Lost in Translation

Speech recognition and memory processes in native and non-native language perception

Lisa Kilman

Linköping Studies in Arts and Science No. 655

Studies from the Swedish Institute for Disability Research No. 74

Linköping University,

Department of Behavioral Sciences and Learning

Linköping 2015

Linköping Studies in Arts and Science • No. 655

Studies from the Swedish Institute for Disability Research • No. 74

At the Faculty of Arts and Science at Linköping University, research and doctoral studies are carried out within broad problem areas. Research is organized in interdisciplinary research environments and doctoral studies mainly in graduate schools. Jointly, they publish the series Linköping Studies in Arts and Science. This thesis comes from the Swedish Institute for Disability Research at the Department of Behavioral Sciences and Learning.

Distributed by:

Department of Behavioral Science and Learning

Linköping University

SE-581 83 Linköping

Lisa Kilman

Lost in Translation

Speech recognition and memory processes in native and non-native language perception

Edition 1:1

ISBN 978-91-7685-972-8

ISSN 0282-9800

ISSN 1650-1128

©Lisa Kilman

Department of Behavioral Science and Learning, 2015

Cover image: Designed and made by Gustav Kilman

Printed by LiU-tryck, Linköping 2015

When you change the way you look at things,

the things you look at change.

Tao Te Ching

Abstract

This thesis employed an integrated approach and investigated intra- and inter-individual differences relevant for normally hearing (NH) and hearing-impaired (HI) adults in native (Swedish) and non-native (English) languages in adverse listening conditions. The integrated approach encompassed the role of cognition as a focal point of interest as well as perceptual-auditory and linguistic factors.

Paper I examined the extent to which proficiency in a non-native language influenced native and non-native speech perception performance for NH listeners in noise maskers compared to native and non-native speech maskers. Working memory capacity in native and non-native languages and non-verbal intelligence were also assessed. The design of paper II was identical to that of paper I, however the participants in paper II had a hearing-impairment. The purpose of paper III was to assess how NH and HI listeners subjectively evaluated the perceived disturbance from the speech- and noise maskers in the native and non-native languages.

Paper IV examined how well native and non-native stories that were presented unmasked and masked with native and non-native speech were recalled by NH listeners. Paper IV further investigated the role of working memory capacity in the episodic long-term memory of story contents as well as proficiency in native and non-native languages.

The results showed that generally, the speech maskers affected performance and perceived disturbance more than the noise maskers did. Regarding the non-native target language, interference from speech maskers in the dominant native language is taxing for speech perception performance, perceived disturbance and memory processes. However, large inter- individual variability between the listeners was observed. Part of this variability relates to non-native language proficiency.

Perceptual and cognitive effort may hinder efficient long-term memory encoding, even when stimuli are appropriately identified at a perceptual level. A large working memory capacity (WMC) provides a better ability to suppress distractions and allocate processing resources to meet assigned objectives.

The relatively large inter-individual differences in this thesis, require an individualized approach in clinical or educational settings when non-native persons or people with

hearing impairment need to perceive and remember potentially vital information. Individual differences in the very complex process of speech understanding and recall need to be further addressed by future studies. The relevance of cognitive factors and language proficiency provides opportunities for individuals who face difficulties to compensate using other abilities.

List of papers

This thesis is based on the following papers, referred to in the text by their Roman numerals.

I. Kilman, L., Zekveld, A. A., Hällgren, M., & Rönnberg, J. (2014). The influence of non-native language proficiency on speech perception performance. *Frontiers in Psychology 5:651. doi:10.3389/fpsyg.2014.00651*

II. Kilman, L., Zekveld, A. A., Hällgren, M., & Rönnberg, J. (2015). Native and non-native speech perception by hearing-impaired listeners in noise and speech maskers. *Trends in Hearing* Vol. 19; 1-12.

III. Kilman, L., Zekveld, A. A., Hällgren, M., & Rönnberg, J. (2015). Subjective ratings of masker disturbance during the perception of native and non-native speech. *Frontiers in Psychology* 6:1065. doi: 10.3389/fpsyg.2015.01065.

IV. Kilman, L., Zekveld, A. A., Hällgren, M., & Rönnberg, J. Episodic long-term memory by native and non-native stories masked by speech. Manuscipt.

List of abbreviations

ANOVA	Analysis of variance
ELU	Ease of Language Understanding model
HI	Hearing-impaired
HINT	Hearing In Noise Test
ICF	International Classification of Functioning, Disability and Health
LTM	Long-term memory
NH	Normally hearing
PTA_4	Pure-tone average at 500, 1000, 2000 and 4000 Hz
RAMBPHO	Rapid, Automatic and Multimodal Binding of PHOnological information
SIC	Size comparison
SPL	Sound pressure level
SRT	Speech reception threshold
WMC	Working memory capacity

Table of contents

Introduction

Many years ago, I attended a party with guests from both Sweden and abroad. The room was crowded, the speech level high and both Swedish and English were spoken. The day after, I could recall the different conversations in Swedish but the English ones seemed to be harder to remember, although at the time, the English language spoken was understood. How could this be?

In Sweden, our native tongue is Swedish and the English language is what most people consider as their primary non-native tongue. We commonly learn English, from the fourth grade (about 10 years of age) in school. However, in adulthood individuals vary widely in how frequent they use the language. Communicating in English is for many people more or less inconvenient depending on experience and proficiency. Although good knowledge of the language is an advantage if we eventually need to communicate in English. Such interactions are usually relatively short and if it means any discomfort, it can be ended fairly quickly.

However, the situation is different from that described above, for immigrants here and in all countries. Living in another country often means speaking a non-native language, and proficiency in the language in question is both desirable and necessary from a societal perspective as well as for the citizen. Therefore, it is notable that the national test results in “Swedish as a second language” for high school students, tend to be the lowest among the national test results in different school subjects. This being the case, what are the consequences of the absence of deeper knowledge in a language, for individuals and for the society? Additionally, what are the implications for an individual attempting to learn a topic of any kind without mastering the language of instruction, and with ongoing background speech?

Communication with others in native and non-native languages and in adverse listening conditions is a complex process, involving perceptual-auditory, linguistic and cognitive factors. Research has been carried out within the emerging field of cognitive hearing science (Arlinger, Lunner, Lyxell, & Pichora-Fuller, 2009), concerning native hearing-impaired (HI) and normally hearing (NH) persons’ abilities to perceive and understand speech in adverse condition (e.g. Alexander & Lufti, 2004; Arbogast, Mason, & Kidd, G. Jr., 2005; Desjardin & Doherty, 2013; Larsby, Hällgren, Lyxell, &

Arlinger, 2005). From a linguistic perspective, native and non-native speech perception in difficult conditions has been extensively studied (e.g. Brouwer, van Engen, Calandruccio, & Bradlow, 2012; Calandruccio, Brouwer, van Engen, Dhar, & Bradlow, 2013; Calandruccio, Dhar, & Bradlow, 2010; Cooke, Garcia Lecumberri, & Barker, 2008; van Engen & Bradlow, 2007). A small number of studies have also involved a cognitive perspective, but these are relatively few in number compared to those conducted within a linguistic framework (Andringa, Olsthoorn, van Beuningen, Schoonen, & Hulstijn, 2012; Mattys, Carroll, Carrie, Li, & Chan, 2010; Olsthoorn, Andringa, & Hulstijn, 2012). However, an integrated perspective, encompassing factors relevant for NH and HI individuals in native and non-native languages and in adverse listening conditions, has not been directly employed before. Therefore, this thesis is taking an integrated approach with cognition as a focal point and examines how native and non-native speech recognition in difficult conditions interacts with perceptual-auditory, linguistic and cognitive factors for NH and HI adults.

The specific aims were:

1) To explore if four different masker conditions, namely stationary and fluctuating noise, and background speech in Swedish and English, affected performance differently.

2) To investigate the extent to which proficiency in a non-native language influences speech perception performance.

3) To explore how working memory capacity and non-verbal intelligence are associated with different listening conditions.

4) To investigate the degree of subjectively experienced disturbance caused by the four masker conditions in native and non-native languages.

5) To investigate how the comprehension of stories (rather than sentences) presented in native and non-native languages with background speech in Swedish and English interact with working memory capacity and long term memory in normal-hearing adults.

The following hypotheses were examined:

- Stationary and fluctuating noise interfere with perception, but listening to intelligible and meaningful speech maskers is cognitively more demanding and will therefore interfere with performance more profoundly. (Papers I and II).
- Perception of a non-native target language will be more difficult than the native target language for both normal-hearing listeners and hearing-impaired listeners. (Papers I and II).
- Proficiency in a non-native language is essential to follow sentences in quiet so therefore, proficiency will be even more important in masked conditions. (Papers I, II and IV).
- Previous research has shown that working memory capacity and non-verbal intelligence are related to speech perception in difficult listening conditions. Consequently, it is likely that working memory capacity and non-verbal intelligence are associated with speech perception and recall in the conditions applied in the current studies as well. (Papers I, II and IV).
- Since the speech maskers probably will interfere more with cognition than the noise maskers, it is likely that they will be experienced as being more disturbing than the noise maskers. (Paper III). The native speech masker will be experienced as more disturbing than the non-native speech maskers. (Paper III).
- It is likely that the native speech masker will affect working memory capacity and long-term recall of the heard speech to a greater extent than will the non-native speech masker. (Paper IV).

The outline of the thesis

The present thesis starts with a description of the relationship between perceptual-auditory factors, including energetic and informational masking, and hearing impairment. This is followed by a discussion of linguistic aspects of speech perception including language, the linguistic similarity hypothesis, and bilingualism. The third section addresses cognition and includes working memory, Baddeley's multi-component model, the Ease of Language Understanding model (ELU), attention, capacity theor, executive functions, assessing working memory capacity, long-term memory, the perception of masked speech, and perceived disturbance. The fourth section addresses the empirical studies I-IV and gives an overview of the participants included, the general procedure, experimental, language and cognitive tests and a brief summary of the papers. This section is followed by a general discussion with the main findings, novelty of the studies and a discussion about the effect of speech maskers in native and non-native language performance. Then, a discussion about cognitive load, working memory capacity (WMC), long-term memory (LTM), and the ELU model and its mismatch function is included. The thesis ends with a methodological discussion about the English Reading Span, the application of non-native sentences versus non-native stories in research, and perceived disturbance. Finally future directions and conclusions are presented.

Background

Gaining access to, and being able to decode speech in native and non-native languages in adverse conditions, requires an approach that recognizes both external an internal (listener-related) factors. The external factors may originate from inadequacies in the communication channel between the speaker and the listener (Mattys, Davis, Bradlow, & Scott, 2012) such as distance, background sounds and/or unclear speech/ accent by the speaker (Arlinger, 2007). Other important factors related to the speaker include articulation, the strength of the voice and word speed.

Regarding the internal factors, it is useful to consider the listeners' hearing acuity, cognitive abilities, language proficiency and phonological/semantic long-term memory representations in native and non-native languages.

Perceptual- Auditory

External factors include competing signals derived from energetic and informational masking (Brungart, 2001). Energetic masking occurs when the intelligibility of the signal is decreased, due to spectro-temporal overlap between the target and the masker, such as multi-talker babble and stationary and fluctuating noise. If a masker has a fluctuating amplitude, limited parts of the target signal may aid recognition (Cooke, 2006; Festen & Plomp, 1990). Informational masking includes any further masking effect, once its energetic effect has been accounted for (Cooke et al., 2008). Consequently, informational masking typically refers to attentional distraction, semantic interference and increased cognitive load (Cooke et al., 2008; Mattys, Brooks, & Cooke, 2009; Mattys et al., 2012).

Hearing impairment

Sensorineural hearing impairment is a partial or more profound loss of hearing, due to noise exposure, genetic factors, certain diseases and/or aging (Arlinger, 2007) and is caused by damage to the cochlea, the brainstem or the auditory nerve.

When there are defects in the cochlea, loudness recruitment and frequency selectivity can occur. Loudness recruitment is a phenomenon where the perceived loudness of sound grows faster with increased sound intensity than for a person with NH, giving a narrow functional range of hearing (Dix, Hallpike, & Hood, 1968). A sound that is just about audible at 70 dB may be intolerable at 110 dB. The frequency selectivity depends on the functional state of the cochlea and is the ability to perceive separately multiple frequency components of a complex sound. HI listeners have a reduced spectro-temporal acuity when compared to NH listeners, resulting in difficulties in separating the different sound components.

Hearing impairment is defined in the current study according to the best ear pure-tone average across the frequencies 500, 1000, 2000 and 4000 Hz (PTA_4). The average degree of hearing loss varied from slight (16-25 dB), through mild (26-40dB), moderate (41-55 dB), moderately severe (56-70 dB) to severe (71-90 dB) (Clark, 1981).

Language

Language and thinking are inevitably connected to each other and represent essential reciprocal functions, necessary in all comprehension and semantic-processing activities. For example, language and thinking have a shared relationship as the language we use can influence how we think (Crystal, 1997) and the language we use can also be a tool for expressing *how* we think (Bloom, 2004).

Language is a system of symbols and rules and by combining those, it is possible to produce an endless amount of messages and meanings (Harley, 2001). In speech, sound is used to create meaning in language and the smallest meaningful unit is the phoneme. Every language has its own set of phonemes, consisting of various consonant and vowel sounds and by combining these we create words and sentences. The rules that govern a language are called syntax which is the order of the words. A native listener immediately perceives if the syntax is violated, or if a phoneme is pronounced inaccurately.

Prior research has shown that native and non-native listeners do not rely on the same cues in processing speech. Non-native listeners are not able to take advantage of syntactic cues to the same extent as the native listeners are (Sanders, Neville, & Woldorff, 2002).

When using the rules and combining the words to produce sentences it is possible to transfer a semantic message to another person. In order to comprehend the message, this person interprets the semantics and forms a mental representation of the inferred meaning and intent, before a selected utterance is communicated back. Besides comprehending and communicating, hearing and listening are two remaining auditory and cognitive functions, outlined by International Classification of Functioning, Disability and Health (ICF; World Health Organization, 2001) and by Kiessling et al. (2003), crucial for participating in daily life. While hearing is passive, relying on sensory input, listening is an active, self-motivated, cognitive process.

In quiet, we perceive speech rather easily. We seem to be especially well-equipped to perceive speech efficiently, given that certain cues can be provided. One important cue is of course semantic context. Without semantic context, the difficulties increase considerably. For example, when words are taken out of their natural linguistic context, they are recognized only half of the time (Lieberman, 1963).

The linguistic similarity hypothesis

The linguistic-similarity hypothesis was proposed by Brouwer et al. (2012). The hypothesis outlines the linguistic aspects of processing target speech in background maskers. The linguistic similarity hypothesis suggests that the more similar the target and the masker speech are, the more difficult it is to efficiently keep the two streams apart. Equally, the hypothesis also suggests that the more dissimilar the target and the masker speech, the easier it is to keep apart the two streams efficiently (Brouwer et al., 2012). The relative similarity between two speech streams will depend on the phonetic, syntactic and/or the semantic content of the target and the masker. When the target and the masker tap into the same levels of processing as for example, semantically meaningful targets and semantically meaningful maskers, the more likely it is that semantic interference will occur than if the target and masker tap into different levels of processing. Listener-related factors such as listeners' knowledge or experience with the target and the masker languages is also important. For example, comprehensible maskers will be more harmful to target recognition than maskers that are incomprehensible (cf., Van Engen & Bradlow, 2007).

Bilingualism

The most extensive classification of bilingualism includes anyone who knows two languages (Baker, 1993). This classification is, however, too wide-ranging to be valuable. The underlying causes, such as when and how language was learned, the acquired language skills (speaking, listening, reading, writing) and how frequently the language is used, all affect the individual language ability (von Hapsburg & Peña, 2002) and are of importance in understanding individual performance.

The reason for becoming bilingual varies for many individuals. In many countries, people routinely speak two or more languages, these languages being commonly acquired in childhood. Others are elective bilinguals who choose to become bilingual and acquire a non-native language through courses or study abroad. Circumstantial bilinguals become bilingual out of necessity and are usually immigrants (von Hapsburg & Peña, 2002).

Of importance for anyone acquiring a language is the age of acquisition. It has been argued that "the earlier the better" and the debate concerns whether there is a biologically based critical or sensitive period for non-native language acquisition and if so, at what age-range it ends. Adolescence has been mentioned but some studies have found that even after mid-adolescence, it is possible to acquire the proficiency of a native speaker, if not with the perfect accent (Bialystok, 2001). The question of a biological sensitive or critical period in the acquisition a non-native language may be followed by the question of whether second-language acquisition is represented in the same area(s) in the brain as a native language. A brain-imaging study (Perani, Paulesu, Galles, Dupoux, & Dehaene, 1998) used positron emission tomography (PET) scans to measure cortical activation patterns in the brains of English-speaking Italians as they were listening to stories in English and Italian. The study showed that highly proficient non-native listeners who learned the language before the age of ten showed representations of the two languages in the same cortical areas. However, less proficient Italians, who learned English later in life showed representations in different areas (Perani et al., 1998). Yet, even for highly proficient bilinguals, some brain regions within this common network became more active when they used the non-native language. This suggests that even if the person is fluent in his/her second language, the person has to exert more conscious effort to process the less dominant language (Marian, Spivey, & Hirsch, 2003).

In Sweden, the current school system consists of compulsory primary school (age 7-16) and an optional secondary school (age 16-19). In this thesis, the participants started learning English, their non-native language, between the age of 9 and 12, in grades 3 to 6. It is most common in Sweden, to begin learning English in the fourth grade and to continue through to the end of secondary school. English has existed as an educational subject in the Swedish school-system since the 1880s, but became the primary foreign language in 1939 (it was German before this time). From the 1950s English was compulsory from grade 5 and has been compulsory ever since then. During the past decades, the teaching time, as well as the age at which students begin to learn English, has varied only marginally.

Additional improvement in the English language in advanced school system as well as the frequency with which individuals use English are undeniably of importance for English proficiency. In this thesis, the frequency with which the participants use English, varied between daily and never. This might partly reflect the great variability in how frequently Swedish people use the English language, and as such also the relatively great variability in English proficiency.

Cognition

Working Memory

Working memory is the capacity to store and manipulate information over momentary periods of time (Baddeley & Hitch, 1974) and an operational system involving separable, simultaneously interacting components that can be used to carry out complex cognitive activities (Gathercole, 2007). In communicating with others, working memory supports us when we actively take part and maintain the focus on the conversation; this comprises turn-taking and following the gist (Rönnberg et al., 2013), while eventually inhibiting irrelevant speech from the surroundings. Working memory is particularly active in targeted work, when we do several things at the “same” time, and also of great importance in all new learning, for example when we learn another language.

Baddeley's multi-component model

The multicomponent model of Baddeley (2000, 2012) describes different components of the memory system, each with different functional roles when it comes to the storing and processing of information. The phonological loop is capable of briefly holding speech-based information and the visuospatial sketchpad provides temporary storage for visual and spatial information. The phonological loop is related to the development of vocabulary in children and the speed of non-native language vocabulary acquisition in adults. The central executive is a third component which has a limited capacity and is responsible for the selection, initiation and termination of domain-general processing resources. It controls attention and also controlling and using the phonological loop and the visuo-spatial sketchpad for specific purposes. Finally, the episodic buffer is also a temporary storage system and integrates information from the phonological loop and the visuospatial sketchpad. The episodic buffer is also supposed to be controlled by the central executive.

Ease of Language Understanding model (ELU)

While the model of Baddeley (2000, 2012) focusses on the *component*s of working memory, the Ease of Language Understanding (ELU) model (Rönnberg, 2003; Rönnberg et al., 2013) focusses on the *function* of working memory in language understanding

The ELU model describes essential cognitive aspects of language understanding in straightforward and challenging listening situations. The model takes into account the processing of multimodal information and the mutual interaction between implicit, bottom-up and explicit, top-down functions. Additionally, the model suggests an episodic buffer for the multimodal input, in which Rapid, Automatic and Multimodal Binding of PHOnological information (RAMBPHO) takes place. As long as the incoming linguistic signal is clear and the phonological stream matches stored representations in semantic long-term memory, the process is expected to progress rapid and implicitly, permitting automatic lexical access and understanding. However, in unfavorable situations when the linguistic signal is inaccurate, due to for example disturbing noise, another language or hearing impairment, a mismatch occurs and an explicit, more effortful process is necessary. The explicit processing requires working memory capacity to interpret and infer meaning in spite of incomplete information. In the new model (Rönnberg et al., 2013) further predictions are outlined, including early attention mechanisms (Sörqvist, Stenfelt, & Rönnberg, 2012) when listening to speech, inhibition of speech maskers and its effect on episodic long-term memory, effort and the effects of hearing-impairment on semantic and episodic long-term memory.

Attention

Attention is a central cognitive process that permits us to focus on relevant speech (Best, Gallun, Ihlefeld, & Shinn-Cunningham, 2006; Best, Gallun, Mason, Kidd, & Shinn-Cunningham, 2010; Cherry, 1953). It has been claimed that the ability to control attention in the presence of interference essentially reflects working memory capacity (Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Bleckley, Conway, & Engle, 2001; Engle, 2002; Kane & Engle, 2000). In fact, a large working memory capacity maps an attentional ability to maintain information in a dynamic and quickly obtainable state, to avoid distraction (Engle, 2002) and to be less disposed to attend to background speech (Beaman, 2004, Sörqvist, Ljungberg, & Ljung, 2010). Consequently, a low working memory capacity involves reduced ability to control attention and to be more sensitive to background interference. This was also shown in a study by Conway, Cowan and Bunting (2001). The individuals with lower spans were more likely to detect their name in an unattended message, than were those with higher spans. Actually, 65 % of the lower memory span individuals paid attention to their name, compared to 20 % of the individuals with higher memory spans.

Even though it appears that higher working memory capacity individuals have a greater ability to control attention and therefore modulate how much information is given access to higher cognitive processing (Rönnberg et al., 2013), there is still an ongoing debate whether the irrelevant information is filtered out early or late in the chain of processing.

Kahneman (1973) assumed that a limited pool of attentional resources can be allocated among mental operations or tasks. The more resource-demanding a certain task, the less resources become available for use somewhere else in the system. Additionally, the more complex and unfamiliar the task, the more mental resources are required for successful performance. However, individuals can select their focus and assign the mental effort in that direction. The mental resources also depend on the overall level of arousal. According to Kahneman (1973), one effect of being aroused is that more cognitive resources are available for various tasks. The level of arousal also depends on task difficulty: We are less aroused when performing easy tasks as they require fewer resources to complete. Additionally, we pay more attention to issues we are interested in, are in the mood for, or have judged as important.

Executive functions

The term executive functions refers to the higher level functions of planning and decision making (Miyake et al., 2000). Three basic aspects have been conceptualized, namely inhibition, shifting and updating. Inhibition denotes the ability to suppress an action or pre-potent response. Shifting involves the ability to switch between tasks, mental sets or operations, and finally, updating, refers to the constant monitoring of a given task. The executive functions are relevant in language understanding (Rönnberg et al., 2013) and particularly in recognizing speech in noise (Sörqvist & Rönnberg, 2012): The updating is important as the listener continually has to replace stored information with newly received information, the inhibition function supports the listener in situations where it is necessary to suppress noise interference, and the shifting enables the listener to focus on the selected conversation and not on the disturbing background noise.

Assessing working memory capacity

Complex span tasks are generally applied to assess working memory capacity. In this thesis, the Reading Span (papers I, II and IV) and Size Comparison Span (SIC Span) (paper IV) tests have been administered. In these complex span tests, sentence comprehension is combined with recall (Daneman & Carpenter, 1980). In both tests, sentences are presented on a screen and judgements are required following each sentence. In the Reading Span, the participant has to indicate whether the sentence is absurd or not, and in the SIC Span, whether an item is larger than another item. In the SIC Span, a recall word is presented on the screen before another comparison sentence is shown on the screen again. After a set of sentences (Reading Span) and a set of sentences/recall words (SIC Span) have been presented, the participants are asked to recall the recall words (SIC Span) or either the first or the last word in each sentence (Reading Span). In both tests, a distractor-activity is present at the same time as the sentences have to be processed semantically. Reading Span and SIC Span both tap into the ability to simultaneously store and process information, which is the core concept of working memory (Unsworth, Redick, Heitz, Broadway, & Engle, 2009). Other operating subprocesses are the executive functions of attentional control, inhibition and task switching ability (Bayliss, Jarrold, Baddeley & Gunn, 2005; Unsworth & Engle, 2007). The exact tradeoff between those functions in Reading Span and SIC Span is not entirely clear. However, previous studies have shown that SIC Span is a good predictor of speech distraction (Sörqvist, et al., 2010) while Reading Span has been shown to be a reliable predictor of energetic masking (Rudner, Foo, Rönnberg, & Lunner, 2009; Rönnberg, Rudner, Lunner, & Zekveld, 2010).

Long-term memory

Working memory connects a moment's thoughts with the next moment's action and long-term memory connects our days and our lives (Klingberg, 2011).

Long-term memory consists of a system of memory functions where any given system is responsible for a specified domain. The properties of one memory system should clearly differ in various ways from other memory systems.

Episodic long-term memory refers to the memory of past personal experiences and semantic long-term memory refers to general knowledge in a language or about the world; as the distinction is between remembering (episodic) and knowing (semantic). The episodic memory is not completely reliable as the mental interpretation depends on emotions and whether an event occurs that affects us strongly (Gathercole & Alloway, 2008).

The more deeply information is processed, the more likely it is to be recalled efficiently (Craik & Lockhart, 1972). If information is stored, using both verbal and visual codes it will presumably enhance memory and increase the possibility that at least one of the codes will be available for later recall (Paivio, 1969). Bower (1970) proposed that visual codes like imagery improves memory but due to more associations between the items to be recalled. However, the storage potentials of long-term memory are dependent on effective manipulation in working memory for further transfer to the long-term memory. Any distraction or other competing tasks will reduce the possibility for working memory to secure information in episodic long-term memory.

The perception of masked speech

If listening in quiet is a rather straightforward process, listening in noise is the opposite. Recognizing and understanding speech in noise is complex, involving central bottom-up auditory functions as well as top-down, cognitive abilities (Pichora-Fuller, Schneider, & Daneman, 1995). In a bottom-up process, the reliance is driven by the sensory system and the brain is recognizing and interpreting the speech sound solely via input. The top-down process on the other hand, is driven by a persons' expectations, knowledge and/or experience. Speech comprehension in noisy surroundings often means that listeners mentally have to put together pieces of missing information. The more degraded a signal, the more top-down processes are used. Contextual information is even more important in noise, providing necessary cues and support to the listener when perceiving ambiguous and distorted auditory signals (Koelewijn, Zekveld, Festen, & Kramer, 2012; Pichora-Fuller et al., 1995; Plomp, 2002; Weber & Cutler, 2004; Wingfield, 1996; Zekveld et al., 2011).

Different types of noise have different characteristics - and do also affect people differently. Individuals with high cognitive capacity are usually good at listening in the dips in fluctuating noise. The ability to combine the accessible fragments from the dips in the fluctuating masker is likely to relate to the processes of integration and inference making (Rönnberg, Rudner, Foo, & Lunner, 2008). However, persons with hearing impairment are not able to listen in the dips, due to frequency selectivity (Moore, 1985).

When people are participating in a conversation and interference from the background consists of speech, it is likely that the relative contribution of cognitive functions are higher than if the background interference consist of noise maskers. The reason for this is threefold: a) difficulties in suppressing the background speech in favor of the conversation b) the semantic content of the background speech is intelligible (Hoen et al., 2007; Van Engen & Bradlow, 2007) c) attention switches back and forth between the conversation and the background speech (Schneider, Li, & Daneman, 2007). The cognitive functions involved in this process are likely to be working memory and executive functions (Rönnberg et al., 2010; Rönnberg et al., 2013). In order to succeed in processing the preferred conversation, it is necessary to prevent irrelevant speech from gaining access to working memory in the first place or remove it from working memory when it does intrude (Hasher & Zacks, 1988). In other words, either use the executive functions of inhibiting, in order to suppress unwanted speech, or the executive functions of updating in order to remove unwanted information and simultaneously replace it with new.

Recognizing a non-native speech in noise is even more demanding than recognizing a native speech, due to different factors. First, speech is continuous, which means that sounds from one word blend into the next word. This continuously changing pattern of sound can make it difficult to perceive where one word ends and the other begins (i.e., this may create a segmentation problem). Secondly, the sound of phonemes depend on the context in which they are heard. More specifically, how a phoneme is pronounced depends on the preceding and subsequent phonemes. This so-called co-articulation effect is defined as the overlapping of contiguous articulations (Ladefoged, 2001). However, each phoneme provides a cue for what is coming next if the cues are perceived and properly interpreted. This brings us to the third factor of importance which is proficiency. The ability to comprehend speech in a non-native language masked with noise is related to the listeners' proficiency in the non-native language (Sörqvist, Hurtig, Ljung, & Rönnberg, 2014). A well-established lexical and phonological representation in long-term memory facilitates word recognition in continuous speech masked with noise.

Perceived disturbance

Human responses to noise or background speech exposure are in general negative. Even so, people seem to evaluate disturbing sounds during speech perception differently, possibly depending on a variety of internal and external factors. Hearing impairment is one such factor as individuals with hearing loss are more likely to suffer in difficult listening situations than NH individuals are (McCoy et al., 2005; Tun, McCoy, & Wingfield, 2009; Zekveld, Kramer, & Festen, 2010). Other aspects of importance may be cognitive functions and age, as well as features of the target language and the background speech or noise. Generally, speech is considered to be more disturbing than other sound sources (Bradley & Gover, 2010; Venetjoki, Kaarlela-Tuomaala, Keskinen, & Hongisto, 2006) and irrelevant speech has been shown to have negative effect on cognitive performance (Knez & Hygge, 2002; Schlittmeier, Hellbrück, Thaden, & Vorländer, 2008).

Summary

Recognizing and comprehending a native as well as a non-native language in speech and noise maskers requires both perceptual-auditory bottom-up and linguistic, cognitive, top-down functions. However, the relative contribution of bottom-up and top-down functions are dependent on individual variability, such as proficiency in a language, working memory capacity and/or hearing impairment. Proficiency in a non-native language is crucial in background noise as it facilitates the access to lexical and phonological representations in semantic long-term memory and therefore increases the possibility of success in decoding the semantics. Working memory capacity and executive functions enable us to focus attention on the required word stream and to suppress the unwanted. Efficient and deep encoding by visual and verbal codes in episodic long-term memory increases the possibility for better recall. Perceived disturbance of noise and speech might depend on a variety of internal and external factors such as hearing impairment, cognitive functions and the characteristics of the disturbing sounds.

Empirical Studies

Method

Participants

Papers I and III

Twenty-three native, Swedish normal-hearing adults participated (13 females and 10 males) with an average age of 49.5 (SD = 9.8, range = 28-64). The participants were recruited from different workplaces in Linköping, Sweden. They had 11 to 21 years (M = 15.8) of education.

Papers II and III

Twenty-three native, Swedish hearing-impaired adults participated (14 females and 9 males) with an average age of 50.1 (SD = 10.2, range = 28-65). The participants were recruited from the audiology clinic at Linköping University Hospital, Sweden. An acquired bilateral sensorineural hearing-impairment and no severe tinnitus complaints were inclusion criteria. The average hearing loss across frequencies (PTA_4) was 46.7 dB HL (SD = 10.7). The PTA_4 ranged from 25.0 dB HL to 71.3 dB HL. The average degree of hearing loss varied from slight (16-25 dB; n = 1) through mild (26-40 dB; n = 6), moderate (41-55 dB; n = 13), moderately severe (56-70 dB; n = 2) to severe (71-90 dB; n = 1) (Clark, 1981). They had 8 to 21.5 years (M = 13.7) of education.

Paper IV

Twenty-three native, Swedish normal-hearing adults participated (12 females and 11 males) with an average age of 47.8 (SD = 11.3, range = 20-65) years old. The participants were recruited from different workplaces in Sweden. They had 11 to 24 years (M = 16.3) of education.

General Procedure

Papers I - III

Participants were tested in one session of approximately 3.5 hours. The test session started with an audiometric test, carried out by the experimenter for the NH participants and by an experienced audiologist for the HI participants. The experimental test followed and after a break, participants completed the cognitive tests. The order of the experimental test and the language- and cognitive tests were counterbalanced. The audiometric test and the experimental test took place in a sound-treated booth and the language- and cognitive tests in a quiet, nearby room. Auditory stimuli were presented over headphones (Sennheiser HD600).

Paper IV

Participants were tested in one session of approximately 3.5 hours, including a break of half an hour. The test session started with collection of background data and an audiometric test, carried out by the experimenter, followed by the experimental test, in a sound-treated room. Participants were instructed to listen to the first part of the story (approximately for 5 min) and when the story paused, judge the audibility of the story. This was repeated, every 5 min until both stories were listened to. After a break, the test session continued with the language and the cognitive tests in a quiet, nearby room. The order of the experimental test and the order of the language and the cognitive tests were counterbalanced except for the long-term memory test which was always performed as the last test in this session.

Experimental tests

Sentence intelligibility in noise was measured using a speech reception threshold test (SRT) developed by Plomp and Mimpen (1979). Swedish and English target speech was presented in the SRT test. Swedish HINT (Hearing In Noise Test) (Hällgren, Larsby, & Arlinger, 2006) target sentences and American English HINT (Nilsson, Soli, & Sullivan, 1994) target sentences were used in the SRT test. The HINT materials in Swedish and English are comprised of short, everyday sentences, judged to be natural by native persons. The phonemically balanced lists of sentences are grouped in 25 lists with 10 sentences in each. The sentences were recorded by male native speakers. Participants performed in eight test conditions: English or Swedish target language, either language combined with four different types of maskers: a stationary masker, a fluctuating masker, two-talker babble Swedish and two-talker babble English (see description below). Every condition counted 20 sentences each and every new condition started with several practice sentences. The first condition had 10 and the following had 5 practice sentences each. The sequence of conditions was counterbalanced across listeners and each sentence was used only once per listener. For the NH listener speech was presented at a fixed level of 65 dB SPL and for the HI listener the level of the target and the masker signals was individually adjusted off-line according to the Cambridge prescription formula (Moore & Glasberg, 1998) and based on the pure-tone thresholds of the best ear. The masker level was changed in a stepwise two-up-two-down adaptive procedure (Plomp & Mimpen, 1979) targeting 50% sentence intelligibility. Masker onset was 3 s before speech onset and masker offset was 1 s after speech offset. The participants repeated each sentence verbally and the experimenter scored whether all words in the sentence were repeated correctly.

The stationary masker consisted of the speech-shaped noises developed by Nilsson et al. (1994) and Hällgren et al. (2006). The spectrum of the masker was shaped according to the long-term average spectrum of the speech material of the corresponding set (same procedure for Swedish and English).

The fluctuating masker was constructed by modulating the speech-shaped noise of the target speech and had the same envelope fluctuations as the two-talker babble in Swedish or English. These envelopes were extracted by applying a low-pass filter with cut-off frequency of 32 Hz (for details see Agus, Akeroyd, Gatehouse, & Warden, 2009). Two modulated maskers were used, spectrally matched to the target language in Swedish or English, respectively, and also matched temporally to the babble in Swedish or English, respectively.

The two-talker babble in Swedish was recorded with one native Swedish female and one native Swedish male reading from Swedish newspapers. The two-talker babble in English was recorded with one native female English/American speaker and one male British English speaker reading from English/American newspapers. By mixing the soundtracks from the female and the male speakers the two-talker babbles were created in Swedish and American/English, respectively. The speech maskers were spectrally matched to the long-term spectrum of the target speech presented (Swedish or English).

Experimental test and episodic long-term memory of spoken discourse

Participants listened to two fictitious stories presented over headphones in a sound-attenuated both. One story was in English (The life of Mary Mason) and one story was in Swedish (The story of Agnes Odencrantz). The stories were written for this experiment. Each story is approximately 15 min long and was divided into three parts of approximately 5 min each. Two parts of each story were masked by either two-talker babble in Swedish or two-talker babble in English. One part was not masked at all, to get a non-distracting baseline. The three background conditions (two different speech maskers and one unmasked condition) were counterbalanced across each story. The order of the English and Swedish stories were also counterbalanced across participants. The long-term memory test was always performed as the last test in this session.

Audibility

Participants were instructed to rate the audibility of the story, ranging from 0 to 10 on a continuous scale, where 10 represented *Perfectly audible* and 0 represented *Not audible at all.*

Episodic Long-term memory

Episodic long-term memory was measured using a multiple choice questionnaire. The total number of questions was 24. Every 8 questions corresponded to 5 min of recorded story, 4 simple questions and 4 more complex questions. The factual multiple choice questions had very different response options, often consisting of one word. The more complex questions had longer and more similar response options and the facts could be collected from different parts of the story (within the 5 min part).

Language and Cognitive Tests

Reading Span

The Reading Span test is a complex span test of working memory capacity (Daneman & Carpenter, 1980; Rönnberg, Lyxell, Arlinger, & Kinnefors, 1989). In the test, short sentences were presented word-by-word on a computer screen. Half of the sentences made sense (e.g. *Prästen läste bibeln*; The priest read the bible) and the other half did not (e.g. *Bastun kokade gröt;* The sauna cooked porridge). Immediately after each sentence, a question appeared on the screen, asking whether the sentence made sense or not. The participants answered by button presses, yes or no. Sentences were presented in sets of 3 to 5, with progressively increasing set sizes. When a set had been presented, the participants were asked to orally recall either the first or the last word in the previous set of sentences. The participants did not know until after the set, which words (the first or the last) they had to report. The number of words correctly recalled were scored as the dependent variable regardless of order.

In papers I and II, the Reading Span test was used in both Swedish and English and the order of the tests was counterbalanced across participants (max score, 23 in both the English and Swedish tests). In paper IV, the Reading Span test was used in Swedish only (max score 28).

Size Comparison Span

Participants were presented with comparison-sentences on a computer screen, for example " Är en mus större än ett lejon" (Is a mouse larger than a lion) and they used button presses to answer yes or no. They had 5000 ms to respond before a recall word was presented on the screen (e.g., elefant; elephant) for 800 ms before a new comparison sentence was presented on the screen. The comparison words and the recall words always belong to the same semantic category (e.g. animals) within a set and the set sizes gradually increase from 2 to 6. The words and the categories were used only once. When a set of comparison sentence and recall words had been presented, the participants were questioned from the screen to type the recall words in the right presentation order. A total of 40 words were possible to recall. Each word correct recalled was scored as the dependent variable, irrespective of order. The SIC Span in Swedish was used in paper IV.

Non-verbal reasoning ability

The Raven standard progressive matrices (Raven, Raven, & Court, 2000) is a multiple choice measure of non-verbal reasoning (fluid intelligence). The test provides sixty matrices divided into five sets, A - E. The task is to identify (from x alternatives) which missing piece best completes a larger pattern. The participants performed sets B – D. The difficulty increased within each set and every new set was progressively more difficult than the previous one (max score = 36). Raven standard progressive matrices was used in papers I and II.

English proficiency test

This test assesses English language comprehension and is a standardized, national test, essentially developed for the optional Secondary School level (http://www.nafs.gu.se/digitalAssets/1193/1193558_last_exp.pdf). The test consists of a text and two sets of tasks. In the first set, participants answered text-based, open questions with their own words and the second set consists of sentence completion; participants are requested to explain one bold printed word in a sentence with one word in the open end of the sentence (example: “Since mountaineering involves many **perilous** activities, this sport is considered both challenging and _______ (dangerous)”) (Max score =12). The English proficiency test was assessed in papers I, II and IV.

Swedish proficiency test

This test assess Swedish language comprehension and is a standardized, scholastic aptitude test (Högskoleprov). The test consisted of reading comprehension with multiple choice questions and sentence completion: The participants read 10 different descriptions of a topic with open parts in which one word or a few words could be selected from multiple choices. No time-limit. The Swedish proficiency test was used in paper IV.

Story Grammar

The fourth paper in this thesis used stories in Swedish and English, masked with speech, for later recall. To ensure that the stories were not heard or read before, they were written and recorded specifically for use in this study. The stories aimed to be narrative, and used short sentences and an illustrative language. Further, the stories were also developed within the same structure to increase the internal consistency between the stories. This was made to improve the possibilities that any differences in recalling the stories were due to language and/or the masker effect. Therefore, a pattern referred to as story grammar, was followed.

Story grammar evolved from folktales which, regardless of age or culture, follow a pattern. Story grammar (Mandler & Johnson, 1977) involves a manifestation of the character's problem or conflict, a description of efforts to solve the problem, and an analysis of the sequence of events that lead to resolution. An analysis of the characters reaction to the events in the story, is also involved (Dimino, Gersten, Carnine, Blake, 1990). Although the story schema, a hypothesized mental structure, moves through a seemingly simple progression of beginning, middle and end, the schema also relies on underlying components of setting, main character, problem/conflict, events, goals, attempts, twist, character information, reaction and resolution. The resolution or ending can be of attainment or non-attainment of the goal by the character and the characters' reaction to the outcome and eventually a moral. The story schema is a means for exploring the characters of the story and also a signal where attention is to be paid and the storing of information for later recall.

Summary of the papers

Paper I

The relevance of proficiency in a non-native language has been claimed in previous research (Brouwer et al., 2012; van Engen, 2010; van Wijngaarden, Steeneken, & Houtgast, 2002; Weiss & Dempsey, 2008) although its plausible role in related research has not been the focal point before. This paper examined the extent to which proficiency in a non-native language influenced native and non-native speech perception in speech- and noise maskers. The paper also examined whether the speech- or the noise maskers interfered most with performance in native and non-native languages. The speech maskers consisted of two-talker babble in Swedish and two-talker babble in English and the noise maskers of stationary and fluctuating noise. In addition, the role of working memory capacity, assessed in the native and non-native language was investigated, as well as non-verbal intelligence.

Method

The sentence intelligibility was measured using speech reception thresholds, SRT test (Plomp & Mimpen, 1979). Swedish (Hällgren et al., 2006) or American English HINT (Hearing In Noise Test) (Nilsson et al., 1994) sentences were applied. The target sentences was presented in 65 dB SPL. The participants performed eight conditions; Swedish and English target language, combined with four background maskers. The maskers consisted of stationary noise, fluctuating noise, two-talker babble in Swedish and two-talker babble in English. The reason for using two-talker babble is because it is known to produce a high masking effect (Brungart, 2001; Calandruccio et al., 2010; Van Engen & Bradlow, 2007). The participants also performed cognitive and language tests, such as Reading Span in Swedish and English, and a non-verbal intelligence test and a proficiency test in English.

Result and Discussion

The results showed that the level of proficiency in English is a highly decisive factor for performance in background speech- and noise maskers. The level of proficiency determined performance in the non-native language. Two groups were created according to the participants performance; high and low proficiency groups. The low proficiency participants had pronounced difficulties in performance in all masker conditions in the non-native language. Yet, when the maskers consisted of speech, the difficulty was even larger. In the native language, the Swedish babble masker was the most interfering masker for all participants. The large interference may be due to the linguistic similarity between the target speech and the masker speech (Brouwer et al., 2012) and/or intelligible words in the masker.

Paper II

Hearing-impaired listeners face three challenges in non-native speech communication in noisy surroundings: The hearing-impairment itself, the noise and the non-native language. How these three factors interact was examined in the second paper. The differential impact of speech- and noise maskers was also taken into account.

Method

The experimental test, stimuli and cognitive tests were identical to those described in the method section for paper I. Regarding the presentation levels of the target and masker signals, they were individually adjusted offline according to the Cambridge prescription formula (Moore & Glasberg, 1998), and based on the pure-tone thresholds of the best ear.

Result and Discussion

Results indicated that the speech maskers were more interfering than the noise maskers in both target languages. Better hearing acuity (PTA) was associated with better perception of the target speech in Swedish and better English proficiency was associated with better speech perception in English. Larger working memory capacity and better PTAs were related to the better perception of speech masked with fluctuating noise in the non-native language. This suggests that both are relevant in highly taxing conditions. A large variance in performance between the listeners was observed, especially for speech perception in the non-native language.

The performance was influenced by the listener-related factors such as hearing-impairment, age, cognitive abilities and proficiency in the non-native language. The, external factors also played a role such as the target language (native versus non-native) and the different masker types. The interactions between these factors had an impact on the complexity of the listening conditions as experienced by the HI listeners.

Comparing NH and HI participants

To find out whether there were differences between the NH group from paper I and the HI group from paper II in performance, two separate analyses of variance (ANOVAs) were conducted.

The first ANOVA was conducted to assess the impact of target language (Swedish and English) and speech maskers (Swedish babble, English babble) on NH and HI participants. The result demonstrated a main effect of language: F (1, 41) = 111.58, p < .001 and a main effect of masker: F (1, 41) = 25.5, p < .001.There was a significant three-way interaction between language, maskers and group: $F_{(1, 41)} = 7.37$, $p < .05$ (Figure 1). Investigating this further, a significant two-way interaction was shown between target-language and maskers for the NH participants: $F_{(1, 21)} = 13.59$, $p < .001$. A post-hoc Bonferroni-corrected t-test was performed to examine the origin of this effect. The effect of target language (Swedish vs English) was largest for babble English t (21) = -7.9, p < .001. The result suggested that the NH participants benefit in the Swedish target when the language of the masker is in another language. This result replicates previous results that when background speech consists of an unfamiliar or less well mastered language, the result is generally a release in masking (Calandruccio et al., 2010; Gautreau, Hoen, & Meunier, 2013; Kilman, Zekveld, Hällgren, & Rönnberg, 2014; Rhebergen, Versfeld, & Dreschler, 2005; Van Engen, 2010; Van Engen & Bradlow, 2007). The HI participants performed equally poorly in the speech maskers in both target languages.

A second ANOVA was conducted to assess the impact of target language (Swedish and English) and noise maskers (stationary, fluctuating) on NH and HI participants. The result showed no interaction effect between language, masker and group $F_{(1, 41)} = .697$, p = .40 = .409

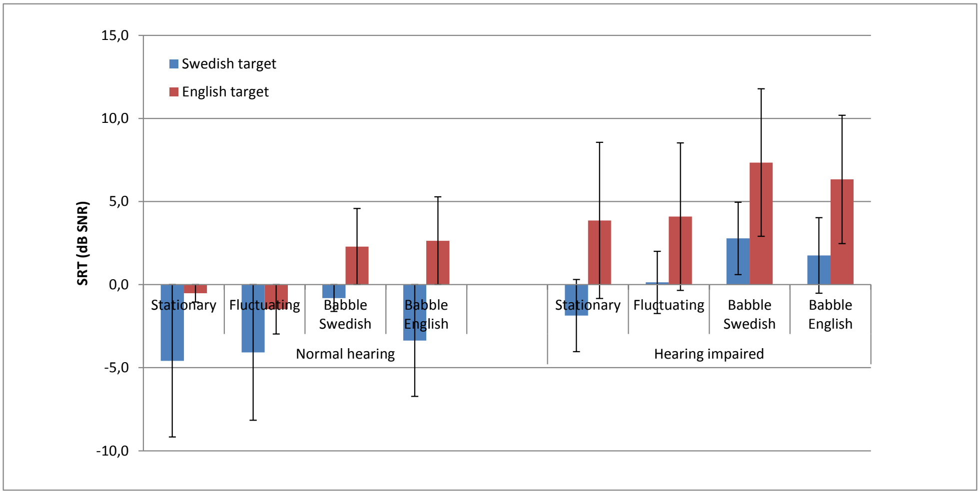

Figure 1. Three-way interaction between target language (Swedish, English), speech masker (Swedish babble, English babble) and group (NH, HI). Error bars represent standard deviations.

Paper III

The purpose of the third study was to assess how NH and HI participants evaluated the perceived disturbance from speech- and noise maskers in the two target languages, English and Swedish.

Method

When participants had performed the adaptive SRT in speech and noise maskers, they evaluated the perceived disturbance immediately after each condition on a disturbance scale, ranging from 0 to 10 where 0 represented “not disturbing at all” and 10 represented “extremely disturbing”.

Result and Discussion

A three-way interaction effect between target language, masker condition, and group (hearing-impaired versus normal-hearing) was the main result in this study. The HI listeners perceived the Swedish speech masker as significantly more disturbing for the native target language (Swedish) than for the non-native language (English). Further, this Swedish speech masker was perceived significantly more disturbing than each of the other masker types in the Swedish target speech. For the NH listeners, the Swedish speech masker was more disturbing than the stationary and the fluctuating noise-maskers for the perception of English target speech. The speech maskers was perceived as more disturbing than the noise maskers for the NH listeners but this was not the case for the HI listeners. However, it is notable that the HI listeners had particular difficulty with the perception of native speech masked by native babble, a common condition in daily-life listening conditions. The results indicate that the characteristics of the different maskers applied in the current study seem to affect the perceived disturbance differently in HI and NH listeners.

Paper IV

Successful perception of degraded speech input may be obtained at the expense of attentional resources and cause episodic memory failures (e.g. Rabbitt, 1968). The purpose of this study was to examine how well Swedish (native) and English (non-native) stories masked by Swedish and English speech were recalled by the NH adult listeners. Processing and interpreting speech masked by interfering speech may increase cognitive load and reduce the resources available for elaborative encoding of the information in memory (e.g. Surprenant, 1999; Wingfield, Tun, & McCoy, 2005). Is it possible that adverse listening conditions can impair long-term memory even though the listeners have heard what has been said? How do speech maskers influence working memory and episodic long-term memory?

Method

Participants listened to one fictional story in English (The life of Mary Mason) and one in Swedish (The story of Agnes Odencrantz). Each story was approximately 15 min long. The stories were divided into three 5-min sections. One was presented unmasked, one was masked with two-talker Swedish babble, and one was masked with two-talker English babble (counterbalanced design). Target speech was presented at 65 dB SPL and maskers were presented at 60 dB SPL. After every 5 min participants rated target speech audibility on a continuous scale from 0-10, where 0 represented *Not audible at all and* 10 represented *Perfectly audible. Long-term memory* (LTM)*:* 24 multiple choice (four-alternative forced choice) questions were used. Every 8 questions corresponded to 5 min of recorded story and included 4 simple and 4 complex questions. Participants also performed cognitive tests and language tests in Swedish and English proficiency. The cognitive tests consisted of the Reading Span test (Daneman & Carpenter, 1980; Rönnberg et al., 1989) and the

Size comparison test (Sörqvist et al., 2010). Both complex span tests of working memory capacity. The *English* and *Swedish proficiency tests a*ssess reading comprehension and sentence completion in English and Swedish, respectively.

Result and Discussion

The stories in quiet were significantly better recalled than the stories masked with Swedish babble. This result is consistent with previous result by Rabbitt (1968) and Surprenant (1999) when words presented in noise were more difficult to recall than words presented in quiet. In the current study when background speech consisted of the intelligible and meaningful native language, the difficulties may have been large for the listeners. Also, regarding the condition with a non-native target (English) and a native masker (Swedish), it is plausible to assume that it is effortful to ignore the well-known masker in favor of the less known target. Providing that the resource capacity is in limited supply (Rabbitt, 1968; Kahneman, 1973), the effortful situation in this study might have reduced the resource capacity, even though the target speech was rated audible.

Larger WMC (Reading Span) was associated with better recall of the English story in the quiet condition and in the Swedish babble condition. Individuals with larger WMC might have a higher ability to focus on the assigned task, suppress distraction from the environment and to inhibit irrelevant speech (Sörqvist et al., 2010). This was consistent with the results in the current study, as also better inhibition ability (SIC Span Intrusion words) was associated to better recall in the Swedish story/English babble condition.

High English proficiency was related to better recall in the English story/quiet condition and the English story/Swedish babble condition, indicating that English proficiency is essential for recollection of the English story. This result was expected as the significance of proficiency was shown in Kilman et al. (2014). However, it was not expected that high Swedish proficiency (sentence completion) should be related to better recall of the English story/quiet condition and that high Swedish proficiency (reading comprehension) should be related to better recall of the English story/ Swedish babble condition. The sentence completion proficiency is a useful top-down ability and may, in the context of storytelling, operate as a substitute for a not completely mastered language.

General Discussion

Summary of the findings

In this thesis, the native speech maskers affected performance (papers I and II), perceived disturbance (paper III), working memory capacity (paper IV) and episodic long-term memory (paper IV). However, the individual variabilities were large between NH listeners and HI listeners and between the target languages. In general, the two speech maskers in native and non-native languages affected performance more than the energetic maskers for both NH and HI listeners.

The importance of proficiency in a non-native language was experienced in paper I (NH), paper II (HI) and paper IV (NH). Proficiency in the native language was also of significance in paper IV.

Performance

NH: In Swedish as target language, the NH listeners had most performance- difficulties in the Swedish speech masker, as compared to the English speech masker and the energetic maskers. In English as target language, the informational maskers (Swedish and English) affected performance more than the energetic maskers (fluctuating and stationary noise).

HI: In Swedish as well as English target language, the HI listeners had most difficulties in performance in the informational maskers as compared to the energetic maskers.

Perceived Disturbance

NH: For English as target language, the perceived disturbance was larger for the Swedish speech masker as compared to the energetic maskers. No other significant differences between or within any of the target languages was observed.

HI: A three-way interaction showed a difference in perceived disturbance between Swedish and English as target language. In Swedish target, the perceived disturbance was larger for the Swedish speech masker as compared to all the other maskers. In English target, no difference between the maskers.

Working memory capacity and episodic long-term memory

The NH listeners recalled the stories (Swedish and English) in quiet significantly better than the stories in the Swedish speech masker. Larger WMC (Reading Span) was associated to better recall of the English story in the Swedish speech masker condition and in the quiet condition. Better inhibition ability (SIC Span intrusion words) was related to better recall of the English story in the Swedish story/English background speech condition.

English and Swedish proficiency were related to the English story/background speech in the Swedish condition and the English story/quiet background condition

When target language is not well mastered as when it is in a non-native language, the interference from background speech in the dominant native language is demanding for performance, perceived disturbance and memory processes.

What is new?

Each variable in the present thesis has been examined in previous research. The innovative aspects of the current study primarily concern the combination of variables and different perspectives, including target language (native and non-native speech perception), hearing, cognition, perception, perceived disturbance and memory aspects.

Proficiency in a language is an indisputable resource for an individual in non-native speech perception performance. Study I focused on its role in speech perception in a variety of conditions. Besides proficiency, study I assessed the role of internal and external factors of importance and examined what variables could possibly explain intra- and inter-individual differences in native and non-native speech perception performance. The approach with proficiency in focus is new.

Study II innovatively assessed the influence of hearing impairment on non-native language perception. In a group of HI listeners, we investigated how the above mentioned variables interacted with hearing-impairment.

A further novel approach was the combination of acquiring performance measures together with a subjective evaluation of the perceived disturbance when native and non-native languages are masked with noise and speech maskers. Previous studies that examined the subjective evaluation of perceived disturbance concerned the effects of environmental noise. The subjective evaluations assessed in the field of speech perception have mainly focused on perceived effort instead of perceived disturbance.

Paper IV investigated how well NH adults recalled native and non-native stories masked by speech in Swedish and English. The approach of recall in combination with a non-native story masked with native and non-native speech has not been investigated before. The application of a native proficiency test was also new in this context.

The effect of speech maskers in native and non-native language performance

In papers I and II, performance was poorer when the masker types consisted of speech maskers as compared to noise maskers.

In paper I the native speech masker was the most interfering in the Swedish target language for the NH listeners. Factors that may explain this strong masking effect are either attentional distraction (Mattys et al., 2010) due to the intelligible and meaningful words in the masker (Hoen et al., 2007; Van Engen & Bradlow, 2007) or strong interference from a familiar language masker (i.e., the linguistic similarity hypothesis, Brouwer, et al., 2012), or both of these factors. The similar linguistic streams embody a difficult condition that challenge attention as well as perceptional processes. Attention is essential to focus on one auditory stream while perceptual processes are required for segregating the streams and inferring the meaning from the favored stream. The NH listeners also had a relative release in masking from the non-native background speech (Calandruccio et al., 2010; Gautreau et al., 2013; Rhebergen et al., 2005; Van Engen, 2010; Van Engen & Bradlow; 2007). In fact, the low proficiency group had the largest release in masking from the non-native background speech, suggesting that reduced intelligibility and thereby reduced understanding of the unfamiliar masker made it easier to suppress the masker, which may have improved the perception of the native target speech. For native Swedish target speech, the HI listeners in paper II performed more poorly in both speech maskers when compared to the noise maskers. This suggests that speech is a source of greater disturbance than noise, irrespective of masker language. This may also suggest that the HI listeners distinguish relatively fewer potentially intelligible words in the maskers when compared to the NH listeners in the Swedish masker. It is also likely that impaired spectral and temporal resolution (Moore, 1985), involving reduced ability to distinguish between different sounds, provides an additional explanation. Even though the HI listeners had the presentation level of the target and the masker adjusted individually according to the Cambridge formula (Moore & Glasberg, 1998) hearing aids cannot compensate for such effects.

For non-native target speech, both the NH listeners and the HI listeners showed a similar pattern of results, with great difficulties for the speech maskers. Recognizing speech in speech is demanding in a non-native language. The results suggest that the difficulties might be due to not sufficiently specified representations of non-native segments and/or incomplete lexical or syntactic knowledge (Bradlow & Alexander, 2007; Mattys et al.,2010; van Wijngaarden et al., 2002). In fact it has been shown that non-native language proficiency is a strong predictor for non-native language comprehension when speech is masked with speech (Kilman et al., 2014, Sörqvist et al., 2014).

Cognitive load, WMC and LTM

In paper IV, the results showed that the English story presented in quiet was better recalled than was the English story presented in the Swedish babble, even though all story-parts were rated as audible by the participants in all conditions. The Swedish background speech put extra demand on the resources available and might have required resources that otherwise could have been used to encode and rehearse the materials (Rabbitt, 1968, 1990; Surprenant, 1999; Wingfield et al., 2005). When masking or distortion does not produce identification errors, a listener may nevertheless have to spend more effort to distinguish what is said.

Speech recognition in adverse listening conditions often results in increased cognitive load (Desjardin & Doherty, 2013; Gosselin & Gagné, 2011; Koelewijn, Shinn-Cunningham, Zekveld, & Kramer, 2014; Zekveld et al., 2011). The underlying inference is that when listening effort increases, the limited cognitive supply must be allocated to speech recognition and therefore, less capacity is available for encoding of the speech information or performance of a secondary task (McCoy et al., 2005; Murphy, Craik, Li, & Schneider, 2000; Rabbitt, 1968; Schneider & Pichora-Fuller, 2000; Surprenant, 1999; Tun et al., 2009). Cognitive load has been defined as any aspect causing extraordinary demands on essential attentional or mnemonic capacities (Mattys & Wiget, 2011). Thus, cognitive load can refer to external as well as internal factors. External factors are, for example, accented or non-native speech and signal degradation (Rönnberg et al., 2010). Individual attentional capacity or WMC can be considered to be internal factors.

In paper IV, larger working memory capacity (Reading Span) was related to better episodic long-term memory performance in the English story/ Swedish speech masker condition. The ELU-model postulates that cognitive abilities and working memory are especially relevant in demanding listening situations. In such listening situations, cognitive processing load will be higher for listeners with lower working memory capacity (Rönnberg, 2003).

It has repeatedly been found that individual differences in working memory capacity predicts the ability to recognize (Rönnberg et al., 2010; Rönnberg et al., 2013) and remember (Kjellberg, Ljung, & Hallman, 2008 Ljung, Sörqvist, Kjellberg, & Green, 2009; Sörqvist & Rönnberg, 2012) masked speech. This is especially true for persons with hearing impairment (Füllgrabe, Stone, & Moore, 2015). In general, it is likely that the more imprecise the speech signal is, the greater the need for cognitive abilities.

When listening to a fictional story as in paper IV, it is likely that the listener unintentionally produces mental images of the events and the characters in the story. The contextual richness in a story may leave imprints/traces on interpretations, increases the associative functions and thereby facilitates the possibility of improved recall. Paivio (1969) assumed that visual imagery increases as a purpose of concreteness; the more concrete and richer the image, the more refined the internal code. An alternative interpretation of imagery was proposed by Bower (1970) who stated that imagery improves memory due to the production of more associations between the items to be recalled.

Unsworth, Fukuda, Awh and Vogel (2014) investigated strategic aspects of recall from semantic and autobiographical LTM. Participants who relied on visualization in self-reported strategies were less inclined to rely on ordered strategies such as alphabetic search. When instructed to use visualization, the participants had the highest levels of performance and the most efficient search, compared to ordered strategies.

A reasonable conclusion is that visualization was of importance in study IV and facilitated retrieval from episodic LTM. Ericsson and Kintsch's (1995) LTM model suggests that variability in WMC is due to dissimilarities in encoding information into LTM and to use retrieval cues for rapid and effective access to essential information. Likewise, a large WMC might facilitate controlled search, which comprises setting up a general retrieval plan, producing retrieval cues to search memory with, and various monitoring decisions (Unsworth, 2010; Unsworth, Brewer, & Spillers, 2009).

ELU and the mismatch function

In the current studies, the mismatch function in the ELU-model was proposed to be activated due to the hearing-impairment (papers II and III), the non-native language (papers I, II, III, IV), the perceived disturbance (paper III), the energetic masking (papers I, II, III) and the informational masking (papers I, II, III, IV). The activation of the mismatch function occurs due to either poor input in the ELU-episodic buffer and/or poor phonological representations in semantic long term memory (Rönnberg et al., 2013).

Generally, successful speech comprehension is the result of both bottom-up processes, relying on sensory, perceptual input, and top-down processes, relying on previous knowledge or expectations. The relative contribution of bottom-up and top-down processes can vary over time and discourse, even within the same sentence as a moment to moment aspect, depending on external (e.g., noise) and internal (e.g., temporarily losing focus) factors (Rönnberg et al., 2010; Rönnberg et al., 2013).

The internal factors discussed in this thesis did not have a variable character as hearing-impairment is probably a permanent condition and the level of English proficiency does not change rapidly. Usually, for a non-native listener, the primary insufficiency is the imprecise representations of non-native language segments (Mattys et al., 2010). It is this imperfect encoding of segments that leads to unsuccessful lexical access and sentential integration, rather than poor lexical or syntactic ability per se (e. g., Bradlow & Alexander, 2007; Bradlow & Bent, 2002) Even so, diffuse lexical activation, as it is when the language is non-native, is likely to lead to more word competition and more instability in the lexical selection process.

The ELU-model assumes that a minimum number of phonological representations needs to be activated in semantic LTM to access a certain lexical item. Above this threshold, the proper lexical representation is selected but below the threshold there are differences depending on how close to the threshold the phonological representations are found. Regarding non-native English language perception, the mismatch function is activated to a higher or lower degree depending on the level of proficiency. In lexical confusability, a high-proficiency individual may activate a number of phonological representations which also lead to the retrieval of phonological neighbors (Luce & Pisoni, 1998). This may eventually activate the proper lexical item. This depends on the degree of contextual support for the target word (e.g. Kalikow, Stevens, & Elliott, 1977). A high proficiency listener may benefit more from co-articulation, and activate more precise phonological and semantic representations in semantic LTM than a low proficiency listener.

For a low proficiency non-native listener, the reliance on bottom-up input is large due to gaps in knowledge. Consequently, a low proficiency listener may be more dependent on contextual cues than a high proficiency listener. Zekveld et al. (2011) showed that semantic cues improved speech intelligibility in noise. However, a low proficiency listener may be easily confused if the contextual cues are not congruent with the target words. This may occur if the speaker changes topic. In such instances, there might be too few phonological representations in the target words to help solving the confusion. (Diffuse lexical activation or no selection at all.)

In many situations, HI listeners depend on the amount of contextual cues for interpreting the surroundings, leading to more top-down activation and explicit processing. Classon, Rudner and Rönnberg (2013) showed that acquired hearing impairment is related to progressive decline of phonological representations in LTM. The longer the time with impoverished auditory input, the less well-defined the mental representations of speech sounds in LTM. This will lead to mismatch in phonologically challenging tasks.

Concerning the external factors presented in this thesis, an important distinction has to be made between the effects of energetic and informational masking. Environmental degradation caused by energetic masking implies that the intelligibility of the target being reduced by physical overlap between the target and masker signals, while informational masking additionally refers to the higher level cognitive implications, caused by semantic interference. The latter commonly affects interpretation of a target and produces attentional dilemmas. With regard to the results presented in this thesis, it is likely that the energetic maskers have required less explicit processing resources than have the informational maskers (see papers I, II, III). The results in papers I, II, III and IV suggest large amounts of interference and a discrepancy between the incoming target signal and the activated phonological/semantic representations in LTM for native informational maskers.

The mismatch functions of the ELU model described above prerequisite a single mismatch activation determined by a single triggering factor. This is seldom the case in a real-life situation, where it is likely that more than one mismatch function is activated. For example, the double disadvantage when an incomplete language capacity is exacerbated by environmental degradation limits the implicit processing. This is exemplified by the ELU-model loop which describes how explicit/implicit processes interact over time. Each explicit loop is activated by a mismatch. How frequently a listener passes through an explicit loop is influenced by the internal and external circumstances at the time. Consequently, for a HI listener who does not master a second language too well, any further disadvantages like speech or noise in the background will result in the any limitations having a cumulative effect, thus creating a greater requirement for explicit processing resources.

In study III, this was exemplified in the perceived disturbance ratings of the eight masker conditions in English and Swedish target speech. Perceived disturbance may be defined as “an interruption of intended listening” and may possibly result in a shift of attentional focus towards the disturbing sound. The extent to which the disturbing sounds interfere depends on internal as well as external factors. The internal factors include hearing-impairment and/or proficiency in English, which may also relate to WMC. It has been claimed that the ability to control attention and/or to ignore background sounds is associated with high WMC (Conway et al., 2001; Kane et al., 2001; Sörqvist et al., 2012; Sörqvist & Rönnberg; 2012). The relevant external factors relate to the actual characteristics of the masker (e.g., energetic versus informational masking aspects). Thus, the perceived disturbance experienced by an individual relates to the degree of activated mismatching internal and external factors and also to the amount of explicit cognitive resources required to adjust the ambiguity between the language input and the stored language representations in the LTM.

Methodological Discussion

English Reading Span

In papers I and II, the English Reading Span was employed as a measure of WMC. As the English sentences were translated directly from the Swedish Reading Span, some of the sentences may have been too difficult for the Swedish participants. Therefore it is difficult to determine the degree to which the English Reading Span was a measure of working memory capacity and/or a measure of English proficiency.

Non-native sentences versus non-native stories

In paper I, the HINT sentences were used in the experimental tests. The HINT consists of short everyday sentences with a length of five to nine syllables each. The participants (NH) listened to a sentence in the non-native language masked by energetic or informational maskers. Immediately after, they *repeated* what they heard. The purpose of paper I was to evaluate performance in the different conditions. In paper IV, participants (NH) listened to a non-native story that was presented unmasked (5 min), or masked with native speech (5 min) or non-native speech (5 min) (counterbalanced order). *Recall* was requested a couple of hours later. The purpose of paper IV was to explore the retrieval of stories masked by native and non-native speech.

The primary difference between the sentences and the stories relates to the contextual circumstances. While the non-native sentences used in the experimental test for paper I were short, leaving little possibilities for contextual interpretations and therefore no top-down activation, the native and non-native stories used for the experimental test in paper IV were relatively long (15 min/each), comprising ample content and substance. The contextual interpretations from the stories may have been derived from the main storyline as well as from smaller episodes, the described environment and characters of the personalities. English proficiency was important for following the English story as well as repeating the English sentences.

Paper I suggested a strong relationship between proficiency in the English language and performance in English target speech in all four conditions. Paper IV suggested a relationship between English proficiency and recall of the English story in two out of three conditions. However, while proficiency was the most important issue in paper I, paper IV also indicated associations between larger WMC (as measured with Reading Span in Swedish) and better recall of the English story. Following a non-native story is challenging and probably loads WMC. In study I, the associations did not suggest any associations between WMC (measured with Reading Span in Swedish) and performance in the four experimental conditions for English target speech, although a tendency was there. So, how could this be explained?

It is well known that complex measures of WMC are usually associated more strongly with higher-order cognitive processes than are simple measures of short-term memory (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Conway & Engle, 1996; Daneman & Merikle, 1996; Kail & Hall, 2001). Also, previous studies have suggested that individual differences in span performance are vital in situations of distraction and interference, or when it is necessary to suppress or inhibit information that is not relevant to the task (Engle et al., 1999). However, other studies have found that the relationship between simple span and measures of higher order cognitive functions can be similar to those found with tests of complex WMC (Bayliss et al., 2005; Shah & Miyake, 1996).

Although paper I included difficult conditions, the current results might suggest that repeating short sentences immediately after hearing them does not merely tap into WMC but possibly into short-term memory (STM). STM usually refers to a limited temporary-storage capacity while WMC refers to a combined limited temporary storage and processing capacity (Baddeley, 2012; Baddeley & Hitch, 1974). The key characteristics attributed to STM are: very limited capacity and susceptibility to distraction that causes forgetting. However, it has been argued that simple span and complex span basically measure the same fundamental processes, like maintenance, rehearsal, controlled search and updating, yet differ in the magnitude to which the processes function during a specific task (Unsworth & Engle, 2007). The degree to which a certain task measures all of these abilities is in part regulated by the scoring procedure and the existence or absence of other processes (e.g. rehearsal) that might influence performance (Unsworth & Engle, 2007).

The result in paper I support previous findings that HINT sentences generally do not show a strong relation with WMC (Rönnberg et al., 2010). In both paper I and paper IV, proficiency in the non-native language was important for the perception of non-native target speech. Furthermore, Paper IV indicated that non-native language proficiency and WMC were equally important for recalling the non-native story. It is not surprising that repeating a short non-native sentence exclusively asks for high English proficiency, as the small amount of contextual information available to aid sentence perception prevents misinterpretations only to a limited extent. It is more likely that the long stretches of speech in a non-native story provide helpful contextual cues that engage top-down processes and thereby tax WMC.

The current conclusion is that listening to and processing a story in a non-native language demands mechanisms other than those required to merely repeating a sentence. Repeating sentences in a non-native language provides less contextual support for target words therefore, proficiency is extremely important. A story in any language provides rich contextual support for perceiving the target words. Language proficiency and WMC are both relevant for recall of the speech information. (Appendix I and II).

Perceived disturbance

In this thesis, perceptual performance (as objectively measured by scoring sentence repetition performance) was assessed in papers I and II which described the performance of NH and HI participants in native and non-native languages with the four different maskers. In paper III, the subjective ratings masker *disturbance* relating to the same conditions were evaluated.

Subjective evaluations of any kind have their benefits and limitations. The limitations derive from the subjective nature, as noise disturbance can be attributed to a range of variables besides the merely acoustical parameters. For example, there is no guarantee that one individual interprets “disturbing” in the same way as another individual. As the concept of disturbance is associated to other words, such as annoyance, concern, bother, displeasure, dissatisfaction, irritation, distress etc, (Koelega, 1987) there may be variations in how individuals define the sense and the meaning of the words. In the study of Guski et al., (1999), noise scientists from different countries (English, German, the Netherlands Japan, USA, Australia, France, Sweden) were asked to suggest the main result of noise and to rate the relationship between the concept of noise annoyance and several related concepts. Annoyance was seen as the main effect of noise, and noise annoyance is a complex psychological concept comprising behavioral and evaluative factors. Even though the scientists agreed upon the core components of the annoyance concept (e.g. disturbance, unpleasantness and nuisance), there were some significant differences between different languages which may be due to the way in which the languages are constructed and/or how the words were weighed against each other (Guski et al.,1999). Whether this was due to the different concepts of annoyance or due to the different connotations of the related words was not analyzed in the study.

The data collection in this thesis was completed in Sweden and the Swedish concept of “störande” was selected as the perceived disturbance is related to the intensity of the noise (and the subsequent feeling of annoyance.) In subjective measurements of noise in English language research, the concept of “annoyance” is commonly used. In Swedish, the main translation of annoyance is “irritated” and this word is associated to a more negative subtext than ‘disturbing’ is.

While self-reported measures in listening disturbance frequently are assessed in environmental studies, another related but not interchangeable measure usually assessed within the field of speech perception is ‘effort’.

As stated, the concepts are associated even though they are not measuring exactly the same underlying construct. Linking the measures to the ELU-model, we propose that perceived disturbance corresponds to the mismatch function and effort corresponds to the degree of explicit processing needed for a given task.

If interpreting and rating disturbance might be difficult for some participants, interpreting effort might even be harder. Previous studies have observed that there is a disagreement in how to conceptualize effort (McGarrigle et al., 2014) and effort has been difficult to validate since it is unclear what exactly has been rated (van Esch et al., 2013).

Another subjective limitation stems from the disturbance thresholds; what one individual finds disturbing may be different from what another individual find disturbing. For example, in this thesis, some individuals rated all conditions on the ten-point perceived disturbance rating-scale between two and four, while others rated all conditions between eight and ten. Additionally, individuals seem to have difficulties in separating task-difficulty from performance accuracy (McGarrigle et al., 2014). To assess the relations between task accuracy and perceived disturbance, the associations between speech reception thresholds (SRTs) and perceived disturbance in the eight conditions were assessed. The results indicated significant associations for NH participants between the SRT for the Swedish language in Babble English condition and the perceived disturbance in the same condition (r = .48), and also between the SRT Swedish language/Babble English condition and the perceived disturbance in the Swedish language/Babble English condition (also r= .48), indicating that higher SRTs (low task performance accuracy) were related to higher perceived disturbance. No other significant associations were found.

The different interpretations, as well as the subjectivity, are drawbacks of self-report measures. Still, subjectivity is a benefit as it provides an understanding of how the individual experiences disturbance in different noise types in native and non-native languages. Therefore, using a combination of objective and subjective measures provides a broader and deeper knowledge of the individuals' experiences and performance in noise.

Future directions

In this thesis, paper IV describes how NH listeners recalled native and non-native stories masked with speech in native and non-native languages.

It would be interesting to further investigate the memory traces in native and non-native languages by including other groups, such as HI listeners, younger (students, 15-16 years) and older participants, above 65 years.

Listening to native and non-native narrative stories may raise a variety of emotions. Therefore, the temporary emotions/mood evoked by condition could be assessed. Additionally, since it is suggested that arousal affects the limited pool of attentional resources (Kahneman, 1973), it would be interesting to evaluate the level of arousal after every masked/unmasked part of the native and non-native stories.

Future studies could employ other stimuli like expository texts. These may tap into the same category as narrative stories but differences may be observed as well. Previous research has shown recall is better for narrative text compared to expository text (Tun, 1989; Zabrucki & Moore, 1999; Panico & Healy, 2009). An expository text can be challenging in itself for memory processing. This would provide an opportunity to test free recall instead of cued recall. Furthermore, what are the consequences for memory processing if for example mathematical tasks, dual tasks and/or factual information are used as stimuli, unmasked/masked with native/non-native speech? It would be interesting to further investigate such issues in school-settings.

Conclusions

- The different perspectives and the extended design in this thesis allowed for an integrated approach aimed at explaining intra- and inter-individual differences relevant for native and non-native speech perception in difficult conditions.
- When the target language is not well established, such as when it is in a non-native language, the interference from background speech in the dominant native language is detrimental to speech perception performance, perceived disturbance and memory processes. However, large variability between the listeners was observed. Part of this variability relates to non-native language proficiency.
- The native babble masker is a relatively intrusive masker for NH listeners due to attentional distraction and the presence of intelligible and meaningful words in the masker. This leads to difficulties in suppressing the native masker in favor of the target signal. However, a non-native or unfamiliar speech masker usually results in a relative release. For HI listeners, both speech maskers provide strong masking effects when compared to the noise maskers.
- Listening to a non-native target in background speech is most likely taxing for the HI listener, already coping with impaired spectro-temporal resolution leading to a reduced ability to separate different sounds and speech segments.
- In general, non-native listeners depend on contextual cues to serve as a compensatory mechanism for a less well mastered language. It is more likely that these cues will be provided in a story than in a short sentence.
- Speech perception performance and perceived disturbance for NH and HI are complimentary measures and are both valuable. The different pattern of results as well as the different relations with other variables provides unique insights into the relevant processes.

- In situations where internal listener-related factors like hearing impairment and/or an imperfectly mastered language, and external factors like noise and speech do not match the expectations, explicit processing is necessary.
- An interfering native background speech may decrease the possibility to secure LTM processes and may therefore impair retrieval.
- Increased perceptual and cognitive effort may reduce memory capacity, even when stimuli are appropriately identified at a perceptual level.
- A large WMC may result in a better ability to inhibit distractions and allocate processing resources to meet assigned objectives.

Some final words

Taking a step back and considering the lessons learned from the thesis: Several issues have been shown to be of importance but some were not. In this context, non-verbal intelligence (Raven) turned out to be of minor relevance. However, of importance are the large individual differences reported in this thesis. This suggests that an individualized approach is important in clinical or educational settings when non-native persons or people with hearing impairment need to perceive and remember potentially vital information.

Finally…this thesis gives us new insight into the story presented at the very beginning. It is very possible that I didn't remember the English conversations at the party due to the non-native English and the background sounds!

Acknowledgements

Att skriva en avhandling är verkligen inte ett en-persons jobb. Denna avhandling hade inte existerat utan stöd och hjälp från många personer.

Mitt första och djupaste tack till alla fantastiska deltagare som delat med sig av sin tid och medverkat på de ibland påfrestande sessioner som föregått de artiklar och manus som finns i den här boken.

Ett Stort Tack också till:

Jerker Rönnberg för din expertis, din djupa kunskap som du delar med dig av. Du är och har i många fall varit den sista ”utposten”, det vill säga den person som har svaret när alla andra går bet. Du hittar alltid en kreativ lösning. För din aldrig sinande energi som du ger ut, för roliga och intressanta diskussioner. För att du är en klippa i skarpt läge. För att jag (efter ett litet tag☺) fick göra den datainsamling som jag ville.

Adriana Zekveld for outstanding skill, so amazingly bright, so trustworthy. You have helped me so much during this years, over and over again with an angels’ patience. This thesis would certainly not exist without you.

Mathias Hällgren för din vänlighet och ett otroligt tålamod med att förklara all teknisk audiologi som skulle ingå i datainsamlingarna. För att du tog hand om alla inspelningar som behövde göras och för snabba svar i krislägen.

Elaine Ng, you have a wonderful personality, a fantastic combination of intelligence and helpfulness and you are truly a good humored company. My deepest thanks for all your help and for being there, again, the very last hours before the printing started…

Cecilia Henriksson, för din öppna dörr, för att du är en så underbar person som ger hjälp och stöd i alla lägen. För din omtanke. Du är verkligen en person som gör skillnad. För din underfundiga humor – inte alltid tolkningsbar…

Cecilia Nakeva von Mentzer för din vänskap och att du delar med dig av din djupa förståelse av språkprocesser, för resan till Brac och för jag därigenom fick en ny vän, Nelli Kalnak.

Josefine Andin som är min nya smarta ”granne”. Du är så bra! Tack för din hjälp med att hitta en lösning i slutstressen när min hjärna hade lagt av.

Tack alla kollegor på avdelningen för att det alltid är roligt att komma dit: Lisa Palmqvist för hjälp, att du är så positiv och att du gör avdelningen till en ännu bättre plats, Mary Rudner för din vänlighet, intresse och goda råd, Rachel for fantastic proof reading and for ”lending” your voice, Håkan Hua för hjälp och support, Emilie Nordqvist för roliga pratstunder. Tack Anna Levén och Carine Signoret för support, Anna Wretman, Björn Lyxell, Wycliffe Jumba och Shahram Moradi.

Tack alla rumskamrater för trevligt prat och ”heja, heja”: Emil Holmer, Jakob Dahl, Michaela Socher och fina Malin Wass. Jag kommer sakna dig. Henrik Danielsson för att du såg och agerade. Örjan Dahlström för vänlighet och pedagogisk, tålmodig hjälp med statistiken.

Maria Hugo-Lindén och Marie Louise Lund Mattson, vilken otrolig tillgång ni är, ni gör ett fantastiskt jobb! Ett stort tack till Maria för all hjälp, ditt glada och tålmodiga upprepande av alla instruktioner när absolut inget går in.

Tack Anja Torsten och Karin Forslund-Frykedal som är så fina vänner och kollegor som någon kan ha. Ett speciell tack till Karin för insatsen de sista skälvande timmarna…

Tack för support, hjälp och goda råd: Gisela Eckert, Eva Hammar Chiriac, Annette Larsson, Gunvor Larsson Abbad, Tove Mattson och Camilla Forsberg för gott samarbete.

Tack Tomas Bjuvmar på US för att du lärde en ovetande!

Ett stort tack till mina makalösa vänner som alla ställt upp och hjälpt till på olika sätt så att denna avhandling har kunnat bli till. Som lyssnat och stöttat och lyssnat igen. Som förundrat genom sitt intresse för avhandlingsarbetet. Ett speciellt tack till fantastiska Ros-Marie Pettersson och Else Hafvenstein som använt sina personliga resurser och begåvningar för att hjälpa till, Zagah Askling, Marie Bergkvist, Agneta Svensson, Annika Gustavsson, Lisbeth Svensson, Margareta Windahl, Åsa och Per Olsson och Kenneth Lundgren.

Lena Svensson och Ingrid Midbäck. Mina barn-och ungdomsvänner som är vänner för livet. Så bra ni är och vilken glädje ni ger!

Min familj som är helt avgörande för mig:

Mamma och pappa för all kärlek och stöd. Varje dag. Alltid.

Min underbara syster Marian Jansson som är den bästa syster som någon kan ha. Som alltid finns där och stöttar. Ulf Jansson min svåger som gjort en liknande resa och förstår, för ditt intresse och insiktsfulla kommentarer. Syskonbarnen Kristin och Lasse båda betydelsefulla samt Kajsa som funnits där som den fina person du är.

Och till sist, den största glädjen, ni är hjärtat i mitt liv och underbara alla tre:

Min son Gustav Kilman som kom på titeln till avhandlingen och som använt sin kreativitet för att göra framsidan. Min svärdotter Dali Kilman som använt sina språkkunskaper för att hjälpa till och mitt ljuvliga barnbarn Elvin som väckt en sådan kärlek, överraskande men självklar.

References

Agus, T. R., Akeroyd, M. A., Gatehouse, S., & Warden, D. (2009). Informational masking in young and elderly listeners for speech masked by simultaneous speech and noise. *J Acoust Soc Am, 126*(4), 1926-1940.

Alexander, J. M., & Lufti, R. A. (2004). Informational masking in hearing-impaired and normal-hearing listeners: Sensation level and decision weights. *J Acoust Soc Am, 116*(4), 2234-2247.

Andringa, S., Olsthoorn, N., van Beuningen, C., Schoonen, R., & Hulstijn, J. (2012). Determinants of success in native and non-native listening comprehension: An individual differences approach. *Lang Learn, 62* (Suppl. 2), 49-78.

Arbogast, T. L., Mason, C. R., & Kidd, G. Jr. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, *117*(4), 2169-2180.

Arlinger, S. E. (2007). *Nordisk lärobok i audiologi* (1st ed). Bromma, Sweden: C. A. Tegnér AB.

Arlinger, S., Lunner, T., Lyxell, B., & Pichora-Fuller, M. K. (2009). The Emergence of Cognitive Hearing Science. *Scand J Psychol, 50*(5), 371-384.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends Cogn Sci, 4*, 417-423.

Baddeley, A. D. (2012). Working Memory: Theories, models, and controversies. *Annu Rev Psychol, 63*, 1-29.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (pp. 47-89). London: Academic Press.

Baker, C. (1993). *Foundations of bilingual education and bilingualism*. Clevdon, Avon, U.K.: Multicultural Matters.

Bayliss, D. M., Jarrold, C., Baddeley, A. D., & Gunn, D. M. (2005). The relationship between short-term memory and working memory: Complex span made simple? *Memory, 13*(3-4), 414-421.

Beaman, C. P. (2004). The irrelevant sound phenomenon revisited: What role for working memory capacity. *J Exp Psychol Learn Mem Cogn, 30*(5), 1106-1118.

Best, V., Gallun, F. J., Ihlefeld, A., & Shinn-Cunningham, B. G. (2006). The influence of spatial separation on divided listening. *J Acoust Soc Am, 120*(3), 1506-1516.

Best, V., Gallun, F. J., Mason, C. R., Kidd, G. Jr., & Shinn-Cunningham, B. G. (2010). The impact of noise and hearing loss on the processing of simultaneous sentences. *Ear Hear, 31*(2), 213-220.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy and cognition.* Cambridge, England: Cambridge University Press.

Bloom, P. (2004). Children think before they speak. *Nature, 430*, 410-411.

Bower, G. H. (1970). Imagery as a relational organizer in associative learning. *J Verb Learn Verb Behav, 9*, 529-533.

Bradley, J. S., & Gover, B. N. (2010). Speech levels in meeting rooms and the probability of speech privacy problems. *J Acoust Soc Am, 127*(2), 815-822.

Bradlow, A. R., & Alexander, J. A. (2007). Semantic-contextual and acoustic-phonetic enhancements for English sentences in noise recognition by native and non-native listeners. *J Acoust Soc Am, 112*(4), 272-284.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *J Acoust Soc Am, 112*(1), 1562-1573.

Brouwer, S., van Engen, K. J., Calandruccio, L., & Bradlow, A. R. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. *J Acoust Soc Am*, *131*(2), 1449-1459.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am, 109*(3), 1101-1109.

Calandruccio, L., Brouwer, S., van Engen, K. J., Dhar, S., & Bradlow, A. R. (2013). Masking release due to linguistic and phonetic dissimilarity between the target and masker speech. *Am J Audiol*, *22*, 157-164.

Calandruccio, L., Dhar, S., & Bradlow, A. R. (2010). Speech-on-speech masking with variable access to the linguistic content of the masker speech. *J Acoust Soc Am*, *128*(2), 860-869.

Cherry, E. C. (1953). Some experiments on the recognition of speech with one and with two ears. *J Acoust Soc Am*, *25*, 975-979.

Clark, J. G. (1981). Uses and abuses of hearing loss classification. *ASHA, 23*, 493-500.

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

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychol Bull Rev, 8*, 331-335.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed and general fluid intelligence. *Intelligence*, *30*, 163-183.

Conway, A. R. A., & Engle, R. W. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory, 4*, 577-590.

Cooke, M. (2006). A glimpsing model of speech perception in noise. *J Acoust Soc Am, 119*(3), 1562-1573.

Cooke, M., Garcia Lecumberri, M. L., & Barker, J. P. (2008). The foreign language cocktail party problem: energetic and informational masking effects in non-native speech perception. *J Acoust Soc Am, 123*(1), 414-427.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing; A framework for memory research. *J Verb Learn Verb Behav*, 11, 671-684.

Crystal, D. (1997). *A dictionary of linguistics and phonetics* (4th Ed). Cambridge, MA: Blackwell.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in integrating information between and within sentences. *J Exp Psychol Learn Mem Cogn, 9*(4), 561-584.

Daneman, M., & Merikle, P. M. (1996). Working memory and Language comprehension: A meta-analysis. *Psychon Bull Rev, 3*(4), 422-433.

Desjardin, J. L., & Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. *Ear Hear, 34*(3), 261-272.

Dimino, J., Gersten, R., Carnine, D., & Blake, G. (1990). Story grammar: An approach for promoting at-risk secondary students' comprehension of literature. *Elementary School Journal, 91*(1), 19-32.

Dix, M. R., Hallpike, C. S., & Hood, J. D. (1968). Loudness recruitment and its measurement with special reference to the loudness discomfort level test and its value in diagnosis. *Ann Otol Rhinol Laryngol, 77*(6), 1131-1151.

Engle, R. W. (2002). Working memory capacity as executive attention. *Curr Dir Psychol Sci, 11*, 19-23.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *J Exp Psychol Gen, 128*(3), 309-331.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychol Rev, 102*(2), 211-245.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *J Acoust Soc Am, 88*(4), 1725-1736.

Füllgrabe, C., Moore, B. C. J., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal-hearing: contributions from auditory temporal processing and cognition. *Front Aging Neurosci*, *6*:347. doi: 10.3389/fnagi.2014.00347.

Gathercole, S. E., (2007). Working memory and Language. In G. M. Gaskell (Ed.). *The Oxford Handbook of Psycholinguistics*. New York: Oxford University Press.

Gathercole, S. E., & Alloway, T. P. (2008). *Working memory and learning: A practical guide for teachers.* London: Sage Publications.

Gautreau, A., Hoen, M., & Meunier, F. (2013). Let's all speak together! Exploring the masking effects of various languages on spoken word identification in multi-linguistic babble. *PLoS ONE, 8*(6). doi: 10.1371/journal.pone.0065668

Gosselin Andersson, P., & Gagné, J-P. (2011). Older Adult Expend More Listening Effort Than Young Adults Recognizing Speech in Noise. *J Speech Lang Hear Res, 54*(3), 944-958.

Guski, R., Felscher-Suhr, U., & Schuemer, R. (1999). The concept of noise annoyance: How international experts see it. *J Sound Vibr*, *223*(4), 513-527.

Harley, T. A. (2001). *The psychology of language: From data to theory* (2nd Ed.). Hove, UK: Psychology Press.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation (*Vol. 22, pp. 193-225). San Diego CA: Academic Press.

Hoen, M., Meunier, F., Grataloup, C-L., Pellegrino, F., Grimault, N., Perrin, F., …Collet, L. (2007). Phonetic and lexical interferences in informational masking during speech-in-speech comprehension*. Speech Commun, 49*(12), 905-916.

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*(4), 227-237.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension – individual differences in working memory. *Psychol Rev, 99*, 122-149.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kail, R., & Hall, L. K. (2001). Distinguishing short-term memory from working memory. *Mem Cogn*, *29*, 1-9.

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J Acoust Soc Am, 61*(5), 1337-1351.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *J Exp Psychol Gen, 130*(2), 169-183.

Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *J Exp Psychol Learn Mem Cogn, 26*, 336-358.

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., … von Wedel, H. (2003). Candidature for and delivery of audiological services: Special needs of older people. *Int J Audiol, 42*(Suppl 2), S92-S101.

Kilman, L., Zekveld, A. A., Hällgren, M., & Rönnberg, J. (2014). The influence of non-native language proficiency on speech perception performance. *Front Psychol, 5*:651*.* doi: 10.3389/fpsyg.2014.00651.

Kjellberg, A., Ljung, R., & Hallman, D. (2008). Recall of words heard in noise. *Appl Cognit Psychol, 22*, 1088-1098.

Klingberg, T. (2011). *Den lärande hjärnan. Om barns minne och utveckling.* Stockholm: Natur & Kultur.

Knez, I.*, &* Hygge, S*.* (2002). Irrelevant speech and indoor lighting: Effects on cognitive performance and self-reported affect*. Appl Cognit Psychol, 16***,** 709-718*.*

Koelega, H. S. (1987). *Environmental annoyance: Characterization, measurement and control. Proceedings of the international symposium on environmental annoyance*. In H. S. Koelega (Ed.). Introduction: Environmental annoyance (pp. 1-7). Amsterdam: Elsevier.

Koelewijn, T., Shinn-Cunningham, B. G., Zekveld, A. A., & Kramer, S. E. (2014). The pupil response is sensitive to divided attention during speech processing. *Hear Res, 312*, 114-120.

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear Hear, 32*, 291-300.

Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sounds of languages.* Oxford: Blackwell.

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. *Int J Audiol, 44*(3), 131-143.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Lang Speech*, *6*, 172-187.

Ljung, R., Sörqvist, P., Kjellberg, A., & Green, A-M. (2009). Poor listening conditions impair memory of intelligible lectures: Implications for acoustical classroom standards. *Building Acoustics, 16*(3), 257-265.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear Hear, 19*(1), 1-36.

Mandler, J. & Johnson, N. (1977). Remembrance of things parsed: Story structure and recall. *Cog Psych*, 9, 111-151.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brainimaging. *Brain Lang*, *86*, 70-82.

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: dissociating energetic from informational factors. *Cogn Psychol, 59*, 203-243.

Mattys, S. L., Carroll, L. M., Carrie, K. W., Li, S., & Chan, L. Y. (2010). Effects of energetic and informational segmentation by native and non-native speakers. *Speech Comm, 52*, 887-899.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse condition: A review. *Lang Cogn Proc, 27*(7-8), 953-978.

Mattys, S. L., & Wiget, L. (2011). Effect of cognitive load on speech recognition. *J Mem Lang, 65*, 145-160.

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. Q*uart J Exp Psychol, 58A*(1), 22-33.

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group "white paper". *Int J Audiol, 53*, 433-445.

Miyake, A., Friedman, N. P., Emerson, M. J., Witziki, A. H., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contribution to complex frontal lobe tasks: A latent variable analysis. *Cogn Psychol, 41*(1), 49-100.

Moore, B. C. J. (1985). Frequency selectivity and temporal resolution in normal and hearing-impaired listeners. *Br J Audiol*, *19*, 189-201.

Moore, B. C. J., & Glasberg, B. R. (1998). Use of a loudness model for hearing-aid fitting. I. Linear hearing aids. *Br J Audiol, 32*(5), 317-335.

Murphy, D. R., Craik, F. I., Li, K. Z., & Schneider, B. A. (2000). Comparing the Effects of Aging and Background Noise on Short-Term Memory Performance. *Psychol Aging, 15*, 323-334.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise*. J Acoust Soc Am, 95*, 1085-1099.

Olsthoorn, N. M., Andringa, S., & Hulstijn, J. H. (2012). Visual and auditory digit – span performance in native and non-native speakers. *Int J Bilingualism*, *18*(6), 663-673.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychol Rev, 76*, 241-263.

Panico, J., & Healey, C. E. (2009). Influence of Text Type, Topic Familiarity, and Stuttering Frequency on Listener Recall, Comprehension and Mental Effort. *J Speech Lang Hear Res, 52*, 534-546.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., & Dehaene, S. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain, 121*, 1841-1852.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J Acoust Soc Am, 97*(1), 593-608.

Plomp, R. (2002). *The intelligent ear*. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Rabbitt, P. (1968). Channel capacity, intelligibility and immediate memory. *Quart J Exp Psychol, 20*, 241-248.

Rabbitt, P. (1990). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngol Suppl, 476*, 167-175.

Raven, J., Raven, J. C., & Court, J. H. (2000). *Manual for Raven's progressive Matrices and vocabulary scales. Section 3: The standard progressive matrices.* San Antonio, TX: Harcourt Assessment.

Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. A. (2005). Release from informational masking by time reversal of native and non-native interfering speech. *J Acoust Soc Am, 118*, 1274-1277.

Rudner, M., Foo, C., Rönnberg, J., & Lunner, T. (2009). Cognition and aided speech recognition in noise: Specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scand J Psychol, 50*, 405-418.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *Int J Audiol, 42*, 68-76.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., … Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical and clinical advances. *Front Syst Neurosci, 7:*31. doi:10.3389/fnsys.2013.00031.

Rönnberg, J., Lyxell, B., Arlinger, S., & Kinnefors, C. (1989). Visual evoked potentials: Relation to adult speechreading and cognitive function. *J Speech Hear Res, 32*(4), 725-735.

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *Int J Audiol, 47*, 99-105.

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

Sanders, L. D., Neville, H. J., & Woldorff, M. G. (2002). Speech segmentation by native and non-native speakers: The use of lexical, syntactic and stress-pattern cues. *J Speech Lang Hear Res, 45*, 519-530.

Schlittmeier, S. J., Hellbrück, J., Thaden, R., & Vorländer, M. (2008). The impact of background speech varying in intelligibility: Effects on cognitive performance and perceived disturbance. *Ergonomics, 51*(5), 719.

Schneider, B. A., Li, L., & Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *J Am Acad Audiol, 18,* 559-572.

Schneider, B. A., & Pichora-Fuller, M. K. (2000). Implications of perceptual deterioration for cognitive aging research. In F. I. M. Craik & T. A. Salthouse (Eds), *Handbook of aging and cognition* (2nd ed., pp. 155-220). Mahwah, NJ: Erlbaum.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *J Exp Psychol Gen, 125,* 4-27.

Surprenant, A. M. (1999). The effect of noise on memory for spoken syllables. *Int J Psychol, 34*, 328-333.

Sörqvist, P., Hurtig, A., Ljung, R., & Rönnberg, J. (2014). High second-language proficiency protects against the effects of reverberation on listening comprehension. *Scand J Psychol, 55*, 91–96.

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

Sörqvist, P., & Rönnberg, J. (2012). Episodic long-term memory of spoken discourse masked by speech: what is the role for working memory capacity? *J Speech Lang Hear Res, 55*(1), 210-218.

Sörqvist, P., Stenfelt, S., & Rönnberg, J. (2012). Visual-verbal cognitive load modulates auditory-sensory gating in the brainstem. *J Cog Neurosci, 24*(11), 2147-2154.

Tun, P. A. (1989). Age differences in processing expository and narrative text. *J Gerontol, 44*(1), 9-15.

Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, Hearing Acuity and the Attentional Cost of Effortful Listening. *Psychol Aging, 24*(3), 761-766.

Unsworth, N. (2010). On the division of working memory and long-term memory and their relation to intelligence: A latent variable approach. *Acta Psychol (Amst), 134*(1), 16-28.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2009). There’s more to the working memory-fluid intelligence relationship than just secondary memory. *Psychol Bull Rev, 16,* 931-937.

Unsworth N., & Engle. R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychol Bull*, *133*(6), 1038-1066.

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control and secondary memory retrieval. *Cogn Psychol, 71*, 1-26.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher order cognition: A latent–variable analysis of the relationship between processing and storage. *Memory*, *17*(6), 635-654.

Van Engen, K. J. (2010). Similarity and familiarity: Second language sentence recognition in first and second-language multi-talker babble. *Speech Commun, 52*(11-12), 943-953.

Van Engen, K. J., & Bradlow, A. R. (2007). Sentence recognition in native and foreign language multi-talker background noise. *J Acoust Soc Am, 121*, 519-526.

Van Esch, T. E., Kollmeier, B., Vormann, M., Lyzenga, J., Houtgast, T., Hällgren, M., ... Dreschler, W. A. (2013). Evaluation of the preliminary auditory profile test battery in an international multi-centre study. *Int J Audiol, 52*(5), 305-321.

Van Wijngaarden, S. J., Steeneken, H. J. M., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *J Acoust Soc Am, 111(4)*, 1906-1916.

Venetjoki, N., Kaarlela-Tuomaala, A., Keskinen, E., & Hongisto, V. (2006). The effect of speech and speech intelligibility on task performance. *Ergonomics, 49*(11), 1068-1091.

Von Hapsburg, D., & Peña, E. D. (2002).Understanding bilingualism and its impact on speech audiometry. *J Speech Lang Hear Res, 45*, 202-213.

Weber, A., & Paris, G. (2004). The origin of the linguistic gender effect in spoken-word recognition: evidence from non-native listening. *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 1446–1451). Mahwah, NJ: Erlbaum.

Weiss, D., & Demsey, J. J. (2008). Performance of bilingual Speakers on the English and Spanish Versions of the Hearing in Noise Test (HINT*). J Am Acad Audiol, 19*, 5-17.

Wingfield, A. (1996). Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory. *J Am Acad Audiol, 7*, 175-182.

Wingfield, A., Tun, P. A., & McCoy, S. L. (2005). Hearing loss in older adulthood, what it is and how it interacts with cognitive performance. *Curr Dir Psychol Sci, 14*(3), 144-148.

World Health Organization. (2001). Hearing aids and services for developing countries. *Rev Panam Salud Publica, 10* (2), 139-142.

Zabrucki, K., & Moore, D. (1999). Influence of text genre on adults' monitoring of understanding and recall. *Educ Gerontol, 25*, 691-710.

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

Zekveld, A. A., Rudner, M., Johnsrude, I. S., Festen, M. J., van Beek, J. H. M., & Rönnberg, J. (2011). The Influence of Semantically Related an Unrelated Text Cues on the Intelligibility of Sentences in Noise. *Ear Hear, 32*, 16-25.

Appendices

Appendix I. HINT sentences in English and Swedish. (Paper I, II, III)

Appendix II. Excerpts from the English story and the Swedish story. (Paper IV)

HINT sentences in English and Swedish:

The boy broke the wooden fence

The boy did a handstand

The cat drank from the saucer

She’s washing her new silk dress

The young people were dancing

The tub faucet is leaking

He found his brother hiding

She looked in the mirror

He broke his leg again

He needs his vacation

De svarta fåren går på ängen

Staketet målades vitt

Fem tjejer sprang på skolgården

Stegen är två meter lång

Flicka spiller sås på tröjan

Han förlorade partiet

Under fönstret stod en säng

Ljuset brinner långsamt

Flicka lyssnade på talet

Flickan har tio väskor

Excerpts from the English story and the Swedish story:

“The path began to rise sharply. It went up and up. The air was thin and it was difficult to breathe. The fog was dense at the time and mercifully covered the depth below the path. I followed my guide closely and began to hear the sound of gongs far away. The sound increased in intensity, the closer we came to the peak of the mountain and on the top I was surrounded by a swirl of sounds. The clouds that encircled our bodies, disappeared and a hidden valley extended below our feet.”

”Jag mötte honom i ett gathörn. Det regnade och blåste den dagen och jag skyndade hem efter lektionerna, våt om fötterna efter att ha trampat i en vattenpöl. Paraplyet vände i blåsten och jag kämpade med det när en man kom springandes runt hörnet på Rue du Temple. Han hade huvudet nerböjt i motvinden samtidigt som han försökte hålla kvar sin hatt på huvudet. Kollisionen var oundviklig och jag tappade balansen. I nästa stund låg jag på gatan, blöt och smutsig. En ström av franska och armar som hjälpte mig upp, hans profil var lugn och ren i motljuset och det var som jag alltid känt honom.”

Studies from the Swedish Institute for Disability Research

1. **Varieties of reading disability**

   Stefan Gustafson

   ISBN 91-7219-867-2, 2000

2. **Cognitive functions in drivers with brain injury – anticipation and adaptation**

   Anna Lundqvist

   ISBN 91-7219-967-9, 2001

3. **Cognitive deafness**

   Ulf Andersson

   ISBN 91-7373-029-7, 2001

4. **Att lära sig leva med förvärvad hörselnedsättning sett ur par-perspektiv**

   Carin Fredriksson

   ISBN 91-7373-105-6, 2001

5. **Signs, Symptoms, and Disability Related to the Musculo-Skeletal System**

   Gunnar Lundberg

   ISBN 91-7373-160-9, 2002

6. **Participation – Ideology and Everyday Life**

   Anette Kjellberg

   ISBN 91-7373-371-7, 2002

7. **Föräldrar med funktionshinder – om barn, föräldraskap och familjeliv**

   Marie Gustavsson Holmström

   ISBN 91-7203-500-5, 2002

8. **Active wheelchair use in daily life**

   Kersti Samuelsson

   ISBN 91-7373-196-X, 2002

9. **Två kön eller inget alls. Politiska intentioner och vardagslivets realiteter i den arbetslivsinriktade rehabiliteringen**

   Marie Jansson

   ISBN 91-7373-568-X, 2003

10. **Audiological and cognitive long-term sequelae from closed head injury**

    Per-Olof Bergemalm

    ISBN 91-7668-384-2, 2004

11. **Att vara i särklass – om delaktighet och utanförskap i gymnasiesärskolan**

    Martin Molin

    ISBN 91-85295-46-9, 2004

12. **Rättvis idrottsundervisning för elever med rörelsehinder – dilemma kring omfördelning och erkännande**

    Kajsa Jerlinder

    Licentiate Degree, 2005

13. **Hearing impairment and deafness. Genetic and environmental factors – interactions – consequences. A clinical audiological approach**

    Per-Inge Carlsson

    ISBN 91-7668-426-1, 2005

14. **Hearing and cognition in speech comprehension. Methods and applications**

    Mathias Hällgren

    ISBN 91-85297-93-3, 2005

15. **Living with deteriorating and hereditary disease: experiences over ten years of persons with muscular dystrophy and their next of kin**

    Katrin Boström

    ISBN 91-7668-427-x, 2005

16. **Disease and disability in early rheumatoid arthritis**

    Ingrid Thyberg

    ISBN 91-85299-16-2, 2005

17. **"Varför får jag icke följa med dit fram?" Medborgarskapet och den offentliga debatten om dövstumma och blinda 1860-1914**

    Staffan Bengtsson

    ISBN 91-85457-06-X, 2005

18. **Modalities of Mind. Modality-specific and nonmodality-specific aspects of working memory for sign and speech**

    Mary Rudner

    ISBN 91-85457-10-8, 2005

19. **Facing the Illusion Piece by Piece. Face recognition for persons with learning disability**

    Henrik Danielsson

    ISBN 91-85497-09-6, 2006

20. **Vuxna med förvärvad traumatisk hjärnskada – omställningsprocesser och konsekvenser i vardagslivet. En studie av femton personers upplevelser och erfarenheter av att leva med förvärvad traumatisk hjärnskada**

    Thomas Strandberg

    ISBN 91-7668-498-9, 2006

21. **Nycklar till kommunikation. Kommunikation mellan vuxna personer med grav förvärvad hjärnskada och personernas närstående, anhöriga och personal**

    Pia Käcker

    ISBN 978-91-85715-88-6, 2007

22. **"Aspergern, det är jag". En intervjustudie om att leva med Asperger syndrom**

    Gunvor Larsson Abbad

    ISBN 978-91-85831-43-2, 2007

23. **Sounds of silence – Phonological awareness and written language in children with and without speech**

    Janna Ferreira

    ISBN 978-91-85895-74-8, 2007

24. **Postponed Plans: Prospective Memory and Intellectual Disability**

    Anna Levén

    ISBN 978-91-85895-57-1, 2007

25. **Consequences of brain tumours from the perspective of the patients and of their next of kin**

    Tanja Edvardsson

    ISBN 978-91-7668-572-3, 2008

26. **Impact on participation and service for persons with deafblindness**

    Kerstin Möller

    ISBN 978-91-7668-595-2, 2008

27. **Approaches to Audiological Rehabilitation with Hearing Aids: studies on prefitting strategies and assessment of outcomes**

    Marie Öberg

    ISBN 978-91-7393-828-0, 2008

28. **Social Interaction and Participation in Activities of Everyday Life Among Persons with Schizophrenia**

    Maria Yilmaz

    Licentiate Degree, 2009

29. **Focus on Chronic Disease through Different Lenses of Expertise Towards Implementation of Patient-Focused Decision Support Preventing Disability:**

    The example of Early Rheumatoid Arthritis

    Örjan Dahlström

    ISBN 978-91-7393-613-2, 2009

30. **Children with Cochlear Implants: Cognition and Reading Ability**

    Malin Wass

    ISBN: 978-91-7393-487-9, 2009

31. **Restricted participation: Unaccompanied children in interpreter-mediated asylum hearings in Sweden**

    Olga Keselman

    ISBN: 978-91-7393-499-2, 2009

32. **Deaf people and labour market in Sweden. Education – Employment – Economy**

    Emelie Rydberg

    ISBN: 978-91-7668-725-3, 2010

33. **Social rättvisa i inkluderande idrottsundervisning för elever med rörelsehinder – en utopi?**

    Kajsa Jerlinder

    ISBN: 978-91-7668-726-0, 2010

34. **Erfarenheter av rehabiliteringsprocessen mot ett arbetsliv – brukarens och de professionellas perspektiv**

    Helene Hillborg

    ISBN: 978-91-7668-741-3, 2010

35. **Knowing me, knowing you – Mentalization abilities of children who use augmentative and alternative communication**

    Annette Sundqvist

    ISBN: 978-91-7393-316-2, 2010

36. **Lärare, socialsekreterare och barn som far illa – om sociala representationer och interprofessionell samverkan**

    Per Germundsson

    ISBN: 978-91-7668-787-1, 2011

37. **Fats in Mind**
    Effects of Omega-3 Fatty Acids on Cognition and Behaviour in Childhood

    Ulrika Birberg Thornberg

    ISBN: 978-91-7393-164-9, 2011

38. **”Jobbet är kommunikation”**
    Om användning av arbetshjälpmedel för personer med hörselnedsättning

    Sif Bjarnason

    Licentiate Degree. ISBN: 978-91-7668-835-9, 2011

39. **Applying the ICF-CY to identify everyday life situations of children and youth with disabilities**

    Margareta Adolfsson

    ISBN: 978-91-628-8342-3, 2011

40. **Tinnitus – an acceptance-based approach**

    Vendela Zetterqvist

    ISBN: 978-91-7393-040-6, 2011

41. **Applicability of the ICF-CY to describe functioning and environment of children with disabilities**

    Nina Klang

    ISBN: 978-91-7668-864-9, 2012

42. **Bringing more to participation**
    Participation in school activities of persons with Disability within the framework of the International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY)

    Gregor Maxwell

    ISBN: 978-91-628-8484-0, 2012

43. **From Eye to Us.**
    Prerequisites for and levels of participation in mainstream school of persons with Autism Spectrum Conditions

    Marita Falkmer

    ISBN: 978-91-637-2091-8, 2013

44. **Otosclerosis, clinical long-term perspectives**

Ylva Dahlin-Redfors

ISBN 978-91-628-8617-2, 2013

45. **Tinnitus in Context - A Contemporary Contextual Behavioral Approach**

Hugo Hesser

ISBN 978-91-7519-701-2, 2013

46. **Hearing and middle ear status in children and young adults with cleft palate**

Traci Flynn

ISBN 978-91-628-8645-5, 2013

47. **Utrymme för deltagande, beslutsprocesser i möten mellan patienter med ospecifika ländryggsbesvär och sjukgymnaster i primär vård**

Iréne Josephson

ISBN 42-978-91-85835-41-6, 2013

48. **”Man vill ju klara sig själv” Studievardagen för studenter med Asperger syndrom i högre studier**

Ann Simmeborn Fleischer

ISBN 978-91-628-8681-3, 2013

49. **Cognitive erosion and its implications in Alzheimer’s disease**

Selina Mårdh

ISBN 978-91-7519-621-1, 2013

50. **Hörselscreening av en population med utvecklingsstörning**
Utvärdering av psykoakustisk testmetod och av OAE-registrering som komplementär metod

Eva Andersson

Licentiate Degree. ISBN 978-91-7519-616-9, 2013

51. **Skolformens komplexitet – elevers erfarenheter av skolvardag och tillhörighet i gymnasiesärskolan**

Therése Mineur

ISBN 978-91-7668-951-6, 2013

52. **Evaluating the process of change:**
Studies on patient journey, hearing disability acceptance and stages-of-change

Vinaya Kumar Channapatna Manchaiah

ISBN 978-91-7519-534-6, 2013

53. **Cognition in hearing aid users: Memory for everyday speech**

Hoi Ning (Elaine) Ng

ISBN 978-91-7519-494-3, 2013

54. **Representing sounds and spellings Phonological decline and compensatory working memory in acquired hearing impairment**

Elisabet Classon

ISBN 978-91-7519-500-1, 2013

55. **Assessment of participation in people with a mild intellectual disability**

Patrik Arvidsson

ISBN 978-91-7668-974-5, 2013

56. **Barnperspektiv i barnavårdsutredningar – med barns hälsa och barns upplevelser i fokus**

Elin Hultman

ISBN 978-91-7519-457-8, 2013

57. **Internet Interventions for Hearing Loss Examining rehabilitation Self-report measures and Internet use in hearing-aid users**

Elisabet Sundewall Thorén

ISBN 978-91-7519-423-3, 2014

58. **Exploring Cognitive Spare Capacity: Executive Processing of Degraded Speech**

Sushmit Mishra

ISBN 978-91-7519-386-1, 2014

59. **Supported employment i en svensk kontext – förutsättningar när personer med funktionsnedsättning når, får och behåller ett arbete**

Johanna Gustafsson

ISBN 978-91-7529-012-6, 2014

60. **Effects of Specific Cochlear Pathologies on the Auditory Functions: Modelling, Simulations and Clinical Implications**

Amin Saremi

ISBN 978-91-7519-365-6, 2014

61. **Children with profound intellectual and multiple disabilities and their participation in family activities**

Anna Karin Axelsson

ISBN 978-91-85835-48-5, 2014

62. **Lexical and Semantic Development in Children With Cochlear Implants**

Ulrika Löfkvist

ISBN 978-91-7549-546-0, 2014

63. **Rethinking sound. Computer-assisted reading intervention with a phonics approach for deaf and hard of hearing children using cochlear implants or hearing aids**

Cecilia Nakeva von Mentzer

ISBN 978-91-7519-270-3, 2014

64. **Assessing cognitive spare capacity as a measure of listening effort using the Auditory Inference Span Test**

Niklas Rönnberg

ISBN 978-91-7519-267-3, 2014

65. **Employees with Aided Hearing Impairment: An Interdisciplinary Perspective**

Håkan Hua

ISBN 978-91-7519-240-6, 2014

66. **Prosthetic and Orthotic Services in Developing Countries**

Lina Magnusson

ISBN 978-91-85835-55-3, 2014

67. **Dealing with digits - Arithmetic, memory and phonology in deaf signers**

Josefine Andin

ISBN: 978-91-7519-235-2, 2014

68. **Time is of the essence in speech recognition: Get it fast or think about it**

Shahram Moradi

ISBN: 978-91-7519-188-1, 2014

69. **Effects of hearing loss on traffic safety and mobility** Aspekter på lärande vid dövblindhet

Birgitta Thorslund

ISBN: 978-91-7519-178-2, 2014

70. **Aspekter på lärande vid dövblindhet - möjligheter och begränsningar för personer med Alström syndrom**

Berit Rönnåsen

Licentiate Degree. ISBN: 978-91-7529-068-3, 2015

71. **Memory and communication in typically developing infants and children with Autism Spectrum Disorder: Behavioral and electrophysiological indices**

Emelie Nordqvist

ISBN: 978-91-7519-078-5, 2015

72. **Cognitive capacities and composite cognitive skills in individuals with Usher syndrome type 1 and 2**

Cecilia Henricson

ISBN: 978-91-7685-999-5, 2015

73. **Functioning and Disability in Adults with Hearing Loss Preparatory studies in the ICF Core Sets for Hearing Loss project**

Sarah Granberg

ISBN: 978-91-7529-086-7, 2015

Papers

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-121034