J Am Acad Audiol 18:618–631 (2007)

Recognition of Speech in Noise with New Hearing Instrument Compression Release Settings Requires Explicit Cognitive Storage and Processing Capacity

Catharina Foo*
Mary Rudner*†
Jerker Rönnberg*†
Thomas Lunner†‡§

Abstract

Evidence suggests that cognitive capacity predicts the ability to benefit from specific compression release settings in non-linear digital hearing instruments. Previous studies have investigated the predictive value of various cognitive tests in relation to aided speech recognition in noise using compression release settings that have been experienced for a certain period. However, the predictive value of cognitive tests with new settings, to which the user has not had the opportunity to become accustomed, has not been studied. In the present study, we compare the predictive values of two cognitive tests, reading span and letter monitoring, in relation to aided speech recognition in noise for 32 habitual hearing instrument users using new compression release settings. We found that reading span was a strong predictor of speech recognition in noise with new compression release settings. This result generalizes previous findings for experienced test settings to new test settings, for both speech recognition in noise tests used in the present study, Hagerman sentences and HINT. Letter monitoring, on the other hand, was not found to be a strong predictor of speech recognition in noise with new compression release settings.

Key Words: Speech recognition in noise, Hagerman sentences, HINT, compression release settings, modulated noise, unmodulated noise, explicit cognitive storage and processing capacity, working memory, reading span, letter monitoring

Abbreviations: AVC = Automatic Volume Control; CI = cochlear implant; CVC = consonant vowel consonant; FAAF = Four Alternative Auditory Feature Test; HINT = Hearing in Noise Test; PTA7 = pure tone average hearing threshold across the seven frequencies 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 3000 Hz, 4000 Hz, 6000 Hz; SNR = signal-to-noise ratio; WDRC = Wide Dynamic Range Compression

Sumario

La evidencia sugiere que la capacidad cognitiva predice la habilidad de beneficiarse de ajustes específicos de liberación de la compresión en instrumentos auditivos digitales no lineales. Estudios previos han investigado el valor de predicción de varias pruebas cognitivas en relación con el reconocimiento amplificado del lenguaje en medio de ruido utilizando ajustes

*Department of Behavioural Sciences and Learning, Linköping University, Sweden, †The Swedish Institute for Disability Research, Linköping University, Sweden, ‡Department of Technical Audiology, Linköping University, Sweden, §Oticon A/S, Research Centre Eriksholm, Snekkersten, Denmark

Address for correspondence: Catharina Foo, MA, Department of Behavioural Sciences and Learning, Linköping University, 581 83 Linköping, Sweden; Phone: +46 13 282133; Fax: +46 13 282145; E-mail: catharina.karlsson.foo@liu.se

de liberación de la compresión que han sido experimentados por un cierto período de tiempo. Sin embargo, el valor de predicción de la pruebas cognitivas con los nuevos ajustes, donde le usuario no ha tenido la oportunidad de acostumbrarse, no ha sido estudiado. En el presente estudio, comparamos el valor de predicción de dos pruebas cognitivas, lapso de lectura y monitoreo de letras, en relación con el reconocimiento amplificado del lenguaje en medio de ruido, para 32 usuarios habituales de dispositivos auditivos, usando nuevos ajustes de liberación de la compresión. Este resultado generaliza hallazgos previos para ajustes de prueba en sujetos con experiencia, pasados a nuevos ajustes, tanto para pruebas de reconocimiento de lenguaje en medio de ruido usadas en este estudio, como para frases de Hagerman y HINT. El monitoreo de letras, por otro lado, no se ha visto que sea un fuerte elemento de predicción para el reconocimiento del lenguaje en medio de ruido con nuevos ajustes de liberación de la compresión.

Palabras Clave: Reconocimiento del lenguaje en ruido, frases de Hageman, HINT, ajustes de liberación de la compresión, ruido modulado, ruido no modulado, habilidad cognitiva explícita de almacenamiento y procesamiento, memoria de trabajo, lapso de lectura, monitoreo de letras

Abreviaturas: AVC = Control Automático de Volumen; CI = implante coclear; CVC = consonante-vocal-consonante; FAAF = Prueba de Cuatro Rasgos Auditivos Alternativos; HINT = Prueba de Audición en Ruido; PTA7 = umbrales auditivos tonales puros en las siete frecuencias: 250, 500, 1000, 2000, 3000, 4000 y 6000 Hz; SNR = tasa señal/ruido, WDRC = Compresión de Rango Dinámico Amplio

Recent papers in audiology and related disciplines have reflected an increased interest in the role of cognition in hearing (Pichora-Fuller and Singh, 2006). Processing a message in noisy environments is always demanding, irrespective of the individual's sensory acuity. Therefore, explicit cognitive resources are required to (a) retrospectively resolve ambiguities of previous speech elements during a dialogue and also to (b) construct expectations of prospective exchanges in the dialogue. The general aim of this paper is to pursue an analysis of the contribution of explicit, higher order cognitive functions to the recognition of poorly specified speech signals.

In terms of cognitive processing, a well-specified speech signal is generally decoded without the engagement of conscious processes. This is denoted implicit processing. A poorly specified signal, on the other hand, often requires the engagement of conscious cognitive processes to be understood, hence the term explicit processing (Rönnberg, 2003a). Speech signals can be poorly specified in a number of different ways. For a deaf person, for example, the speech signal loses its auditory component and becomes purely visual. Visual features such as lip movements do not uniquely specify speech sounds (e.g. Summerfield, 1987). For a person fitted with a cochlear implant (CI), the auditory component is present but degraded (e.g. Lyxell et al, 1996).

When the speech signal is purely visual – as in speechreading – it has been repeatedly found that explicit tests of cognitive functions such as complex working memory (i.e., reading span) and verbal inference-making tests (i.e., sentence completion tasks) predict sentence-based speech understanding (Lyxell and Rönnberg, 1989, 1991, 1992). When the signal is audiovisual – as in CI-mediated speech understanding – again explicit tests of complex storage and processing functions in working memory appear to provide important information about the cognitive prerequisites for speech understanding (Lyxell et al, 1998; Lyxell et al, 1996). In both the above two examples, the quality of the signal is compromised in relation to a clear and audible speech presentation.

The extensive body of studies on lipreading, speechreading and speech understanding conducted at our laboratory, including around 1000 participants, has revealed a small set of cases of extremely skilled speechreaders, belonging to the top five of the whole sample. This set of cases of (visual or visual-tactile) speechreading demonstrates that the importance of higher-order complex storage and processing functions is invariant across communicative forms and habits, irrespective of age of onset of hearing impairment. These higher order, explicit and individually well-developed functions are combined with an individual threshold which sets the limit for what can implicitly be extracted from the speech signal (Rönnberg et al, 1998). The existence of such a threshold is presumed to be due to the efficiency of visual, phonetic and lexically supported perception being constrained by visual-neural and lexical access speed, as well as by the quality of phonological representations and the speed with which they are accessed (Pichora-Fuller, 2003; Andersson, 2001). Thus, we have shown in this series of case studies of extreme speech understanding skill (Lyxell, 1994 [the case of SJ, visual strategy]; Rönnberg, 1993 [the case of GS, visual-tactile strategy]; Rönnberg et al, 1999 [the case of MM, sign-speech bilingual background]) that individual performances on implicit parameters are generally on par with the control group. However, what really seems to account for their extreme communicative skills are their highly capacious, explicit storage and processing functions in working memory. This value of explicit functions is generalizable to both children (e.g. Lyxell and Holmberg, 2000) and adults (Rönnberg et al, 2000; but see Andersson and Lidestam, 2005, for the case of AA).

Rönnberg (2003a) formulated a more dynamic view of the interplay between implicit and explicit functions in language understanding. With a poor specification of the speech signal, a general mismatch mechanism is assumed to come into operation. The perceived signal is matched with long term memory representations phonologically and semantically. When a match occurs, lexical access proceeds smoothly and implicitly. When a mismatch between input and stored representations is present, a mismatch signal is produced in the system, calling for explicit processing and storage resources. The mismatch function will determine the ratio between implicit and explicit functions during speech understanding, and a highly capacious explicit, working memory function is assumed to contribute to the ease with which language is understood (a working memory system for Ease of Language Understanding [ELU], Rönnberg, 2003a).

A next logical step in the study of cognition in hearing, would be to assess whether explicit cognitive processing in persons with hearing impairment, measured by tests of complex working memory function, contributes to processing of speech in noise mediated by hearing instruments. In particular, we focus not on impoverished signals per se, but on signal processing in the presence of noise. One of the first "existence proofs" of a contribution of cognitive function in hearing impaired individuals to different amplification algorithms during signal-in-noise processing was demonstrated for cognitive functions assessed by a letter monitoring test and for benefit scores using the Four Alternative Auditory Feature Test (FAAF; Foster and Haggard, 1987) while comparing the potential benefit of fast Wide Dynamic Range Compression (WDRC) amplification settings with slow Automatic Volume Control (AVC) in noises with a temporal envelope (Gatehouse et al, 2003). The results suggest that a cognitively skilled person could better utilize the temporal dips in the noise to uncover the signal. The importance of relevant cognitive skills was again demonstrated by Lunner and Sundewall-Thorén (this issue) where the benefit of fast over slow compression release settings showed a high association between speech recognition in noise and cognitive function assessed by the letter monitoring test.

In a related vein, Lunner (2003) found direct evidence to suggest a relatively robust correlation (r = -.61) between performance on the reading span test (Daneman and Carpenter, 1980; Rönnberg et al, 1989) and aided speech recognition in noise for sentences from the Hagerman test (Hagerman and Kinnefors, 1995). The negative value of the correlation reflects the fact that while better performance on the reading span test results in a higher score, better performance on the Hagerman test results in a lower score in terms of signal-to-noise ratio (SNR) measured in dB. Gatehouse and colleagues (2003, 2006), Lunner (2003) and Lunner and Sundewall-Thorén (this issue) all computed correlations between speech recognition in noise performance and performance on cognitive tests after the participants had experience of the particular hearing instrument compression release settings used during testing.

In order to obtain a more complete picture of the role of cognition in hearing instrument use, it is important to examine the relationship between explicit cognitive processing and speech recognition in noise with new hearing instrument settings, before the users have become accustomed to them. Ultimately, this may lead to the development of reliable clinical tests of cognitive capacity that can be administered in connection with hearing instrument fitting to enhance quality assurance.

In the present study, we investigate how cognitive skill is related to the ability of hearing impaired persons to recognize speech in noise with new hearing instrument settings. We use the two cognitive tests used previously to investigate cognitive contributions to speech recognition, the letter monitoring test and the reading span test, to determine which is the better predictor of speech recognition in noise with new hearing instrument settings.

The letter monitoring test (Gatehouse et al, 2003; Knutsson et al, 1991) and the reading span test (Daneman and Carpenter, 1980; Rönnberg et al, 1989) both tap cognitive capacity but analytically address different components of cognitive processing. The letter monitoring test requires identification of lexical items of the form Consonant-Vowel-Consonant (CVC) in a continuous stream of presented letters. This procedure demands the memory capacity to store three letters, the phonological ability to combine three letters into a CVC item and the linguistic ability to determine whether the CVC item is stored in the lexicon. These steps can be performed serially. The reading span test, on the other hand, is based on semantic judgments of sentences and their subsequent recall, and requires parallel processing and storage. This procedure demands the memory capacity to store a number of words, the semantic ability to determine whether they combine to form a meaningful sentence and the additional memory capacity to store an increasing number of sentences while the first two processing steps continue. Thus, while the letter monitoring test taps memory capacity, phonological ability and lexical access serially, the reading span test is more clearly a dual task, tapping memory storage capacity and semantic processing abilities in parallel.

Beyond the empirical demonstrations, relatively little is known at this stage as to the nature of the cognitive functions involved in speech recognition under taxing conditions. However, Rönnberg's ELU framework (Rönnberg, 2003a) predicts that more explicit processing, particularly in the form operationalized by the reading span test, will be required with new hearing instrument settings, than with settings of which experience has been gained. This is because with a new hearing instrument setting, the probability of a mismatch signal is increased. In this paper, we investigate whether the predictive power of the reading span test and the letter monitoring test in relation to speech recognition in noise with experienced settings can be generalized to new settings and the relative predictive power of these two tests.

Two different tests of speech recognition in noise are used in the study to investigate possible generalizations across type of test, the Hagerman sentences (Hagerman and Kinnefors, 1995; cf. Lunner, 2003) and a Swedish version of the Hearing in Noise Test (HINT, Hällgren et al, 2006). Both these tests involve determining SNR required for recognizing sentences of similar length. However, they differ both as regards the nature of the sentences used and the way in which they are scored.

The Hagerman sentences constitute a speech in noise recognition test that is well established in Sweden. The HINT sentences represent a new paradigm that has recently been introduced in Sweden based on a test developed in the US. The Hagerman sentences have a stereotyped structure. They all consist of five words, each representing a different word class, in the following order: proper noun, verb, number, adjective and plural noun. The number of words within each word class is restricted to 10 items. In other words, the Hagerman sentences are constrained and thus predictable. The HINT sentences are varied in length, structure and content. They consist of between three and seven words, word class order is not predictable and neither is the semantic content of the words. Thus, there is no upper limit on the number of sentences acceptable within the HINT paradigm. In other words, the HINT sentences are highly unconstrained and thus less predictable. The HINT sentences have been developed to provide a more ecologically valid test of speech recognition in noise. They are considered to be more ecologically valid precisely because of their naturalistic variation which distinguishes them from the Hagerman sentences. On the other hand, they are low on internal context because none of the constituent words predicts the semantic content of any of the following words. It is likely that the lack of constraint in the form and content of the HINT sentences and their consequent unpredictability may make speech recognition in noise harder than with the Hagerman sentences. Similarly, different types of noise may interfere differently with recognition of the different sentence types.

MATERIAL AND METHODS

The methods and data reported here represent a subset from a wider experiment aimed at studying the role of cognition in aided speech recognition in noise. The subset reported here refers to the role of cognition in aided speech recognition in noise with new compression release settings.

Participants

Thirty-two hearing impaired participants (20 women and 12 men) with mild to moderate sensorineural hearing loss took part in the study. The participants were chosen from the patient population at the Linköping University Hospital and were invited to participate by letter. All gave their informed consent. The mean age of the participants was 70.3 years (SD = 7.7) and pure tone average hearing threshold across the frequencies 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 3000 Hz, 4000 Hz, and 6000 Hz (PTA7) was 46.0 dB (SD = 6.5) for both ears. All participants were habitual hearing instrument users; that is, they had used their hearing instruments for at least six hours a day and for at least one year.

The study was powered to detect a within-subject between-fitting difference of 0.5 dB on the mean scores across conditions on the speech recognition test described subsequently for $p< 0.05$ at 80% power. This required at least 23 complete data sets.

The study was approved by the regional ethics committee and all testing took place at Linköping University Hospital.

Hearing Instruments

The participants had previously (at least one year before the experiment) been bilaterally fitted and fine-tuned with behind-the-ear or in-the-ear digital two-channel, nonlinear signal processing hearing instruments, Oticon Digifocus 2. The instruments had the ASA2 fitting rationale (Elberling, 1996; Lunner et al, 1997; Schum, 1996), which means an attack time of 20 ms and a release time of 80 ms in the low frequency, and attack time of 10 ms and a release time of 320 ms in the high frequency channels. Hearing aid output was at the time of fitting verified by real-ear measurements. Thus, because own well-fitted hearing aids were utilized, no further fine-tuning was necessary prior to the experiment. In this experiment, the only setting that was manipulated experimentally was the release time of the compression; release time was either fast or slow. Theoretically attack time settings can also be manipulated, but this was not done in the present study. Altering time constants affects the dynamic behaviour of the hearing instrument. Slow release time gives quasi-linear amplification, which preserves syllable characteristics to a high degree; whereas fast release gives nonlinear amplification which results in syllabic compression and thus a somewhat distorted sound at syllable level (see e.g. Dillon, 2001). In the present study, speed of attack (10 ms) was held constant while speed of release was either fast, i.e. 40 ms in both the low frequency and high frequency channels, or slow, i.e. 640 ms in both channels.

Cognitive Tests

Reading span. In the reading span test (Rönnberg, 1990) sentences are presented word by word on a computer screen. After each sentence, the participant is instructed to determine whether the sentence is reasonable or not. Half the sentences are nonsense sentences e.g. "The horse sang loudly" and half the sentences are reasonable e.g. "The girl played on the beach". After a number of sentences have been presented, the participant is asked to recall a particular word that occurred in each of the sentences presented, either the first word or the last word, in serial order. The experimenter records the number of words correctly recalled in any order. This procedure is repeated with increasing numbers of sentences. First three sets of three sentences are presented, followed by three sets of four sentences, followed by three sets of five sentences and three sets of six sentences. Thus, there are 54 sentences in all. The participant's reading span score is the aggregated number of words correctly recalled across all sentences in the test and thus the maximum possible score is 54.

Letter monitoring. In the letter monitoring test (Gatehouse et al, 2003; Knutson et al., 1991) letters are presented individually in sequence on a computer screen for one second each. The letters are alternately consonants and vowels and some of the consonant-vowel-consonant (CVC) sequences form words. The task of the participant is to press a button each time one of the presented letters forms the final letter in a valid CVC sequence. A new Swedish version of the letter monitoring test was developed specially for the present study in order to carefully control the cognitive components tapped by the test. In order to control level of familiarity, the target words were common Swedish words chosen from the Swedish Academy Word List (http://spraakdata.gu.se/saol/) with similar frequency levels according to published corpus material (http://spraakbanken.gu.se). In order to control for material dependence, two different letter sequences were used and half the participants performed the test with one sequence and the other participants performed the test with the other. Each sequence contained 15 CVC targets randomly spaced. Responses were recorded automatically and a score was calculated for each participant based on rate of valid button presses minus the rate of invalid button presses according to the following formula:

$(w/y)-(x/(z-y))$

Where

w = valid button presses

x = invalid button presses

y = number of targets in list

z = number of letters in list

Speech Recognition Tests

Hagerman sentences (*Hagerman and Kinnefors*, 1995). This test is based on a set of sentences which consist of five words each, for example, "Ann had five red boxes." The sentences are prerecorded by a female speaker and are presented in randomized order for each participant in the presence of noise. For each sentence, the participant is instructed to repeat as many of the words as possible. The experimenter records the number of correct words repeated on a computer terminal. On the basis of this score the SNR is automatically adapted on the basis of an algorithm (Brand, 2000) to determine the SNRs required for 50% and 80% levels of performance respectively. The 80% level of performance has previously proved to be most sensitive to cognitive capacity (Lunner, 2003) and we therefore chose to report this level in the present study.

HINT (*Hällgren et al, 2006*). This test is based on a set of sentences which consist of three to seven words, for example, "The old man read a book" and "She got shampoo in her eyes". The sentences are prerecorded by a female speaker and are presented in randomized order for each participant in the presence of noise. For each sentence, the participant is instructed to repeat as many of the words as possible. The experimenter records the number of correct words repeated on a computer terminal. An automatic adaptive up-down procedure determines the SNR. The first sentence in each trial is presented below threshold (-8 dB SNR) and is increased by 2 dB steps until the whole sentence is correctly repeated. The following sentences in the list are presented once each, with presentation level dependent on the result of the previous response. Sentence presentation levels are reduced by 2 dB after a correct response and increased by 2 dB after an incorrect response. Scoring is based on correctly reported whole sentences.

Noise

During the speech recognition tests, the speech signal is held constant at the level of 65 dB, and SNR is adjusted by varying the level of background noise. Background noise is either unmodulated or modulated. The unmodulated speech-weighted noises have the same long-term average speech spectrum as their respective sentence material, Hagerman and HINT. The modulated noise, on the other hand, is generated according to different principles for the two sentence types. Generally speaking, noise modulation is designed to simulate conversational interference during speech recognition. In the present study, the degree of conversational interference is manipulated by generating modulated noise in different ways. For the Hagerman sentences, the modulated noise is the fully modulated noise as used by Hagerman (2002). This means that the speech and noise are locked to each other before editing, which involves smoothing the noise at the splices, resulting in the same local SNR for any specific word in all lists. This avoids different overall levels of difficulty for different lists of Hagerman sentences. The modulated noise used with the Hagerman sentences is based on speech produced by one person and thus represents the degree of interference attributable to a single speaker. This modulated noise has the same long-term average spectrum as the speech material, Hagerman sentences. Regarding the Swedish HINT sentences, no dedicated modulated noise was available so the two-talker modulated noise from the ICRA-CD (ICRA-track 6; Dreschler et al, 2001) was chosen. The ICRA-6 noise does not have the same long-term average speech spectrum as the HINT sentences; the high-frequency range (above 2 kHz) for the ICRA-6 noise is a few dB above that of the HINT sentences. The modulated noise used with the HINT sentences is based on speech produced by two speakers and thus represents the degree of interference attributable to up to two speakers. This means that the modulated noise used with the HINT sentences can be experienced as more cognitively demanding.

Test Administration

Participants were tested in three sessions held on three separate days, each one week apart. In session one, background data was recorded, PTA7 measured and the reading span test administered. In session two, speech recognition was tested, and in session three the letter monitoring test was administered. When testing took place, the hearing instrument settings that were manipulated experimentally were new to the participants. In other words, they had not had the opportunity to become accustomed to the settings.

The two speech recognition tests were administered with fast and slow release times and in modulated and unmodulated noise. Order of conditions was randomized. Testing took place in a sound-isolated chamber. The equipment in the chamber comprised a loudspeaker and a chair. The subjects were seated one meter in front of the loudspeaker (Tannoy System 800). PC equipment outside the chamber generated the acoustic signal via a 16-bit soundcard (Analog Devices SoundMax, HINT sentences) or through a CD player 288 followed by attenuators (Tucker-Davis PA-4 System 2, Hagerman sentences) and an audiometer (Interacoustics AC40) which was fed into a power amplifier (Rotel RB-03) and the loudspeaker.

Data Analysis

Test data was preprocessed as described under the various tests. For two out of the total of 128 individual scores on the Hagerman sentences, the slope of the individual performance curve between the 50% and 80% levels of performance was less than 2%, indicating test failure, perhaps due to fatigue. These scores were excluded. This accounts for the differences in n and degrees of freedom for some of the results reported. ANOVAs were computed for each sentence type to analyze speech recognition in noise with two types of noise and two compression release settings. Correlations were computed to identify the extent to which cognitive processing components tapped by the two cognitive tests, reading span and letter monitoring are involved in speech recognition under the various conditions. Where significant correlations were found, partial correlations were computed to control for the effects of PTA7 and age. Previous work has shown that patterns of aided hearing understanding under various conditions may differ for groups with different cognitive capacity (Gatehouse et al, 2003; Lunner, 2003). In view of this, we also computed ANOVAs for speech recognition with a between-groups factor based on median split of the participants for a) reading span performance and b) letter monitoring performance.

Results

Cognitive Tests

Reading span. The average aggregate reading span score was 23.8 (SD = 5.8). This level is similar to that shown using the same test for the experimental group in the study by Lunner (2003).

Letter monitoring. The average score for letter monitoring was 0.4 (SD = 0.3). There was no significant difference in performance levels between the two lists, indicating that performance is not material-specific on this newly developed version of the test.

Speech Recognition

Mean SNR (SD) for the various conditions is shown in Table 1.

Hagerman sentences. Performance was better with modulated noise than with unmodulated noise ($F(1, 28) = 20.4$, $MSE = 2.4$, $p < 0.001$). ANOVAs including cognitive grouping as a between-subjects factor showed an interaction between hearing instrument setting and cognitive grouping ($F(1, 28) = 4.7$, $MSE = 3.3$, $p < 0.05$), revealing a simple main effect of hearing instrument setting for the group that performed worse on the reading span test such that performance was poorer with the fast setting than the slow setting ($t(60) = 2.2$, $p < 0.05$). For the group that performed better on the reading span test, no such difference was found. No effect of letter monitoring performance was found.

HINT. Performance was better with unmodulated noise than with modulated noise ($F(1, 30) = 43.4$, $MSE = 0.7$, $p < 0.001$). It should be noted that the effect

Table 1. Mean (SD) SNR for Speech Recognition in Noise

Speech recognition test	Compression release setting	Noise	n	Mean SNR dB (SD)
Hagerman	Fast	Unmodulated	31	0.4 (2.7)
		Modulated	31	-0.2 (4.1)
	Slow	Unmodulated	32	1.0 (4.0)
		Modulated	32	-1.3 (3.2)
HINT	Fast	Unmodulated	32	2.4 (3.1)
		Modulated	32	3.1 (3.1)
	Slow	Unmodulated	32	2.5 (3.4)
		Modulated	32	3.8 (3.2)

of noise was different for the two different sentence types. While modulated noise gave better performance with the Hagerman sentences, unmodulated noise gave better performance with HINT. However, as different types of modulated noise were used for the two sentence types, a direct comparison is not possible. ANOVAs including cognitive grouping as a between-subjects factor showed an interaction between hearing instrument setting and cognitive grouping ($F(1, 30) = 5.1$, $MSE = 3.1$, $p < 0.05$), revealing a simple main effect of hearing instrument setting for the group that performed worse on the letter monitoring test such that performance was poorer with the slow setting than the fast setting ($t(60) = 3.8$, $p < 0.001$). For the group that performed better on the letter monitoring test, no such difference was found. No effect of reading span performance was found. It should be noted that the effect of cognitive grouping was different for the two different sentence types. While reading span interacted with hearing instrument setting, differentially affecting performance with the Hagerman sentences, letter monitoring interacted with hearing instrument setting, differentially affecting performance with HINT. Furthermore, although it was the weaker cognitive group that reacted differently to hearing instrument settings with both types of sentences, they generally performed worse with the fast setting on the Hagerman sentences and worse with the slow setting on HINT.

Correlations

There was a significant positive correlation between age and PTA7, as expected, and a strong negative correlation between age and reading span, which is also in line with previous results (Gatehouse et al, 2003, 2006; Lunner, 2003; Lunner and Sundewall-Thorén, this issue); see Table 2. The negative correlation between age and reading span agrees with findings that episodic memory capacity, on which working memory relies, declines with age (Nilsson, 2003). Reading span also correlated significantly with PTA7 (cf Lunner, 2003). Although letter monitoring correlated significantly with reading span, this measure did not correlate with either age or PTA7.

Reading span correlated negatively with speech recognition irrespective of hearing instrument setting, sentence type and type of noise; see Table 3. The correlation coefficient ranged from -0.41 to -0.67. Letter monitoring only correlated with speech recognition of the Hagerman sentences when a hearing instrument was used with a slow setting; see Table 3.

The correlation coefficient with modulated noise was -0.36, and with unmodulated noise it was -0.41.

Partial correlations between reading span and speech recognition were computed, controlling for PTA7, and age and PTA7 respectively. When PTA7 was controlled for, all but one of the correlations still held good; see Table 4. When both

Table 2. Correlations between Age, PTA_7 and Cognitive Measures, Pearson's r

	Age	PTA_7	Reading span
PTA_7	0.39*		
Reading span	-0.65**	-0.37*	
Letter monitoring	-0.28	0.04	0.38*

*Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

Table 3. Correlations between Cognitive Measures and Speech Perception, Pearson's r

Speech recognition test	Hagerman				HINT			
Compression release setting	Fast		Slow		Fast		Slow	
Noise	Unmod	Mod	Unmod	Mod	Unmod	Mod	Unmod	Mod
Reading span	-0.67**	-0.65**	-0.41*	-0.61**	-0.53**	-0.47**	-0.60**	-0.65**
Letter monitoring	-0.19	-0.27	-0.41*	-0.36*	-0.51	-0.38	-0.31	-0.29

*Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

Table 4. Partial Correlations between Reading Span and Speech Perception, Controlled for PTA_7 and PTA_7 and Age, Pearson's r

Speech recognition test	Hagerman				HINT			
Compression release setting	Fast		Slow		Fast		Slow	
Noise	Unmod	Mod	Unmod	Mod	Unmod	Mod	Unmod	Mod
PTA_7	-0.58**	-0.59**	-0.44*	-0.50**	-0.42*	-0.34	-0.51**	-0.58**
PTA_7 and Age	-0.47*	-0.47*	-0.43*	-0.32	-0.44*	-0.30	-0.47*	-0.62**

*Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

PTA7 and age were controlled for, only one more of the correlations lapsed, despite the fact that age is strongly correlated with reading span (r = -0.65); see Table 2.

DISCUSSION

Reading span predicted speech recognition performance under all conditions: irrespective of hearing instrument setting, type of noise and speech recognition in noise test, Hagerman sentences and HINT. This finding is in line with previous results which have shown the reading span test to be a good predictor of aided speech recognition in noise with experienced compression release settings (Lunner, 2003) and generalizes them to new settings; in other words, settings to which the user has not had the opportunity to become accustomed. Thus, aided speech recognition in noise with new compression release settings, along with aided speech recognition in noise with experienced compression release settings, can be added to the list of language processing abilities that are predicted by reading span. This list includes abilities such as reading comprehension (Daneman and Carpenter, 1980), primary school achievement (Gathercole et al, 2004), the ability to follow a telephone conversation after cochlear implantation (Lyxell et al, 1998), speechreading (Rönnberg, 2003a, 2003b) and visual-tactile speech recognition (Rönnberg, 1993). It is interesting to note that some of the correlations found in the present study for new compression release settings are stronger than those found previously between aided speech recognition in noise with experienced compression release settings and reading span (cf Lunner, 2003). Although strong conclusions cannot be drawn from this observation, it fits in with the prediction of the ELU framework (Rönnberg, 2003a) that more explicit cognitive processing should be involved in aided speech recognition in noise with new compression release settings than with experienced settings, reflected in stronger correlations with cognitive tests that tap explicit functions.

Reading span held good as a predictor of performance on the Hagerman sentences with hearing instruments when PTA7 was partialed out of the correlation. To some extent, the correlation held good when both PTA7 and age were partialed out, even though there was a significant correlation between age and reading span. This strongly suggests that the simultaneous storage and semantic processing capacity tapped by the reading span test is a key factor in unlocking the meaning of this type of sentence presented in noise with the assistance of a hearing instrument, irrespective of hearing loss, and, to some extent, age.

For the HINT sentences, with their more diverse form, reading span was also a good predictor across the board. This suggests that speech recognition in noise with HINT also taxes simultaneous storage and processing, even though they, unlike the Hagerman sentences, are not characterized by predictability and low ecological validity. The only exception to this rule was when release time was fast and noise modulated when corrections are made for PTA7 and age. The combination of fast compression release times and modulated noise has been identified as a particularly taxing speech recognition condition (Lunner and Sundewall-Thorén, this issue). One possible interpretation of the lack of significant correlation between reading span performance and HINT performance with fast release and modulated noise is that this condition is simply so taxing that greater cognitive capacity is of little benefit. This interpretation of the results can be further supported by the fact that different types of modulated noise are used for the two types of test. The modulated noise used with the HINT sentences, based on two speakers, can be experienced as more difficult, and thus more cognitively taxing than the noise used with the Hagerman sentences, based on one speaker. However, generally speaking, reading span is also a good predictor of HINT, indicating that the predictive power of the reading span test may be generalized to the more ecologically valid HINT test.

The other cognitive test included in the study, letter monitoring, did not show the same consistent pattern of strong correlation with aided speech recognition in noise as reading span. Although performance on the letter monitoring test correlated significantly with performance on the reading span test, it did so only relatively weakly, suggesting that the two tests are not tapping exactly the same cognitive functions. The letter monitoring test is a serial task that taps working memory capacity, phonological ability and lexical access. The reading span test is a dual task, tapping parallel memory storage capacity and semantic processing abilities. Performance on the letter monitoring test correlated significantly, and then only weakly, with performance on the Hagerman sentences with a slow release time. This shows that the version of the letter monitoring test used in the present study is not a good predictor of speech recognition in noise for a clinical population with new compression release settings. This finding complements previous results which have shown a predictive value of letter monitoring in relation to aided speech recognition in noise with compression release settings that the participants are used to, and suggests that testing letter monitoring ability in connection with hearing instrument fitting may not assure quality of fitting in the same way as testing reading span.

However, there may be other reasons why letter monitoring has proved to be less predictive in the present study. A new Swedish version of the letter monitoring test was developed for use in the present study with careful attention to frequency levels of targets. This descriptive detail has not been published in previous studies. It may be the case that versions of the letter monitoring test used previously are less carefully controlled on this point, resulting in greater cognitive load involved in adjudicating borderline cases of lexical decision. It should also be taken into consideration that this is the first time, to our knowledge, that a Swedish version of the letter monitoring test has been used. English (Gatehouse, 2006) and Danish (Lunner, 2003; Lunner and Sundewall-Thorén, this issue) versions have been used in previous studies and it is possible that the cognitive load involved in carrying out the test varies between languages. In either case, greater cognitive load would necessitate the engagement of parallel cognitive processes more similar to those involved in solving the reading span test than the more serial and phonological elements of the test which come to the fore when lexical decision is more clearcut. However, as the letter monitoring test is simple to administer in clinical situations, future work should focus on identifying the features of earlier versions of the test that have made them good predictors of speech recognition in noise.

In the present study, we found that, with unmodulated noise, performance on speech recognition in noise was superior for the Hagerman sentences compared to HINT. As the unmodulated noise used in the present study was generated according to the same principles for both speech recognition tests, and was thus comparable across tests, this result reflects the way in which the differences in form and content of the two types of sentences affect performance.

The modulated noise used in the present study differed for the two speech recognition tests. The two-talker ICRA noise used with HINT is deemed to be more cognitively taxing than the single-talker fully modulated noise used with the Hagerman sentences. Indeed, for HINT, speech recognition was more disrupted by modulated noise than unmodulated noise, while for the Hagerman sentences, the effect was the opposite. Noise modulation, which is designed to simulate conversational interference in speech recognition, results in noise level variation throughout testing which means that while some words may be obscured by noise, other words can be clearly perceived. The results in the present study, showing that performance on HINT was poorer than performance on the Hagerman sentences can be explained in terms of differences in the modulated noise. This difference is probably enhanced by the unpredictability of HINT.

Previous work has shown that cognitive skills are a major factor in aided speech recognition in noise and this is confirmed by the reading span correlations in the present study. When our experimental group was split on the basis of the two cognitive tests, interesting interactions were found with speech recognition that further our knowledge of the contribution of specific cognitive factors to speech recognition in noise. The subgroup that performed weakly on the letter monitoring test also performed more weakly on speech recognition with HINT with the slow setting. One of the key aspects of the letter monitoring test is phonological analysis. The slow hearing instrument setting gives less distortion of the speech signal than the fast setting thus allowing better opportunities for phonological analysis. Thus, we may be seeing the effect of poor phonological abilities affecting both letter monitoring and speech recognition with a slow hearing instrument setting. Further, this effect is only found for the more ecological HINT sentences which have greater phonological variation.

The subgroup that performed weakly on the reading span test also performed more weakly on speech recognition with the Hagerman sentences with the fast setting. The key aspect of reading span is parallel storage and semantic processing. When simultaneous storage and processing abilities are low, as indicated by poor performance on the reading span test, the ability to cope with a fast hearing instrument setting is also low for the Hagerman sentences. The Hagerman sentences are characterized by their low level of internal sentence context, which means that understanding these sentences in noise places extra demands on semantic processing at sentence level, which in turn demands parallel storage of individual items. When the hearing instrument setting is fast, the speech signal is distorted, which places further strain on semantic processing. Thus, reading span is a particularly good predictor of speech recognition in noise with the Hagerman sentences with a fast hearing instrument setting. This finding is supported by the correlation pattern and fits in with the theoretical model presented by Rönnberg (2003a) which predicts that explicit processing capacity, as measured by reading span, comes into play under particularly taxing conditions for language understanding.

CONCLUSION

The main finding of the present study is the robustness of the correlation patterns between reading span and aided speech recognition in noise for hearing impaired, habitual hearing instrument users with new compression release settings. This applies irrespective of hearing instrument setting, type of noise and speech recognition in noise test, Hagerman sentences and HINT. This demonstrates, once again, the role of explicit cognitive processing and storage in unlocking a poorly specified speech signal. More specifically, it suggests that the explicit cognitive processing and storage previously found to be associated with aided speech recognition in noise with experienced settings generalizes to new settings. This finding is in line with the ELU framework (Rönnberg, 2003a).

We have also found that the letter monitoring test does not have the same predictive value as regards aided speech recognition in noise with new settings as has previously been found with experienced settings. However, this may be due to the fact that the new version of the test used in the present study does not tap cognitive load to the same extent as versions used in previous studies.

An overall appraisal of the results regarding performance in unmodulated noise suggests that the HINT test taxes phonological processing abilities in a way that the Hagerman sentences do not. This is probably due to lack of constraints on sentence structure associated with ecological validity. Further, reading span predicts performance on HINT as well as performance on the Hagerman sentences, even though HINT, unlike the Hagerman sentences, is not characterized by predictability and low ecological validity. In keeping with Hällgren and colleagues (2006) we suggest that HINT may be a more clinically relevant test due to its ecological validity. By implication, the reading span test is potentially clinically relevant, both because of its predictive value in relation to HINT, and because of its general potency as a predictor of speech recognition in noise.

Finally, when the group was split on the basis of letter monitoring, an interaction was obtained with hearing instrument setting for HINT sentences, suggesting that letter monitoring may have predictive value particularly in relation to HINT. However, the overall evaluation of the letter monitoring test as a potential cognitive predictor must await further studies.

Acknowledgments. The authors would like to thank all the participants who were so generous with their time. We would also like to thank Gunilla Wänström for assistance with data collection and Elisabet Sundewall-Thorén for comments on an earlier version of this manuscript.

REFERENCES

Andersson U. (2001) Cognitive deafness. The deterioration of phonological representations in adults with an acquired severe hearing loss and its implications for speech understanding. PhD diss. Studies from the Swedish Institute for Disability Research, No. 3.

Andersson U, Lidestam B. (2005) Bottom-up driven speechreading in a speechreading expert: the case of AA (JK023). *Ear Hear* 26(2):214–224.

Brand T. (2000) *Analysis and Optimization of Psychophysical Procedures in Audiology*. Oldenburg: Bibliotheks- und Informationssystem.

Daneman M, Carpenter PA. (1980) Individual differences in integrating information between and within sentences. *J Exp Psychol Learn Mem Cogn* 9:561–584.

Dillon H. (2001) *Hearing Aids*. Thieme: Stuttgart.

Dreschler WA, Verschuure H, Ludvigsen C, Westerman 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.

Elberling C. (1996) A new digital hearing instrument. *Hear Rev* May:38–39.

Foster JR, Haggard MP. (1987) The Four Alternative Auditory Feature Test (FAAF) linguistic and psychometric properties of the material with normative data in noise. *Br J Audiol* 21:165–174.

Gatehouse S, Naylor G, Elberling C. (2003) Benefits from hearing aids in relation to the interaction between the user and the environment. *Int J Audiol* 42(Suppl. 1):S77–S85.

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

Gathercole SE, Pickering SJ, Knight C, Stegmann Z. (2004) Working memory skills and educational attainment: evidence from national curriculum assessments at 7 and 14 years of age. *Appl Cog Psychol* 18(1):1–16.

Hagerman B, 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 Audiol* 41(6):321–329.

Hällgren M, Larsby B, 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 JF, Schartz HA, Gantz BJ, Tyler RS, Hinrichs JV, Woodworth G. (1991) Psychological change following 18 months of cochlear implant use. *Ann Otol Rhinol Laryngol* 100:877–882.

Lunner T, Hellgren J, Arlinger S, Elberling C. (1997) A digital filterbank hearing aid: three DSP algorithms—user preference and performance. *Ear Hear* 18:373–387.

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

Lyxell B. (1994) Skilled speechreading—a single case study. *Scand J Psychol* 35:212–219.

Lyxell B, Arlinger S, Andersson J, Bredberg G, Harder H. (1998) Phonological representation and speech understanding with cochlear implants in deafened adults. *Scand J Psychol* 39:175–179.

Lyxell B, Arlinger S, Andersson J, Harder H, Näsström E, Svensson H, Rönnberg J. (1996) Information-processing capabilities and cochlear implants: pre-operative predictors for speech understanding. *J Deaf Stud Deaf Educ* 1:190–201.

Lyxell B, Holmberg I. (2000) Visual speechreading and cognitive performance in hearing impaired and normal hearing children (11–14 years). *Br J Educ Psychol* 70:505–518.

Lyxell B, Rönnberg J. (1989) Information-processing skills and speechreading. *Br J Audiol* 23:339–347.

Lyxell B, Rönnberg J. (1991) Visual speech processing: word decoding and word discrimination related to sentence-based speechreading and hearing-impairment. *Scand J Psychol* 32:9–17.

Lyxell B, Rönnberg J. (1992) The relationship between verbal ability and sentence-based speechreading. *Scand Audiol* 21:67–72.

Nilsson L-G. (2003) Memory function in normal aging. *Acta Neurologica Scandinavica* 107(Suppl. 179):7–13.

Pichora-Fuller MK. (2003) Processing speed and timing in aging adults: psychoacoustics, speech perception, and comprehension. *Int J Audiol* 42:S59–S67.

Pichora-Fuller MK, Singh G. (2006) Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiological rehabilitation. *Trends Amplif* 10(1):29–59.

Rönnberg J. (1990) Cognitive and communicative function: the effects of chronological age and "handicap age." *Eur J Cog Psychol* 2:253–273.

Rönnberg J. (1993) Cognitive characteristics of skilled Tactiling: the case of GS. *Eur J Cog Psychol* 5:19–33.

Rönnberg J. (2003a) 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. (2003b) Working memory, neuroscience and language: evidence from deaf and hard of hearing individuals. In: Marschark M, Spencer P, eds. *The Handbook of Deaf Studies, Language, and Education*. Oxford: Oxford University Press, 478–489.

Rönnberg J, Andersson J, Samuelsson S, Söderfeldt B, Lyxell B, Risberg J. (1999) A speechreading expert: the case of MM. *J Speech Lang Hear Res* 42:5–20.

Rönnberg J, Arlinger S, Lyxell B, Kinnefors C. (1989) Visual evoked potentials: relation to adult speechreading and cognitive function. *J Speech Hear Res* 32:725–735.

Rönnberg J, Samuelsson E, Borg E. (2000) Visual cognitive tests, central auditory function and auditory communication. *Scand Audiol* 29:196–206.

Rönnberg J, Samuelsson S, Lyxell B. (1998) Conceptual constraints in sentence-based lipreading in the hearing impaired. In: Campbell R, Dodd B, Burnham D, eds. *Hearing by Eye: Part II. Advances in the Psychology of Speechreading and Audiovisual Speech*. London: Lawrence Erlbaum, 143–153.

Schum D. (1996) Adaptive Speech Alignment: a new rationale made possible by DSP. *Hear J* 49:25–30.

Summerfield AQ. (1987) Some preliminaries to a comprehensive account of audiovisual speech perception. In: Dodd B, Campbell R, eds. *Hearing by Eye: The Psychology of Lipreading*. Hove, UK: Lawrence Erlbaum, 3–51.