capacity may also play a role in determining the benefit of directional microphones for a given individual in a given situation. Consider, for example, figure 2, in a situation with 0 dB SNR (dash-dotted line). If we assume that the individual SRT in noise reflects the SNR at which WM capacity is severely challenged, figure 2 indicates that the WM capacity limit is challenged at about -5 dB for a high WM capacity person. At 0 dB SNR, the person with high WM capacity probably possesses the WM capacity to use the omni-directional microphone, while at -5 dB this person may need to sacrifice the omni-directional benefits and use the directional microphone to release WM resources. However, for the person with low WM capacity, even the 0 dB situation probably challenges WM capacity limits. Therefore, this person is probably best helped by selecting the directional microphone at 0 dB to release WM resources, thereby sacrificing the omni-directional benefits. Thus, it may be the case that the choice of SNR at which the directional microphone is invoked should be a trade-off between omni-

directional and directional benefits and individual WM capacity, and that inter-individual differences in WM performance may be used to individually set the SNR threshold at which the hearing aid automatically shifts from omni-directional to directional microphone.

Noise reduction systems

Noise reduction systems, or more specifically, single microphone noise reduction systems, are designed to separate target speech from disturbing noise by using a separation algorithm operating on the input. Different amplification is applied to the separated estimates of speech and noise, thereby enhancing the speech and/or attenuating the noise (e.g. Chung, 2004; Bentler & Chiou, 2006).

There are several approaches to obtaining separate estimates of speech and noise signals. One approach applied in current hearing aids is to use the modulation index (or modulation depth) as a basis for the estimation. The idea is that speech includes more level modulations than noise (see e.g., Plomp, 1994) and thus that the higher modulation index the greater the likelihood that a target signal has been identified. Algorithms to calculate the modulation index usually operate in several frequency bands. If a frequency band has a high modulation index, it is classified as including speech and is given more amplification, while frequency bands with less modulation are classified as noise and thus attenuated (see e.g., Holube, Hamacher & Wesselkamp, 1999). Other noise reduction approaches include the use of the level-distribution function for speech (Ludvigsen, 1997) or voice-activity detection by synchrony detection (Schum, 2003). However, the relative estimation of speech and noise components on a short-term basis (milliseconds) is very difficult, and misclassifications may occur. Therefore, commercial noise reduction systems in hearing aids are typically very conservative in their estimation of speech and noise components, and only give a

rather long-term estimation (seconds) of noise or speech. Such systems do not seem to aid speech recognition in noise (Bentler & Chiou, 2006). Nevertheless, typical commercial noise reduction systems do give a reduction in overall loudness of the noise compared to the target signal, which is rated as improving comfort (Schum, 2003) and thus may reduce the annoyance and fatigue associated with using hearing aids.

Noise reduction systems with more aggressive forms of signal processing are described in the literature, including 'spectral subtraction' or weighting algorithms where the noise is estimated either in brief pauses of the target signal or by modeling the statistical properties of speech and noise (e.g. Ephraim & Malah, 1984; Martin, 2001; Martin & Breithaupt, 2003; Lotter & Vary 2003; for a review see Hamacher *et al.*, 2005). The estimates of speech and noise are subtracted or weighted on a short-term basis in a number of frequency bands, which gives a less noisy signal. However, this comes at the cost of another type of distortion usually called 'musical noise' (Takeshi, Takahiro, Yoshihisa & Tetsuya, 2003). This extraneous signal may increase cognitive load during listening since it is a competing, and probably distracting signal, the suppression of which may consume WM resources. Thus, in optimizing noise reduction systems there is a trade-off between the amount of noise-reduction and the amount of distortion.

Sarampalis et al. (2006; 2008; in press) investigated the WM capacity of listeners with normal hearing and listeners with mild to moderate sensorineural hearing loss, using the dual-task paradigm described earlier. Auditory stimuli were presented with or without a short-term noise reduction scheme based on the algorithm proposed by Ephraim & Malah (1984). For people with normal hearing there was some recall improvement with noise reduction in low-context sentences. The authors interpreted this as demonstrating that the algorithm mitigated some of the deleterious effects

of noise by reducing cognitive effort. However, the results for the listeners with hearing impairment were not easily interpreted. More research is needed with regard to individual WM differences and short-term noise reduction systems to determine the circumstances under which these systems may release WM resources.

Another recent approach to the separation of speech from speech-in-noise is the use of binary timefrequency masks (e.g. Wang, 2005; Wang, 2008; Wang, Kjems, Pedersen, Boldt & Lunner, 2009). The aim of this approach is to create a binary time-frequency pattern from the speech/noise mixture. Each local time-frequency unit is assigned to either a 1 or a 0 depending on the local SNR. If the local SNR is favorable for the speech signal this unit is assigned a 1, otherwise it is assigned a 0. This binary mask is then applied directly to the original speech/noise mixture, thereby attenuating the noise segments. A challenge for this approach is to find the correct estimate of the local SNR. Ideal binary masks (IBM) have been used to investigate the potential of this technique for hearing impaired test subjects (Anzalone, Calandruccio, Doherty & Carney, 2006; Wang, 2008; Wang et al., in press). In IBM-processing, the local SNR is known beforehand, which is not the case in a realistic situation with non-ideal detectors of speech and noise signals. Thus, IBM is not directly applicable in hearing aids. Wang et al. (2009) evaluated the effects of IBM processing on speech intelligibility for listeners with hearing impairment by assessing the SRT in noise. For a cafeteria background, the authors observed a 15.6 dB SRT reduction (improvement) for listeners with hearing impairment, which is a very large effect. If individual SRTs reflect the situation where the WM capacity is severely challenged, applying IBM processing in difficult listening situations would release WM resources. However, IBM may produce distortions that increase the cognitive load, especially in realistic binary mask applications where the speech and noise are not available

separately, but have to be estimated. Thus, a trade-off may have to be made between noise reduction and distortion in a realistic noise reduction system.

In situations where the listener's cognitive system is unchallenged, using a noise reduction system may be redundant or even counterproductive, since distortion of the signal could outweigh any possible gain in SNR. However, since realistic short-term noise reduction schemes (including realistic binary mask processing) will rely on a trade-off between amount of noise reduction and minimization of processing distortions, the use of such systems may be dependent on the interindividual WM differences, suggesting that persons with high WM capacity may tolerate more distortions and thus more aggressive noise reduction than persons with low WM capacity in a given listening situation.

Fast acting wide dynamic range compression

A fast-acting wide dynamic range compression (WDRC) system is usually called fast compression or syllabic compression if it adapts rapidly enough to provide different gain-frequency responses for adjacent speech sounds with different short-term spectra. This contrasts with slow-acting WDRC (slow compression or automatic gain control). These systems maintain near constant gain-frequency response in a given speech/noise listening situation, and thus preserve the differences between short-term spectra in the speech signal. Hearing-aid compressors usually have frequency-dependent compression ratios, because hearing loss generally varies with frequency. However, WDRC can be configured in many ways, with different goals in mind (Dillon, 1996; Moore, 1998). In general, compression may be applied in hearing aids for at least three different reasons (e.g., Leijon & Stadler, 2008):

- 1. To present speech at comfortable loudness level, compensating for variations in voice characteristics and speaker distance.
- 2. To protect the listener from transient sounds that would be uncomfortably loud if amplified with the gain-frequency response needed for conversational speech.
- 3. To improve speech understanding by making also very weak speech segments audible, while still presenting louder speech segments at a comfortable level.

A fast compressor can to some extent meet all three purposes, whereas a slow compressor alone can fulfill only the first objective. Fast compression may have two opposing effects with regard to speech recognition: (a) it can provide additional amplification for weak speech components that might otherwise be inaudible, and (b) it reduces spectral contrast between speech sounds. It has yet to be fully investigated which of these effects has the greatest impact on speech recognition in noise *for the individual*, with regard to individual WM capacity. The first studies that systematically investigated individual differences in coping with the speed of compression were those of Gatehouse, Naylor and Elberling (2003, 2006a, 2006b). These studies indicated that both cognitive capacity and auditory ecology have explanatory value as regards individual outcome of e.g. speech recognition in noise and subjectively assessed listening comfort. In a study that replicated some of the findings of the Gatehouse *et al.* studies (figure 3, Lunner & Sundewall-Thorén, 2007), the cognitive test scores of listeners with hearing loss were significantly correlated with the differential advantage of fast compression versus slow compression in conditions of modulated noise. Other studies have shown that cognitive performance is related to the ability to cope with new compression settings (Foo *et al.*, 2007, Rudner *et al.*, 2008).

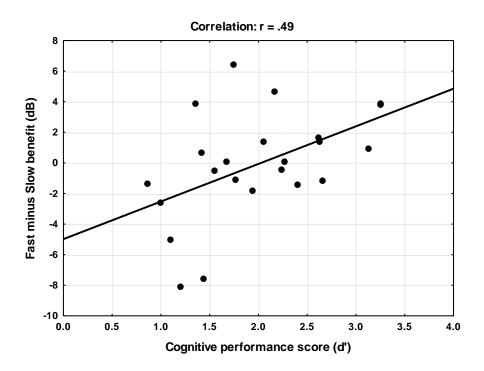


Figure 3. Scatter plot and regression line showing the Pearson correlation between the cognitive performance score and differential benefit in speech recognition in modulated noise of fast versus slow compression. A positive value on the Fast minus Slow benefit (dB) axis means that fast compression obtained better SRT in noise compared to slow compression (Replotted from Lunner & Sundewall-Thorén, 2007).

Figure 3 shows a scatterplot and regression line showing the Pearson correlation between the cognitive performance score and differential benefit of fast versus slow compression in speech recognition in modulated noise. The correlation in figure 3 is plausibly explained as an interaction between cognitive performance and fast compression as suggested by Naylor and Johannesson (2009). These authors have shown that the long-term SNR at the output of an amplification system that includes amplitude compression may be higher or lower than the long-term SNR at the input, dependent on interactions between the actual long term input SNR, the modulation characteristics of the signal and noise being mixed, and the amplitude compression characteristics of the system under test. Specifically, fast compression in modulated noise may increase output SNR at negative input SNRs, and decrease output SNR at positive input SNRs. Such shifts in SNR between input and output values may potentially affect perceptual performance for users of compression hearing

aids. The compression-related SNR shift affects perceptual performance in the corresponding direction (G. Naylor, R.B. Johannessen & F.M. Rønne, personal communication, December 2008); a person performing at negative SNRs may therefore be able to understand speech better with fast compression while the same may not be true for a person performing at positive SNRs. Thus, the relative SNR at which listening takes place is another factor which determines if fast compression is beneficial or not. A person with high WM capacity and SRT at a negative SNR would probably benefit from fast compression in that particular situation, while a person with low WM capacity and SRT at a positive SNR might be put at a disadvantage.

Cognition-driven signal processing

From the examples above it seems that inter-individual and intra-individual WM differences should be taken into account in the development of hearing-aid signal-processing algorithms and when they are adjusted for the individual hearing-aid user. Often it will be a case of balancing the trade-off between opposing effects in relation to the individual's WM capacity. For directional microphones the trade-off is between omni-directional and directional benefits; for realistic short-term noise reduction schemes it is between amount of noise reduction and processing distortion and for fast-acting versus slow compression it is between absolute performance levels in SNR and the choice to invoke fast compression to improve output SNR.

In less challenging situations, individual differences in WM are possibly of secondary importance for successful listening. The individual peripheral hearing loss is the main constraint on performance, and the most important objective for the design of hearing aid signal processing is to make sounds audible (e.g. by slow acting compression, e.g. Dillon, 1996).

In more challenging listening situations, hearing aid signal processing systems such as directional microphones, noise reduction systems, and fast compression should be activated on an individual basis to release WM resources, taking into account the above mentioned trade-offs between signal processing benefits and drawbacks. Below, a new and rather radical concept is suggested where knowledge of individual WM resources is combined with knowledge on hearing aids signal processing concepts.

One way to conceptualize the hearing aid processing requirements would be as a 'hearing aid with cognition-driven signal-processing', where the hearing aid signal processing is designed to take individual cognitive capacity into account to optimize speech understanding. The construction of such a cognition-driven hearing aid requires monitoring of the individual 'cognitive workload' on a real-time basis, in order to determine the level at which the listening situation starts to challenge WM resources. WM resources are challenged differently depending on listening situation, and different individuals may have different cognitive resources available to handle such specific workloads. Therefore, there is a need to develop monitoring methods for estimating cognitive workload. Two different lines of research can be foreseen; indirect estimates of cognitive workload and direct estimates of cognitive workload.

Indirect estimates of cognitive workload would use some form of cognitive model that is continuously updated with environment detectors that monitor the listening environment (e.g., level detectors, SNR detectors, speech activity detectors, reverberation detectors), as well as the conversational situation; the identity, mood and behaviour of the conversational partner as well as the purpose of the communication (social, information exchange) and feedback pattern. The cognitive model should produce at least two states, indicating cognitive High load or cognitive Low load. If cognitive High load is detected, hearing aid signal processing systems, such as directional

microphones, noise reduction systems and fast compression, should be invoked to release cognitive resources. The cognitive model needs to be calibrated with the individual cognitive capacity (e.g., WM capacity, verbal information processing speed), and connections between listening environment monitors, hearing aid processing system, and cognitive capacities have to be established. Inspiration might be found in the ease of language understanding (ELU) model of Rönnberg *et al.* (2008), which has a framework for suggesting when a listener's WM system switches from effortless implicit (bottom-up) processing to effortful explicit (top-down) processing.

However, a more direct way to assess cognitive workload would be through physically measurable correlates (e.g. Kramer, 1991). Given direct estimates of cognitive load, measures of cognitive High and Low load could be established. However, relations between environment characteristics, signal processing features and cognitive relief would still have to be established. A straightforward, but technically challenging example of a direct estimate of cognitive High and Low load could be obtained by electroencephalographic measurements (EEG, Gevins *et al.*, 1997). A wearable system has been proposed by Lan *et al.* (2007), which could be used to produce a cognitive state classification system based on EEG measurements. This could possibly be used to control the parameters of hearing aid signal processing algorithms to individually reduce cognitive 'workload' in challenging listening situations.

In summary, the concept of cognition-driven hearing aid signal processing is at the meeting point between the audiological and cognitive psychology disciplines, and mutual research is of great benefit to the development of our understanding of how hearing aid signal processing interacts with cognitive abilities. In the long term, a cognition-driven hearing aid could be beneficial not only for

optimizing signal processing but also for minimizing the negative impact of sensory impairment on cognitive function.

Acknowledgements

I would like to thank Stig Arlinger, Kathleen Pichora-Fuller, Graham Naylor, and two anonymous reviewers for their helpful and insightful comments on earlier versions of this manuscript.

References

- Akeroyd, M.A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology, 47* (Suppl. 2), S125-S143.
- Andersson, U., Lyxell, B., Rönnberg, J. & Spens, K-E. (2001). Cognitive Correlates of Visual Speech Understanding in Hearing-Impaired Individuals. *Journal of Deaf Studies and Deaf Education*, 6, 103 116.
- Anzalone, M. C., Calandruccio, L., Doherty, K. A. & Carney, L. H. (2006). Determination of the potential benefit of time-frequency gain manipulation. *Ear and Hearing*, 27, 480-492.
- Baddeley, A.D. (1986). Working memory. Oxford: Oxford University Press.
- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in* Cognitive Science 4(11), 417–423.
- 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.
- Bentler, R. & Chiou, L-K. 2006. Digital noise reduction: An overview. Trends in Amplification, 10(2), 67-82.

- Chung, K. 2004. Challenges and recent developments in hearing aids. Part I. Speech Understanding in Noise, Microphone Technologies and Noise Reduction Algorithms. *Trends in Amplification*, 8(3), 83-124.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185.
- Craik, F. I. M. (2007). Commentary: The role of cognition in age-related hearing loss. *Journal of the American Academy of Audiology*, 18, 539-547.
- Daneman, M. & Carpenter, P. A. (1980). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 9, 561-584.
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A metaanalysis. *Psychonomic Bulletin and Review*, *3* (4), 422-433.
- Dillon, H. (1996). Compression? Yes, but for low or high frequencies, for low or high intensities, and with what response times? *Ear and Hearing*, 17, 287-307.
- Dubno, J. R., Dirks, D. D. & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of America*, 76(1), 87–96.
- Engle, R. W., Cantor, J. & Carullo, J.J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 972-992.
- Ephraim, Y. & Malah, D. (1984). Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32(6), pp. 1109–1121.

- Feldman Barrett, L., Tugade, M. M. & Engle, R. W. (2004). Individual Differences in Working Memory Capacity and Dual-Process Theories of the Mind. *Psychological Bulletin*, 130(4), 553–573.
- 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, 553-566.
- Gatehouse, S., Naylor, G. & Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *International Journal of Audiology*, 42: (Suppl 1), S77-S85.
- Gatehouse S, Naylor G, Elberling C. (2006a) Linear and nonlinear hearing aid fittings 1. Patterns of benefit. *International Journal of Audiology*, 45, 130-152.
- Gatehouse, S., Naylor, G. & Elberling, C. (2006). Linear and non-linear hearing aid fittings 2. Patterns of candidature. *International Journal of Audiology*, 45, 153-171.
- Gevins, A., Smith, M. E., McEvoy, L. & Yu, D. (1997). High resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, 7(4), pp. 374–385.
- Hagerman, B. & Kinnefors, C. (1995). Efficient adaptive methods for measurements of speech reception thresholds in quiet and in noise. Scandinavian Audiology, 24, 71–77.
- Halford, G. S., Wilson, W. H. & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, 21, 803–865.
- Hamacher, V., Chalupper, J., Eggers. J., Fischer, E., Kornagel, U., Puder, H. & Rass U. (2005).Signal Processing in High-End Hearing Aids: State of the Art, Challenges, and Future Trends.EURASIP Journal on Applied Signal Processing, 18, 2915–2929.

- Holube, I., Hamacher, V. & Wesselkamp, M. (1999). Hearing Instruments: noise reduction strategies, in Proc. 18th Danavox Symposium: *Auditory Models and Non-linear Hearing Instruments*, Kolding, Denmark, September.
- Humes, L.E., Wilson, D.L. & Humes, A.C. (2003). Examination of differences between successful and unsuccessful elderly hearing aid candidates matched for age, hearing loss and gender.

 *International Journal of Audiology, 42, 432-441.
- Just, M.A. & Carpenter, P.A. (1992). A capacity theory of comprehension—individual differences in working memory. *Psychological Review* 99:122–149.
- Kemper, S., Herman, R.E. & Lian, C.H.T. (2003). The costs of doing two things at once for young and older adults: Talking, walking, finger tapping, and ignoring speech or noise. *Psychology and Aging*, 18, 181–192.
- Kramer, A.F. (1991). Physiological metrics of mental workload: A review of recent progress. In D.L. Damos (Ed.), Multiple-task performance. (pp. 279-328). London: Taylor & Francis.
- Lan, T., Erdogmus, D., Adami, A., Mathan, S. & Pavel, M. (2007). Channel Selection and Feature Projection for Cognitive Load Estimation Using Ambulatory EEG. *Computational Intelligence* and Neuroscience, Volume 2007, Article ID 74895, 1-12.
- Larsby, B., Hällgren, M., Lyxell, B. & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing impaired subjects. *International Journal of Audiology*, 44(3), 131–143.
- Leijon, A & Stadler, S. (2008). Fast amplitude compression in hearing aids improve audibility but degrades speech information transmission. Internal report 2008-11; Sound and Image Processing Lab., School of Electrical Engineering, KTH, SE-10044, Stockholm, Sweden
- Li, K.Z., Lindenberger, U., Freund, A.M. & Baltes, P.B. (2001). Walking while memorizing: Agerelated differences in compensatory behavior. *Psychological Science*, *12*(3), 230-237.

- Lotter, T. & Vary, P. (2003). Noise reduction by maximum a posteriori spectral amplitude estimation with super gaussian speech modeling," in *Proc. International Workshop on Acoustic Echo and Noise Control* (IWAENC '03), pp. 83–86, Kyoto, Japan, September 2003.
- Ludvigsen, C. (1997). Schaltungsanordnung für die automatische Regelung von Hörhilfsgeräten, Europäische Patentschrift, EP 0 732 036 B1.
- Lunner, T. (2003). Cognitive function in relation to hearing aid use. *International Journal of* Audiology, 42 (Suppl 1), S49-S58.
- Lunner, T., Hellgren, J., Arlinger, S. & Elberling, C.(1997). A digital filterbank hearing aid: Three DSP algorithms user preference and performance. *Ear and Hearing*, *18*, 373-387.
- 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. *Journal of the American Academy of Audiology, 18*, 539-552.
- McCoy, S.L., Tun, P.A., Cox, L.C., Colangelo, M., Stewart, R.A. & Wingfield, A. (2005). Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Quarterly Journal of Experimental Psychology A*, 58, 22-33.
- Magnusson, L., Karlsson, M. & Leijon, A. (2001) Predicted and measured speech recognition performance in noise with linear amplification. *Ear and Hearing*, 22(1), 46-57.
- Martin, R. & Breithaupt, C. (2003). Speech enhancement in the DFT domain using Laplacian speech priors, in *Proc. InternationalWorkshop on Acoustic Echo and Noise Control* (IWAENC '03), pp. 87–90, Kyoto, Japan, September 2003.
- Martin, R. (2001). Noise power spectral density estimation based on optimal smoothing and minimum statistics. IEEE Transactions on Speech and Audio Processing, 9(5), pp. 504–512.
- Miyake, A. & Shah, P. (1999). Models of Working Memory. Cambridge, UK: Cambridge University Press.

- Moore, B.C.J. (1998). A comparison of four methods of implementing automatic gain control (AGC) in hearing aids. *British Journal of Audiology*, 22, 93-104.
- Naylor, G. & Johannesson, R.B. (2009). Long-term Signal-to-Noise Ratio (SNR) at the input and output of amplitude compression systems. *Journal of the American Academy of Audiology*, 20 (3).
- Pearsons, K.S., Bennett, R.L., & Fidell, S. (1977). Speech levels in various environments (EPA-600/1-77-025). Washington, DC: Environmental Protection Agency.
- Pichora-Fuller, M.K. (2007). Audition and cognition: What audiologists need to know about listening. In C. Palmer & R. Seewald (eds.) *Hearing Care for Adults*. Stäfa, Switzerland: Phonak, pp. 71-85.
- Pichora-Fuller, M. K., Schneider, B.A. & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America*, 97, 593–608.
- Plomp, R. (1988). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *Journal of the Acoustical Society of America*, 63(2), 533–549.
- Plomp, R. (1994). Noise, amplification, and compression: considerations for three main issues in hearing aid design. *Ear and Hearing*, *15*, 2-12.
- Rabbitt, P. (1968). Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology*, 20, 241–248.
- Rabbitt, P. (1990). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Oto-laryngologica*, 476(Suppl), 167-176.
- Ricketts, T.A. (2005). Directional hearing aids: Then and now. *Journal of Rehabilitation Research* and Development, 42(4), 133-144.
- Rönnberg, J. (1990). Cognitive and communicative function: The effects of chronological age and "handicap age". *European Journal of Cognitive Psychology*, 2, 253-273.

- 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), S171-S177.
- 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, S68-S76.
- Rudner, M., Foo, C., Sundewall-Thorén, 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 (Suppl. 2), S163-S170.
- Rudner, M., Lunner, T., Behrens, T., Sundewall Thorén, E. & Rönnberg, J. (2009). Self-rated effort, cognition and aided speech recognition in noise. EFAS.
- Sarampalis, A., Kalluri, S., Edwards, B. & Hafter, E. (2006). Cognitive effects of noise reduction strategies. *International Hearing Aid Research Conference (IHCON)*, Lake Tahoe, CA, August.
- Sarampalis, A., Kalluri, S., Edwards, B. & Hafter, E. (2008). Understanding speech in noise with hearing loss: measures of effort. *International Hearing Aid Research Conference (IHCON)*, Lake Tahoe, CA, August 13-17.
- Sarampalis, A., Kalluri, S., Edwards, B. & Hafter, E. (In press). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language, and Hearing Research*.
- Schulte, M., Vormann, M., Wagener, K., Büchler, M., Dillier, N., Dreschler, W., Eneman, K., Froehlich, M., Grimm, G., Harlander, N., Hohmann, V., Houben, R., Jansen, S., Leijon, A., Lombard, A., Luts, H., Mauler, D., Moonen, M., Puder, H., Spriet, A., Wouters, J. (2009). Listening Effort Scaling and Preference Rating for Hearing Aid Evaluation. HearCom Workshop on Hearing Screening and new Technologies.

http://hearcom.eu/about/DisseminationandExploitation/Workshop.html

- Schum, D.J., Matthews, L. J. & Lee, F.S. (1991) Actual and predicted word-recognition performance of elderly hearing-impaired listeners. *Journal of Speech Hearing Research*, *34*, 636-642.
- Schum, D.J. (2003). Noise-reduction circuitry in hearing aids: (2) Goals and current strategies. *The Hearing Journal*. *56*(6), 32-40.
- Shinn-Cunningham, B.G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, 12(5), 182-186.
- Shinn-Cunningham BG & Best V (2008). Selective attention in normal and impaired hearing.

 Trends in Amplification, 12(4), 283-299.
- Stone, M.A. & Moore, B.C.J. (2004). Side effects of fast-acting dynamic range compression that affect intelligibility in a competing-speech task. *Journal of the Acoustical Society of America*, 116(4), 2311-2323.
- Stone, M.A. & Moore, B.C.J. (2008). Effects of spectro-temporal modulation changes produced by multi-channel compression on intelligibility in a competing-speech task. *Journal of the Acoustical Society of America*, 123(2), 1063-1076.
- Takeshi, H., Takahiro, M., Yoshihisa, I. & Tetsuya, H. (2003). Musical noise reduction using an adaptive filter. *Acoustical Society of America Journal*, 114 (4): 2370-2370
- Tun, P.A. & Wingfield, A. (1999). One voice too many: adult age differences in language processing with different types of distracting sounds. *Journals of Gerontology Series B:*Psychological Sciences and Social Sciences, 54(5), 317-327.
- van Boxtel, M.P., van Beijsterveldt, C.E., Houx, P.J., Anteunis, L.J., Metsemakers, J.F. & Jolles, J. (2000). Mild hearing impairment can reduce verbal memory performance in a healthy adult population. *Journal of Clinical Experimental Neuropsychology*, 22, 147-154.

- van den Bogaert, T., Doclo, S., Wouters, J. & Moonen, M. (2008). The effect of multimicrophone noise reduction systems on sound source localization by users of binaural hearing aids. *Journal Acoustical Society of America*, 124(1), 484-97.
- Wang, D. L. (2005). On ideal binary mask as the computational goal of auditory scene analysis. InP. Divenyi (Ed.), *Speech separation by humans and machines* (pp. 181-197). Norwell, MA:Kluwer Academic.
- Wang, D. L., Kjems, U., Pedersen, M. S., Boldt, J. B. & Lunner, T. (2009). Speech intelligibility in background noise with ideal binary time-frequency masking. *Journal of the Acoustical Society of America*, 125 (4), 2336-2347.
- Wang, D.L. (2008). Time-Frequency Masking for Speech Separation and Its Potential for Hearing Aid Design. *Trends in Amplification*, *12*, 332-353.
- Wilken, P. & Ma, W.J. (2004). A detection theory account of change detection. *Journal of Vision*, 4:1120-1135.
- Wingfield, A. &Tun, P.A. (2001). Spoken language comprehension in older adults: Interactions between sensory and cognitive change in normal aging. *Seminars in Hearing*, 22 (3), 287-301.